

# Tools and Indicators for Community Research Evaluation and Monitoring

## Final Report

*Volume 1: Summary of findings and recommendations*

**July 2009**

*Prepared for*  
**European Commission - DG Research**  
Unit A.3 – Evaluation and Monitoring of Programmes

*Authors:*

*Michael Braun  
Benedetto Lepori  
Emanuela Reale  
Stig Slipersaeter  
Aris Kaloudis  
Ghislaine Filiatreau  
Philippe Larédo*

Submitted by



Bad Camberg,  
Germany



Paris, France



Oslo, Norway



Rome, Italy



Paris, France

***Project coordination and contact:***

Proneos GmbH  
Limburger Strasse 33  
D-65520 Bad Camberg  
Germany

*This study contributes to the Research Directorate-General's initiative to develop a new generation of evaluation and monitoring instruments for the Framework Programmes (FPs). It explores the potential of Science and Technology (S&T) indicators for FP evaluation and monitoring and proposes a methodology for their systematic use. Based on an analysis of FP context, current practices and future requirements improvement potentials are identified. The resulting proposed methodological framework includes relevant indicators and indicator categories, principles for their application and pathways and actions for an efficient implementation.*

*This final report summarises the study findings. It is organised in two volumes:*

*Volume 1 summarises the study findings and conclusions. Based on a brief analysis of current situation, future needs and indicator potential, an indicator framework and principles of its application are proposed. To implement this framework successfully, necessary preconditions are identified and recommendations are formulated.*

*Volume 2 provides a complementary in-depth discussion of the framework and of its key elements. Its core is a detailed description of identified indicator categories and indicators and an in-depth discussion of their status, application, methodological potential and limitations. Further elements are an analysis of the FP's current evaluation and monitoring practices and a discussion of the principles of indicator application, especially in evaluations, based on the state-of-the-art in indicator science.*

*This is necessary because both the complexity of FP evaluation and the nature of indicators make it impossible to design ex-ante a single set of indicators which is suitable for all types of future FP evaluations. To provide the necessary evidence, indicators must be chosen according to specific evaluation questions. Main methodological choices and issues are discussed in each indicator domain and possible options and alternatives, as well as their implications, are highlighted.*

*The study has been conducted and the report has been created by a team consisting of Michael Braun (Proneos GmbH, Bad Camberg), Benedetto Lepori and Ghislaine Filliatreau (Observatoire des Sciences et Techniques, Paris), Stig Slipersaeter and Aris Kaloudis (NIFU STEP Innovation Studies, Oslo), Emanuela Reale (CNR CERIS, Rome) and Philippe Larédo (LATTIS, Université de Paris Est).*

*This work has been made possible by the support, background information and feedback provided by the involved European Commission services, especially DG Research's Unit A3 – Evaluation and Monitoring of Programmes – and by many inputs from Bianca Poti (CERIS, Rome) and other colleagues from the scientific community and practitioners from research and industry. We are particularly indebted to Yann Cadiou (former member of OST). His contributions have been vital for the Chapter on Scientific Output Indicators. Thank you very much to all.*

*All views expressed in this report are those of the authors. The study does not represent the official view of the European Commission.*

## Table of contents

### Volume 1

Summary and recommendations .....	1
1 Situation and challenges .....	5
1.1 The overall system of FP evaluation practices .....	5
1.2 FP monitoring and related reporting .....	7
1.3 Strategic Planning and Programming (SPP) cycle .....	8
1.4 Available data for indicator production .....	9
1.5 Challenges and requirements for the proposed indicator-based framework .....	9
2 Proposed indicator framework .....	12
2.1 Systematic use of indicators for FP evaluation and monitoring .....	12
2.2 Proposed indicator domains and indicators .....	13
2.3 State-of-the-art and recommendations for further development of the methodological framework .....	19
3 Application of the indicator-based framework in evaluations .....	22
3.1 Integration of indicators in evaluation logics and workflow .....	22
3.2 Identification of programme rationale and establishment of the link between policy/programme objectives and achievements .....	23
3.3 Formulation of evaluation questions .....	25
3.4 From evaluation questions to evidence and indicators .....	27
3.5 Integrated, life cycle-oriented framework for FP evaluations and management .....	30
3.6 Recommendations for implementation .....	31
4 Application of indicators in FP monitoring .....	33
4.1 Development options for the FP reporting ‘landscape’ .....	33
4.2 Indicators for FP monitoring .....	34
4.3 Complementary ‘health check’ function .....	37
5 Necessary prerequisites .....	39
5.1 Priorities for further development of the data base for evaluation and monitoring .....	39
5.2 Process and structural prerequisites .....	41
Literature quoted and suggested readings .....	42

### Volume 2

(separate document)

*(Detailed analysis of FP context, current evaluation and monitoring practices, methodological background and in-depth description of core elements of the proposed framework, in particular of the proposed indicator-based framework and indicators)*

## Summary and recommendations

The Framework Programmes (FPs) are the key instruments for realising Europe's increasingly ambitious research policy objectives. Therefore, progress in their implementation and achievement of their objectives must be thoroughly followed. This requires an extension and enhancement of current evaluation and monitoring practices.

This study shows that S&T indicators can make substantial contributions to the improvement of evaluation and monitoring practices. Their wider and systematic use would certainly contribute to transparency, efficiency and credibility of FP evaluation and monitoring.

For this purpose, an indicator-based framework is proposed. It builds on six indicator domains, covering the areas of programme management; FP participation and funding, scientific results, technological results, economic and social impacts and structuring of the ERA. To apply the indicators which are identified in these domains efficiently, important prerequisites are (1) a good understanding of the role and value added of indicators in the evaluation process, (2) the integration of indicator production and application in the evaluation process, (3) a sound methodological background and (4) the availability of high-quality data and necessary support structures.

### **Role of indicators in evaluations**

There is no lack of possible indicators which could contribute to FP evaluation and monitoring. But especially indicators for evaluation purposes are very different from traditional, routinely produced generic indicators. They are mostly ad-hoc indicators which need to be customised for this highly specific application, depending on evaluation context and data availability. Accordingly, indicators should be chosen not only as a function of their technical quality, but also on the basis of their ability to provide new (partial) evidence and to nurture the evaluation debate. This may imply that under certain conditions indicators of a partial and contestable nature which are available on time and which provide new 'food for thought' can contribute more than sophisticated indicators which are produced too late or are too far from the core of the evaluation debate. As a consequence, a close integration of evaluation activities and indicator design, production and application as well as a careful management of the interface between both processes are necessary.

This means that an enhanced indicator use may even require an additional resource investment to produce indicators, to integrate and exploit them in the overall evaluation exercise and to collect and process the necessary data.

### **Design of tailor-made indicators**

Indicator design must be driven by precise and explicit evaluation questions. A major effort should be devoted to specifying these questions, checking their relevance in view of main programme objectives and key evaluation questions and to deriving relevant indicators.

#### **← Recommendation 1:**

*Those carrying out or commissioning evaluations should make a dedicated effort at the beginning of their work to specify the evaluation questions to a level of detail amenable to quantitative measurement and to devise the possible approaches to provide necessary quantitative evidence.*

From a longer-term perspective, the application of such indicators is still limited by methodological and data gaps. Therefore, this study lists requirements for further development of indicators, related methodologies and underlying data. But this requires a longer-term effort which takes too long to support upcoming evaluations. Here, the focus should be on using available indicators and data which complement other available types of evidence. This requires a pragmatic perspective with a focus on indicators – mostly from existing data sources – which help to answer a specific evaluation question, taking directly into account limitations

of concepts, data, resources and time and finding suitable trade-offs between them. Under this perspective, the main judgment criterion is the added value for the evaluation process.

