Mid-term Programme Management Evaluation of Austrian Genome Research Programme GEN-AU (Austria)

1. Introduction

This is an interim programme management evaluation of the Austrian Genome Research Programme GEN-AU, conducted in 2005. Its aim was to advise if the programme should be continued and in what ways if could be improved in terms of management and design and how it could develop a future impact assessment.

The GEN-AU programme was set up in 2001 in order to boost excellence, collaboration, development of young researchers and the visibility of Austrian genome research. It is long term oriented and invests in basic research activities, i.e. it does not expect immediate economic returns for firms, but rather builds the basis for future applications.

The programme evaluation was highly useful for the programme management as it allowed a sound justification for prolongation and improvement of the measure. The evaluation was a critical milestone, to support a 'go or no-go' decision for the measure.

2. Description of the evaluated measure

GEN-AU seeks to strengthen and network genome research activities in Austria. It promotes science, creates critical mass, boosts existing strengths, supports international linkages, creates new knowledge that is application oriented in the long run, supports technology transfer to industry (through the means of patents, creation of start ups, by strengthening companies, etc.) and thus also contributes to economic opportunities in the future (by creating jobs, etc.). It also seeks to build up a strong new generation of genome researchers and increase the public acceptance of genomics research in Austria.

As noted in the Evaluation Report: "In 2001 the Austrian Federal Ministry for Education, Science and Culture, launched the 'Austrian Genome Research Programme GEN-AU' (GENome Research in AUstria). The genome research programme was planned for a period of nine years. It receives funding every three years (consequently, the running time of GEN-AU is made up of three different phases), and approximately €10.7 million are spent on GEN-AU each year. The mission is to strengthen genome research in Austria and to foster networking among all relevant stakeholders and actors. In order to achieve this goal a variety of project types were developed: Large cooperative projects, network projects, pilot projects and projects addressing accompanying research in the social sciences. The project types differ in terms of the number of involved partners, their running time and their funding volume. In phase I of GEN-AU €27.8 million have been allocated to 23 projects run by 91 partner organisations. The majority of the funded institutions are located in Vienna, but a number of scientists in Graz, Innsbruck and Linz received financial support, too".

The Austrian Genome Research Programme GEN-AU is a top down research programme with a focus on collaboration, networking and the support of a younger generation of researchers.

The initial budget for the phase covered by the evaluation (2001 to 2005) amounted to €31.7 million plus €1.85 million for ELSA projects (analysing **E**thical, **L**egal and **S**ocial **A**spects).

The start date was in 2001 with the end of the first phase in 2005; the programme was terminated in 2011. The evaluation budget for this evaluation was €70,000 and the

evaluation study duration was 6 months. The evaluation study was undertaken by Joanneum Research with inputs from TIA Consulting (USA) and KMU Forschung Austria (the Austrian Institute for SME Research).

The programme was a cornerstone of the build up of new structures to support biotechnology in Austria as part of the science driven strategy. It also sought to build up capacity and practices in order to access better the EU Research Programmes and to broaden and sharpen industry activities in biotechnology in Austria.

3. Designing evaluation study

The evaluation was commissioned because of the need to legitimate further funding in support of the measure. The measure was funded through a special budget held by the Austrian Council. The Council demanded that an interim evaluation should be undertaken. Equally, the ministry wanted to check if the programme's implementation and management structures are appropriate (for the purposes of learning, adjusting, etc.).

The Terms of Reference were designed using a moderated discussion process. An external evaluation expert moderated this discussion, taking advantage of the established Evaluation Standards of the Austrian Platform Research and Technology Policy Evaluation as a baseline.[1] This moderation made clear that the evaluation was about the design and basic approaches as well as the management: it was not concerned with the scientific results of the research supported under the programme.

