



# Report

**Workshop:  
Exchange of experiences between ICT evaluators in FP7**

21<sup>st</sup> January 2009

Budapest

## Introduction

In 2007 more than 1300 experts from 52 countries were involved in the process of evaluating European ICT projects. Each of them offered valuable perspectives by representing diverse background and experience during the process.

The goal of the experts' workshop on 21 January 2009 in Budapest was, to bring together a large number of experts from different countries for an exchange of experience. The results of this interactive discussion were intended to help

1. the experts in future evaluation processes,
2. the European Commission to improve the evaluation process where necessary and
3. the National Contact Points in giving advice to proposers for the improvement of their proposals.

In 2007 a similar national workshop with German evaluators took place in Cologne. To extend the scope to an international level, the workshop in Budapest was organised.

36 evaluators from 14 countries took part in the workshop (Fig. 1), of which one fourth was from industry, three fourth from academic. One fourth of the evaluators participating in the workshop was female.

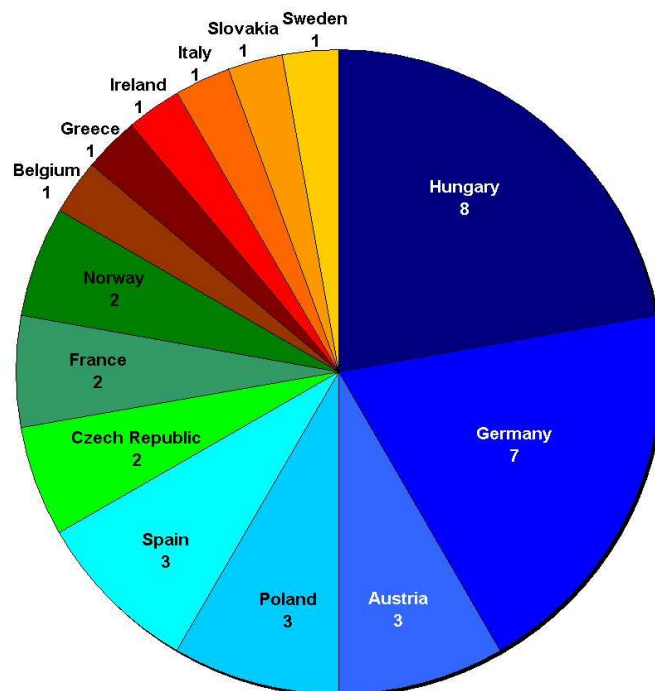


Fig. 1: Distribution of experts by country

More than half took part in more than one evaluation process and one third took part in more than 3 evaluation processes with different topics. Additionally 50 ICT NCP members from 33 countries attended to pass the gathered information on to their proposers.

The preparation of this workshop begun by examining prior experience accumulated within its own ranks: information already gathered from interactions with evaluators, the European Commission (EC) and the European ICT research constituencies.

Particular attention was given to new aspects of FP7 submission and evaluation procedure (e.g. the value of the extension of the “pre-proposal check”, the effects of remote evaluation, the benefits/drawbacks of the 1-step and 2-step evaluation systems) allowing NCPs to understand the implications of these and ultimately provide more advice to proposers.

The results of these activities led to the definition of the workshop sub-topics which were discussed in detail with a selection of evaluators to share personal experience.

The 36 evaluators were divided into 2 moderated groups which both discussed the same list of questions, to bring about a consensus. The main discussion topics were the evaluation process and the most important dos and don'ts for applicants. The latter were presented to the proposers by Idealist the next day during the European ICT Proposers' Day,.

The attending NCPs were not allowed to participate actively in the evaluator's discussion in order to get a real opinion about the evaluation process as such.

## **Executive Summary**

### **Time frame**

The discussion started with the time frame given per proposal. Here both groups agreed that the time allowed for the evaluation (2 hours for STREPs and 3 hours for IPs) is sufficient. By contrast the time given to read the proposals prior to hearings or panels was seen as too short. Furthermore it was agreed that the briefing of evaluators is much too long for experienced evaluators.

### **Hearings**

The opinion concerning hearings between both groups differed. While one group stated that hearings are good because they improve the quality of the judgement, but almost never result in a change of ranking, the other group saw no added value, even if they changed the results of the prior evaluation and may in contrary be uncomfortable for the evaluators because of the loss of anonymity.

### **Remote-evaluation**

Both groups stated pros and cons about remote-evaluation. Pro: it is easier to involve female evaluators and to acquire more evaluators from industry; one can look at different proposals in parallel; the internet can be used (here new guidelines from the Commission are needed). Con: lack of social and team building aspects between the evaluators, spending of more time on each proposal.

