European Commission

MLE on Performance-based Research Funding Systems (PRFS)

Third Stream Metrics in PRFS

Thematic Report No 4

July 2017

Research and Innovation

Third Stream Metrics in PRFS – MLE on Performance-based Research Funding Systems

European Commission
Directorate-General for Research and Innovation
Directorate A – Policy Development and Coordination
Unit A4 – Analysis and monitoring of national research and innovation policies
Contact Marta Truco Calbet
E-mail marta.truco-calbet@*ec.europa.eu*
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in July 2017.

This document has been prepared for the European Commission however it reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

© European Union, 2017.
Reuse is authorised provided the source is acknowledged. The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

EUROPEAN COMMISSION

MLE on Performance-based Research Funding Systels (PRFS)

Third Stream Metrics in PRFS

Thematic Report No 4

Jack Spaapen

Directorate-General for Research and Innovation

2017 EN

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

This paper is part of a sequence of papers written in the context of the Mutual Learning Exercise (MLE) in 2016-17 on Performance-Based Research Funding Systems (PRFS) for institutional funding. The focus of this paper is on the use of or demand for **third stream metrics** in relation to such funding systems. The intention of the challenge paper is to support discussion and mutual learning among the participating countries on different aspects, issues and questions related with third stream metrics and their use in evaluation systems.

Currently, only a few countries in the world have experience with third stream metrics as a routine element in their evaluation systems (ie the UK, the Netherlands). The experience of the countries participating in this MLE with third stream metrics is limited (see Table 1, below). Nevertheless, there is a rising interest and demand for knowledge and practical guidelines regarding third stream metrics. The reason is that more and more governments require their research institutions not only to deliver results that are regarded as high quality by the academic community, but also results that bear relevance for the major socio-cultural and economic challenges in society. This becomes visible in the changes in funding instruments that stimulate collaboration between academics and societal partners, where evidence of the work's relevance for societal problems is required.

The response of academic institutions in most European countries so far has been not only a changing rhetoric in their policy papers and mission statements, but also initial steps have been taken to include third stream metrics in regular evaluation procedures (LERU 2017). These new developments in evaluation range from single indicators – for example Croatia uses a third-stream metric to identify research institutions´ popularisation efforts through publications for a broader audience – to wider experimental approaches (Italy, Norway for social sciences), to system approaches (the UK, the Netherlands).

A questionnaire has been sent out to all countries participating in this MLE concerning the changes in policy demand specific use of and experiences with or demand for third stream metrics serves as an important source of information for the preparation of this paper. See in particular Chapter 5 for the results of this survey.

2 WHAT ARE THIRD STREAM METRICS?

Evidence of the contribution of academic research to questions related to societal challenges can be delivered in various ways, some quantitative, other qualitative. Taken together we refer to the methods used in this area of research evaluation as 'third stream metrics'. Metrics, thus, may refer to both quantitative and qualitative instruments.

Having said that, the development of robust third stream metrics is not well advanced in most countries, yet "difficult but necessary" as the editor of the British Medical Journal, Richard Smith wrote back in 2001. Smith's article was triggered by an experiment of the Dutch Academy of Arts and Sciences to develop methods to evaluate the societal impact of applied health research. He succinctly summed up the problem with measuring impact:

> *"Much research that scientists judge of high quality has no measurable impact on health—often because the lag between the research and any impact may be decades. Thus, scientists would think of the original work on apoptosis (programmed cell death) as high quality, but 30 years after it was discovered there has been no measurable impact on health. In contrast, research that is unlikely to be judged as high quality by scientists—say, on the cost effectiveness of different incontinence pads—may have immediate and important social benefits."* (Smith 2001)

Smith goes on to say that because most funding systems concentrate on the quality of research and not on relevance, this leads to a serious imbalance in the research portfolio. Smith wrote this article more than 15 years ago, when societal impact measurement was in its infancy. But what we can learn from this is that there are at least two kinds of problems with measuring societal impact, and thus with third stream metrics. One is that scientists and funding systems are (were?) primarily oriented towards original work of high quality; the second is that there is an issue of temporality. Between original research and applications, the time span easily can be one or two decades.

Between the beginning of this century and now, discussions about the measurement of societal impact have branched out into a wide variety of topics and came to include not only applied research, but all academic research. Moreover, and closely tied to the emergence of the 'knowledge society', both internal scientific and external socio-political developments have caused a change in the way science is organised and funded: more room for larger endeavours that bring together researchers and experts with a non-academic background, and these endeavours are often of an inter- and transdisciplinary nature, which makes it more difficult to assess via a singular approach or method (Thompson Klein 1990, 2010). Accordingly, evaluation and instruments of measurement have changed too, notwithstanding the fact that there is large uncertainty how to do this.

In the discussion about what societal impact is and how it can be measured, many concepts float around. Bornmann (2013) mentions the following: third stream activities, societal benefits, societal quality, usefulness, public values, knowledge transfer, and societal relevance. While these concepts by and large may refer to the same phenomenon, what makes it difficult to measure them is that they are – as Bornmann notices - concerned with "the assessment of social, cultural, environmental, and economic returns (impact and effects) from results (research output) or products (research outcome) of publicly funded research". To be short, societal impact may refer to a whole area of social sectors and a wide variety of challenges that need input from different knowledge areas and other expertise.

The problem with finding robust measurement instruments that help assessing the effects of research on this wider societal context is rather different from measuring scientific quality alone. While scientific quality is judged in a relative limited social context of the international scientific community where more or less the same norms and values are shared and instruments are used mainly based on the output in a more or less stable set of journals, societal impact is depending on a wide variety of contexts in which different stakeholders operate, each with their own expertise, goals, values and expectations. And these all may change over time too!

Gibbons, Nowotny and colleagues (1994 and 2001) who made a distinction between traditional academic research – mode 1 - and research in the context of application - mode 2 -, noticed this and referred to the first kind of knowledge production as discipline-oriented and set in a homogeneous context, while the second kind of knowledge production takes place in a heterogeneous and socially distributed context. They see it as inevitable that novel quality control mechanisms need to be developed in the latter context. They refer to the kind of knowledge developed in and with the wider context as socially robust knowledge.

Nowotny explains what the authors mean with that concept of socially robust knowledge: it must link up with diverse practices, institutions and actors, and it addresses audiences that are never solely composed of fellow-experts, whose expectations and modes of understanding reflect the heterogeneous experience of mixed audiences (Nowotny 2003). She also warns against the tensions that go with what she calls the "democratisation of expertise", especially on the institutional level where people's conception of knowledge and knowledge production are not always ready for the new context of 'pluralistic expertise' in which academics operate these days.

Most individual researchers will work in mode 1 or 2 alternatively, sometimes writing an article for a scientific journal, other times collaborating with other individuals inside or outside academia. When it comes to assessing societal impact of individuals this leads to a problem of attribution because in collaborative endeavours results (impact) will always be the product of many different contributions. An individual researcher working in 'mode 1' type of research can be assessed on his or her impact in the scientific community (eg via citation analysis), but if the researcher works in 'mode 2' and is for example part of a team trying to find sustainable energy solutions, it often will be hard to trace impact back to that individual.

This kind of impact is best assessed by looking at the effects of the collaboration in the relevant societal context, such as the changing attitude of users, or different ways of producing energy. To be short, the evaluation of societal impact lends itself more for a higher aggregation level than the individual researcher.

3 CHANGING CONTEXT OF ACADEMIC RESEARCH

The demands coming from national and supranational governments for academic research to be more relevant for society and the economy clearly relate to the emergence in recent decades of the 'knowledge society', the paradigm that connects academic knowledge development with an innovative and sustainable economy and a high quality of life (Afgan and Carvalho, 2010). But this demand has a much longer history. We could go back to the 19th century and the land grant universities in America, or to the second half of the 20th century when governments in many European countries introduced their first science policy documents, to indicate the long standing public need for more practical proceeds from higher education institutes.

In the context of this report it suffices to say that the current demand shows at all levels of research and higher education policy. Whether we look at the European level to the Horizon 2020 framework programme, the national level in most European countries and the institutional level, most if not all the policy papers that address research policy contains mission statements that not only mention high quality research as a prime goal but also societal relevance (LERU 2017).

3.1 Big demands from society: grand challenges and more

The Horizon 2020 Framework programme – €80bn for seven years with as main slogan: taking great ideas from the lab to the market - is perhaps the best example of the changing policy context of academic research.[1] It is divided into three large sections, each more or less of the same size: (i) excellent science, (ii) industrial leadership and (iii) societal challenges. While the first one targets unfettered – 'blue sky' – research, the other two parts clearly are primarily directed to the economy and society. They both strive towards the improvement of the connections between academic research and the wider society. The industrial leadership programme focuses on enabling technologies in sectors like space, ICT and nanotechnology. The goal is to have businesses set the agenda for investments in research and innovation and to fund key enabling technologies that help create growth and jobs. The societal challenge programme addresses a number of broad societal areas and problems like health, food, energy, transport, climate and inclusive and secure societies. The idea of the societal challenges is to bring together resources and knowledge across different fields, technologies and disciplines, including social sciences and the humanities, with a new focus on innovation-related activities, such as piloting, demonstration, test-beds, and support for public procurement and market uptake.