### **Integration of indicators in the evaluation process**

This type of indicator production and application cannot be separated from other evaluation work. It requires a good understanding of evaluation needs as well as indicator competencies.

At the level of the overall *FP evaluation process*, not all specific needs and requirements of individual evaluation exercises can be foreseen. But many evaluation questions will be clear already from general policy and FP goals. This permits the preparation of a core set of indicators in anticipation of these issues. To cover further needs of upcoming evaluations, complementary indicator production and/or studies can be launched ad-hoc in critical areas.

#### ☛ **Recommendation 2:**

*Future FP proposals should include a detailed planning of indicator requirements as a part of evaluation planning. This includes a proposed core set of indicators, of specific studies to be launched and resulting resource and data requirements.*

At the level of a *specific evaluation exercise*, indicators must be supplied under tighter time and resource constraints. Therefore, they will be mostly limited to the use of existing indicators or to smaller-scale elaborations which use available data. To exploit their potential, those carrying out or commissioning evaluations should identify early in their work the best possible use of quantitative indicators as sources of evidence, based on specific evaluation questions.

#### ☛ **Recommendation 3:**

*In planning their work, those carrying out or commissioning evaluations should ensure the necessary indicator competency and reserve the necessary time for discussing and integrating indicator-based findings. If panels are used, they should include a member with an explicit mandate concerning S&T indicator use in the evaluation.*

### **Establishment of a set of core indicators**

A large spectrum of indicators can contribute to FP evaluations. This study identifies such potentially useful indicators, grouped in six indicator domains. Their potential value added, methodological status and data sources are discussed in detail in Volume 2 of this study. From this large choice, a set of potential priority indicators, which are likely to be particularly useful, is identified.

This proposed set of indicators is based on the study team's perception. As other experts and stakeholders might come to slightly different priorities, a debate should be launched by the Commission which leads to a consensus on the range of indicators whose production can and should be organised in advance of major evaluation events.

#### ☛ **Recommendation 4:**

*The Commission should launch an initiative to establish a set of core indicators beyond those used today. This includes the launch of a debate on the appropriate indicator set and, where required, the launch of studies and other implementation measures.*

### **Establishment of a methodological framework for indicator use in FP evaluations**

To ensure the necessary methodological consistency, indicator use in all evaluations should build on a sound common methodological base. Even if indicators must be customised for the specific goals and questions of each evaluation, their production and use are based on common principles and concepts. For this purpose, the study proposes a methodological framework, consisting of a number of consecutive steps.

#### ☛ **Recommendation 5:**

*After examination of the proposed indicator framework, including its discussion with stakeholders and experts, the Commission should use it to formulate and disseminate a set of guiding principles for indicator use in evaluations.*

### **Enhanced use of indicators in FP monitoring**

Today, FP monitoring is based on a range of proven instruments. Recent monitoring and related reports, which are based on a profoundly revised approach to FP7 monitoring, have been important steps towards a more systematic use of indicators for FP governance. Additional information is available from other sources. But the overall reporting 'landscape' is still heterogeneous and does also not yet cover all relevant aspects in the desired depth.

Therefore we propose to launch a review of FP-oriented monitoring and reporting practices. Existing reports should be better aligned. The range of applied indicators should be extended to cover all relevant FP aspects. For this, the proposed six indicator categories may provide a general framework. But the resulting set of indicators is different because of the different type of value which monitoring adds for decision makers and stakeholders and because of the limited availability of certain types of data while the programme is still in progress.

A further possible enhancement might be the introduction of a complementary 'health check' reporting function which provides decision makers and stakeholders with a pragmatic, decision-oriented snapshot of the FP's implementation and target achievement status and of possible areas where corrective action is necessary.

#### **← Recommendation 6:**

*The Commission should initiate a review of FP-oriented monitoring and reporting practices, involving all units which produce and use FP monitoring elements, to identify possibilities for enhanced coordination and alignment of reports, to agree on an extended core set of indicators for FP reporting and to discuss a possible introduction of a complementary FP 'health check'.*

### **Necessary support structures and competencies**

An enhanced use of indicators in FP evaluations requires the availability of competencies and resources for their design, production, maintenance and application support. Appropriate resources and organisational structures are necessary. Some work like basic methodological studies, dealing with feasibility and design of specific indicators and their production, can be subcontracted. But the coordination of indicator-based evaluation practices in the complex FP structures, the production of core indicators and the maintenance of databases require continuity. Furthermore, capacities for the support of ongoing evaluations should be available. The question if this can best be done in existing structures or if it requires the creation of a new dedicated unit cannot be answered by this study. This issue requires further elaboration.

#### **← Recommendation 7:**

*To ensure the necessary coherence and support for indicator use in FP evaluations and monitoring, appropriate resources should be made available. Alternatives for the best possible organisational set-up of such a support function should be worked out and evaluated in the implementation of this study's recommendations.*

### **Improved data infrastructure**

The analysis of currently available data and future requirements shows a considerable mismatch. Even if many data are available, possibilities for indicator production are limited by unavailability of data in certain domains, by data quality and structures and by the absence of intermediate data layers. But this is not an issue of collecting ever more data. One might even argue that less data should be collected from the projects. In most cases, the problem is more related to the concepts behind data collection. Programme administration and evaluation have different needs; therefore available data are only of limited use for the latter.

To improve this situation, a system of dedicated databases for indicator production should be set up. This requires exploiting, cleaning and structuring of available information. The key to achieving this is *simplification*: indicator databases do not have to include all kind of data required for programme management. They should concentrate on those data which are

critical for indicator production. These data should be based on taxonomies and closed lists. Free field data are in most cases of little use, especially for large programmes, where cleaning by hand is too resource-consuming. Such databases should be designed in collaboration with indicator experts.

The detailed discussion of indicators shows that many of the proposed indicators require a *common participation database* for the whole FP, possibly maintained on a cross-FP base in the future. Secured data quality and fast identification of participants *at laboratory level* are required for most meaningful indicators, except for the most aggregated ones. This task requires the identification of individual research units, careful tracking of organisational changes over time and data validation. Without such a database, it will be difficult to perform in-depth evaluations of FP participation patterns with the help of quantitative data.

← **Recommendation 8:**

*A dedicated effort should be launched to develop a participation database allowing identification and tracking of FP participants at the level of individual research groups. Priority should be given to tracking large units and regular FP participants.*

Two further improvements are possible. A *common system of logging of administrative events* would be very useful to analyse the programme management and administration. This system should be restricted to a core of key events required for monitoring and evaluation. A *joint system to survey participants and projects* can track the scientific and technological production of projects, participants' opinions and feedbacks on programme management and project impacts. For both, pilot projects should be launched as soon as possible.

The target should be to have functioning systems in place for the next FP, even if the next FP7 evaluations will presumably not fully benefit from them because of development time.

## 1 Situation and challenges

### 1.1 The overall system of FP evaluation practices<sup>1</sup>

All of the specific research activities which together constitute the FP are subject to detailed evaluations, typically carried out once during the life of each FP. How this is carried out depends on the DG which is involved. DG RTD has a largely decentralised approach and the evaluation work at the level of research activities is the responsibility of the individual Commission services which have responsibility for their part of the programme. In DG Information Society and Media by contrast, the evaluation work is centralised.

