Evaluation of the Research and Technology Development Projects and Programmes

Constanta BODEA, Narcisa CIOBOTAR, Vasile BODEA
Academy of Economic Studies, Bucharest, România

The paper presents the specific ways in which indicators and artificial intelligence methods and tools can be used for the evaluation of research projects and programmes. The author's research purpose is to improve the programme ex post evaluation and ex ante impact assessment thought the development of a improved set of strong integrated research performance indicators, structured according to the results chains and comprehensively described using a standard indicator template; the development of data sets and databases for projects and programmes evaluation and, finally the development of projects and programmes evaluation techniques, based on database and machine learning technologies. Using these methods a new and better understanding of the scientific, technological, human resources, structuring, economic, social, environmental etc impacts of national and European research programmes is possible. The research is financed by the Minister of Education and Research, IDEI programme.
Key words: *KPI, RTD project, RTD programme, data mining, project portfolio, indicator template.*

1 Introduction

The evaluation of the research projects and programmes is a data intensive endeavour and for this reason the reliable and comparable research and technology development (RTD) statistics and indicators are of crucial importance. Eurostat, Statistical Office of the Europe an Communities, Unit F4: Education, Science and Culture statistics assures RTD data collection, integrity, and access by the public, quality and dissemination. The RTD data are compiled in accordance to the guidelines laid down in the [1].

According to Frascati manual, in the context of the knowledge-based economy RTD statistics need to be examined within a conceptual framework that relates them both to other type of resources and to the desired outcomes. "While indicators of RTD output are clearly needed to complement input statistics, they are far difficult to define and produce" ([1], pg. 17).

The improvement of RTD performance indicators might be achieved by adopting a results-based and project portfolio approach and assuring the indicator documentation through a standard and comprehensive indicator description, named the indicator template.

2. The European practices for the RTD programmes evaluation

The starting point of the European Commission's approach to evaluation of research and technological development (RTD) programmes goes back more than twenty years. Before the 1980s, a series of experimental activities had established the peer review as the EU approach. Independent experts, constituted into panels met usually every 6-8 months to review the results of surveys and interviewed the key stockholders. The main evaluation criteria were scientific and technological quality of research, effectiveness of programme management and contribution of results to the general progress of science and technology.

The institutionalization of the evaluation activities has started in early 1980s when a centralised evaluation unit was established by DGXII. This unit was in charged of the development of improved evaluation schemes and their implementation ever since. It had defined the first multi-annual plan of action for evaluation in 1983, which initiated the operational evaluation system setting up process. The second plan of action (1987-1991) consolidated this system. Each Spe-

cific Programme within the Framework Programme (FP) was evaluated by external experts at least once during the implementation period. Not all the FP activities fell under the aegis of DGXII. For example, the programmes related to Information and Communication Technologies (ICT) is administrated by DGXIII. The DGXIII had developed its own evaluation system and the first evaluation was a mid-term review of ESPRIT in 1985. A panel of experts had held face-to-face meetings with relevant organisations and conducted a questionnaire survey of participants. Other similar evaluations followed, focusing on the objectives and priorities reviews. DG Research has a long-established Evaluation Unit which has acted over the years both as a focus for methodological and strategic development of evaluation, with initiatives such as the SPEAR research programme and as the node for a succession of networks of evaluators and responsible officials, and as the unit responsible for the practical implementation of the evaluation of the Framework Programme as a whole and for those sub-programmes within the ambit of DG Research. External scrutiny of the evaluation approach for the Framework Programme is carried out by an Evaluation Sub-Committee of CRET (Committee for Scientific and Technical Research) the Commission and Council's advisory committee. The Sub-Committee consists of representatives of the Member States and the Commission.

In 1994, the Commission introduced an evaluation scheme based on continuous monitoring reporting annually and five-year assessment carried out midway through programme implementation. This evaluation time line assures two previous programmes to be included (the ex post evaluation of the previous programme and the mid-term appraisal of the current programme) and produce results for the next FP proposal preparation. Figure 1 presents this time type for FP5, FP6 and FP7.