Targeted programmes to connect academic research with industry or society at large were set up in many countries. The Dutch top sector programme[2], launched in 2012 with the slogan knowledge-technology-cash, may serve as an example on the national level. In this programme, the Dutch government selected 9 different economic sectors deemed of vital importance for the economy such as high tech, agriculture, health and logistics. The quintessence of this programme was twofold: public and private money should each provide half of the initial investment of three billion euro, and the National research council (NWO) and the universities did not get extra funding but were forced to reallocate existing funding to this top sector programme. The latter obviously mentioned a shift from academic research mode 1 type to mode 2 research. Academic researchers were to collaborate with partners in society and somehow attune their academic goals to those of societal stakeholders. Consequently, the way these new endeavours were going to be evaluated needed to change, too.

3.2 Response from the academic sector: cross-roads

One of the consequences of this growing demand coming from policy and society is that gradually scientific research and researchers are becoming part of larger entities, often inter-

---

[1] And this development most likely will become even stronger in the next framework program where according to commissioner Moedas impact will be one of the three main pillars.

[2] A variation on this slogan can be found in a recent advisory report to the European Commission which is called LAB-FAB-APP, referring to the same process from academic laboratory to applications in the market (High Level Group 2017)

or transdisciplinary by nature, because the challenges that need to be addressed are wide ranging (global warming, healthy aging, secure societies) and need, next to scientific input contributions from other experts and public organizations. This means that researchers must learn to collaborate with other researchers from different disciplines and or non-academic researchers working in industry or public organisations, and consequently that different goals, norms and values should be attuned. This affects evaluation procedures that come to include the societal relevance of academic research.

At the same time, however, the pressure on academics to publish in high impact journals has grown immensely compared to some decades ago when the phrase 'publish or perish' started to gain momentum, and with it, the importance of impact factors and h-indexes in academic evaluation culture. For academics and their institutions, this means that they have to operate at the crossroads of international academic competition (fed by international rankings such as the Shanghai or the Times higher Education ranking, and for individuals also by the risk of losing their job) and the demands coming from policy and society. More and more, evaluation systems seek to mediate between these two pressure systems, like the UK REF or the Dutch SEP (see Section 4.3).

Thus, it is the ultimate challenge for universities and faculties to find new ways to monitor and evaluate the work of their researchers in the context of global problems, local demand, and international competition between academic institutions. A new balance needs to found between the need for high quality research and the demand to produce research that addresses urgent societal problems. In a recent LERU paper, the collected universities of that organisation announced that finding that new balance is one of their main goals for the near future (LERU 2017).

4 THIRD STREAM METRICS IN PRFS AND EVALUATION

As noted above, not many countries so far have included third-stream metrics in their PRFS systems. There are several reasons why this is difficult to realise. The first is how to do this for so many different disciplines that work in so many different societal contexts. Clearly, working in health research, for example, with a context of academic hospitals, pharmaceutical industry and patient groups is a different ball game compared to working in demography and migration studies where the context is made up by politicians, policy makers, human rights activists, refugees, NGO's and the public at large.

The second question is what kind of metrics to use and how to weigh them against metrics that are used to assess the scientific quality of research. A major problem is that in the latter area, instruments to measure quality are overall well established, both in terms of peer review and quantitative methods, while this is not the case when it comes to measure societal impact, though there are promising developments to spot (4.3).[3]

The third issue is what kind of consequences the outcome of such evaluation should have. While evaluations regarding the scientific merits of work directly or indirectly lead to reward or punishment (involving money, people, equipment, etc.), for evaluations of societal impact other considerations are at stake because the success or failure of these broad projects depends on many factors outside the academic realm. And, moreover, they often can be measured only in the long term.

4.1 Consequences for societal impact evaluation

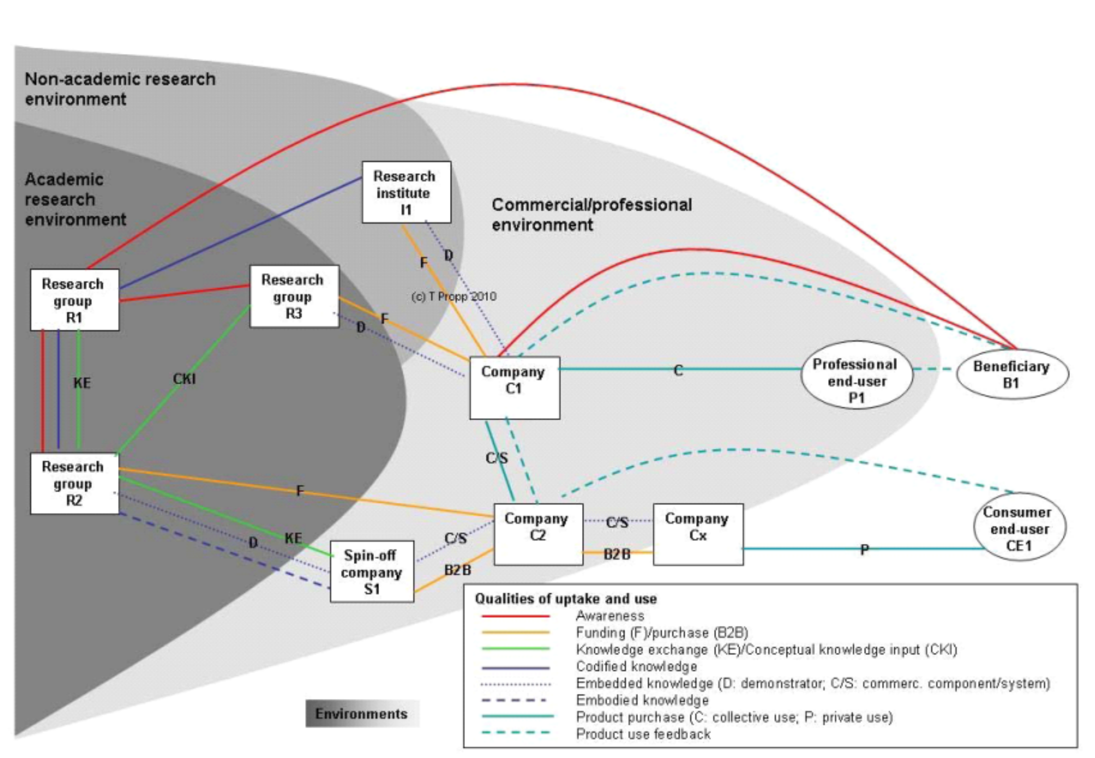
We briefly clarify the above three questions before giving some practical examples of integrating societal impact / third stream metrics in evaluation systems.

How to deal with the contextual character of impact evaluation?

The fact that societal impact is dependent on the context in which researchers operate, and that this context differs per field and per topic, and furthermore varies over time, make it necessary to approach this kind of evaluation with a lot of flexibility. But that does not mean that no systematic methodology can be followed. Unlike the assessment of scientific quality, however, there is no one- size-fits-all method. What is possible, nonetheless, is an approach that is based on an overall framework in which the relations with the societal contexts and the exchanges that take place in that context, form the basis of the evaluative exercise. For each of these relations, quantitative or qualitative metrics can be devised, as long as they are representative for the area at hand. To give an example coming from the field of nano studies. The network around a number of academic groups in the Netherlands and France was reviewed through interviews with key players from both the academic and the societal side. The main questions were what kind of relations were important for participants in this network and what kind of exchanges took place between the different stakeholders. The resulting graph is given in Figure 1, below.

[3] Notice however that there is also a growing critical stream with regard to the use of metrics to assess the scientific quality of research: DORA2013, Wilsdon 2015: 82)

Figure 1: Societal context of nano research (SIAMPI 2011) © Tilo Propp



As becomes clear from this graphical representation, societal impact is generated via a variation of interactions between variegated actors in a network collaborating in more or less narrow relationships. The network consists of academic and non-academic stakeholders, coming from public and private backgrounds, working on a common cause, but from different perspectives. To identify these interactions as related to societal impact, one needs to understand the knowledge exchange within such a network. For example, one of the research groups produces in joint projects with firm X demonstrators for analogue-digital converter (ADC) integrated circuits. Firm X then, after up scaling, sells these as components to their customers. The one academic research group rarely interacts with these customers. Other research groups may work closely with industry and their customers. It is very difficult to decide what relationship is more important or what kind of exchange is, and each contribution, no matter how small, can be of vital importance for the success of the network as a whole.