The evaluation and implementation monitoring schemes at the level of the FP itself have undergone continuous updates since they were introduced in the early 1980s. A major revision took place in the mid 1990s. The features of this system were in place until the end of FP6.<sup>2</sup> In the last years, there have been considerable changes. The traditional system required that before submitting the proposal for the following FP, an external evaluation had to assess the scientific and technological objectives set out and the implementation and achievements of Community activities carried out via the specific programmes during the five years preceding that assessment.<sup>3</sup> This system was based on the so-called *Five-Year Assessment*, which combined elements of an ex-post evaluation of the previous programme with a mid-term appraisal of the on-going programme and recommendations for future activities.

Cascades of panels were part of this set up for the 1996 and 2000 Five Year Assessments. These were abandoned for the 2004 Five-Year Assessment and replaced by a single panel. Multiple panels were also used for monitoring up to 2003 and then replaced by a single panel.

In response to the growing size and complexity of the FPs, to the 'Better Regulation' initiative<sup>4</sup> and to other administrative reforms, this evaluation and monitoring approach was revised when FP7 was launched. An ex-ante impact assessment was carried out in conjunction with the FP7 proposal (European Commission (2005b)). In the renewed system, the previous focus on the Five Year Assessments was replaced by an extended portfolio of evaluations of the FP and of its key elements at defined milestones.<sup>5</sup>

The new monitoring and evaluation requirements for FP7 were actually set out in the ex-ante Impact Assessment of FP7, with some of the key details confirmed in the formal FP7 Decisions (European Parliament (2006)). The implementation of FP7 and its Specific Programmes shall be continually and systematically monitored. An evidence-based interim evaluation<sup>6</sup> shall be carried out no later than 2010, preceded by a progress report. Two years following the completion of FP7, an external evaluation of its rationale, implementation and achievements shall be carried out by independent experts. Table 1 summarises the resulting new FP evaluation scheme in comparison with its predecessors.

A recent review of the Commission's evaluation system by the European Court of Auditors, covering the period from FP4 to partly FP6, took a very critical view although did acknowledge that the basic requirements had been met. It highlighted a significant number of areas in which improvement could be made.<sup>7</sup> These include a higher transparency of the interven-

---

<sup>1</sup> Overview, for details see Volume 2, Chapter 1.

<sup>2</sup> For a review of the EC's traditional FP evaluation framework, see for example Williams et. al. (2002) or Georgiou and Polt (2004).

<sup>3</sup> This type of scheme is described for example in European Parliament (1999).

<sup>4</sup> For details see European Commission (2008b) and literature quoted therein.

<sup>5</sup> See Table 1. However, this system is still in motion, For certain elements like the ERC or the JTPs, whose governance is still under development, detailed evaluation principles still have to be worked out. See also European Commission (2007a) for planned further amendments.

<sup>6</sup> Which shall cover the quality of the research activities under way, as well as the quality of implementation and management, and progress towards the objectives set.

<sup>7</sup> Source: European Court of Auditors (2007); see this source for details.

tion logic in programme definition and of links between scientific, technological and socio-economic objectives. From these, programme objectives and performance measures should be derived which permit to assess the relevance, effectiveness and efficiency of programme execution at any stage. In the context of the decentralised FP management and of the evaluation system which mirrors this, coherent standards must be applied to all monitoring and evaluation work. For this, a reinforced methodological framework and guidance must ensure that evaluations follow consistent standards, gather and use relevant data and address all relevant issues appropriately. This requires strong coordination and support for indicator use in all evaluations and an improved access to reliable data for their production. Experience with recent evaluation panels (used extensively by the Commission for the FP level evaluations) has shown that their efficiency and willingness to use indicators intensively depends on the accessibility, ease of use and quality of necessary data and indicators.

<b>FP 4 (1995-1998)</b>	<b>FP 5 (1999-2002)</b>	<b>FP 6 (2003-2006)</b>	<b>FP 7 (2007-2013)</b>
<b>two evaluation exercises</b> (mid term and final), in each case for its 20 sub-programmes plus the FP; final evaluation overlaps with mid-term of FP5	<b>two simultaneous evaluation exercises (mid-term)</b> , in each case for its 7 Specific Programmes plus the FP	<b>one evaluation exercise</b> for the FP and its 3 Specific Programmes (mid-term), but not specifically for all 20 sub-programmes, complemented by specific evaluation and regular progress (i.e. monitoring) reports	<b>at least five evaluation exercises</b> (ex-ante, mid-term and final) including specific progress report before interim evaluation
<b>Ex-ante:</b> No	<b>Ex-ante:</b> No	<b>Ex-ante:</b> No	<b>Ex-ante:</b> Part of 'Impact Assessment of the Commission's proposal for FP7' (2005)
<b>Mid-term:</b> Five-year Assessment mid-way through FP, prior to subsequent FP (1997)	<b>Mid-term:</b> Five-year assessment prior to presenting proposal for subsequent FP (2000); progress review mid-way through FP (2001)	<b>Mid-term:</b> Evaluation of effectiveness of instruments (2004); five-year assessment, prior to presenting proposal for subsequent FP (2004); regular progress (i.e. monitoring) reports	<b>Mid-term:</b> Ex-post, 2 years after the end of previous FP (2008); progress report before interim evaluation (before 2010); interim evaluation (2010)
<b>Final:</b> On completion on FP (1999)	<b>Final:</b> No	<b>Final:</b> No [modified by FP7 Decision]	<b>Final:</b> Two years after FP completion (2015), supported by specific studies, the interim evaluation and other evaluation activities carried out throughout the FP period

Table 1: Key requirements for monitoring and evaluation in FPs 4 to 7<sup>8</sup>

For the development of the proposed indicator-based framework, this raises three important issues. Firstly, indicators and methodologies must account for the growing complexity of the FP and its objectives. Traditional linear, input-output-oriented indicator concepts are not sufficient for this purpose. Secondly, indicator use in evaluations needs to be strengthened and institutionalised. In recent evaluations, a multitude of instruments<sup>8</sup> have been applied in very

<sup>8</sup> Including scientific peer review, expert /consultant review, participant questionnaire, interviews, EC data/document review, control groups and national research council data (Examples taken from Arnold (2005. This source reviews evaluations carried out until 2005.)).

different forms and mixes. This has led among other things to a very heterogeneous – and often only marginal – use of indicators. Even if some evaluations have used indicators to provide evidence and to support conclusions and recommendations<sup>9</sup>, their potential has by far not been fully exploited. Currently there is also no binding, consistent and comprehensive guideline for indicator use in FP evaluations. Existing methodological guidelines mention the necessity to provide quantifiable evidence and indicators, but they do not specify in detail how this should be done, even if elements have been reviewed and updated continuously.<sup>10</sup>

Thirdly, the necessary preconditions for an efficient and consistent use of indicators do not fully exist. The importance of translating high-level policy goals into more tangible quantified or otherwise measurable objectives has been highlighted already in the 2001 ex-ante evaluation guidelines (European Commission (2000)). The 2008 Strategic Review of Better Regulation in the European Union states: “Impact assessment should quantify impacts where this is feasible and proportionate — without glossing over underlying uncertainties. Despite progress (for example, administrative costs are systematically considered and, if significant, measured with the EU Standard Cost Model), insufficient or unreliable data often prevents quantification” (European Commission (2008b)).

## 1.2FP monitoring and related reporting

Annual monitoring has traditionally been carried out by the Commission services with the assistance of external experts. Monitoring has been intended to examine programme implementation. It has provided a quick response mechanism to programme development, designed to provide a constructive independent critique on key issues.

