AIR and performance review

Seminar
March 26-27, 2019 Amsterdam[1]

Day 1

Result of the mapping exercise

An initial 'mapping exercise' invited the participants to indicate sections in the reporting template which pose questions or might be challenging to complete (each participant had three red dots to indicate it on the reporting template).

The result of the mapping exercise is as follows (the 4 most critical ones):

No	Section	Number of hits	Comment
5.	Issues affecting the programme performance and measures taken	7	To be seen in conjunction with section 14. In section 5. all issues having an effect in the reporting period should be tackled.
9.2	Specific actions taken to promote equality between men and women, non-discrimination, access for persons with disabilities	12	Proposal for the approach see ppt 1, slide 18
13.	Smart, sustainable and inclusive growth	22	Proposal for the approach see ppt 1, slide 19
14.	Issues affecting the programme performance and measures taken - performance framework	12	To be seen in conjunction with section 5. In section 14. only issues having an effect on indicators which are part of the performance framework should be tackled.

Kind reminder for eMS users

In our last seminar on AIR in March 2018 we have reached the following agreement among programmes using eMS:

[1] For more information please also check the three presentations of the event (Interact AIR and Performance Review, AIR _David Alba, Performance Review _David Alba) under: http://www.interact-eu.net/library#2534-presentations-air-and-performance-review

Agreement on definitions for programmes using eMS

A common definition of the terms 'Selected' and 'Output delivered' is highly appreciated when using a common monitoring system! The Group has agreed on the following definitions:

Selected = Contracted (projects counted as selected are the contracted ones)

Output delivered = Outputs are included in a report which has been approved by JS

It is obvious that the use of a shared definition is beneficial in order not to aggregate projects in different stages of the project life cycle. The definition for selection was chosen to be on the safe side since upon selection by the MC project applications might still drop out e.g. in case applicants refuse to comply with conditions imposed by the MA/MC.

However, the definition is a recommendation for the sake of data consistency.

AIR: David Alba, DG REGIO, Evaluation and European Semester Unit

See presentation 'AIR David Alba': http://www.interact-eu.net/library#2534-presentations-air-and-performance-review

Key issues

- The unit has set-up a system of plausibility checks for all data reported to the SFC (and thus also to the Open Data Platform) – according to these checks 15% of data suffer from basic plausibility issues (which does not necessarily mean that the reported data are wrong)
- The system used for plausibility checks will be provided by the evaluation unit to the desk officers (DO) – so the desk officers will come back to you with 'warnings' generated by the system [these 'warnings' do not per se indicate an error but point at data which might require additional explanation (or correction)]
- Data reliability is a major concern
    - A general one since the Open Data Platform is becoming increasingly popular and data are used for manifold purposes when reporting to stakeholders such as the European Parliament, the Council, European Court of Auditors (ECA), other DGs
    - A specific one for this AIR since EC decision on the performance framework will be based on the data reported in the AIR and any later modifications will lead to rather cumbersome and lengthy proceedings
- The warnings produced in SFC should also be understood as an indication to reflect on; but if your approach is deemed to be correct – go for it
- Most frequent phenomenon is that values for 'Selected' are higher than 'Target' – which is understandable from a technical, internal perspective but obviously poses questions from the perspective of external readers
- For section 4 (synthesis of evaluation reports) there will be optionally a structured grid for inputting – it might be useful to structure the section
- There are two new features in the Open Data Platform (ODP):
    - The option to briefly present projects

- A new feature is under development and will be available in the coming months: the option to embed your data on the ODP – that means data changes will be uploaded automatically to the ODP

Question & Answer (Q & A)

Q: Are there any other reporting requirements stemming from IPA which go beyond the programme reports? And - is there any concise guidance related to the issue?
A: There is no chances since last year related to the reporting requirements for IPA programmes. However the Interreg template was adapted and there is a different SFC module this year.

Q: Is it recommendable to change the target at programme level? Especially in case of underachievement?

A: if significant under- or overachievement it is in the end due to a wrong assumption; the reasons for the wrong assumption might be diverse. If the output indicators are not part of the performance framework the fact deserves an explanation but targets might remain unchanged.