Having said that, not every result or product of scientific research is part of such a complex network. Sometimes a policy makers comes to a researcher or a department with a more or less clear question, for which the research options are rather straightforward. In those cases, the impact of report or other product demanded can be measured through simpler techniques.

How to use metrics?

Looking at the second question, regarding the kind of metrics and how to use them, it is important to distinguish between the two main functions of evaluation: summative and formative evaluations. The first one is backward looking and regards evaluations primarily conducted to account for the work done, the second type is more forward looking and aims at mutual learning (Scriven 1991; Scriven 1996). Summative evaluations are usually characterised by a unilateral relationship between a funding actor (a government, a research council, a university) and a research entity and focus on the assessment of output of scientific work. Formative evaluations generally are more process oriented and focus on exchanges that take place in research and innovation networks, and on the progress that is made in the actions, often via small steps forward that can be measured through so-called intermediate indicators. Summative evaluations generally use more quantitative metrics, formative more qualitative. Nevertheless, the distinction between the two is not always so sharp, in particular regarding the balance between looking backward and forward, and when it comes to using specific metrics, often a mix of methods are used.

What kind of consequences?

Regarding the third question, what the consequences of societal impact assessment can or should be, the first thing to remember is that mutual learning between the participants is arguably the most important goal of impact evaluation. The main question is not who is the best but what can we do better to improve the societal issue at hand, that is, the evaluation and indicators should be informative about the progress that is being made with the issue at hand. Clearly, sound and robust scientific research is most important input in such a process, and thus, consequences should be considered both in terms of scientific contributions and of societal demand. This means that a balance need to be sought between different areas of expertise and the different communication exchanges between these areas. To some extent, this parallels the assessment of interdisciplinary work, also something that is 'difficult but necessary', and where different production and communication modes and different interests should be attuned. At the same time, one also should bear in mind that whatever the contribution of scientific research is, the fate of any piece of research is not in the hand of the scientist, but depending on how this is received in the societal context, where particular interests sometimes do not coincide with scientific outcomes.

4.2 Developing new metrics for societal impact

When we take a closer look at methods being developed for societal impact evaluation, we can distinguish two main strands: (i) discrete measurements where impact is seen in terms of a linear process, in which messages are conveyed from a sender (the researcher) to a receiver (the user). Often, these methods tend to emulate what is done in bibliometrics, for example patent statistics or counting popular publications; (ii) systems approaches. Here the focus is on multi-lateral exchanges in research and innovation networks, using a mixture of quantitative and qualitative methods like narratives, case studies or some form of SWOT analysis.

There is not always a strict distinction between the two strands, in the sense that system approaches sometimes make use of linear methods, for example when they focus on short or medium term advancements in the network through intermediate indicators. And linear methods often also use system elements, for example feedback loops. Policy makers usually prefer the first strand type of indicators because they produce quantitative metrics, which arguably are ensure transparency, consistency, and comparability across disciplines (HEFCE 2013), which make it easier to make 'fair' policy decisions because one can see how individual researchers or groups perform in the competition with other researchers (Cozzens and Snoeck, 2010).

(i) Quantitative methods for societal impact

While in some fields quantitative methods such as patent or license statistics might be used to assess societal (economic) impact, in most fields they do not make sense for lack of relevant or meaningful data. In some cases, quantitative methods are used or being developed, especially where the focus is on the economic impact of research in a given sector, for example by counting the returns achieved for the research funds invested. There are not many examples looking at other aspects of societal impact, but note Croatia where popularisation publications are used in an indicator that directly (but weakly) influences funding in the PRFS.

Altmetrics

Altmetrics is the new kid on the block when it comes to finding out what the impact of research is. The gist of altmetrics is that there are all kind of traces in the digital world that may inform researchers about the uptake of their research. With altmetrics it is possible to track main stream media, public policy documents, blogs, social media, online reference managers (Mendelay f.e.), citations, etc. Through the combination of these methods it is possible to get information of both the uptake in the scientific world and in society at large. (Holmberg 2017; see also https://www.altmetric.com/)

(ii)Systems approaches

Most countries that have PRFS at a national level focus on three aspects, scientific quality, societal impact and some kind of management criterion. The ratio between these elements may differ considerably. In the Dutch SEP system, scientific quality and societal impact have equal weight; the same is in Estonia (although with different criteria). In Croatia, impact is weighted higher (30%) than quality (25%); in the UK, scientific output is much more important (65%) than impact (20%).

Below, we briefly explain two national systems where impact assessment is integrated, the Dutch SEP 2015-2021 and REF UK 2014, and then we give some examples of methodological developments in impact assessment.

SEP 2015-2021

In the Netherlands, all academic research, and most other publicly funded research, is evaluated through the Standard Evaluation Protocol (SEP), a system running since 2003. Every six years a new version is published, and the current one runs from 2015 till 2021. The SEP is self-organised by the universities and based on international peer review via site visits. Its prime goal is not (re)allocation of funds (although that might be an indirect effect), but accountability for quality and impact. There are three main criteria, scientific quality, societal impact and viability of the research unit. For both quality and impact, three types of indicator categories are mandatory: output, use and recognition (http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/SEP2015-2021.pdf, see especially p. 25 of this document)

Table 1, below, shows the six categories that form the basis of any SEP evaluation. Indicators (both quantitative and qualitative) can be used to underpin the different categories. Together, this information is used to write a so-called self-evaluation report that is offered to the review committee.

Table 1: Six categories for research production.

<table>
<tr><td></td><td></td><td colspan="2">Quality domains</td></tr>
<tr><td></td><td></td><td>Research quality</td><td>Relevance to society</td></tr>
<tr><td rowspan="3">Assessment dimensions</td><td>Demonstrable products</td><td>Research products for peers</td><td>Research products for societal target groups</td></tr>
<tr><td>Demonstrable use of products</td><td>Use of research products by peers</td><td>Use of research products by societal target groups</td></tr>
<tr><td>Demonstrable marks of recognition</td><td>Marks of recognition from peers</td><td>Marks of recognition by societal target groups</td></tr>
</table>

(based on table D1, p25 of SEP 2015-2021)

Productivity, which used to be a separate 4th criterion in the previous SEP, is left out of the current version as a reaction to the growing critique that too much focus on producing articles has perverse effects on both the quality and relevance of scientific research (see for example Science in Transition 2013). Review committees are asked to value scientific quality and societal impact on an equal basis. Much like the REF in the UK, research units can present case studies or narratives to underpin their societal impact. However, there is no mandatory structure for these cases, unlike the UK. The first results of evaluations according to the new SEP are available, but they have not yet been analysed. An interesting effect is that all the Dutch humanities faculties gathered into a joint project to find an adequate structure for narratives, and the evidence to support them. This humanities adaptation of the SEP includes information about indicators for impact assessment (https://www.qrih.nl).

REF UK 2014

The Research Excellence Framework (REF UK 2014) introduced impact case studies in which research units could demonstrate their impact on the economy and society at large. The studies are structured according to five chapters:

1. summary of research

2. underpinning of research

3. references to research

4. details of impact

5. sources to corroborate impact

These case studies counted for 20% in the assessment (as opposed to 65% for scientific research output and 15% for vitality). The impact case studies are short (max. 4 pages). They are assessed by two criteria: *Reach* – "the spread or breadth of influence or effect on the relevant constituencies"; and *Significance* – "the intensity or the influence or effect". £1.6 billion of funding over the next five years will be determined by impact case studies. The results of the assessment of case studies were analysed by researchers from King's College London and they found that much of the research was multi- and interdisciplinary, that the societal effects were diverse and wide-ranging, with over 60 unique 'impact topics' identified, and more than 3,700 unique pathways leading from research to impact (King's College: 2015).

4.3 Examples of third stream metrics

So, while the development of third stream metrics is still in its infancy, as is also shown in the limited feedback we received from the participating countries regarding the question about methods in use, there are indications of a growing conformity and concurrence in the emerging field of third stream metrics. One example of growing cooperation is the COST action ENRESSH which focuses on the development of indicators for both research quality and impact in the social sciences and humanities (http://enressh.eu/). Almost all European countries participate in this action, and there is also interest from outside the EU (South Africa, China, Mexico).

What follows are examples of (experimental) projects that offer a few different ways to think about societal impact measurement, some of which have been tested in practical situations.