The questions addressed in the ToR were (source: evaluation report, executive summary, slightly paraphrased):

- Should the programme be continued and what room for improvement is there?
- Can the strategic goals of the programme be achieved?
- How is the programme positioned within the wider Austrian system?
- Is the programme sufficient in supporting a next generation of scientists?
- Are hot topics in genome research addressed?
- Are networking and communication processes functioning and are they sufficient?
- How can female researchers be better mobilised?
- Should the programme application guidelines be changed to include structural and cooperation issues (rather than only technical and scientific content)?
- Are the tools in the programmes (cooperation, networking, pilot projects) the best/most appropriate ones to achieve the goals? Could any further tools be suggested?
- Are sufficient public relations activities undertaken?
- Is there an adequate networking of policy around genome research?
- Where on an international level does the collateral research programme on Ethical, Legal and Social Aspects (ELSA) stand?

The Terms of Reference did not specify the evaluation methods and were thus entirely open. They were oriented towards an "appropriate method mix", asking for more than

[1] http://www.fteval.at/cms/en/home/standards/standards-platform.html

one method to be used, but did not specifically mention any particular methodology that should be used. This followed the normal operating evaluation standards in Austria.

Equally, the data sources and baseline were left open to the evaluators and the data needs were not discussed in the terms of reference.

In retrospect, despite the fact that one of the tasks was to understand the role of the programme in the overall science and innovation system and the relative meaning of communication, collaboration and networking in the genomics research area, this aspect was not followed up and deepened as it should have been. During the evaluation, the focus both of evaluators and client turned much more to the actual effectiveness and efficiency of the programme itself, rather than its overall systemic role.

As regards the networking requirements and effects themselves, the ToR as the programme itself indicate an understanding whereby more networking is good for genome research. This underlying assumption, however, seems to have led to an overemphasis on the issue of networking requirements in the programme and then subsequently in the evaluation. Consequently, the evaluation did not tackle systematically the question of potential negative effects of pushing too hard for networking (see below).

4. Implementing evaluation: methodology and process

The evaluation was externally conducted and there was an interactive definition of the terms of reference within the ministry and with an external consultant on the basis of which an open tender procedure was started. The external consultant was not asked to support the aims of the evaluation, but rather comment on the requirements and methodological issues for an evaluation that would be asked to answer those questions. The external consultant was a respected and well known specialist in the field of evaluation in Austria.

In accordance with the multi-aim, multi-instrument nature of the programme, the evaluation used a mixed method approach. The open formulation in the Terms of Reference helped to design a methodology that was fit for purpose, but it was agreed upon in the early stages of the process by the client.

The evaluation started off with a desk based research and interview-supported analysis of the pre-history of the measure. This was carried out in order to understand the motivation, the perception of needs and opportunities in the area (politically, scientifically, economically, institutionally) and to establish the external validity of the programme.

For the internal validity, i.e. the match of goals, instruments, activities, outputs, outcomes and (desired) impacts, a logic chart of the programme was developed. This was done in cooperation with the programme management team to establish a shared understanding as a basis for the evaluation design.

This logic chart was instrumental (see annex to this document). It connects the objectives of the programmes with the various activities and the (potential) outcomes. This allowed a clear communication about the programme and its impact logic as well as a clear definition of evaluation steps and dimensions.

On that basis, further evaluation methods employed included desk research, compiling and analysing monitoring data and other programme management data. The evaluation was understood to be very interactive and thus the evaluation team conducted a broad interview programme, interviewing 72 actors representing the programme management, policy context, economic impacts, participants (established and young researchers) and non-participating researchers and firms and finally international genomics research experts.

Those interviewed proved to be very cooperative, all interview wishes were fulfilled, signalling a significant interest on the part of the research and broader stakeholder communities to report on their attitudes, the effects or their specific criticisms of the programme.

In order to obtain a deeper insight into the programme's mechanisms and processes a set of case studies was conducted, with one in-depth case for each type of instrument:

- Collaboration
- Network
- pilot project
- accompanying ELSA projects (ethical, legal, socio-economic aspects of the programme).

Furthermore, the study looked at some comparable international programmes in order to comparatively assess GEN-AU in terms of its design and management characteristics.