### **Evaluation criteria**

Both groups agreed with the number of evaluation criteria (3) but stated that they should be weighted differently. The range of scores (0 – 5, with half scores) was found not to be enough to differentiate proposals. But most important was that the lack of weighting of criteria is not appropriate, because some criteria are found more important than others.

### **Consensus meetings**

Regarding consensus meetings one group was of the opinion that that in some cases dominant evaluators might influence scoring, but only if the moderator not is able to deal with this. Therefore it was seen as important that the moderators get trained.

### **Ranking of proposals**

Panels are not seen to alter the scores given in the consensus meeting, but a single evaluator might, by 'defending' a score affect the ranking. Therefore a strong need was expressed for an impartial predefined process of how to rank proposals with the same score.

One group stated that neither a merged ranking of different instruments (IP vs. STREPs) is realistic or desirable nor the setting of minimum budgets for IPs & STREPs. Furthermore it was seen as a deficit by one group, that the work programme does not provide any guidance to decide on supporting/rejecting a single IP vs. several STREPs?

### **Dos and don'ts**

Finally all participants from both groups mentioned the most important do's and don'ts for applicants.

# Structured Report

## 1. Evaluation process

### *a) Evaluation methodology*

#### **I. Time-frame per proposal**

The evaluators stated that they are spending 2 hours evaluating STREPs and 3 hours for IPs on average, which is sufficient and does neither have to be extended nor to be shortened. Even if there was more time per proposal, that would not influence the evaluation outcome.

#### **II. Hearings**

Concerning the hearings the outcomes in the two groups differed. While one group said that hearings are a good thing because they are improving the quality of the judgement and are making it possible to meet the proposers face to face and get an immediate feedback, the second group stated that hearings do have a very small added value only, they might be uncomfortable for the evaluators, because they are no longer anonymous and that there is not enough time prior to the hearings to read the proposals carefully.

Furthermore one group stated that hearings do almost never change the results of the evaluation, while the second group held the view that there might be a change in the results of the evaluation, e. g. if the consortium is weak. But if so, the marks are mostly going down for the proposal after the hearing. Besides it was criticised that the opinion formation is mainly based on personal impressions, because of the insufficient time for reading the proposal in advance. One evaluator advanced the view, that there should be a hearing for all proposals, if they passed the threshold, because many proposals have tied scores thus it is difficult to make a fair judgment.

#### **III. Remote-Evaluation**

Both groups stated pros and cons about remote-evaluation. As pros were seen that remote-evaluation might make it easier to involve female evaluators and to acquire more evaluators from the industry which cannot afford to be out from their work for one week to do evaluation. Furthermore it was seen positive that one can look at different proposals in parallel. A disadvantage might be the lack of social and team building aspects between the evaluators.

There was a consensus in both groups that evaluators are spending more time on each proposal since the introduction of remote evaluation. Furthermore it was agreed, that evaluators are using the internet to verify statements in proposals and backgrounds of proposers and it was found useful, as one cannot be an expert on all aspects of the proposal. Moreover there was no concern, that the wide availability of academic citations vs. industrial research information on the Internet will impact evaluation results. But it was mentioned by one group, that the current guidelines of the European Commission specify to base evaluation solely on the proposal, which is why new guidelines are needed.

The risk with respect to confidentiality posed by remote-evaluation was appraised variably by the groups. As one group had no mayor concerns, the other group was concerned about the loss of confidentiality.

Eventually there was a consensus, that remote-evaluation did influence the quality and content of proposals.

### *b) Evaluation criteria*

#### **I. Number of criteria / range of scores**

Both groups agreed with the number of criteria (3) but stated that they should be weighted differently. Furthermore there was a consent, that the range of scores (0 – 5, with half scores) is not enough to differentiate proposals and a finer scoring is needed. On this one evaluator commented, that consensus finding might become more difficult with a broader range of scores. A supposed

solution was to have a 'bulk' ranking first and afterwards let the proposals with the highest scores undergo a further more accurate scoring.

## **II. Interpretation of criteria**

It was agreed in one group, that in general the evaluation criteria are clearly defined and that there are no overlaps, but that some rely on one another, e.g. impact and science.

There was disagreement within the group regarding the effect of the lack of anonymity on the evaluation scores. While one opinion was, that it has no effect, because even as the proposers were anonymous, one could guess who was involved from the profiles, the other opinion was that the evaluation could be influenced by a prejudice for different organisations or countries.

Subsequently it was consent in one group, even if it would be desirable, that there is no common understanding of the evaluation criteria as shown in consensus meetings.