Basic reporting requirements are defined by Article 173 of the EC treaty. As a part of their fulfilment, the *Annual Report on Research and Technological Development Activities of the European Union* (European Commission (2007b)) describes from a high-level perspective achievements related to community research and technological development activities and defined objectives, in particular the Lisbon and Barcelona targets. It is usually complemented by a Commission Staff Working Document (European Commission (2007b1)) which summarises European research support. Its arguments are supported by statistical data where appropriate<sup>11</sup>. A Statistical Annex (European Commission (2007b2)) provides detailed information about FP proposals and contracts.

---

<sup>9</sup> For example, the panel in charge of the Five-Year Assessment of the 1999 -2003 FP activities (European Commission (2004); ‘Ormala report’) used extensively data that were organised for the exercise, including specific studies; independent experts’ and Commission staff inputs and a specially developed database of evaluation and policy documents. Panel members also undertook own research and data gathering. Another example is the final evaluation of FP6 (European Commission (2009)). The impact assessment and ex ante evaluation in preparation of the FP7 proposal (European Commission (2005b)) made extensive use of indicators to assess the impact of proposed measures on Europe’s capability to meet economic, social and environmental challenges and on its capability to cure the weaknesses of the its research system. A range of indicators was used to compare scenarios under the proposed FP7 with other policy options. An appendix provides complementary background information.

<sup>10</sup> For details see European Commission (2000), European Commission (2002), European Commission (2005a). Examples for available criteria are those for project selection which may be also useful for assessing relevance. These include (1) Criteria related to the Community “value added” and the subsidiarity principle (Overall Objective: Select activities which are more efficiently pursued at the Community level by means of research activities conducted at that level); (2) Criteria related to social objectives (Overall objective: Further major social objectives of the Community reflecting the expectations and concerns of its citizens) and (3) Criteria related to economic development, scientific and technological prospects; Source: Source: European Parliament 1999).

<sup>11</sup> For a more concise presentation and a better assessment of achievements towards the Lisbon agenda, 14 structural indicators will be covered in the Update of the Statistical Annex to the 2008 Progress Report to the European Council., see [http://epp.eurostat.ec.europa.eu/portal/page?\\_pageid=1133,47800773,1133\\_47802588&\\_dad=portal&\\_schema=PORTAL](http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1133,47800773,1133_47802588&_dad=portal&_schema=PORTAL).

The approach to monitoring FP implementation has been revised for FP7. In FP6 and previous FPs, monitoring had been implemented through annual panels of independent experts. These selected specific areas of FP implementation and performance to analyse and report on. The FP7 monitoring system is designed as an internal management tool, based on a system of indicators.<sup>12</sup> Under this new system, Monitoring Reports shall be structured along the same principles, thus allowing for the first time a kind of longitudinal analysis of the FP7 implementation. This is reflected in the first FP7 Monitoring Report which covers the year 2007 (European Commission 2009a).<sup>13</sup>

For the Specific Programmes, this exercise examines progress in relation with the original objectives and whether these objectives, priorities and financial resources are still appropriate. The FP level exercise compiles the results.

### 1.3 Strategic Planning and Programming (SPP) cycle

It is important to understand that the research evaluation and monitoring scheme is only one part of the overall package of measures which seek to assess performance.

As part of a range of measures which together constitute the SPP, each DG across the Commission as a whole produces its own *Annual Management Plan (AMP)*. This document defines the Directorate General's objectives and provides a planning framework for resource deployment. Planning is broken down into the individual activities defined at budget level. For each DG involved in FP management, these map approximately those FP areas which it covers. For each activity, objectives are formulated, main outputs are defined and resulting indicators are identified. For all defined indicators, targets are formulated and compared with actual achievement levels.<sup>14</sup> At the overall policy area, a limited number of objectives and indicators are also defined.

Further information about the FP implementation status and about the relevant context is provided by a number of statistical and other complementary publications (see Chapter 1.4).

Implications for the indicator-based framework proposed in this study are: (1) Current reporting uses mostly traditional linear indicators. Extending these to include for example composite indicators which describe complex issues like the ERA status would add value to reporting. (2) Reporting activities are scattered over a range of sources. Some of these are very comprehensive, but they address the needs of specialists in great depth. There is no focused, indicator-based reporting form which provides policy makers and other stakeholders in a comprehensive, user-friendly way with a rapid overview over the main parameters.

---

<sup>12</sup> See Volume 2 of this report, Table 2 for details. Monitoring Reports before FP7 can be found under [http://ec.europa.eu/research/evaluations/index\\_en.cfm?pg=pre-fp7&showtoo=&show=2006report#2006\\_report](http://ec.europa.eu/research/evaluations/index_en.cfm?pg=pre-fp7&showtoo=&show=2006report#2006_report).

<sup>13</sup> The new monitoring system is based on Article 7(1) and 6(1) of the EC and Euratom FP7 Decisions (European Parliament 2006) which states that "The Commission shall continually and systematically monitor the implementation of the Seventh Framework Programme and its specific programmes and regularly report and disseminate the results of this monitoring." Further information on the detail of the new system is provided by the Ex ante Impact Assessment on FP7: "Monitoring of implementation management would be ensured by operational senior management within the Commission on a continuous basis with annual check points and using a common set of management performance indicators. Adequate resource would be given to this process. The annual results of this exercise will be used to inform senior management and as an input to the ex post assessment exercise."

<sup>14</sup> In the 2008 version of the Annual Management Plan, the column 'latest known results' recognises however that for a range of defined indicators, data are not yet available.

## 1.4 Available data for indicator production

From EC-internal sources, many data on FP-funded research activities are available. Most of these are collected from the research project applicants or participants.<sup>15</sup> All information from tenders and projects can be accessed through two data bases: E-CORDA provides authorised users with comprehensive online-information about the outcome of concluded FP7 calls for proposals, including FP7 participation and performance statistics. The Commission's online reporting tool for research and technological projects, SESAM, is used for submission of all reports and questionnaires throughout project execution. It collects all resulting project information. Further information is accessible through CORDIS.

Beyond this, a number of Commission services provide also data and indicators which describe the Framework Programmes' relevant (external) context (e.g. international R&D expenditure and related issues). Sources of such data include *Eurostat*,<sup>16</sup> *CORDIS*, *ERA-WATCH*, *PRO INNO* with its *European Innovation Scoreboard* (EIS) and *Innobarometer* or the *2007 EU Industrial R&D Investment Scoreboard* (European Commission (2007d)), presenting information on industrial R&D behaviour and the relative position of European industrial R&D and innovation. Specific data are also generated by other Directorates General which are responsible for thematic areas. For example, DG Information Society and Media (DG INFSO) has developed an ICT-specific monitoring system, initiated a variety of studies, for example to assess progress towards the realisation of the ERA in Information Society Technologies and pursues ICT-specific evaluation exercises.<sup>17</sup>

Further data about global research, innovation, patents, economy, etc., which complement the Commission data are available from a multitude of sources, e.g. national policy makers statistical institutions, international organisations like the OECD or industry associations.<sup>18</sup>

Implications for the proposed indicator-based framework are: (1) the current FP data base provides a large pool of data which is potentially relevant for FP evaluation and monitoring. But so far, these have been exploited only partially because of the discussed quality and structural limitations. (2) Despite this, not all data exist today which would be necessary to produce an extended indicator set representing the increasingly complex FP context.