One of the recurring phenomenon is overachievement during implementation. Forecast should comprise open and closed projects. It would be more prudent to adapt values under 'forecast': 'Forecast' should comprise implemented and delivered outputs for partially implemented or finished projects in case values implemented are much higher than forecasts of the project – that would support the consistency of data.

If you already know that the information is not correct (forecast) please correct it (e.g. not to have more target population reached than persons in the programme area). Admittedly it is difficult to filter double counting but, in the end, it is not plausible that projects support more people than the population in the programme area.

Q: What if figures for 'selected' are now lower than in the previous report? Obviously the error was due to wrong counting by the system – how to tackle that?

A: It should be explained in the observations; the change of data for previous years is possible and the reason should be given in the field for observations

Q: The method chosen for some of the result indicators is quite time consuming and resource demanding. Could a change in the methodology be envisaged?

A: An adjustment of the methodology is possible provided the current methodology poses a disproportionate burden on resources. It should be discussed with the desk officer.

The obvious advantage of statistical data is that even if an indicator or the underlying statistical method for data collection are being changed historical data will be available. In contrast to that a survey cannot be repeated retroactively.

Q: For a result indicator measurement resulting from a survey has been proposed three times (2018, -20 and -23) during the programme period. The first measurement has been done –could the subsequent measurement be skipped since the results will not differ substantially?

A: Such a change of the methodology can be justified provided that the expectable added-value of the interim measurement is most likely to be quite limited.

It was intended that the assessment of result indicators should be done in close connection with the impact evaluation. It should be the evaluator's task to elaborate qualitative effects or to distinct intended from unintended effects etc. Take the reported case where a result indicator has fallen below the baseline. It should be the task of the impact evaluation to make understandable what happened.

Q: The impact evaluation has come to important findings but underlying data collection has gone into 2019. The outcomes of the evaluation would be quite useful for the AIR but the desk officer stated that the cut-off date for inputs to the AIR is end 2018.

A: It is ideal if timing allows to use findings of the evaluation for this report. It underlines also that evaluation is an important tool for programme management. It is recommended to discuss it once again with the desk officer – the use of the evaluation findings would be advantageous.

Q: The statistical data initially used are no more available due to a change of the respective administrative units.
A: In such cases a change of the result indicator could be envisaged. It might be an opportunity to change to a more representative result indicator which shows measurable effects, and can be collected with reasonable effort.

Q: For one of the objectives a statistical indicator has been used but the underlying method has been changed – now the baseline is higher than the target in the Programme and the MC refuses to adjust the target. What could be done to avoid meaningless and void explanations in reporting?
A: It is recommended to change to a qualitative indicator which reflects the initial intent.

Wrap-up

- Conservative planning of outputs and milestones at programme start together with conservative approaches taken by projects when planning often results in tremendous overachievement – which obviously does not look good when targets are exceeded multiple times in the middle of the funding period

- It is evident that certain numbers in ODP look implausible for external stakeholders even if there are technical internal reasons for it – policy analysis and communication requires consistent and plausible aggregate data for Cohesion Policy and the contribution of individual programmes is important

- With a view to post2020: according to the draft regulations the AIR is no more required but regular updates of data will become increasingly important – thus programmes should re-think current routines which might be an obstacle to regular adjustment of data (such as linking it to contract changes)

I Panel: the role of Monitoring Systems (MS) in reporting

Q: What are your key technical challenges when preparing the data for the annual report?

A:

- The technical link between outputs and system work packages (i.e. if an output is in multiple work packages, then it is counted multiple times).
- We thought that each output should contribute to the OI – of course they might have
- main outputs which do not contribute to OI

Q: What is your approach to ensure that the data in the MS are reliable and as recent as required? How do you ensure that all persons in charge of data input know what to do?