SIAMPI

The first example has been developed in an FP7 project called SIAMPI, and its central concept is productive interactions between stakeholders. SIAMPI, which stands for Societal Impact Assessment Methods through Productive Interactions, developed an analytical framework for societal impact evaluation that arguably can be used for all research fields (Spaapen and Van Drooge 2011). It analyses the productive interactions in research and innovation networks. Productive interactions are the mechanisms through which research activities lead to a socially relevant application. An *interaction* entails a particular contact or exchange between a researcher and a stakeholder, mediated through various means, as diverse as a research publication, a policy report, a prototype, a guideline, a website, a design, a protocol, a membership of a committee, shared use of facilities or financial contributions by a stakeholder. Three main types of interaction are distinguished:

- direct or personal interaction
- indirect interaction through a medium
- financial or material exchanges

The interaction is *productive* when it leads to efforts by stakeholders to apply research results to societal goals, i.e. when it induces behavioural change.

For each of the three types of interaction a variety of indicators can be developed, some quantitative, others qualitative. In Table 2 some examples of indicator categories are presented, based on interviews with researchers in 4 different fields in 4 different countries.

Table 2: Examples of indicators for societal impact in three categories

Direct interactions via people	Indirect interactions via media	Material interactions
• face-to-face meetings • double functions, other mobility arrangements • phone/skype conferences • email • social media • videoconferencing • public debate • radio, tv, internet	• academic journals • professional journals • non-academic journals • popular media • exhibitions • artefacts, models • films • master theses, graduate projects • standards, protocols • social media	• research contracts • facility, instruments sharing • start ups • contribution "in kind" (people) • IPR arrangements, patents, licenses • Professional training • Other stakeholder interest

For some of these indicators it makes sense to develop quantitative indicators, especially in the media category. The main problem will be the availability of robust data. For others, mostly in the other two columns, it makes more sense to develop qualitative indicators, for example via narratives, or case studies. A current example of a project that is aiming at the development of more robust methods, is the Dutch humanities project (https://www.qrih.nl/nl/), A necessary pre-condition for this is the existence of reliable databases, both at the institutional and at the national level. That is what the project is working on at this moment. In another European project, based on a similar threefold division, an attempt was made to develop a composite indicator for societal impact. They distinguished three categories indicating institutional cooperation, training people and commercialisation (Finne et al. 2011: 27).

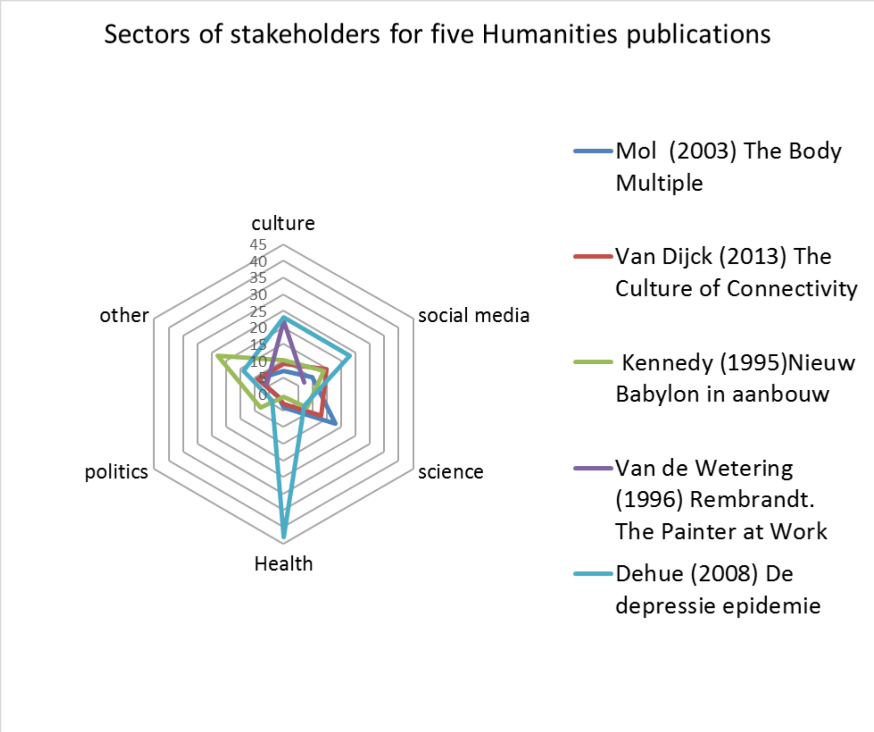
Contextual Response Analysis (CRA)

This method has been developed in particular for research institutes that work in a policy context, for example in health or environmental research. It is an example of altmetrics and focuses on so-called hybrid output of research that is directed towards both the scientific community and stakeholders in society. It combines the assessment of scientific and societal impact of research. The first kind of impact is reviewed in terms of citations found in google scholar via regular methods. For societal impact, CRA reviews the uptake of research in other media, social media included.

The CRA focuses on identifiable and unique traces of publications on the internet using google, or Bing or other search engines, and specialized databases. In that way, it identifies relevant users of research. Identifying users offers the opportunity to place the response into a context of use, such as characterized by the domain in which the response emerges (e.g. news media), or by the characterization of the user based on social function or social-economic sector (e.g. education, individuals (bloggers), for-profit services etc.). The CRA method focuses on the variety of use, comparing both the diversity of use of various products of one institute and the diversity among users among different institutes. Results are presented in a radar graph that shows both the variety of uptake or use, and level of interest. The example below (Figure 2) shows the review of five hybrid publications in social and human sciences in the Netherlands (Prins and Spaapen, 2017).

The main goal of this method is to inform institutional management and researchers about who in their context is interested in their output, and with what intensity. The level of interest is based on counting of downloads by the different societal sectors.

Figure 2: Contextual Response Analysis of five hybrid publications



Impact pathways

The best-known example of impact pathways are the impact case studies of the REF UK, which we introduced in Section 3.2, above. More information can be found on the website (http://www.hefce.ac.uk/rsrch/REFimpact/) Here we shortly present two lesser known examples of impact pathways as a method for the assessment of societal impact, one of which is in the process of implementation (INRA France), the other is in an intentional stage (Imperial College London)

INRA

The INRA (l'Institut National de la Recherche Agronomique) impact pathway policy is inspired by the ASIRPA project[4], which was launched by INRA's general directors in January 2011. The INRA method uses standardized case studies, making cross-case comparisons and broader conclusions possible. Each case study contributes unique information and is associated with different types of impacts. Some case studies involve research whose impacts are largely economic. For instance, in France, the economic value of genomic selection in dairy cattle is estimated to be 1–2 billion euros. Another example is the more judicious use of nitrogen fertilizers, which has resulted in savings of 4 billion euros over 25 years and has helped protect the environment. However, most cases have more than one type of impact. In the fight against scrapie in sheep, there were both political and health-related impacts; genetic tools prevented the need for massive culling when the BSE epidemic hit.

One of the findings of the ASIRPA project was – not surprisingly - that the pathways between academia and society are often long and seldom straight, but sometimes rather circuitous.

[4] The ASIRPA (*Analyse Socio-économique des Impacts de la Recherche Publique Agronomique*) project analyzed the impacts of publicly funded agricultural research. It was launched in 2011 and carried out by INRA scientists from two research units, the Sciences and Society Unit (SENS) and the Joint Research Laboratory for Applied Economics (GAEL) in Grenoble, as well as by collaborators at the French Institute for Research and Innovation in Society (IFRIS).

The leaders of the ASIRPA project, Pierre-Benoit Joly and Laurence Colinet, explain on the INRA website why they are developing the impact pathway methodology: "There are two general methods for evaluating the applied importance of research. The first focuses on the research's economic impacts in a given sector, by estimating the returns achieved for the research funds invested. This method is quite useful, and it shows that rates of return are usually very high. However, this technique remains myopic because it only focuses on the research's economic benefits. Furthermore, it cannot be used to uncover the mechanisms that are generating impacts. The second method uses case studies. It can be utilized to flesh out the details of the paths that lead to impacts (i.e. impact pathways). However, its disadvantage is that it relies on the analysis of a collection of different research 'stories' which can make it difficult to draw more generalized lessons. We have developed an approach that uses standardized case studies, making cross-case comparisons and broader conclusions possible" (INRA: 2014).[5]

Joly's co-leader, Laurence Colinet, explains further why impact pathways through case studies are a valuable approach for evaluating the societal impact of research, much better than quantitative methods: "Our research has allowed us to scientifically confirm some ideas that were already more or less accepted or that seemed intuitive. First, impact pathways for agricultural research are long: on average, 19 years elapse between the beginning of a project and the manifestation of its impacts. That is why we need to proceed with caution: asking for rapid returns is sure to be counterproductive. Indeed, economists realized early on that the government should provide research funding because it takes such a long time for research to yield impacts. Our results also underscore the importance of research infrastructure, such as experimental facilities, collections of genetic material, livestock, and databases, as well as partnership schemes. Most of the case studies we examined involved interdisciplinary collaborations".

