EUREC
The Association of European Renewable Energy Research Centres

Feedback on two years of Horizon 2020

Prescriptive calls for solving oversubscription, raising evaluation quality and making the Work Programme more strategic

The silver bullet that could improve Horizon 2020 has already been fired. Engraved on it are three words, “More prescriptive calls”, and in Energy Work Programme 2016-17[1] EUREC is confident it will achieve three things: reduce subscription, improve evaluation quality, and increase enthusiasm for roadmapping exercises, which will lead, ultimately, to the next Work Programme being more strategic.

In opting to define Innovation Action (IA) topics more closely (by splitting them into their own topic headings instead of lumping them together under one heading called LCE 3) and in spelling out low-carbon energy Research and Innovation Actions (RIAs) in roughly twice the space as in the Work Programme 2014-15 (10-15 lines versus 6-8), the EC has shrunk the field of possible eligible proposals in Work Programme 2016-17 compared to Work Programme 2014-15. It is reasonable to suppose that in response subscription rates will fall back to levels closer to FP7’s, which contained prescriptive calls. Oversubscription is the flaw in Horizon 2020’s design that most urgently needs correcting.

Evaluation quality needs a boost. Some Evaluation Summary Reports received by consortia that entered proposals under Work Programme 2014-2015 have been criticised for being superficial, not objective or inaccurate. This turns applicants off Horizon 2020. With fewer proposals arriving for evaluation, evaluation quality is expected to improve. This is because rules limiting the number of hours that one evaluator may work have led to evaluators with less experience being hired to deal with the large numbers of proposals. Also, all evaluators (both experienced and less experienced) had to handle a workload judged as ‘high’ by one of the observers recruited by the EC to report officially on the running of this year’s ‘Energy’ evaluations. The observer followed evaluations in competitive LCE topics and wrote, “In general longer time for remote evaluations is needed, especially when over 10 million € investments are in question. The time constraints are very high.” The evaluators had two weeks for 12 proposals. The maximum length of a proposal is 70 pages. EUREC has specific proposals for improving evaluation quality – see Annex.

Finally, prescription implies choices, and a strategy needs to exist if those choices are not to be arbitrary. Prescriptive topics therefore encourage strategic

[1] Work Programme 2016-2017 'Secure, Clean and Efficient Energy'

thinking and give meaning to the work of bodies like European Technology and Innovation Platforms.

Evaluation Summary Reports – insufficient for accessing European Structural and Investment Funds

Evaluators, aided by the rapporteurs in the 'Consensus groups', have a short space of time in which to write an ESR for each proposal. Even if they are well-written, they are brief, and comment on the proposal without giving guidance (as there is no substantial negotiation before the signing of the grant agreement under Horizon 2020).

This is a problem if bidders want to use that document to unlock funding from other programmes. Energy Work Programme 2016-2017 specifically identified the European Structural and Investment Funds as one of the opportunities they might want to explore. The EC has created opportunities for researchers to find the right local person to lobby for funding (Eye@RIS3 tool) but that person will often not be able to judge if the project being presented, likely to contain new and unusual features, is technically and financially sound. A simple ESR is not sufficient. The project would need to be redesigned for a particular regional fund and a deeper evaluation would be needed. This will need to be accompanied by a quick, lean State Aid clearance procedure to allow the ESIF money to flow to the project without delay or expense.

Competition within a technology topic is fine, across technologies not

The silver bullet does not quite slay all of Horizon 2020's werewolves. Competition within a technology is to be encouraged (e.g. all ocean technologies compete together for a defined budget), but between technologies it is counterproductive. Technologies can end up starved of funding, which prevents them making progress on their R&D strategy. In some calls, it has been difficult to find a reassuring explanation for the difference in average score (Figure 1).

