

Towards Performative and Responsive Evaluation (*Open Evaluation 2016 Conference,* *Vienna, November 2016*)



Prof. Miloš Milenković, PhD
University of Belgrade, Faculty of Philosophy
Department for Ethnology and Anthropology

Conference "[Open Evaluation 2016](#)", recently held in Vienna, organized by three of the leading European institutions that are dedicated solely to evaluation and policies of research, technology and innovation - the Manchester Institute of Innovation Research (MIOIR), the Austrian Platform for Research and Technology Policy Evaluation (FTEVAL) and the Institute for Research and Innovation in Society (IFRIS) examined the current trends and recent changes in international standards of evaluation of science, technology and innovation. The event was highly relevant for the scientific sector in Serbia whose evaluation standards lag behind best comparative practices and need urgent reshaping.

I had the opportunity to participate at the conference thanks to the support of PERFORM - Performing and Responsive Social Sciences, a project of Swiss Agency for Development and Cooperation implemented by HELVETAS Swiss Intercooperation and University of Fribourg, which has been implemented in the Republic of Serbia since 2015.¹

Evaluation as a Profession is in Deep Crisis

What makes this series of conferences specific is the fact that from the outset of this century it brings together administrators, evaluators and researchers. In accordance with its central topic, this year the organizers endeavoured to get many stakeholders involved in its work, as well as several hundred lecturers, panellists and discussants, coming from 34 countries and from all of the continents. As already stated in the keynote speech² and at the plenary lecture³, evaluation as a profession is in deep crisis, characterized by: the necessity for making a shift towards social functions and extra-academic purpose of the evaluation of research, technology and innovation; profound mutual incomprehension with the academia and the economy alike; and the resulting current ethical crisis of the profession.

Thus, Jakob Edler, as the keynote speaker, argued for the "opening of the evaluator's mind", whereas Steven Hill, as a plenary lecturer, explicitly placed emphasis on the search for social legitimacy of the profession of the evaluator in times when it has become evident that, impact as a measure of the quality of science and research has been unjustifiably reduced to influence within the academia (in the first place, measured by the number of citations of an author or a paper in journals), while social impact of the results of scientific research has remained almost completely ignored. Instead, indicators of the social impact, in the first place evaluation criteria for the selection and appraisal of scientific projects defined by participants at selection panels, who include stakeholders not coming from scientific and higher education institutions, are taking the place of well-established standards of influence-as-citation, taken from natural sciences to be applied on all sciences. Among many topics dealt with in the keynote speeches and the plenary lecture, special emphasis was given to the problem of discrepancy in the way results of scientific research are kept

¹ On this occasion, I also wish to express my gratitude to the organizers of the conference and the PERFORM program for their outstanding professionalism, excellent reception and the opportunity I had, as an anthropologist of science and education, to learn, through ethnographic work, about practices, values and goals of a relatively new profession, which is the profession of evaluators. Detailed program of the conference is available at: http://fteval.at/upload/OE2016_booklet_web.pdf.

² Katharina Warta, FTEVAL - Österreichischen Plattform für Forschungs- und Technologiepolitikevaluierung and Jakob Edler, MIOIR - Manchester Institute of Innovation Research.

³ Steven Hill, HEFCE – Higher Education Funding Council of England.

record of, in as well as outside Europe. However, there is a prevailing awareness that records kept on the subject matters of evaluation and the results of evaluation itself must remain qualitative, rather than reduced to quantitative indicators usually subject to misuse (which is the main reason why the profession of the evaluator has fallen into disrepute, in particular among SSH scholars).



Systematic Evaluation – Yes or No?

The Conference was divided into panels, of which most relevant to the problems that concern debates over Serbian scientific policy were: "Trends and Challenges on Systematic Impact Evaluation in Science and Innovation Funding Agencies" and "Evaluating Impacts of Social Sciences and Humanities".

On the first panel, dedicated to the problems of systematic evaluation, the very notion of a scientific system was questioned in a serious, well-documented and very critical manner. Long-prevailing ideology according to which the evaluation of scientific research should be performed in unison and consistently, according to the criteria that should be applied identically to all scientific fields forming "a scientific system", is being abandoned throughout Europe, slowly but steadily, even by top professionals in the field of evaluation. Hence, Erik Arnold⁴ presented results according to which agencies responsible for funding cannot be deemed agents who are trustworthy in their evaluation of their own practices, whereas Wolfgang Polt⁵ pointed out that the research of evaluation practices leads to the conclusion that the agencies have too much decision-making power. Following a vibrant discussion, the first panel ended with a conclusion that a majority of participants at the conference were not absolutely confident, even though they were inclined to think so, that the agencies or similar bodies that decide on research funding should not be simultaneously entrusted with the task of evaluation of the results (of such research). Thus, Sergio Salles-Filho⁶ pointed out that the crucial ongoing tension is the one between administrators who advocate not only for external evaluation, but also for external definition of the criteria for its implementation. As he points out, evaluators who are in their practice aware that if scientific administration would be fully open for the extra-scientific criteria for evaluation, the evaluation itself would be permanently alienated from the academic community (where, in general, it doesn't enjoy a good status).