The method was tested by some of INRA's research divisions and the feedback was so positive that they plan to roll it out during the institute's next five-year evaluation. They are convinced that this methodology can be applied in other research institutes as well.

Imperial College

Imperial College London is planning to introduce a systems approach to wider impact evaluation. The model they propose "encapsulates application of research and experimental education across the private sector, public sector, third sector (charities, foundations, trusts and NGOs) and broader community. This is a dynamic system in which exchange of ideas occurs through interactions and flows of people, knowledge and technology. We call these the pathways to societal impact" (Dave Gann a.o. 2016). The report distinguishes three 'pathways to societal impact' that resemble the trio of interaction channels of the SIAMPI model: (1) People: developing, educating and engaging talented people is the largest direct impact that the College has on society, perhaps followed by treating patients at our hospitals; including full-time and part-time students, permanent and temporary staff (professional services and academic), as well as internships, Adjunct Professorships, those in further education, alumni, partners, clients (e.g. of executive education), donors, advisers, and friends; (2) Knowledge: dominant through scientific publishing, albeit this may have less direct or immediate impact on society; includes pathways such as consulting and problem solving, data sharing, conferences, influencing policy, outreach, and defining new research domains and (3) Technology: the core mission of the College's TTO includes pathways such as patent filing, licensing, entrepreneurial start-ups and spin-outs, as well as less common routes of standards setting. The report aims at a comprehensive policy for "developing, encouraging, measuring and rewarding participation in these pathways.

User survey and impact case studies in Norwegian social sciences

---

[5] See http://www.inra.fr/en/Scientists-Students/Economics-and-social-sciences/All-the-news/Translating-research-into-impacts-30-case-studies.

Inspired by the REF impact case studies, Fridholm and colleagues conducted a user survey and a number of impact studies for the Norwegian Social Science institutes in 2016. The user survey consisted of web survey and telephone interviews.

The impact assessment was predominantly based on 71 impact cases submitted by the institutes at RCN's request. RCN asked for societal impact, defined as any impact except impact on other R&D and impact on the institute's own organisation. Following REF2014, the impact cases were classified in terms of topics, beneficiary types and geographical reach of the impact. For 15 impact cases, additional data were collected through document studies and interviews to be able to complement and elaborate the cases further, as well as to verify the information provided by the institutes. As a complement to the impact cases, the web survey was used to collect data on user impact from working with the institutes.

A main finding was that the R&D activities of an institute generally benefits a user in two main ways; through delivery of knowledge outputs (reports, datasets, software, etc.) and through enhancement of the skills of individuals of the user organisation.

Another important result of this study regards the occurrence of intermediate impact. A ministry or a government agency may present a new policy or policy instrument that in turn leads to societal impact when it affects actors in society. Similarly, a company may introduce a new product on the market that leads to societal impact when its customers use it. Other intermediate societal impact may be dissemination of material to inform or educate societal actors, or knowledge spillovers from the user organisation, such as mobility of staff, publications etc. Throughout this gradual development, the initial contribution of the institute is blended with input from a range of internal and external sources in a process that is also influenced by general societal and economic developments.

A conclusion here can be that it is important to pay attention to these intermediate impacts, also as a part-solution for the temporality problem.

5 COPING WITH SOCIETAL CHALLENGES

All societies nowadays are facing major challenges that need the input of scientific research, and many of these challenges cross the borders of the nation state. The challenge for the research community is how to address these challenges in the best possible way. Major funding programmes at the national and international level ask the research community to apply for support with research of the highest quality that also has the highest impact on the economy and society. If we look at the European Horizon 2020 framework programme, we see that about two third of the total of €80bn is reserved for society-oriented applications. This two third is divided almost equally between industrial leadership and key enabling technologies and the so-called Grand Societal Challenges (GSC): (1) Health, demographic change and wellbeing; (2) Food security, sustainable agriculture and forestry, marine and maritime and inland water research, and the Bioeconomy; (3) Secure, clean and efficient energy; (4) Smart, green and integrated transport; (5) Climate action, environment, resource efficiency and raw materials; (6) Europe in a changing world - inclusive, innovative and reflective societies; and (7) Secure societies - protecting freedom and security of Europe and its citizens.

In this chapter we focus on how countries participating in this MLE cope with their national societal challenges (overlapping sometimes with the European) in terms of their research and funding policies and in terms of their monitoring and evaluation systems. The idea of the GSC is to bring together knowledge institutions and public and private partners from society across different fields, technologies and disciplines, including social sciences and the humanities. It includes establishing links with the activities of the European Innovation Partnerships (EIP).

The following two sections are based on a questionnaire that was sent out to the countries participating in this MLE.

5.1 Addressing the challenges through research and funding policy

Most participating countries have set up national R&D programmes that target challenges relevant for the country and promote the collaboration between academic research, industry, and / or other stakeholders in society. Here we give a number of examples of how countries (governments, universities) respond to these challenges

The government of the **Czech Republic** adopted the "THÉTA" programme of R&D support for the energy sector, which aims at supporting public-private partnerships. The government also approved the "National Competence Centres" programme aiming at the development of long-term strategy public-private-partnerships. Furthermore, a number of key national challenges has been defined by the "National Smart Specialisation Strategy of the Czech Republic" focusing on the following topics: (1) Mechanical Engineering, Energy and Metallurgy; (2) Electronics, Electro-technics and ICT; (3) Transport Vehicles Development; (4) Medicine, Biotechnology and Life Sciences; (5) Culture and Creative Industries; and (6) Agriculture and Environment. The Czech Academy of Sciences has developed the "Strategy AV21" to address socioeconomic challenges. The Strategy AV21 is based on a number of problem-oriented research programmes bringing together academic and other research institutions with industry partners. The Strategy AV21 research programmes are aimed mainly at networking activities, which shall strengthen mutual interactions of public and private research sectors and help them to develop their joint research agendas.

To improve the connection between research and industry, **Austria** established in 2004 the Austrian Research Promotion Agency (FFG) as the national funding agency for industrial research and development in Austria. FFG has a range of funding instruments for programmes, people, facilities, services etc. Among them also new funding instruments such as "Innovation Labs" that support interaction with lead- or end-users and stakeholders to ease the path to solutions for challenges the society faces. Even in single firm projects, involvement of end-users is encouraged to facilitate innovation. Most of these instruments aim at fostering cooperation science - industry (cooperative RDI projects, and of course competence centres / centres of excellence and the other instruments in the field of "structures and networks"). FFG addresses a wide range of challenges and thematic fields like mobility, energy, communications, new technologies, ICT, Life sciences, nano research, safety, aeronautics. It also aims at positioning innovative enterprises and R&D organizations in transnational cooperations (ERA-NET, ERA-NET plus, JTIs, Art. 185 EGV)

The Federal Ministry of Science, Research and Economy took the initiative of Strategic Networking Platforms targeting mission-oriented research topics like Personalized Medicine, Smart Cities, Demographic Change, Sustainable Water Systems. The goal is to bundle and strengthen strategic collaboration between science and research institutions. Funding is also provided by the Ministry of Transport, Innovation and Technology. This policy initiative is part of the Austrian Process for a "Common Alignment Position" and constitutes one of the measures within the Austrian ERA-Roadmap Priority 2a – Jointly addressing grand challenges (https://era.gv.at/directory/236).[6]

At the European level, Austria took a leadership role within JPI Urban Europe as strategic cooperation between EU Member States for supporting Europe's cities in their transition towards a sustainable and resilient future (http://jpi-urbaneurope.eu/).

Estonia has several schemes running to promote societal impact of research. First, the institutional baseline funding promotes patenting activities in R&D institutions and cooperation with private and public organisations. Second, the Estonian Research Council (ETAg) supports the pursuit of socio-economic applied research guided by the needs of the Estonian state to increase the role of the state in the strategic steering of research and the capabilities of R&D institutions in carrying out socially relevant research. As part of the activity 'Strengthening of sectoral R&D (RITA)', the government aims to strengthen the capabilities of R&D institutes to conduct applied research needed by the state. Applied research supported by ETAg means both support for strategic, interdisciplinary R&D activities to tackle social problems. Support is also meant for knowledge-based policy formulation, which aims to tackle issues of national importance faced by the state that require speedy intervention.

Via NUTIKAS, applied research in smart specialization growth areas is stimulated. The support aims to contribute to growth in the research-intensity of the Estonian economy, supporting collaboration between R&D institutions and companies. Based on the analysis of the Development Fund, the RD&I Strategy 2014–2020 and the Entrepreneurial Growth Strategy 2014–2020 specify the growth areas where enterprises have higher-than-average potential for growth and opportunities to achieve a competitive advantage through investment in R&D.