## 1.5 Challenges and requirements for the proposed indicator-based framework

A key issue to obtain optimal leverage from the FP is a better-informed decision making. To be able to set realistic objectives, to focus resources on priority research areas and to follow if the desired outcomes and effects are realised, policy makers need actual and reliable information. Indicators for inputs, outputs, effects, etc. can provide this transparency in many cases.<sup>19</sup> But to realise this value, a number of challenges must be overcome.

---

<sup>15</sup> These documents are submitted according to the Commission's standards, using firm templates to ensure their coherence. For details see the Electronic Proposal Submission Service (EPSS) User Guide (<http://cordis.europa.eu/documents/documentlibrary/91055671EN6.pdf>) and the SESAM user guides (<http://webgate.ec.europa.eu/sesam/index.do;jsessionid=LbGMb5hLpk6DDf757BZYx2wpxjt4G2nh0y6mHQDQJnnTVp31TSp!-1182492268!-110860345>).

<sup>16</sup> A key publication is Science, Technology and Innovation in Europe (last edition: Eurostat (2008).

<sup>17</sup> See [http://ec.europa.eu/dgs/information\\_society/evaluation/studies/index\\_en.htm](http://ec.europa.eu/dgs/information_society/evaluation/studies/index_en.htm) for a list of studies carried out under the auspices of DG Information Society and Media; [http://ec.europa.eu/dgs/information\\_society/evaluation/impact\\_assessment/index\\_en.htm](http://ec.europa.eu/dgs/information_society/evaluation/impact_assessment/index_en.htm) for impact assessment studies carried out in the ICT area and [http://ec.europa.eu/dgs/information\\_society/evaluation/data/pdf/evaluation\\_plan\\_2008.pdf](http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/evaluation_plan_2008.pdf) for a list of evaluations to be launched by DG INFSO in the period 2008-10.

<sup>18</sup> See the OECD science, technology and patents statistics portal, [http://www.oecd.org/topicstats-portal/0,3398,en\\_2825\\_497105\\_1\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/topicstats-portal/0,3398,en_2825_497105_1_1_1_1_1,00.html). On the private sector side, many industry associations maintain statistical data bases for their respective sectors, including R&D activities.

<sup>19</sup> For an example of a comparable quantification initiative, see the EU common methodology for assessing administrative costs imposed by legislation (European Commission (2005c)). Chapter 4

A first challenge is created by the FPs' *complex structure*. Under the FP umbrella, Specific Programmes and action lines pursue different sub-objectives, addressing different types of issues and target groups with different instruments. Accordingly, outcomes are also heterogeneous. They complement each other to some extent. But they cannot easily be 'added up' in a sense of '... Specific Programme X contributes ...% to the achievement of the FP's overall targets...'. As a further consequence, there is also no single, universally applicable 'one fits all'-type set of criteria to describe the success and achievements of all FP elements. Instead, each FP element requires its own specific set of criteria.<sup>20</sup>

A further level of complexity is added by the FPs' *extended range of objectives and their often rather general formulation*. The traditional objective to strengthen European research has been complemented by expected research contributions to reaching larger EU economic goals and by the inclusion of the policy goal of building the ERA. This leads to an ever more complex portfolio of FP objectives and of expected contributions to various overarching policy goals without a visible, spelled-out single overall FP target. This reflects the complex reality of European research and permits the FP to react to the core issues of the European research system in a flexible way. But it makes it also difficult to assess the 'success' of the FP and its elements and to identify the relationships between individual action lines or instruments and observed outcomes. In addition, especially broader policy and overall FP objectives are typically expressed in qualitative, but not in quantitative terms. Therefore, quantifiable success criteria and baselines, against which achievements could be measured, are not readily available. This opens a wide field for (potentially subjective) interpretation and makes it difficult to measure the programme's performance against its initial objectives, the relevance of these initial objectives (in the light of actual situation) and its performance against possibly revised criteria. A higher degree of measurability would be desirable.<sup>21</sup> A step into this direction has been made with the ex-ante impact assessment of FP7 (European Commission (2005b)), which has formulated a number of such quantifiable objectives.

Due to *methodological limitations*, progress and success cannot be described reasonably well in all domains by the indicators which are available today. In certain cases, the methodological indicator base (i.e. definition of indicators, agreed specifications and rules for their production) is not sufficiently advanced and needs further progress.<sup>22</sup> Certain concepts, e.g. 'competitiveness' as an economic objective, are also criticised for being too elusive to permit an accurate assessment of success. In the case of ERA, even eight years after the Commission's first ERA communication (European Commission (2000a)) the debate about the ERA's

---

of this document highlights benefits of such a EU common methodology, including better assessment of measures from the point of view of those affected, a higher transparency (quantifying makes trade-offs more transparent if costs and benefits are both investigated), support for priority setting, facilitated communication, facilitation of EU-wide comparison of performance and the identification of best practices, and a basis which ensures that national data can be easily added up in view of assessing individual acts and/or cumulative effort/benefit.

<sup>20</sup> For example, the 'success' of application-oriented research in the 'Collaboration' programme can be measured in some cases against criteria derived from the specific context (e.g. solution of burning, scientific issues, expected impact on competitiveness, etc.). In difference to this, basic research under the 'Ideas' priority cannot be 'measured' against such criteria because of its exploratory nature. The same is true for the 'People' and 'Capacities' priorities, which are devoted to the development of European research's human resource base and infrastructure.

<sup>21</sup> In its 2000 ex-ante evaluation guidelines (European Commission (2000)), the Commission highlighted already the importance of translating high-level policy goals into tangible, measurable objectives. The 2005 Impact Assessment Guidelines (European Commission 2005a) request that objectives should be 'SMART' (Specific, Measurable, Accepted, Realistic, Time-dependent).

<sup>22</sup> Especially in areas which cannot be described by simple, linear indicators. For example, the 'competitiveness of European industry' cannot be expressed in a single figure. It is influenced by various factors (of which innovation activities are only one) and expresses itself in more than one variable (e.g. market share, growth of revenues, relative technology position, etc.). For such purposes, innovative composite indicators are needed which are not yet available.

main objectives and action areas persists. At the same time, further methodological development in the indicator area and clearer definition of relevant policy concepts are desirable.

*Limited availability of relevant data* restricts the possibilities to construct indicators for some domains. The multitude of FP-internal data which is collected during the proposal phases and project life has never been fully exploited in more sophisticated analytical ways. *Current data structures* do also not always support the necessary processing of raw data and their compilation to obtain compound information about more complex issues which cannot be obtained from a single data set. A third limitation is the *quality of data* which requires in certain cases laborious cleaning of available data sets. To cope with this issue, improved quality assurance mechanisms in data collection processes are currently implemented.

A last improvement area is the necessary *consistent application of methodologies and processes* in the decentralised FP structures. A range of Commission documents provide methodological guidelines. But in recent evaluations, these have been applied to a highly varying extent, also in the degrees of indicator use. A heterogeneous evaluation and monitoring 'landscape' has emerged which calls for a higher degree of coherence and coordination. The outcomes of evaluations of different FP elements should become more compatible.

## 2 Proposed indicator framework<sup>23</sup>

### 2.1 Systematic use of indicators for FP evaluation and monitoring

In research programme evaluations, qualitative information (e.g. assessment by experts and panels) and quantitative information, including indicators, complement each other.<sup>24</sup> Prominent examples of such indicators include bibliometric tools to characterise scientific output and patent analysis to characterise technological output.