A:

- Central Baltic: Using clear language, talking to applicants during the application phase, checking numbers reported (e.g. people reached by a TV commercial) are key.
- MED programme: Main methods and tools: approval of deliveries, clean-up after application phase, online platform where projects upload their deliverables/outputs. This comes with a questionnaire and is subject to quality check. If deliverables/outputs do not qualify, projects are asked to change their numbers reported. The programme provides guidelines and trainings (both for the MS and the web platform). The problems the programme had had with outputs come from the HIT template that assumes that each output corresponds/is linked to an output indicator even though in real life this is not always the case. In these cases, the programme considers these as deliverables rather than outputs.
- Greater Region: 11 contact points and the JS help to ensure that that the data in the MS are reliable and recent. The programme requires an annual report from all projects (to be submitted before the annual meeting). These reports are also collected in external spreadsheets and the content of those are cross-checked with the MS.

Q: Would you recommend using your MS also post 2020?

A:

- Greater Region: Wouldn't recommend it, mainly because of the language issue (i.e. need to translate everything to German).
- MED programme: The management for post-2020 needs to be simplified based on actual needs. Maybe a MS for CBC programmes and a separate one for TN programmes would be beneficial. Would welcome linking the Commission ECAS database to the MS, allowing Commission to retrieve data easier. Also suggests not to focus on financial monitoring only but creating a system that is useful for decision-making and that also focuses on qualitative aspects.
- Central Baltic: Simplification and harmonisation are key for the future – at the end programmes are not that specific. The programme would welcome a community MS coordinated by Interact.

Interact Monitoring System (MS) post 2020

- Interact is currently preparing the fundament for the new MS- the decision should be taken in the MC in May 2019; preparatory steps to develop the technical specifications and to prepare the procurement are currently being done
- Letters have been sent to Mas kindly requesting to confirm the interest in a new system in order to make sure that the Interact-MS post 2020 is an efficient and meaningful investment
- Please understand: It is a big project for Interact; its volume including development, maintenance and staffing is comparable to the current budget of one of the Interact offices; to go for it is a major decision in terms of budgetary commitment and staffing

Harmonised Implementation Tools (HIT)
Peter Racz (Interact)

- The so-called Core Group has been set-up – more than 20 programmes are on board
- The first step is to review the current tools and see for the adjustment needs
- Development of HIT and the new MS will be done in close cooperation

MS post 2020 - Collection of ideas

Current eMS is perceived as a reliable and user-friendly system meeting the requirements and information needs. But it has become quite large in all regards; one is home-made since there is an inherent tendency to overrate the differences between programmes and thus making things look overly sophisticated.

- A firm timeline in development would be utmost important to have planning safety; e.g. it is very burdensome to start with paper applications and transfer these later on to the system
- The new system should allow computing (e.g. aggregating data) so that manual calculations can be avoided
- The new system should allow to reflect also work flows and provide options to send notifications in order to track the stages – that would be extremely useful for its daily use as programme management tool
- It is important that assigning user rights is made easy
- Use of flat rates should allow for flexibility since there might be numerous options – thus no pre-fixed values but installed as adjustable tool
- The presentation of budget should include for simplified aggregate overviews
- Outputs at partner level are adding up when it is part of several work packages thus leading to wrong aggregate numbers – the link between outputs and work packages should be re-designed
- Option to enter payment dates for expenditure items to be able to group it

- As differences between programmes are often overrated, a system that is adapted to core procedures would be appreciated

II Panel: How to collect data from beneficiaries more efficiently?

Application stage

- Ensure common understanding of indicators (guide on indicators, application guide, discussion with beneficiaries, annexes to the Cooperation Programme (CP) and/or Call; using illustrative examples what to count and what not to count) – this is a process which needs to be started when programming and obviously needs recurring training!
- Outputs and deliverables should be clearly separated
    - Number of outputs should be kept small; number of deliverables might be large
    - Try to adapt to the specificities of the projects in order to report realistic figures
    - Distinction is important in order to prevent applicants from reporting even minor deliverables

Assessment / selection stage

- A graph outlining the intervention logic and therein the role of indicators has helped to build a common understanding among the MC members; the graph is used as a recurring reminder to the MC (please see the example in Annex 1)