Finally, and importantly, the evaluation conducted a broad network analysis to assess the extent to which the programme managed to achieve its major goals to support both networking and the better exchange of research-related information. The evaluators introduce the SNA approach critically. They elaborate on the meaning of networks for scientific endeavours and dissemination of results. The Social Network Analysis approach looked at social exchange (communicative relationships) and the exchange of resources relevant to research. This analysis was based on an online survey of all participants in all types of projects using a standardised questionnaire. The questions centred around interaction patterns, such as scope and level of communication, and resource exchange questions. This enabled the depiction of communication networks and functional networks (exchange) at various levels (sub-projects, projects) and subsequently at programme level (to see the shape and functioning of networks across all cast of actors). The target group were all participants in the project, both in the scientific projects and in the ELSA projects. The response rate was 52% on average, whereas networking and collaboration projects showed a higher response rate than ELSA projects.

The evaluators derived their recommendations from a thorough analysis of all data gathered, both qualitative and quantitative. The tools for this analysis were in-built in the methods used, i.e. interpreting the interviews and cases as well as the quantitative data of the Social Network Analysis. There was no in-depth discussion as to the implications of the programme for the innovation system more broadly.

Further, the analysis did not rely on any baseline data (level of cooperation and communication prior to the programme) as information on the baseline attributes was not sufficient. Also, it did not discuss in depth the need for and limits of networking in the first place.

The evaluators derived their recommendations from a thorough analysis of all data gathered, both qualitative and quantitative. The tools for this analysis were in-built in the methods used, i.e. interpreting the interviews and cases as well as the quantitative data of the Social Network Analysis. For time and scope reasons, the approach did not allow for an in-depth discussion as to the implications of the programme for the innovation system more broadly.

Lastly, one sensitive issue concerned the release of evaluation results to the media, as one newspaper apparently obtained some of the results and published them before the official launch of the report.

5. Effectiveness and efficiency of evaluation

The design of the evaluation Terms of Reference was state of the art, especially the moderated process which allowed for the provision of external advice. Furthermore, the Terms of Reference were open, broad and invited methodological variety. Such an approach is appropriate in a system with a rich evaluation community. However, for systems in which the evaluation of innovation and research policy matters are not that advanced, a slightly more prescriptive approach would have been more appropriate.

The basic mix of methodology that was finally chosen was highly appropriate for what the Terms of Reference asked. Most of the questions could not have been answered without a large programme of interviews.

The limitations of the study concerned the budget and the timing. It was felt that €70,000 was insufficient to cover the mix of approaches and the large number of interviews required. The timing (6 months) was also too ambitious (i.e. too short). Together, these factors led to a rush in terms of developing the conceptual basis of the programme as a framework for the assessment itself. A stronger discussion of the merit and limits of networking and cooperation in genomics research as well as a discussion of the status of genomic research in the Austrian system would have improved the validity of the results as it would have made assumptions explicit or even questioned them (e.g. is more networking really better?). Such a discussion would then have improved the interpretation of the network analysis, clarifying the conditions under which more exchange or broader exchange is better, and when exchange and collaboration, when imposed by the funding, might be detrimental.

As for the methodology, one limitation concerned the number of case studies. One case per type of measure cannot give a clear picture as to the typical benefits or problems encountered. Here, a focus group approach bringing together actors funded under the same scheme to discuss the support and mechanisms could have been a better option, at least as a complement. Alternatively, and with a higher budget, a broader case study approach could have been conducted (which then, however, would have been quite a burden for the field).

The data access through the monitoring system was generally adequate and extensive. However, as in many programmes, the monitoring system itself was not appropriately designed to support all the evaluation questions in a systematic way.

A principle consideration, with any evaluation, is the match of timing for the evaluation on the one hand and the political decision-making cycle on the other hand. In this case, the evaluation did not quite fit the decision-making cycle at higher policy levels. This means at the point in time when the political decisions had to be prepared the evaluation results could not be fed into the process. If the evaluation process had been in-built in the programme cycle at the outset, it could have been conducted in order to fulfil the need for discussions on prolongation and changes at the right point in time.