## **III. Weighting of criteria**

As mentioned above, it was agreed by one group, that the lack of weighting of criteria is not appropriate, as proposers spend most of their time for the technical criteria (about 90 %). In addition it is often an issue for proposers how to structure the management criteria.

Referring to the above-mentioned, the group also agreed, that some criteria are more important than others; especially the impact criterion should have less weight, as there might be large discrepancies and misjudgements in estimating the impact of a project.

Further on the group was consent, that weighting of the criteria would probably facilitate ranking.

### *c) Consensus & ranking*

#### **I. Consensus meetings**

Regarding consensus meetings one group was of the opinion that language is no problem, as long as the rapporteur is able to speak English very well.

It was also agreed in one group, that in some consensus meetings there are dominant evaluators that might influence scoring, but as long as the moderator is able to handle this, does not become a major problem.

Even though it is hard to say if it influenced the outcomes of the consensus meetings, it was consent in one group, that the Commission staff is not always neutral.

#### **II. Panels**

According to the opinion of the evaluators in one group, the scores set during the consensus meeting are not altered by the panel.

The moderators, likewise in the opinion of the evaluators of one group, are not always efficient at managing panel discussions. Therefore it was supposed that the moderators should get trained in advance.

Furthermore this group agreed that the ability of an evaluator to "defend" a score affects the ranking, but the rapporteurs do probably not affect the outcome of panel discussions.

It was consensus in one group, that panels do not always arrive at fair rankings, thus more preparation time would be useful. Furthermore, as not all evaluators are present it might be difficult for some proposals to be defended. In addition some experienced evaluators, particularly from small countries, supporting proposals from their own country because they know the proposers personally.

Besides, one group was of the opinion that there is a need for an objective predefined process of how to rank proposals with the same score. In reply to this a representative of the European Commission noted that such a guideline is included in the upcoming handbook for evaluators.

#### **III. Superpanels**

According to the opinion of one group, it is not realistic and desirable to have a merged ranking of different instruments (IP vs. STREPs) or different sub lists for large topics, as IPs and NoEs get a 2<sup>nd</sup> chance at a hearing and STREPs do not.

The same group agreed in not setting minimum budgets for IPs & STREPs, as that does create an artificial lack of flexibility.

It was seen as a deficit by one group, that the work programme does not provide any guidance to decide on supporting/rejecting a single IP vs. several STREPs?

#### *d) Evaluation process in general*

##### **I. Briefing and preparation by commission**

Several evaluators stated that for those who have been evaluators before, the briefing by the Commission is much too long. Moreover the presentation of topics by the Commission may influence the evaluation by telling more than written down in the work programme.

##### **II. Balance of evaluators**

It was stated by one group of evaluators, that the balance of industrial (large & small) and academic evaluators is mostly sufficient, but that it would be desirable to have more experts from the industry as some technical aspects can only be evaluated improperly from the academic experts, as they require more technical experience. Also the representation of Member States and Associated States was seen as satisfactory.

### **3. The most important do's and don'ts for applicants**

(The outcomes are displayed differently: **Consensus in both groups (bold)**, Consensus in one of the two groups (regular).)

What are your most important recommendations for writing excellent proposals?

- **Read the work programme thoroughly.**
- **Be very specific and focused rather than too broad and too ambitious and give concrete examples.**
- **Be careful when 'recycling' a former proposal (copy/paste), you might end up with something inconsistent.**
- **Do not underestimate the importance of the criterion 'impact'. The impact section should be elaborated and covered very well.**
- Show (in 10 pages or less) you know the state of the art and show you go beyond it.
- Proofread your own proposal and let a colleague read it.
- Try to read and score your proposal in 2 hours like the evaluators do.
- Write a proposal because you have an idea – not because there is an open call.
- Give it to an English speaker for linguistic corrections.
- Be up front and do not try to hide anything.
- Show you really can achieve the objectives
- Elaborate what each partner is doing.
- Expose that the project is European and not national and how it will add value to the EU.

What aspects of a proposal prompt you to "deduct" points?

- If someone feels too confident on his own impact and importance and constitutes that he will solve everything and the project will be good for everyone.
- If the technical part is not well supported by literature.
- If the consortium is poor or too big and fails to show how the members will work together.
- If one is being too generic and gives unsubstantiated statements. That leads to a lack of clarity.
- If the project is too low-risk.
- If the scope of the proposal is too narrow and covers only a very small part of the work programme.
- If there is a lack of detail of what will be done after the project.
- If there are discrepancies between the work plan and what they have said they will do.

Subsequently it was consent in one group that with remote evaluation only the support by recent literature has become more important, as the evaluators can find the most recent literature on the internet.