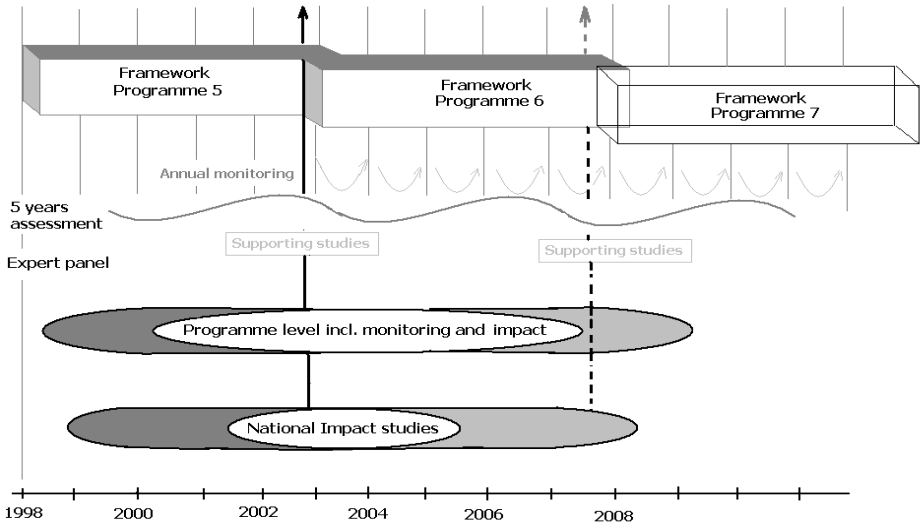

Fig.1. Time cycle for Community RTD Evaluation

Using this five-year assessments model specific programmes/key actions were assessed by panel supported by different tools. Two important studies ([2] and [3]) have focused in particular on new or more theoretically grounded approaches to evaluation.

The EPUB Toolbox was compiled by a network of evaluation specialists in the European Union with financial support from the STRATA programme within the EU Framework Programme. It had been observed that EU policy-making processes, and also those in many Member States, applied evaluation in a way that very much resembled what the EPUB authors understood as "monitoring". For other steps of the evaluation process, peer reviews and expert groups have mainly been used as described above. However, the EPUB Network argued that analysis of available evaluation techniques and experiences

together with methodological advances showed that the evaluation process can offer more to decision-makers than currently is made use of. To address this problem, the Network analyzed the methodologies for evaluating the socio-economic impact of Research and Technological Development (RTD) policies over a two-year period. The resulting toolbox provides policy-makers, scientists and practitioners with an overview of the main evaluation concepts and methodologies, outlines their strengths and limitations, and sets them in relation to the policy context. Emphasis is set on a practice-oriented presentation.

Legal basis for present EU evaluation practices includes articles of the legal texts, financial regulation and rules for implementation of financial regulation; communication on evaluation standards and communication on better regulation.

The levels of evaluation are:

– *European Commission/Policy level*. The relevant documents are the reports to EP, Council, CoA. Evaluation is done according to the evaluation framework (including standards).

– *DG Research/RTD Policy and programme level*. The relevant documents are the reports to EP, Council, and Commission. The evaluation is done as ex ante impact assessment, five-year assessment, ad hoc studies at specific programme level, Annual Activity Report/Annual Management Plan, RTD evaluation framework, and member state impact evaluation.

– *DG RTD/Project level*. The relevant documents are the reports to DG, Commission, MS, and Programme Committees. The evaluation is done according to the evaluation manual, using evaluators database

A study [4] has looked at the issue across all areas including some reference to RTD. Key findings included:

- Direct take up and use of results of evaluation occurs but is not the norm
- Evaluation use is more likely to become cumulative across evaluations
- Evaluation is more likely to be influential than sole cause of subsequent action
- The process of evaluation is seen as a useful opportunity fir reflection and clarification of frameworks
- There is no single model of good practice
- The nature and degree of use of evaluation appears more determined by overall organizational arguments for dissemination, cross-DG consultation and routine liaison between those in the evaluation function and operational policy colleagues

Many problems were identified with the current evaluation approach. It is considered to rely excessively upon questionnaires to participants and the questionnaires continue to be given more prominence in reporting than the simple filter role envisaged by the ETAN panel. The problem is exacerbated by a continuing decline in response rates which call into question the validity on the resulting data. Thinking is continuing on how to develop new approaches, particularly in the light of the broader range of instruments now being used in the Framework Programme.

3. A results-based approach applied to the evaluation of RTD projects and programmes

The results-based management is a life-cycle approach to management that integrates strategy, people, resources, processes and measurements to improve decision-making, transparency, and accountability. The approach focuses on achieving outcomes, implementing performance measurement, learning and changing, and reporting performance. We consider the results-based research evaluation as very important for the evaluation of research policies and programmes, assuring management improvement (focusing interventions to achieve results), marketing success (showing general program progress and demonstrating results, you can win public support), and increased accountability (demonstrate results against money spent). A good integration of ex ante with ex post impact analysis is possible.