During discussions, attention was paid to the comparative analysis of the policies and the practices of evaluation outside Europe, especially in Japan, Russia and the United States. Particularly important feature for our academic community, that many of the participants stressed, is that the systematic evaluation is being abandoned in the USA. That is contrary to the presently established policies and the dominant global trend. In the USA, significant reduction is currently made in the

⁴ Erik Arnold, Technopolis Group.

⁵ Wolfgang Polt, Joanneum Research.

⁶ Sergio Salles-Filho, University of Campinas, São Paulo.

funding for the evaluation of scientific research in favor of the funding for the research itself, as reported by Nicholas Vonortas.⁷ This led to further discussion, also relevant to the current state of the Serbian science and higher education policy, on the situation in the UK (which is traditionally taken as a standard). In the very cradle of the supervisory style of evaluation, there are voices opposing allocation of enormous funds for the evaluation itself, while researchers themselves have been denied such funds for years. Naturally, as is the case in academic circles for years now, this discussion wasn't spared from the remark that the concepts and methods of evaluation are biased and bent so as to favor natural sciences and technology, and that these scientific fields set out and defined the indicators by which all other fields are evaluated.

Additional arguments in favor of the opening of science to the society were given by Shinano Hayashi,⁸ who pointed out that even in a technologically highly advanced society such as Japanese the public has extremely unfavorable perception of science and technology, particularly in situations when scientists are unable to predict huge social changes or natural disasters (such as Fukushima). "The people do not trust the scientists" is a thesis valid in both highly developed and developing economies. In Serbian case, this is a particularly challenging issue, which should be urgently thought over on national academic and science policy forums, before the decision-making power about the future of science is, pursuant to the paradigm of open evaluation, allocated to non-scientists. It is my opinion that this is the way our Centre for the Promotion of Science should think about its future, too.



Institutional vs. Project Funding

Audience was especially impressed by the lecture given by Jonkers,⁹ who presented a comparative analysis of the (non)existing performance-based financing (in our country, it is considered to be the most important, if not the only "global" option there is). In most countries, including the most developed ones, research projects are not the dominant option, and institutional funding serves to protect their own scientific infrastructure. This is essential point in which this conference is relevant to further examination of the situation of the Serbian science and higher education policies, as we were systematically duped that institutional funding is a "thing of the past" and that it is quite "natural" for it to be replaced by project financing.

⁷ Nicholas Vonortas, George Washington University, Washington D.C.

⁸ Shinano Hayashi, Japan Science and Technology Agency, JST.

⁹ Koen Jonkers, Joint Research Centre, European Commission.

Ethical Regulation of Management and Evaluation in the Area of Science

On the second day of the conference, the keynote speech "Accelerating Science, Understanding its Impact: The Promise of Open Science" was given by Liz Allen,¹⁰ who reminded us that the openness of the evaluation as a profession is, by definition, sacrificed on the altar of efficiency. Yet, is openness valuable in itself? Does competition inevitably contribute to quality? How is the evaluation of scientific research to be opened and democratized without being surrendered to incompetent "general public"? How to prevent huge losses of research results in the course of publication process under the dictate of extra-academic planning and financial cuts, which in turn results with the most of the research being never published? How to coordinate the communication of scientific results with the transformation of the audience, especially their habits (bearing in mind that most of the readers do not read scientific journals, but rather require that such results are presented through the media that resemble, or actually are, social networks)? How to prevent huge losses not only in research results, but also in time and energy spent by the researchers on applications and reports, instead on the research itself? How to protect the very notion of evaluation and the profession of the evaluator from the anger of the community of scientific researchers, who mostly see the evaluation as, by definition, immoral, totalitarian and wasteful activity? Could it be that the evaluators themselves are to blame for the fact that the majority of scientists considers them enemies, responsible for the reduction in funding for research and shutting down entire disciplines, Allen asks. The answer lies in the call for ethical regulation of management and evaluation in the area of science and higher education. And that is precisely what our science policy is in need of.



Evaluation of the Impact of Social Sciences and Humanities

"Evaluation of the impact of social sciences and humanities" was a panel of crucial importance for recognizing the relevance of the work of PERFORM itself, as well as the lessons that domestic administrators and evaluators of scientific research can draw from this conference. Thus, at the outset, Stampfer¹¹ highlighted the apparent paradox - although most evaluators have their educational background in SSH, it is precisely this academic field that puts up most organized resistance and the highest number of objections to evaluation as a profession, its concepts and methods. Throughout the globe a strong feeling is shared that the methodology of the evaluation of scientific research, by definition, does not suit SSH, and that is the biggest challenge for a reform

¹⁰ Liz Allen, F1000, King's College London.