Apart from state policy, the academic institutions position themselves regarding the global, economic and social challenges in their Strategic Development plans. And they have created a network (ADAPTER, https://adapter.ee/en/), providing a quick and reliable link for companies and organizations to the research and development community. And Tallinn University of Technology has created the TTU Innovation and Business Centre Mektory (https://www.ttu.ee/?id=31155), a networking environment where businesses, students, professors, schoolchildren and investors from all over the world meet to generate new ideas and innovation together as one team.

In **Norway**, the government presents every four year a White Paper on the long-term perspectives for the economy and society. The most recent paper identified as key challenges: Ageing population and the sustainability of the welfare state; Migration and integration of immigrants; Raising productivity growth – improved performance in the private and public sectors; and the "Green Shift" (reduce oil & gas dependency and transitioning to a greener economy). As long term priorities for the research and higher education sector the following six are mentioned: Seas and oceans; climate, environment and clean energy; public sector renewal, better and more effective welfare, health and care services; enabling technologies; innovative and adaptable industry; and world-leading academic groups. To bring together science and society, the "21-processes" were initiated: actor-driven processes

[6] As an example may serve the platform CCCA "Climate Change Center Austria" which successfully has been established as the national coordinating body for the promotion of climate research in Austria. The platform is a contact point for researchers, politicians, the media, and the public for all questions concerning climate research in Austria (https://www.ccca.ac.at/en/home/)

with participants from research, government, business and NGO's together elaborating strategies for R&D in specific areas, ranging from climate and energy to health and health care. Klima21 (Climate 21) is an example of these strategies (https://www.regjeringen.no/contentassets/e10e5d5e2198426788ae4f1ecbbbbc20/en-gb/pdfs/stm201420150007000engpdfs.pdf).

Individual universities address a variation of societal challenges, most often in inter- and transdisciplinary research projects

Regarding the grand challenges in **Sweden**, several policy instruments have been established. The recent research policy bill further increases academia's drive to address challenges. The Prime Minister's Innovation Council has pointed out five innovation partnership programmes aimed towards societal challenges and the Government's recent research and innovation bill also focuses on societal challenges, and on collaboration. The Swedish Innovation Agency, Vinnova, has launched a Challenge Driven Innovation (CDI) programme already before Horizon 2020, and Vinnova (together with also other research funding agencies) have set up a number of Strategic Innovation Programmes, aimed towards responding to challenges. These initiatives complement other R&I investments (i.e. strategic research areas) and create good synergies with the challenge based approach in H2020, in which SE participant are successful. The challenge-based programmes at Vinnova (CDI and SIP) have shown a changed collaborative pattern compared to traditional programmes, where the public sector participated to a much larger extent.

There is also a support office for small and medium sized companies to help them get involved in Horizon 2020 activities. The main challenges Sweden identified are (according to the Research bill): Climate, Health including life science; Digitalisation; A safe, inclusive and sustainable society; and School results. The research councils have specific long term programmes to support research in these areas.

Many Swedish universities were created to meet material societal demands, e.g. business schools or examination of teachers or health sector personnel. That means that cooperation with societal partners is a normal part of their activities. Successful examples can be found in academic cooperation with the automotive industry in the Gothenburg area or the Life science cooperation in the Stockholm – Uppsala area.

In **Italy**, the National Research Programme for the years 2015-2020 (available on line at http://www.istruzione.it/allegati/2016/PNR_2015-2020.pdf) of the Ministry for University and Education (MIUR), 12 research and innovation areas are identified as national priorities. The areas are the following: 1. Aerospace; 2. Agrifood 3. Cultural Heritage 4. Blue growth 5. Green chemistry 6. Design and Made in Italy 7. Energy 8. Smart industry 9. Sustainable mobility 10. Health 11. Smart, Secure and Inclusive Communities 12. Life environment technologies

To evaluate third stream activities in some of these areas, ANVUR (National Agency for the Evaluation of Universities and Research Institutes) introduced an experimental activity that is summarised in appendix A. Third stream activities have been evaluated separately from research performance and have not been counted in the funding formula. In fact, in the evaluation of these activities several issues need to be addressed, such as context and discipline influence, relation with the institutional strategy, measurement problems, quality of data used in metrics, expected and unexpected outcome of this specific kind of evaluation.

In **Turkey**, the societal and economic contributions of universities have been emphasized in the resolutions adopted at the 22nd meeting of Supreme Council for Science and Technology. Resolutions include advancement of the capacity of technology transfer offices and incubation centres, redesign of criteria for academic promotion, fostering a culture of entrepreneurship among students and researchers as well as a systematic measurement of the performance of universities in terms of entrepreneurship and innovativeness. In this context, the **Entrepreneurial and Innovative University Index** was developed to realize the aforementioned SCST resolutions in support of the realization of the 2023 targets of the national STI ecosystem. The index introduces a dedicated performance evaluation and monitoring system for tracing the activities of universities in favour of the private sector. More specifically, the aims of this index are to increase the entrepreneurship and innovation oriented competition between universities, to measure the performance of universities in entrepreneurship and innovation, and to contribute to the development of entrepreneurship and innovation.

Conclusions

From this short overview, we can draw some conclusions as to how countries deal with the challenges in terms of research and innovation policy measures. The main conclusions are:

For many countries, societal impact of research regards primarily the relationship between research and industry, and most funding instruments are developed to stimulate university – industry collaboration. Furthermore, many of the challenges identified by the different countries show overlap with the Grand Societal Challenges of the Horizon 2020 programme (energy, health, secure societies), but some are country specific (the green shift in Norway, design made in Italy). This not only means that there is ample ground for European collaboration, but also that countries should have a shared interest in finding new ways to evaluate the societal impact of research in these areas.

Second, in most countries, there is a shift in funding instruments and goals towards on the hand industry and on the other hand the societal challenges. Here also, countries could learn from each other to discover more or less effective ways of financing the challenges.

And third, it seems now a general trend that the strategic plans of the universities all mention societal impact / challenges as a main policy objective. It is interesting to see how the different schemes, programmes and policies work out in the near future in finding solutions for these challenges in collaboration with stakeholders in society.

5.2 Methodological developments in participating countries

When asked what kind of methodological developments can be distinguished in the countries with respect to third stream metrics, very few mentioned anything concrete, but many expressed the need for such instruments. However, a number of countries came up with interesting descriptions of provisional attempts to develop a more systematic approach to somehow evaluate the contributions of science to society. So, while concrete methodological developments hardly exist, it is still interesting to summarize what countries reported.

In the **Czech Republic,** there is no experience with third stream metrics. However, the description of the SHARE-CZ infrastructure may serve as an example of a societal impact narrative:

SHARE-CZ is the **Czech** national node of the pan-European distributed research infrastructure SHARE (*Survey of Health, Ageing and Retirement in Europe*) being a multidisciplinary and cross national longitudinal database of micro-data on health, socio-economic status, social and family networks and other issues collected from more than 85,000 individuals coming from 20 European countries aged 50+ and their partners. The result is a free and unique data collection that provides information about the state, history and the future of the Czech and European society. SHARE allows researchers and state administrations to understand the consequences of demographic changes and formulate optimal policies for public finances, labour market, health care or pension systems. Since 2004, SHARE is a longitudinal survey that is repeated every 2 years. The main goals are to create a main questionnaire for 6,000 respondents aged 50+ and their partners in each country, every 2 years collect data on the same individuals and store the collected data and its documentation in a user friendly, free and open access database accessible to all users. SHARE combines 3 unique and innovative strengths. It is ex-ante harmonized across the countries, multi-disciplinary and longitudinal. SHARE also provides publications on the methodology and data and organizes every year international conferences, workshops, user conferences and summer schools at the central level of the SHARE-ERIC coordinator and at each SHARE national node. In the Czech Republic, SHARE-CZ cooperates with the Ministry of Labour and Social Affairs, Expert Commission on the Pension Reform of the Government of the Czech Republic and with more than 20 other universities and research organisations.

In **Austria**, university financing of the 2016-2018 performance agreement period accounted for basic budget of €7,5bn. The indicator-based part accounted for €750m and was distributed via 4 traditional indicators (three students-based and one based on revenues from Austrian Science Fund) and a grant indicator (for Cooperation and Excellence in Research, Teaching and Administration). For the period 2016-2018, this competitive project funding mainly focussed on research infrastructure and basic research projects.

The creation of incentives for third-party funding in the area of EU programmes is considered to be a priority, as it achieves a number of desirable objectives which would otherwise have to be achieved through separate incentive instruments: Increased reflux of EU funds to Austria; high co-operation with industry (due to their size, EU projects can only be achieved meaningfully in cooperation with companies/industry); international competition; global networking; efficient knowledge production through labour and cost sharing; focus on societal challenges.