Principles of results-based management are: focusing on results in all the management

phases, aligning policies and programmes according to the results, keeping measurement & reporting simple and cost effective, managing for, not by, results and using results information for learning and decision making. The planning and implementation processes are integrated by the following results chains (see figure 2):

a) planned results chain (developing causal chain, defining appropriate programme objectives and outcomes to be achieved, developing the results indicator plans);

b) actual results chain (reporting on outputs and outcomes, adjusting the programme to achieve outcomes, review and validate the causal chain, reporting on the programme outcomes)

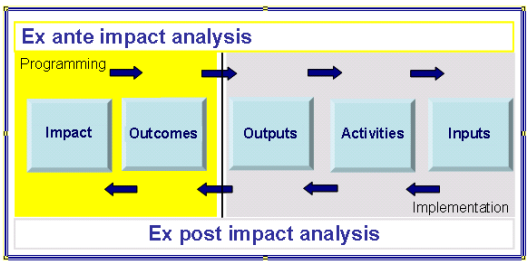

Fig.2. Results-based approach - The results chains

The main results-based evaluation approach requirements are (see figure 3):

- Identification of the program objective and intermediate outcomes which are critical to achieved it
- Development the entire hierarchy of objectives showing the causal linkages (objectives -> sub-objectives)
- Clarifying the evaluation purpose
- Identification of evaluation question
- Select appropriate methods based on the questions to answer
- Defining the performance indicators
- Prepare data collection and analysis plan

The central concept of the results-based approach is the key performance indicator. The Key Performance Indicators (KPI) are financial and non-financial metrics used to quantify objectives to reflect strategic performance of the policy or the programme. For example, key performance indicators for evaluating the university international collaboration (research internationalisation) might be selected from the following: research mobility, externally funded programmes, (total international/bilateral project number, total project value, incoming funds, co-funding, geographic coverage of partners, number of project coordinated, success rate, number of patents, etc), effective participation in ERA (number of created research posts, number of other posts created, n umber of outgoing researchers, number of returning researchers, number of spin off companies, number of centre of excellence, number of publications in international scientific journals etc).

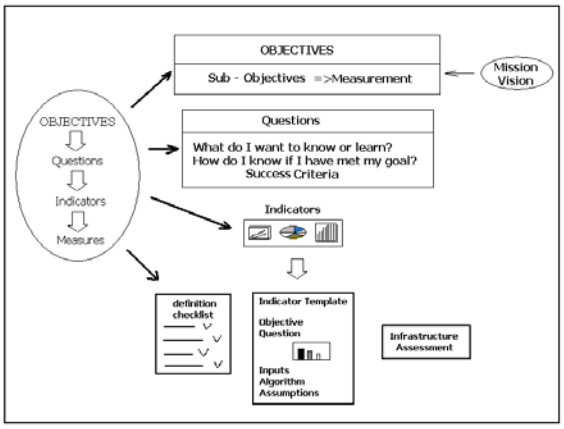

Fig.3. The results-based approach

Performance indicators should be SMART (specific, measurable, attributable, realistic, and targeted). Each indicator must be properly defined and used. This is meant we should define an indicator template, specifying: definition, baseline & targets, data acquisition method; data analysis & reporting method and critical assumptions. It is difficult to decide the performance indicators to be used, because it is used a wrong question ("What indicators should I use?"), instead of: "What decision do I want to make"? Other questions of interest for the policies makers are: "How will I know if I have achieved the goal?" and "How I know if I am making progress toward the goal?" Therefore, several indicators classes should be differentiated, according to the results chains (see figure 4).

Using the results-based approach and a set of research performance indicators, structured according to the results chains it will is possible to estimate the impacts of research ef-

forts and programmes on high-level policy goals.

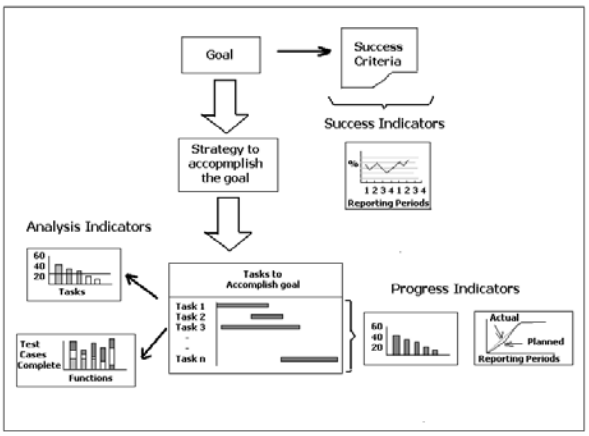

Fig.4. The indicators classes

4. A project portfolio approach applied to the evaluation of RTD projects and programmes

The alignment of the research programme to the policy goals can be assessed by applying project portfolio approach to develop measures and associated indicators to measure the effectiveness, efficiency, costs and benefits, and economic and sustainability impacts of research activities. We propose to apply for RTD evaluation the following related methods:

- Investment portfolio score card
- Project portfolio score card
- Portfolio matrix models
- Project portfolio dynamics reports
- Networks of projects - methods for networking of projects
- Chains of projects – project chains management methods

5. Artificial intelligence methods for the evaluation of RTD projects and programmes

The RTD indicators analysis might be based on the analytical methods, coming from business intelligence, especially data mining. Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees). Consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction.