6. Results of the evaluation

On the basis of the multi-method approach described, the evaluation concluded that the programme, by and large, was a success in funding highest level research and establishing novel networks and collaborations. It must be stressed that the evaluation did not actually measure the impact of the research on knowledge production or even societal or economic impact, but management and design appropriateness. Here, the management, by and large, was assessed to be effective and efficient and the basic design and the tools offered were assessed to be fit for purpose. The network analysis indicated stable and trustful partnerships having been created with no detectable major lock-ins. The evaluation clearly recommended prolongation.

However, it also suggested a range of improvements. Those centred around operational issues (e.g. transparency of review procedures, timeline an deadlines in the management, agencification of the management), strategic issues (e.g. GEN-AU should more explicitly complement and interact with other measures and should establish a better hierarchy of strategic goals that reflect more realistic expectations as to what economic and societal impact a scientific programme can have) and monitoring and evaluation suggestions (asking for a real ex post *impact* analysis and better continuous monitoring).

Out of those management oriented recommendation, the most important one, agencification of the programme management, was implemented, the GEN-AU programme and the management team are now part of a Federal Agency (FFG). Further, an impact evaluation is being prepared along the lines suggested in the evaluation.

7. Conclusions and lessons learned

This evaluation is concerned with a research programme rather than an innovation programme. However, it offers lessons that are of a more general nature. It is an excellent example of a sound match of evaluation methods and processes on the one hand and evaluation questions and the nature of the measure (goals, instruments, target groups, context) on the other hand. The multi-method approach enabled the evaluators to gather qualitative and quantitative insights to improve the management. Most importantly, the programme management is now shifted to an innovation policy agency rather than done within the Ministry. Further, this evaluation is the basis for a sound impact evaluation that is now in preparation. In addition, due to the logic chart approach that included the impact dimension, it enabled the evaluation to set up some preliminary ideas for the design of an impact assessment later on in the process. The logic chart approach applied was exemplary, to be used as communication tool but also as a tool to establish what is measured and why and what the measurement represents in the programme flow. This also allowed a critical appraisal of the programme logic itself.

While the mix of methods was fully appropriate for the evaluation question, there are two lessons on how these types of evaluations might be improved:

a) Qualitative case studies in this evaluation, because of the limitation in numbers (early stages) are always limited in terms of the scope of information provided. As multiple case studies were not possible due to time and budget constraints, a focus group approach might have been a better option, bringing together actors within programme lines, but from different projects/collaborations, to exchange ideas and discuss issues.

b) a social network analysis needs a clear baseline and a rationale for the interpretation of the findings. This is a general problem with SNA, and thus monitoring data and baseline date is paramount when SNA is to be applied. While the report shows a very advanced understanding of the limits of network analysis, it could not discuss in depth the need for and limits of networking in the area of life science research in Austria in the first place. This comes back to the shortcoming of a limited conceptual discussion on the need for and values of networks and the best level and scope of networking in the first place. In the case of GEN-AU, the findings are highly interesting and can be (and were) used for further discussion in the community. However, this is a basis for discussion and interpretation with stakeholders, as there is now meaningful way to establish in quantitative terms how far away the community funded by GEN-AU has been from some kind of optimum networking is not possible. Luckily, the vast amount of interviews and qualitative data qualified the SNA findings and enabled interpretation.