As their name suggests, indicators are constructs which allow quantitative measurement of specific features of reality which cannot be directly observed. This is achieved through the combination of different data and the use of suitable proxies which are supposed to be related to the observed feature.<sup>25</sup> But in a certain sense, indicators are arbitrary and their appropriateness is related to the specific perspective of their users. Thus, there are no indicators without a clear definition of the underlying objectives and questions. Indicators have to be considered as supporting evidence for the evaluator's assessments and for policy debate rather than as 'objective' descriptions of a reality per se.

A second feature of indicators is that they are explicitly meant to condense information. A comprehensive assessment is replaced by less precise information which demands less effort and resources and which is easier to interpret because of its quantitative nature. But the price of this is a loss of detail and of a thorough analysis of underlying causalities. This implies a danger that indicators may be misinterpreted. Therefore, a central criterion for a good indicator is its ability to provide a reasonably precise picture of the reality with an acceptable effort. Indicator feasibility, data availability and effort required for indicator production have to be carefully considered when choosing between indicators and other types of evidence.

Indicator application in evaluations should always start from the generic programme logic.<sup>26</sup> Broader socioeconomic goals must be converted into specific programme objectives and operational mechanisms (e.g. deployed resources, instruments, selection criteria, etc.). This requires that the evaluation identifies the causal relationships between programme objectives, design, activities, instruments and execution and its observable (or anticipated) outcomes and effects, thus providing also insight into the relevance and achievement of objectives. Beyond this traditional linear, input-output-oriented perception, the indicator framework must also account for the increasingly important structuring effects of FP-funded research in the case of its application in the Community context.

The programme's intervention logic can be described with the help of logic charts, differentiating between overall objectives, programme objectives, activities, outcomes, intermediate impacts and ultimate impacts. The underlying evaluation questions determine to which extent each of these levels is addressed in a specific evaluation. Questions related to broader policy goals – like the FP's contribution to the ERA development – require in most cases the use of macro-level indicators (aiming at characterising the broader context in which the programme's scientific production takes place and its effects on this context), whereas evaluations of programme effectiveness and efficiency will concentrate on programme-level indicators (aiming at measuring the programme's activities and results as such).

Based on these considerations, we propose the methodological framework described in Figure 1 to structure identified indicators and relate them to the underlying evaluation issues.

---

<sup>23</sup> This chapter summarises the proposed indicator framework, for details see Volume 2, Part B.

<sup>24</sup> The selection of appropriate evidence and factors which determine if indicators are the method of choice are discussed in Volume 2, Chapter 2.2.3.

<sup>25</sup> For example, the number of citations of scientific publications is taken as a proxy for the outcome of research activities because scientific studies have confirmed that recognition by other scientists, expressed by their citing these works in own papers, is a reasonable measure of research quality.

<sup>26</sup> See the detailed discussion of this issue in Volume 2, Chapter 2.2.

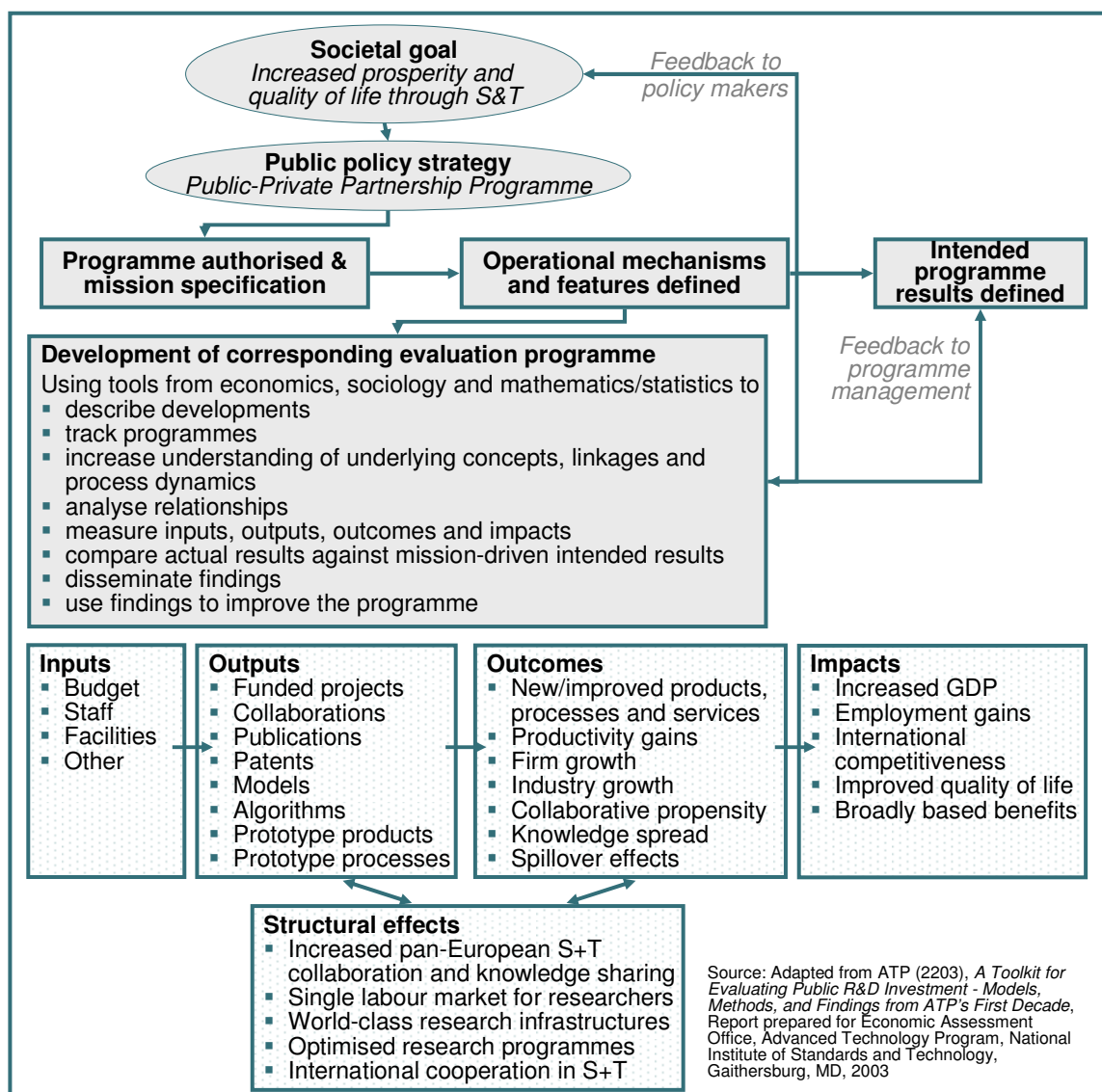


Figure 1: Evaluation logic model and indicators domains

The generic descriptive roles of the indicators depicted in the bottom part of Figure 1 complement each other. Thus, the whole flow of the evaluation logic can be covered by appropriate indicator-based evidence.

## 2.2 Proposed indicator domains and indicators

Based on the possible indicator roles which were defined in Figure 1 and on practical considerations of how available indicators are produced and how data are collected, we propose the following six indicator domains (see Figure 2):

- Indicators related to the *implementation and management of the programme*;
- indicators related to *participation in the FPs*, as well as to funding and its distribution by participants, sectors, regions, countries, etc (input indicators);
- indicators related to *project and programme outputs*, distinguishing between
  - scientific products and results, and
  - technological products and results,

- socio-economic, environmental and other relevant indicators measuring the *longer-term impact* of the FPs, and
- indicators related to the FPs' contribution to the *structuring of the ERA*.