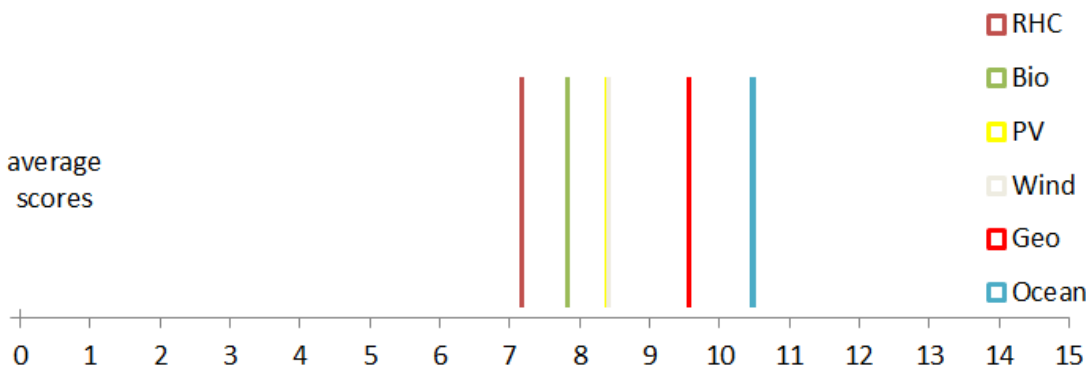

Figure 1 EUREC has seen data for the scores given to proposals under the topic LCE 3,12 (Innovation Actions, one-stage evaluation) with deadline 5/5/2015. The 14 ocean proposals on average scored 3.3 points higher than the 16 RHC (renewable heating and cooling) proposals, with the other technologies somewhere in between. Only for wind and PV are the average scores virtually the same, appearing as one line. Sample size ranged from 6 proposals (Geo) to 17 (Bio). It is surprising that RHC proposals should score so systematically poorly relative to ocean ones. Nothing from the Panel Report of the evaluation, EUREC has been told, suggests that the panel queried this result. The report says only that the standard methodology for tie-breaking was applied, detailed in Annex H of Part 18 of the General Annexes to the Work Programme.

In the Energy Work Programme calls, as elsewhere in H2020, a standard process is followed to evaluate proposals: consensus reports are drafted in consensus groups. The members of consensus groups are evaluator experts specialised in particular areas: if a call invites topics in, say, wind, geothermal and hydropower it is normal for each consensus group to focus on one of these technologies.

Imagine one consensus group covers ocean proposals. The consensus group members know that this year and the next are make-or-break for the industry, which has to deploy at scale and demonstrate potential for huge cost reduction. Every euro of funding it gets is valuable. Out of a sense of loyalty to their industry, therefore; they might mark all proposals generously, while still distinguishing the excellent from the good. While the members would have no conflict of interest in respect of any particular proposal, they may have a bias in favour of the industry that they are experts in.

Horizon 2020's rules are ill-equipped to address this. A step to compare scores across consensus groups is included in the evaluation process, known as a 'panel review'. The panel review members would need very special ability to declare that a renewable heating and cooling proposal judged 'good' by its consensus group in fact had equal merit to an ocean proposal judged 'excellent' by its. Indeed, such analysis is not requested of them. Short of finding panel reviewers with both a deep and broad technical knowledge of energy technologies, the only practical option for the EC is to reduce the scope of topics

and increase their number, with a budget identified for each of the new topics. This is already happening in the WP 2016-2017 for IAs in renewable energy (but not RIAs, which are still bunched together in the LCE-07 topic).

Experimentation in H2020 – what worked and what did not

The Work Programme of 2016-17 is more prescriptive than that of 2014-15 – this has been discussed already. When Innovation Actions and Research and Innovation Actions competed against each other in 2014 for the LCE 8 and LCE 10 topics and in 2015 for the LCE 6 topic, one kind of action vastly dominated over the other in the retained proposals. The reasons are not clear, but this is further evidence of the difficulty in comparing projects that are dissimilar, in this case because they concern technologies at different readiness levels.

In the Energy Work Programme, the two-stage calls in 2014-15 have been replaced with one-stage calls. This is welcome because it hastens time-to-grant and because a lot of the work to prepare a second-stage proposal is already done for the first-stage one.