Contracting stage

- Options for indicator clearing at the contracting stage are:
    - Mandate of the MC to the JS which allows to make small rather technical amendments among others also to indicators
    - Changes to indicators as part of conditional approval

Implementation

- In case of bilingual reporting in CBC programmes it is a help that the report is first drafted in one language and being checked by the JS; only upon the informal o.k. translation is done and the report is submitted officially
- It is important to focus reporting e.g. on key activities which lead to major outputs (instead of getting every single small deliverable listed)
- Underperformance – depending on the general perception and approach of the programme it does not necessarily pose a problem or has to lead to sanctions or financial cuts; obviously a small number of projects allows for closer contact to beneficiaries and thus also for a more case-by-case approach to change management
- At programme level underperformance in one investment priority and related output indicators might be balanced by incentives in subsequent calls (e.g.

potentially higher scores for projects contributing to certain specific objectives resp. delivering certain types of output)

Plausibility checks

- In case of proper front-of-pipe investment in clear-cut templates (application, reporting) and indicator systems, training and building common understanding, throughout all stages of the project cycle the 'ex-post' plausibility checks do not pose a major work burden
- Data quality checks (e.g. to avoid double counting) are rather a recurring element during the progress reporting, some projects might need more 'nudging' whereas others report compliantly after the initial briefings

General issues

- Data quality and availability differs between Member States (MS), IPA and ENI countries – this poses an additional challenge for programmes having an external dimension; this might require to develop specific checklists related to indicators (in order to check the reliability of data sources

Outlook post 2020

- More frequent reporting of data (3 to 6 times per year; frequency is subject to negotiations of the draft regulations) should not pose an extra burden for programme management: most recent data should be reported and data transfer should work via interfaces; it is not meant that beneficiaries should be asked 3 or even 6 times per year to update their data; still people in JS/MA are afraid that it will impose new, additional work routines
- The current system with the AI means that in practice the EC discusses at the Council information which dates back to 2017 (if periods for reporting and management verification are taken into account the data basis for reporting is even more outdated)
- It is important to see the greater picture – even data changes stemming from one programme are very small the data aggregate for all ETC programmes will show constant movement respectively change

III Panel: How to cope with result indicators?

The more tricky ones …

- DE-NL: uses statistical as well as survey-based result indicators (RI):
  - the result indicator on share of SMEs which introduce either product or process innovation (DE-NL) has proven to be one of the tricky examples of a statistical indicator: the share has been de-creasing; now it is a task of the impact evaluation to investigate on the reasons for the development but findings will be available only for the next report;
  - the second survey-based one is on the perception of the border as barrier – the repeated survey has shown that in the broader public the border is perceived increasingly as a barrier – thus also a lot of

qualitative arguments will be brought up in the discussion of these developments (as a reminder – there is the saying of the 'paradoxon of cooperation' meaning the more you cooperate the more obstacles you will identify …)

- Alps-Lake Constance-Upper Rhine (ABH): out of the 9 RIs, some of the indicators are statistical ones some are survey-based; the MA/JS has done an in-house look at result indicators and impact; it provides a preliminary idea what the programme has achieved so far (although – due to its limited scope – it cannot be considered as impact evaluation)
    - Statistical indicators: e.g. share of scientists active in cross-border R&D, R&D expenditure of enterprises in the programme area, number of skilled (academic, vocational education) staff, number of visitors to world heritage sites in the programme area– data based on surveys done by regional statistical offices and other bodies; some of the statistical indicators have posed challenges since measurement / territorial units are not fully comparable
    - Survey-based indicators: e.g. level of improvement of institutional cooperation (qualitative)

- Central Europe: the programme has set-up 10 RIs (one per specific objective); baseline has been set-up in expert panels (altogether 45 experts, nominated by MSs) and also for the update (survey) these panels have been used: method for the baseline is the same as for the update
    - The system is built on qualitative descriptions and scores (scores are used in order to see trends)
    - The expert survey has been done internally – it has been a lot of work but is seen as a good experience: the discussion itself with the experts was very good
    - Challenges: experts nominated by the MC change from time to time thus inherent issues of consistency, difficulty to agree on dates for brief online meetings to share observations (focus group)
    - It feeds into the operational evaluation as additional evidence that the programme is on track; the impact evaluation is supposed to start next year