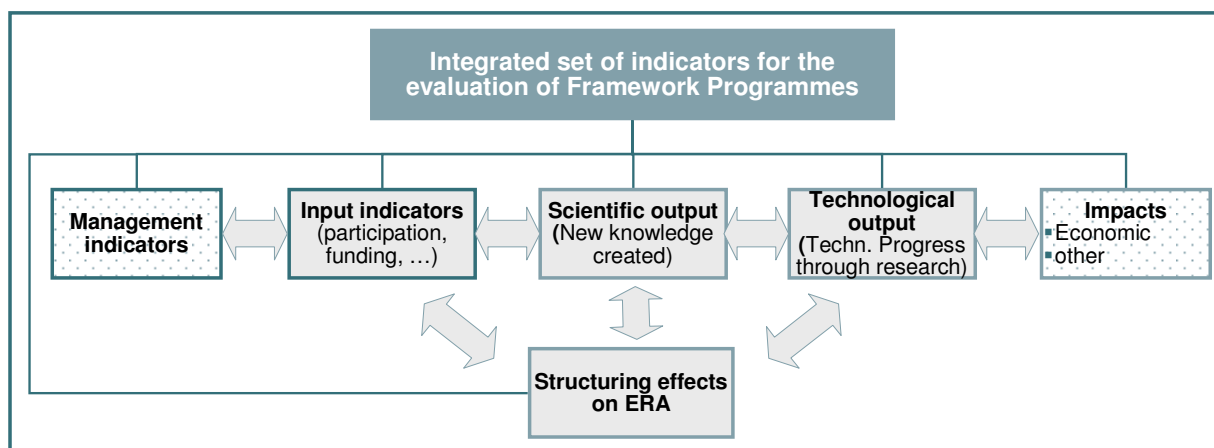


Figure 2: Proposed indicator structure and domains

These indicator domains correlate with the categories of evaluation questions proposed in Chapter 3.3. Each of them contains a comprehensive set of potentially useful indicators from which evaluators can choose the most appropriate ones for their specific purpose, depending on the evaluation's main questions. They are briefly characterised in the following sections<sup>27</sup>:

#### (1) Programme management indicators

Programme management is a central concern in FP evaluations. Inadequate management would not only entail avoidable proposal and project cost and inefficiencies. Even more severe consequences might be that important European research actors are not involved appropriately, that important scientific themes are not addressed, that possible research outcomes are not fully realised or that participants are de-motivated to participate in future activities. To avoid this, indicators must help to answer the following questions:

- Is the thematic priorities selection process efficient? Does it identify the right priorities?
- Do chosen themes address research actors' needs and motivate them to participate?
- Is the FP implemented efficiently?

To answer these questions, the following four classes of indicators are potentially useful:

- *Indicators for monitoring the thematic priorities selection process* analyse the demography of the consultation process in order to assess the reaction of the scientific and policy community to the proposed priorities.
- *Indicators for the demography of proposers and evaluators* provide an understanding of the acceptance and relevance of the proposed research themes, based for example on the demography of proposal submitters (as well as of non-submitters).
- *Indicators for FP management efficiency* indicate whether the programme is executed in the most efficient way, based on an analysis of proposal handling times, administrative burden on project participants, etc.

<sup>27</sup> The corresponding in-depth description of indicator areas, including conceptual background, main evaluation questions, main indicator categories and possible indicators, data sources and methodological issues can be found in Volume 2, Chapters 3 – 8. These sections include also recommendations for further development of the respective indicator areas.

- *Indicators reflecting participants' opinions on management* indicate perceived bottlenecks and inefficiencies which may limit the efficiency of FP-funded research and/or participants' motivation.

Identified indicators in these domains are described in Volume 2, Chapter 3. From this indicator portfolio, the following might be particularly helpful in upcoming evaluations:

- (1.1) *Indicators for proposals demography* (Structuring effects on institutions participating in FP proposals);
- (1.2) *Indicators for success rates* (Overall success rates, identification of domains/countries where success rates are above or below average);
- (1.3) *Indicators for proposal submission resources and cost* (workload of submitting a proposal);
- (1.4) *Indicators for management efficiency* (Efficiency of FP implementation);
- (1.5) *Indicators for administrative bottlenecks in the implementation of FP-funded projects* (resources share for non-research activities, administrative bottlenecks).

## (2) *Indicators for Framework Programme participation and funding*

Information about the (different forms of) participation of European research actors and the repartition of funds among beneficiaries provides important information to assess if the FP has reached the important groups of actors in European research and if it contributes to enhancing their knowledge and to integrating European research. Indicators must help to answer in particular the following questions:

- Are FP participation and the allocation of funding consistent with the need to involve Europe's best researcher from the different actor groups?<sup>28</sup>
- Does the FP's thematic repartition match its research and larger policy objectives?
- Do participation patterns support formation of pan-European research collaboration? Is there a core of regular FP participants, including leading research institutions?
- Are there synergies (or redundancies) between FP funding and funding from other sources (e.g. national programmes)?

This leads to the following four classes of indicators which are potentially useful:

- *Participation and funding statistics* provide information about the involvement of European research actors and highlight main research and structural trends.
- *Indicators for the demography of FP participants and market structure* provide information about its capability to involve leading research institutions and important stakeholder groups.
- *Indicators for synergies and integration with national research programmes* indicate if the programme adds European value, complementing other (e.g. national) initiatives.
- *Indicators for research structures and for structuring effects* show if sustainable research collaborations and networks are formed through FP funded research.

Identified indicators in these domains are described in Volume 2, Chapter 4. From this indicator portfolio, the following might be particularly helpful in upcoming evaluations:

- (2.1) *Indicators for FP participation* (repartition of participation, broken down by country, region, research performer, FP activity);
- (2.2) *Indicators for FP funding allocation* (repartition of funds, broken down by country, region, research performer and FP activity);
- (2.3) *Indicators for participant demography* (*Repartition of participants by number of FP participations over FP6/7*; broken down by research area/activity and country).

---

<sup>28</sup> Including stakeholders with particular needs, e.g. SMEs or new member states.

### (3) Indicators for scientific products and results

Information about (different forms of) direct outcomes of FP-funded research<sup>29</sup> is crucial to assess if the FP reaches its scientific goals. Beyond the creation of new knowledge, this includes also the performance of Europe's research structures and its position in the competitive global environment. The following questions are particularly important:

- Is FP-funded research relevant (i.e.: Does it address Europe's main research and policy objectives? Does it lead to a sound knowledge base for addressing these issues?)?
- Is FP research funding effective (i.e.: Does it have wide-ranging positive impacts on Europe's scientific, technological and socioeconomic performance? Does it stimulate/support the formation of strong scientific capacities?)?
- Is FP research funding efficient (i.e.: Are desired/achieved scientific outcomes obtained at the best cost and effort?)?

This leads to the following three classes of indicators which are potentially useful:

- *Indicators for the relevance of FP-funded research* assess if scientific outcomes are linked with and contribute to the FP's thematic objectives;
- *Indicators for measuring effectiveness* assess if FP-funded research reaches the desired high level of peer recognition and critical mass, and if it provides the basis for a successful transfer of knowledge and skills;
- *Indicators for measuring efficiency* assess if the FP enables its participants to perform better in their scientific knowledge generation and dissemination, e.g. on the basis of co-publication or research team composition patterns.