It is a shame not to see the max-25%-of-total-budget-to-one-technology rule from Work Programme 2014-2015's LCE 2-LCE 11 re-appear in the LCE 7 topic in Work Programme 2016-2017. This is an important safeguard if technologies have to compete against each other, giving a fair chance to all to make progress on their R&D strategies with EU money. It was somewhat misapplied in 2014-15 when all the various forms of renewable heating and cooling technology open in a particular call were bracketed as 'one technology'. The alternative would be to make sure that a fully open competition is complemented by an amount of RIA funding earmarked for all the technologies that are in competition with each other (or at least, several of them) elsewhere in the Work Programme.

We are happy with two-year Work Programmes. A Work Programme that is largely fixed up to three years in advance would build too much rigidity into Horizon 2020 for the fast-evolving renewable energy sector.

The Impact criterion, important and weighing heavily (appropriately) in evaluations, is also the hardest to address correctly in a proposal. Bidders can write sensible text in response to the 'Expected Impacts' laid down in the Topic description, but then struggle with the generic Impact success factors from Part H of the General Annexes, specifically, "Enhancing innovation capacity and integration of new knowledge; Strengthening the competitiveness and growth of companies ...". Potentially explorable to an arbitrary level of detail, the EC should give guidance on how these success factors will be evaluated, and/or be selective in the 'Impacts' that bidders must address.

There is no grant agreement negotiation step in Horizon 2020. This is a sacrifice worth making for quicker time-to-grant. ESRs should not, however, be made

less informative as a consequence, for example by omitting advice from them. Advice could be useful if a project is re-proposed. In two-stage calls, the ESR of the first stage should be made available to the consortium behind the proposal.

We look forward to the 'ocean array' call LCE-15-2016 with its novel feature of a decision in the middle of the project for go-/no-go on deploying the array. The EC has always had the power to cut off funding to a poorly performing project, but this is slightly different: the preparation for the deployment might have been done well, but economics conspire against the project through no fault of the consortium. This is a situation that projects focused on deploying innovative technology could routinely face. For such projects, go/no-go gates should become the norm. In future, two proposals could compete for one pot of build-out money.

The IT tools and policies developed for exchanging information between the EC and its grantees are working well and are a simplification compared to FP7.

Annex – proposals for higher quality evaluations

If the EC is running short of genuinely expert evaluators, then it may need to spend more money, for example to advertise the opportunity to work as an evaluator, or to increase the honorarium that each evaluator is paid. There are also three zero-cost ways to increase quality, which all involve making evaluators more accountable to the consortia whose proposals they evaluate.

1) Demand that evaluators back up the comments they write in ESRs with specific references to passages in the proposal, at the level of page numbers.

2) Expose evaluators to feedback from the consortium. Our suggestion in Box 1 (below) is a much lighter way of getting that feedback to the evaluators that by activating the official evaluation review procedure.

- Proposers would be given the opportunity to write 250 words to the team most closely involved in the evaluation of their proposal (probably those who have written the Individual Evaluation Reports) to tell them what they thought of their ESR. The identities of the team members would not be revealed – the message would be forwarded through the Participants' Portal without being copied to the Commission. The feedback would have no legal bearing on the evaluation outcome.
- Evaluators would need to opt in to the process by ticking a box to agree to accept feedback.
- A result of the feedback loop could be to cause some evaluators to reconsider how they approach their work. In extreme cases, evaluators who feel they receive a lot of justified criticism may choose to de-register.

Box: Might adding a feedback loop be a low-cost way to increase quality? About the only way today in which evaluators can learn how to improve their work is to use the consensus meetings to compare their grasp of a proposal with that of their fellow evaluators.

The danger of a low-quality evaluation round is that it changes bidders' strategies for future calls. If ESRs are perceived to contain incomprehensible or unfair comments, the Horizon 2020 selection process starts to look like a lottery, making the winning strategy from the proposer's point of view one where he churns out as many proposals as possible, compromising on quality if necessary. This overloads the prime evaluators and increases the number of evaluations going to the less experienced, meaning the quality of evaluation stays low and the cycle is repeated.

3) Disaggregate the names of evaluators in the annual published list. At present all evaluators in Energy Societal Challenge calls are listed together. It would be possible to split out the names of the evaluators involved with specific technologies without connecting each name to a particular proposal. This would make it an easier task for bidders to check out the people who might have been involved in evaluating their proposals.

19 Feb 2016