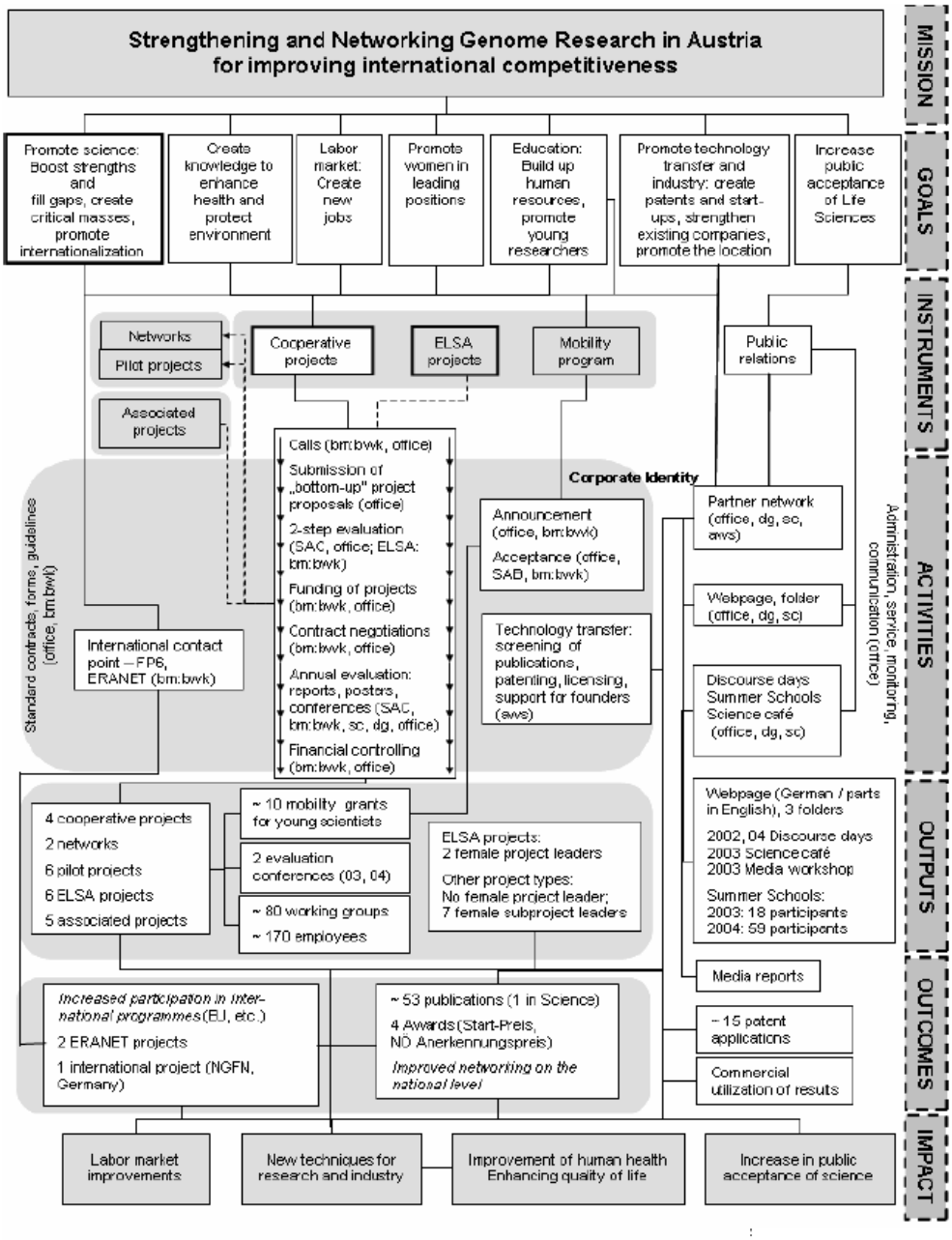
The general problems encountered in the evaluation are mainly related to the lack of time to question the underlying assumptions of the programme more thoroughly. Further, it was impossible, in the time and within the budget given, to thoroughly establish what the relative role of the programme within the broader innovation system of Austria was.

The evaluation could take advantage of the formal standards for evaluation that are established in Austria. This reflects the high standards in policy and in the analytical community as regards evaluation in Austria. In countries with a less well-established evaluation culture and practice, these standards (or some variant of them) could be used as well as guidelines for evaluations. This would be especially important to create a common language for policymakers, stakeholders and evaluators in the ERDF funded innovation policy schemes, as this brings together different communities with less experience and familiarity with innovation policy and its evaluation.

Interviewer and case study author: Jakob Edler, Manchester Institute of Innovation Research

Case study completed: October 2011

Annex: The Logic Chart of the GEN-AU evaluation



Source: Zinoecker, Klaus et al. (2005): 'Austrian Genome Research Programme GEN-AU': Mid Term Programme Management Evaluation. Final Report, Vienna; p.22.