¹¹ Michael Stampfer, Vienna Science and Technology Fund, WWTF.

of the evaluation as a profession by means of openness paradigm. Thus, Reale¹² pointed out that, despite significant differences between scientific systems of France, Spain and Germany, as countries with three different types of political system and three different traditions of state-society relationship, the stance that SSH scholars take on the evaluation in these countries is very similar.

Within the same section, Robinson-Garcia¹³ stressed the necessity of regionalization and localization of indicators for the evaluation of scientific research (which goes far beyond the specifics of the SSH). He questioned, and this is particularly valuable, the idea that there is one state and one society, bringing back into the discussion what most SSH scholars, unlike most evaluators, believe to be a truism – dissimilarity of interests and values of diverse social agents, which we must take into consideration when selecting future agents of open evaluation. Therefore, social sciences and humanities are not only addressing different communities, given that they are largely non-cosmopolitan and oftentimes nationally, regionally and globally oriented, they also produce outcomes not susceptible to unification, which cannot be documented, let alone measured in unison. These and similar observations sparked an interesting and at times amusing discussion about insusceptibility of indicators for SSH evaluation to any kind of generalization, while highlighting the complex and multidimensional nature of engagement of these disciplines, which goes far beyond their visible (and measurable) impact, as in, for instance, altmetrics (which took the form of bibliometrics, transferring it from scientific journals to social networks).

Spaapen and Ochsner¹⁴ submitted a report on the broadest pan-European initiative undertaken so far to examine the prevailing SSH evaluation criteria and to recommend legitimate and viable alternative to scientometrics, which I personally also take part in (COST ENRESSH), which exemplifies modernity and relevance, as opposed to archaism and insignificance, in the long struggle waged by Serbian SSH scholars against attempts to shut down the field of our scientific and educational expertise by devaluing our work, underestimating both our institutions and us as individuals, even denigrating the very purpose of our existence in the society. The report presented preliminary results of a comparative analysis, and pointed to trends that domestic administrators of the sector of science and higher education would be unwise to ignore. Particularly important among these trends is *de-internationalization* of the evaluation of SSH.

Science With and For the Society

As a particularly important conclusion, I emphasize that this conference showed that, at present, it is the evaluators themselves who have to justify their role, their purpose and the costs of their profession to the society. The times when we the researchers had to justify that to them are now dark ages that we have left behind. However, instead of rejoicing over the plight in which the evaluators have found themselves, faced with the requirement to justify their existence, as they once paternalistically required scientists to do, I plea to embrace them, to reach common ground and mutual understanding with them, and to demonstrate together a sense of partnership perspective on science, which must not be reduced to anyone's particular goals and interests in the future. This is precisely how an initiative like PERFORM can be helpful to us, given its commitment to dialogue instead of conflict, inclusion instead of rivalry, and in particular promotion of a sense - mostly lost in the transitional fog – provided by SSH as a contribution to prosperity and preservation of civilized standards in every society, including ours.

Therefore, I urge all colleagues who are dealing with history, politics and ethics of SSH, organizational and administrative aspects of scientific research, promotion of science and its connection with the process of governmental decision-making and the preservation of national identity, and especially those who still find this to be "irrelevant," "shameful" or "unscientific", to pay attention to the fact that the notion of open science, and the resulting notion of open evaluation, is actually beneficial to social sciences and humanities, and by such means, to the society and culture. Our scientific disciplines, as an inalienable part of our public good and our cultural heritage, will be preserved thanks to the very notion of open evaluation, and by fostering of the dialogue between scientists, scientific administrators and stakeholders.

¹² Emanuela Reale, Research Institute on Sustainable Economic Growth, NRC.

¹³ Nicolas Robinson-Garcia, Universidad Politécnica de Valencia.

¹⁴ Jack Spaapen, Royal Netherlands Academy of Arts and Sciences; Michael Ochsner, ETH Zurich.

The general message sent from this conference dedicated to the opening of the evaluation of scientific research is that the shift from the "science and society" paradigm to the "science in society" paradigm is insufficient, because what we really need is the "science with and for the society." These gaps, existent not only in domestic scientific policies, but also in the spheres of management in science throughout Europe and beyond, are filled in by the PERFORM platform, which insists on the dialogue between the interested parties both within and outside the scientific institutions. Dialogue and openness with regard to evaluation and funding, as opposed to imposition, even disregard for the results of the work of scientists and scientific institutions, are the values promoted by PERFORM, which endeavors to follow current changes in trends in the evaluation of scientific research and scientific policy in general, as well as to anticipate on its own what are the common needs and shared interests of the scientific community and the society at large. This serves as an example that the highest standards of Swiss democracy can be implemented into domestic process of decision-making on major social issues.