Data mining can be performed on data represented in quantitative, textual, or multimedia forms. Data mining applications can use a variety of parameters to examine the data. They include association (patterns where one event is connected to another event), sequence or path analysis (patterns where one event leads to another event, such as the birth of a child and purchasing diapers), classification (identification of new patterns), clustering (finding and visually documenting groups of previously unknown facts, such as geographic location and brand preferences), and forecasting (discovering patterns from which one can make reasonable predictions regarding future activities.

As an application, compared to other data analysis applications, such as structured queries (used in many databases) or statistical analysis software, data mining represents a *difference of kind rather than degree*. Many simpler analytical tools utilize a verification-based approach, where the user develops a hypothesis and then tests the data to prove or disprove the hypothesis. In contrast, data mining utilizes a discovery approach, in which algorithms can be used to examine several multidimensional data relationships simultaneously, identifying those that are unique or frequently represented.

Reflecting this conceptualization of data mining, some researchers consider data mining to be just one step in a larger process known as knowledge discovery in databases (KDD). Other steps in the KDD process, in progressive order, include data cleaning, data integration, data selection, data transformation, (data mining), pattern evaluation, and knowledge presentation. Data mining has become increasingly common in both the public and private sectors. In the public sector, data mining applications were initially used as a means to detect fraud and waste, but they have grown also to be used for purposes such as measur-

ing and improving program performance.
In order to apply the data mining techniques for the RTD indicators analysis we should describe the content of the existing RTD datasets and databases using an indicator template. As starting point we consider the indicator template developed by SEI, *Carnegie Mellon University* (figure 5).

Date	
Indicator name/Title	
Objective	
Questions	
Visual display	
Perspective	
Input(s)	
Data Elements	
Definitions	
Data Collection	
How	
When/How often	
By Whom	
Forms	
Data Reporting	
Responsibility for reporting	
By/To Whom	
How often	
Data Storage	
Where	
How	
Security	
Algorithm	
Assumptions	
Interpretation	
Probing questions	
Analysis	
Evolution	
Feedback Guidelines	
X-reference	

Fig.5. The indicator template

According our knowledge it will be for the first time that these kinds of techniques will be applied on RTD metadata. Usually these techniques are applied on data in order to discover hidden patterns. What we expect to find out are the indicators similarities and differentiations, the indicators clusters, the association between indicators, the most important input factors of indicators definition. According the results-based and project portfolio approach we will evaluate the discovered patterns.

6. Conclusions

According our opinion, for the research policies and programme evaluation process, the following requirements has to be solved as soon as possible:

– more results oriented;
– based on limited number of verifiable objectives with SMART indicators to monitor the achievement (build robust hierarchy of logically interdependent outcome objectives, and a limited number of realistic and appropriate indicators);
– well defined and integrated evaluation information system (a clear documentation of the indicators plan, using an indicator template);
– strengthening evidence base (a systematic data collection approach, monitoring based on common performance indicators, independent assessment of scientific and technical quality and progress mid and end of term);
– strong connection between ex ante and

ex post evaluation and coordination between different member states;
– new methods and tools;
– coordinated evaluation studies (horizontal assessment and assessment of impact & achievement at portfolio, programme and higher levels against strategic objectives & indicators set in clear programme logic.
Addressing these requirements our research will have an important contribution to go beyond the state-of-art. The results-based approach assures a *consistent performance indicators structure*, according to the results chains and a strong connection between ex ante and ex post impact evaluation. The project portfolio approach assures a *tight integration* of the research performance indicators, especially between policies goals and programme results. Some important project portfolio management methods, like the investment portfolio score card, project portfolio score card, portfolio matrix models, project portfolio dynamics reports, networks of projects - methods for networking of projects, chains of projects – project chains management methods might be adapted to the research programme evaluation. And, finally defining a comprehensive indicator template allows a better understanding, the development of a detailed analysis and a common platform for sharing datasets and databases.

References

[1] *Proposed standard practice for surveys of research and experimental development* (2002), Frascati manual, OECD.
[2] *Assessing the Socio-economic Impacts of the Framework Programme* (2002), Gheorgiou L, Cameron H and Rigby J (eds) http://www.les.man.ac/prest
[3] *RTD Evaluation Toolbox - Assessing the Socio-Economic Impact of RTD-Policies* (2002), Gustavo Fahrenkrog, Wolfgang Polt, Jaime Rojo, Alexander Tubke, Klaus Zinocher (eds) http://epub.jrc.es/docs/EUR-20382-EN.pdf
[4] *Use of Evaluation in the Commission Services* (2002), Technopolis France (http://europa.eu.int/comm/budget/evaluation/pdf/use_of%20_evaluation_final_report.pdf)
[5] *Managing National Innovation Systems* (1999), OECD, Paris.