Estonia uses a baseline funding formula from 2017 where publications and patents count for 38%, contract research for 47,5%, PhD for 9,5% and work on topics of national importance for 5%.

In the process of assessment of different grant applications in different instruments socio-economic impact is included to evaluation conditions. The importance of the research theme to science and to the economy and culture of Estonia and the EU, including its conformity with national strategic development plans is among the criteria of evaluation. Also, National evaluation of R&D institution includes the Assessment criterion: Societal importance of research. Evaluation and selection of applications of the activity "Support for applied research in smart specialisation growth areas" includes criteria considering project contribution to the cooperation between public R&D institutions and establishments, direct contribution to the development of smart specialization growth areas and economic impact and efficiency of the project. And finally, in selection process of the projects in the frame of the activity „Support for sectoral R&D – RITA" criterion „Socio-economic and social conformity" has weight 60% and criterion „excellence" 40%

There is a need to better understand how impact from research is generated. That is why **Norway** recently established research centre at the University of Oslo will investigate the impact of science. The centre is called OSIRIS (Oslo institute for research on the impact of science), and it is funded by a Research Council programme on research for research and innovation policy. This centre applies both qualitative and quantitative approaches to their research questions (http://www.sv.uio.no/tik/english/research/projects/osiris/About/research/).

Commercialisation indicators are counted in the national database for higher education statistics: business disclosures, patent applications, licence agreements and start-ups. In the performance-based funding system for HEIs we include an indicator on income from business, NGOs and public sector. We also include categories for popularization of research in our national CRIS. However, registration is on a voluntary basis. The question is whether these indicators can capture the real societal impact of research, or that it is rather a measurement of activities. Impact outside academia is included as a criterion in the subject specific evaluations carried out by the Research Council Norway.

There is no formal impact evaluation on national level using quantitative data. But there is a strong political demand for third stream measurements and the government (and parliament) has through the recent R&I bill said that it should influence the PRFS in the future. A government enquiry will look into the governance of Swedish universities. Developing principles for impact evaluations and incentives for third stream activities will be one the tasks for the committee.

Sweden has national statistics that cover different aspects of impact. They are used to analyse specific aspects of impact, not to create an aggregated picture. A non-metrics model has been developed and piloted by the innovation agency Vinnova, in close collaboration with HEI and others. For the forthcoming years, this model will be used for distributing 1/3 of the performance based part of basic funding, together with the indicators Bibliometry (1/3) and External funding (1/3). The piloting work is reported in: http://www.vinnova.se/upload/EPiStorePDF/vr_16_09T.pdf.

Vinnova got in 2013 the task to create a model for evaluation of third stream activities. They have done two pilot projects. See http://www2.vinnova.se/en/Publications-and-events/Publications/Products/Evaluating-the-Role-of-HEIs-Interaction-with-Surrounding-Society/.

How different countries looks into these matters and to what extent they try to evaluate them on national level. Also, to find the most appropriate metrics for third stream activities, and learnings how well such metrics would correspond to a review-type model.

In **Italy** ANVUR is also developing, on an experimental basis, a new system of metrics designed to support evaluation of third stream activities and impact of academic research. Third stream activities have been divided in two main areas, respectively involving research economic valorisation and the production of public and social goods. As for research valorisation, indicators are produced concerning intellectual property management (patents and vegetal varieties), academic entrepreneurship (spin offs), third party activities, intermediation activities; as for the production of public and social goods, we consider indicators concerning the management of cultural activities and the cultural heritage (museums, archaeological excavations and cultural heritage), clinical trials, continuous education and public engagement. Evaluation is based on peer review, informed on the aforementioned information. (see appendix A).

In **Portugal** the criteria normally revolve around a) productivity, meant as the total output of the group in its many different forms, including publications, patents, prototypes or products; b) Relevance, meaning the scientific, technical and/or socio-economic impact of the work carried out by the group, taking into account research choices in view of current trends at the international scene; c) Feasibility, reflecting on the capacity of the group of transforming interesting plans into practical projects that are relevant at the international level; and d) Training of PhDs and master students and participation in graduate programmes. Metrics related to scholarly quality and impact have been used frequently, but framed within broader evaluation criteria.

In **Croatia** criteria for societal impact are part of an overall funding scheme which looks as follows: scientific productivity – 60%; national and international competitive research projects and research mobility – 25%; collaboration between research and business sector, as well as collaboration with the units of local and regional governance and non-governmental sector – 10%; popularization of science – 5%. Analysis is based on the individual performance of each researcher based on the following criteria: scientific productivity; national and international competitive research projects and research mobility; collaboration between research and business sector as well as collaboration with the units of local and regional governance and non-governmental sector; activities of the popularization of science

Moldova does not have a coherent system with clear implications for Performance Based Funding, so also not for societal impact. Despite that, they use some education metrics for the evaluation of the quality of teaching activities: indicators regarding teaching process, teaching staff, research, etc. As well as for research institution's accreditation. This may remind us of the fact that educating young people is an important impact on society. The ambition is to put more focus on other quality aspects like gender equality, young researchers and a how to get a more even distribution among institutions.

In **Armenia,** there are three main funding instruments, (1) basic funding (with criteria like novelty of the research, importance, feasibility, human and technical resources, expected results, etc.), (2) Thematic funding for projects (with similar criteria), and (3) targeted programmes (novelty etc., but also criteria related state policy, expected results and impact).

In **Turkey**, as mentioned above, the Entrepreneurial and Innovative University Index was developed. The index introduces a dedicated performance evaluation and monitoring system for tracing the activities of universities in favour of the private sector. The theoretical framework, indicators, data availability and accessibility, calculation and normalization method, and weighting schemes are defined in consultation with relevant governmental organisations and universities. This process ensures that the final index scores reflect a meaningful distillation of the available information. A High Level Group and a Technical Group were constituted as advisory groups with the representatives of 10 institutions that have key roles in the national innovation system. These institutions are The Scientific and Technological Research Council of Turkey (TÜBİTAK), The Council of Higher Education, the Ministry of Science, Industry and Technology, Ministry of Development, TurkStat, Ministry of Finance, Turkish Academy of Sciences, Small and Medium Sized Enterprises Development Organization, Turkish Patent Institution, and the Technology Development Foundation of Turkey.

In the finalised index, 75 data components are collected mainly from administrative records, which uphold principles of data quality, as well as 178 universities and 49 technoparks. In addition to the values of the indicators, detailed information and descriptions are collected via forms to enable cross data validation. Since communication with university contact points is vital for the quality of data collection and validation, a permanent support desk is established at The Scientific and Technological Research Council of Turkey (TÜBİTAK). Experts provide guidance to the contact points throughout the data collection process. In addition, each university is contacted to guarantee the quality and correct scope of data as well as the values of the indicators, which further removes any missing data.

Conclusions

In summary, countries do not have third stream metrics fully integrated in PRFS, but some have partial quantitative methods, and others are developing and experimenting.

Table 3: Third stream metrics in use or emerging

Third stream metrics in countries participating				
Part of PRFS, institutional or national level		Emerging demand outside PRFS		
Quantitative measurements	Qualitative methods	National demand	Institutional level	Disciplinary level
Croatia, Estonia, Turkey Italy (experimental)		Sweden Norway/Osiris Czech Republic	Austria/universities	Norway (social sciences)

We can see a need in most countries for funding instruments that aim at narrowing the gap between academic research and the wider society. In most cases, the focus is either on business and industry or state policy. There are no cases where comprehensive third stream metrics are integrated in PFRS or institutional evaluation methods.

But we can see a number of approaches to evaluate specific elements of societal impact. They range from sometimes very complex and demanding systems (Turkey, Italy in development) to fairly simple indicators (Croatia, Estonia) and everything in between. We also see experimental methods, the most elaborate in Italy. In some cases, a case study approach appears to be very informative. Some countries express a clear need for better instruments to measure societal impact.

6 THE WAY FORWARD

We have seen that in most countries new funding schemes have emerged to improve the impact of research on business, industry and society at large. Many of the societal challenges are not country specific, but some are. The ways in which countries stimulate academics to reorient their research agenda differs, sometimes through targeted programmes, sometimes via more generic measures. How to measure the impact of these programmes is another matter. In this report, we have given a review of different methodological developments that may help countries to find approaches that help them to assess the results of the programmes targeting societal goals.

In Chapter 4 we noted that so far not many countries have included third stream metrics in their PRFS. We mentioned three reasons for this: (1) the contextuality of societal impact, (2) how to balance third stream metrics against more traditional quality measurements, and (3) what kind of policy consequences the outcome of such evaluation should have.