Day 2

Performance review
David Alba, DG REGIO, Evaluation and Semester Unit

See presentation 'Reformance Review David Alba': http://www.interact-eu.net/library#2534-presentations-air-and-performance-review

- Achievement if:
    - 2 indicators in the PF: all indicators achieved at least 85% of the milestone value
    - 3 or more indicators in the PF: all indicators except one achieved at least 85%, 1 indicator at least 75% of the milestone value
    - Key implementation step (KIS) in the performance framework counts as an indicator!
- Please check all figures in the AIR carefully since there is no room for correction; EC Decision is issued automatically– if it includes a comment it means in the end a review/modification of the CP
- When data comes in checks will be run; the decision will be issued automatically – given the overall number of programmes there will be no time for quality checks; but now tools have been opened to desk officers for checks prior to submission
- There is no danger of re-allocation to other programme – the mechanism of the performance reserve is not applied for ETC
- In case there is the risk of a serious weakness in implementation it is best to start discussion informally with the desk officer (DO) since any decision on consequences is in the discretion of the DO; the DO will have also the tool for the check – please ask your DO for an informal check
- There is the option to extend the certification period; still time to submit payment claims to EC until the submission of the AIR to the EC
- Financial indicators refer always to total expenditure

Question & Answer (Q & A)

Q: What happens if a project is withdrawn at a later stage?

A: Basis for the decision of the EC are figures reported in the AIR.

Q: Is it possible to send a payment request to EC prior to submission of the AIR?

A: It is o.k. if the expenditure has been paid by the beneficiary before the end of 2018.

Q: What should be reported on issues affecting the programme performance?

A: The minimum required is to report along all indicators included in the performance framework meaning financial, output indicators and/or key implementation steps; but it could be more in order to provide consistent arguments.

A level of achievement below 65% means a serious weakness which needs to be explained – its root cause and the counter measures taken. This is best explained in the AIR (in section 15 of the template for ETC.

Q: What is the role of desk officers (DO)?

A: As has been explained a data quality checking system has been developed; it will issue automatic warnings (see slides on plausibility checks) on inconsistencies such as financial indicators higher than the sum of certified expenditure in payment claims, it will also check if data are complete; or overachievement of final target.

DO are the main discussion partners for programmes which encounter serious weaknesses.

Q: My programme is now in the middle of the impact evaluation; we have results dating from spring 2019 and we intended to submit it with the AIR; now we have been told that the cut-off date for the AIR is the end of 2018 meaning that we could not use anything from the evaluation report in the AIR?

A: As long as it is well explained it would be a pity not to have the information in the AIR thus pragmatically speaking the best possible use of the evaluation findings should be sought. It is also particularly valuable since the MC discusses the AIR.

Q: My programme just starts implementation; so we have no achievements related to output or result indicators?

A: You have to report what is the situation – so you will have to enter zero … even if you don’t like it.

Q: Should we report also on a key implementation step even if achievements for the corresponding output indicators are already there?

A: Yes, if the key implementation step is part of the performance framework you to report for it in order to avoid misunderstandings (key implementation steps were meant as a proxy).

Q: What will be the contents of the EC decision on the performance framework?

A: The EC decision will conclude on performing and non-performing priorities. It is one decision per programme on the performance review. There is no reserve but there is the statement of a serious deficiency (weakness).

All consequences (such as for example a financial cut or a suspension or the need for an action plan) is in the discretion of the DO. It is a manual case-by-case exercise to take a decision on that. Legal basis for suspension is not clearly visible.

You should take note of the decision and discuss it in the MC meaning to decide how to address the situation and how to deal with weakly or non-performing priorities – but there is no immediate consequence in legal terms.

Q: What might be consequences if a priority is in the grey zone meaning having achieved between 65 and 85% of the target in the performance framework?

A: It is for you to decide but it boils down to the question if – in the remaining programme period - the absorption deficiency can be changed or not or if a re-allocation of funds is required.