Based on what we have learned from the literature and answers of the participating countries we can now give some practical answers regarding how to develop methods to assess the contribution of research to society. The fact that societal impact means something different in different contexts leads to two conclusions. The first is that qualitative methods (narratives, case studies) should have priority, because they can describe in a meaningful way how research results are used in a societal environment. As was concluded in the ASIRPA project, the pathways between academia and society are often long and seldom straight. But secondly, one can use quantitative measurements to underpin the qualitative description. A number of approaches we described have developed metrics in three broad categories, people, media and commercial/practical activities. An example is the Contextual Response Analysis described above, or various forms altmetrics. Regarding the second issue, the question of balance, this can be solved via new forms of "peer" review, in which there is room for other expertise that is relevant for the topic at hand. The third issue might be the most difficult to handle. Consequences of evaluations will have to be discussed with partners outside academia, in as far as these stakeholders play a vital role in the project.

We have also noted that societal impact is not always the product of a complex network of stakeholders. While this is the case for the many grand societal challenges, sometimes researchers are asked to deliver input on a smaller scale, like for example a new protocol for a particular treatment of a disease. Then, the impact of the researchers can be measured through simpler techniques.

Taken together, the conclusion is that the assessment of societal impact presupposes that from the start of multi-level and multi-actor projects, evaluative goals and methods need to be discussed with stakeholders. Evaluation, then becomes a formative instrument (how can we improve in the future) instead of a summative instrument (judging what has been done in the past).

Practically speaking, countries can learn the following for the future:

There are promising methodological developments regarding third stream metrics. In the first place, these regard qualitative methods like impact path ways, case studies and narratives, Furthermore, there are developments in network approaches, more generic like the SIAMPI method, or more specific like the CRA. These metrics focus mainly on intermediate indicators, that is on small steps in a larger innovation process. And finally, some new quantitative measurements are developed, especially in the area of altmetrics, and popularizing publications.

For fruitful development of third stream metrics, the following points are important:

- Give priority to the development of qualitative methods, given all the insecurities still existing
- Stimulate the development of network indicators and intermediate indicators
- Pay attention to the development of altmetrics for societal impact

- Work on building more robust data collection systems
- Involve stakeholders in method development and data collection

7 ANNEX A: REFERENCES

Afgan, Naim Hamdija and Maria G. Carvalho, 2010, The Knowledge Society: A Sustainability Paradigm, volume 1, issue 1, october 2010

Bornmann: 2013 - L. Bornmann, 'What is Societal Impact of Research and How Can It be Assessed? a Literature Survey', *Journal of the American Society for Information Science and Technology*. 64, pp. 217–233.

Cozzens and Snoek: 2010 – S. Cozzens, S. and M. Snoek, 'Knowledge to Policy; Contributing to the Measurement of Social, Health, and Environmental Benefits', *Workshop on the Science of Science Measurement*, Washington, DC.

DORA: 2013 – Dora, http://www.ascb.org/files/SFDeclarationFINAL.pdf.

Finne et al, 2011, A Composite Indicator for Knowledge Transfer, European Commission's Expert Group on Knowledge Transfer Indicators (2011).

Fridholm, Tobias, Tomas Åström, Emma Ärenman and Lena Johansson de Château, 2016, User survey and impact assessment of the Norwegian social science institutes, Technopolis Group

Gann, David, Miles Tackett and Claire Thorne (2016), Pathways to Societal Impact; a Review of Imperial College's Technology Transfer, Translation and Related Activities, Imperial College London.

Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott and Martin Trow (1994), The New Production of Knowledge; the Dynamics of Science and Research in Contemporary Societies. Sage.

HEFCE 2016, Publication patterns in research underpinning impact in REF2014, A report to HEFCE by Digital Science

Hicks Diana M., Paul Wouters, Ludo Waltman, Sarah De Rijcke and Ismael Rafols, 2015, 'Bibliometrics; The Leiden Manifesto for Research Metrics', *Nature* 520, pp. 429-431.

High Level Group, LAB – FAB – APP — Investing in the European future we want, Report of the independent High Level Group on maximising the impact of EU Research & Innovation Programmes, European Commission, Directorate-General for Research and Innovation

Holmberg, Kim, 2017, Different types of Altmetrics, Thematic Report 1

INRA: 2014 – INRA: Translating Research into Impacts; 30 Case Studies. (http://www.inra.fr/en/Scientists-Students/Economics-and-social-sciences/All-the-news/Translating-research-into-impacts-30-case-studies)

King's College: 2015 – King's College: The Nature, Sscale and Beneficiaries of Research Impact; an Initial Analysis of Research Excellence Framework (REF) 2014 Impact Case Studies, King's College London and Digital Science

LERU 2017, Productive Interactions: Societal impact of research in the knowledge society, position paper, March 2017, Brussels [main authors Wiljan van den Akker, Jack Spaapen]

Nowotny a.o.: 2001 – Helga Nowotny, Peter Scott and Michael Gibbons, Re-Thinking Science; Mode 2 in Societal Context, Polity Press.

Nowotny, Helga, Democratising expertise and socially robust knowledge, *Science & Public Policy*, Jun 2003, Vol. 30 Issue 3, p151-156

Prins, A.A.M., and J.B. Spaapen (2017), Serving variegated audiences: from ranking oriented evaluation to mission oriented evaluation, acceptedby *fteval Journal for Research and Technology Policy Evaluation*

Science in Transition: 2013 - *Science in Transition*. Position paper. (http://www.scienceintransition.nl/english)

Sivertsen, G. (2016). Publication-Based Funding: The Norwegian Model. In M. Ochsner, S.E. Hug, H.D. Daniel (Eds.), Research Assessment in the Humanities. Towards Criteria and Procedures (pp. 79-90). Zürich: Springer Open.

Sivertsen, G. (2016). A bibliometric indicator with a balanced representation of all fields. In Ràfols, I., Mollas-Gallart, J., Castro-Martínez, E., Woolley, R. (Eds.), Proceedings of the 21ST International Conference on Science and Technology Indicators (pp. 910-914). Valencia: Editorial Universitat Politècnica de València.

Smith, Richard, Measuring the Societal Impact of Research, *BMJ*, 2001 Sep 8; 323(7312): 528

Spaapen,J., L. van Drooge, Tilo Propp a.o. (2011), SIAMPI Final Report, Social Impact Assessment Methods for Research and Funding Instruments Through the Study of Productive Interactions Between Science and Society http://www.siampi.eu/Content/SIAMPI_Final%20report.pdf)

Spaapen, Jack and Leonie van Drooge (2011), Introducing "Productive Interactions" in Social Impact Assessment, *Research Evaluation* 20 (3), pp. 211-218.

Thompson Klein: 1990 - Julie Thompson Klein, Interdisciplinarity; History, Theory and Practice, Wayne State University Press.

Thompson Klein, Julie & Carl Mitcham (eds.), *The Oxford Handbook of Interdisciplinarity*. Oxford University Press (2010)

Volberda, Henk, Justin Jansen, Michiel Tempelaar and Kevin Heij (2011), Monitoren van Sociale Innovatie: Slimmer Werken, Dynamisch Managen en Flexibel Organiseren, Rotterdam School of Management, Erasmus University Rotterdam.

Weingart, Peter, Impact of Bibliometrics Upon the Science System; Inadvent Consequences?, *Scientometrics*, vol. 62, no. 1, pp. 117-131.

Wissenschaftsrat: 2016 – Wissenschaftsrat: Wissens- und Technologietransfer als Gegensand Institutioneller Strategien. Positionspapier.

Wilsdon, James a.o. (2015), The Metric Tide; Report of the Independent Review of the Role of Metrics in Research Assessment and Managemen*t*, DOI: 10.13140/RG.2.1.4929.1363.

Getting in touch with the EU

IN PERSON
All over the European Union there are hundreds of Europe Direct Information Centres.
You can find the address of the centre nearest you at: http://europa.eu/contact

ON THE PHONE OR BY E-MAIL
Europe Direct is a service that answers your questions about the European Union.
You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by electronic mail via: http://europa.eu/contact

Finding information about the EU

ONLINE
Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU PUBLICATIONS
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact)

EU LAW AND RELATED DOCUMENTS
For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

OPEN DATA FROM THE EU
The EU Open Data Portal (http://data.europa.eu/euodp/en/data) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

in most countries, new funding schemes have emerged to improve the impact of research on business, industry and society at large. However, so far few countries have found ways to assess the results of such programmes targeting societal goals and/or included third stream metrics in their PRFS. This report provides examples of qualitative methods to assess societal impacts like impact pathways, case studies and narratives, discusses developments in network approaches that focus mainly on intermediate indicators, i.e. on small steps in a larger innovation process, and presents some new quantitative measurements that are being developed, especially in the area of altmetrics and popularising publications.

Studies and reports