Identified indicators in these domains are described in Volume 2, Chapter 5. From this indicator portfolio, the following might be particularly helpful in upcoming evaluations:

- (3.1) *FP participants' publication volumes* (e.g. as share of world volume or as share of Top 10% journals; by major research area);
- (3.2) *Impact factors of researchers involved in FP-funded research* (as a whole, broken down by research areas);
- (3.3) *Networking indicators*:
  - (3.3.1) *Correlation indices between international co-publications of FP-participants and impact factors* (as a whole, broken down by research areas);
  - (3.3.2) *Indicators for bibliometric networks (Co-publications and citations)* (by research area, country);
- (3.4) *Excellence indicators: Size-corrected 'productivity index'*, combining bibliometric impact indexes and input data.

### (4) Indicators for technological products and results

Technological outcomes link the scientific results of FP-funded research to socioeconomic and other effects which are created through their use. This includes the generation of successful new products and services, usually by private sector companies, as well as other forms, e.g. the formation of spin-offs. To assess the FP's contribution to this process, indicators must help to answer the following questions:

- Does new knowledge from FP-funded research lead to the desired level of technological outputs, especially in those sectors which are particularly important for Europe?
- Are important research and technology actors sufficiently involved in FP-funded activities and benefiting from their results?

---

<sup>29</sup> Here, scientific products and results of research are defined as the numbers of scientific research publications and citations, number of scientific co-publications, number of non scientific publications, number of industry-university relations, etc.

- Is new knowledge, created in FP-funded research, transferred efficiently?
- Does FP-funded research generate clusters of technologies with a high (commercial, environmental, etc.) potential? Does the FP add European value to other initiatives?

Based on these considerations and on existing indicator concepts, e.g. from the OECD, the following three classes of indicators are potentially useful:

- *Indicators for technological output and its distribution* identify the FPs' technological results on the basis of output counts (e.g. number of patents, licences, joint ventures, spin-off, software stemming from FP-funded projects; per sector, country and originating institution).
- *Indicators for technological output producers* identify the FP participation and technology production levels of relevant technology producers and highlight differences in their demography and contributions.
- *Macro-level indicators* measure synergies between FP outputs and with other (e.g. national) initiatives.

Identified indicators in this domain are described in Volume 2, Chapter 6. From this indicator portfolio, the following might be particularly helpful in upcoming evaluations:

- (4.1) *Indicators for FP participants' technological output volume* (*Patent counts* by main research area, sector, country; *counts of licensed patents* by main research area, sector, country);
- (4.2) *Indicators for new human resources trained in FP-funded research* (i.e. the FP's capability to produce new human resources and knowledge);
- (4.3) *Indicators for innovation activities at firm level* (Indirect information about impact of FP participation on innovation behaviour of firms);
- (4.4) *Indicators for European and national innovation* (Information about evolution of EU/national innovation capabilities, correlated with FP participation);
- (4.5) *Technology products exchanges* (Technology balance of payments; see OECD definition).

#### (5) *Indicators for economical, social and environmental impact*

To know if the FPs make the desired contributions to achieving overarching policy goals, their impact on economic growth, employment, etc. (and more specifically on the participating firms) must be assessed. Indicators can help to answer the following questions:

- Which socioeconomic, environmental and other relevant impacts result from FP-funded research? How are these related to research (especially FP-funded research)?
- How does FP-funded research affect the innovation behaviour of participants?
- How does the FP influence research and related policies on other levels in Europe?

To answer these questions, the following four classes of indicators are potentially useful:

- *Indicators for socioeconomic impacts* measure the 'return on R&D investment', differentiating between return on private sector R&D, societal return on private sector R&D and return on public sector R&D.
- *Indicators for impacts on innovation behaviour* measure input and output additionality.
- *Indicators for societal and environmental impacts* measure changes of the quality of life, prosperity and health which are attributable to FP-funded activities.
- *Indicators for impacts on public policies* measure to what extent the FPs contribute to dynamising other (e.g. national) research, innovation and related policy levels.

Chapter 7 of Volume 2 describes identified indicators, but also current methodological and data constraints which limit the measurability of effects in this domain and their attri-

bution to research under the FPs considerably. From the currently feasible indicators, the following might be helpful in upcoming evaluations:

(5.1) *Country, region and sector participation intensities* compared to key economic data such as TFP, labour productivity, economic growth, sectoral value added, exports;

(5.2) *Labour productivity, employment (growth), economic growth, value added (for sectors, exports);*

(5.3) Correlations between economic performance evolution and FP participation (e.g. time lags over a longer period, at least two years).

(6) *Indicators for integration and structuring of the ERA*

Integrating and structuring the ERA is a central FP objective. The FPs are the Commission's major instrument to achieve this. Therefore, measuring the ERA development and FP contributions becomes a central evaluation issue. Indicators can help to answer three fundamental questions:

- Are the formulated ERA targets valid? Does the ERA achieve the desired effects?
- Is the ERA implemented efficiently? Is progress achieved in particular in the action areas which are considered necessary to realise the ERA?
- Does the FP make the desired contributions to integrating and structuring the ERA? Is it deploying its resources for this purpose in the best possible way?

The elusiveness of policy and programme objectives and the complex interactions of various instruments, initiatives and policy domains make it difficult to define quantifiable measures for the FPs' structuring and integrating effects and for the FPs' contributions to these. Answers depend also on research system conceptions. At the same time, the review of indicators in Volume 2 shows that indicator concepts are still in their infancy. As of today, robust 'ready-to-use' indicator sets are not yet available.

This gives the following proposed main types of indicators the character of a first draft which still needs to be elaborated further:

- *Indicators to assess ERA effects and rationale* analyse achieved vs. expected effects of the Europeanisation of research. We suggest focusing these on four main issues: Fragmentation of European research, knowledge generation and exploitation, strengthening of innovation capacities and solving Europe's Grand Challenges.
- *Indicators to assess progress in ERA implementation* measure – still on a non FP-specific base – progress in the identified main action areas. To avoid redundancies with other indicator areas, we propose to focus on the following three<sup>30</sup>: Indicators for progress in (1) building a European internal market for knowledge, research and innovation, (2) in coordinating/aligning European research policies and (3) in building valuable initiatives implemented and funded at a European level.
- *Indicators to assess FP contributions to realising the ERA* analyse how the FPs influence the fabric of European research in its complex context.<sup>31</sup> They should provide information about ERA-relevant achievements and contributions of the relevant elements of each Specific Programme.

In view of the infancy of this indicator domain's methodological basis, proposed indicator types should be considered as a stimulus for driving conceptual development in this area with a long-term perspective. As this requires time, upcoming evaluations should be supported by available indicators in a pragmatic way. These may include for example:

<sup>30</sup> Alternatively, these could build on the six action areas defined in the 2007 Green Paper.

<sup>31</sup> As these effects cannot be directly measured, especially not on an aggregate level, we propose to start by identifying relevant elements of each of the Specific Programmes and to formulate reasonable proxies for their contributions to realising the ERA.

- (6.1) *Excellence indicators in bibliometrics* can display the position of European research in the high-tail of scientific quality by domain. If a reliable identification of FP participants is possible, from this some inference can be drawn about specific FP contributions.
- (6.2) *Indicators for scientific collaborations (e.g. co-authorships)* can help to trace the development of a European research market. If identification of FP participants is possible, these maps can be related to FP projects.
- (6.3) Indicators for *collaboration in FP projects* can help to identify repeated collaborations at the participants' level (compared to case-by-case participations).
- (6.4) Indicators for the production of scientific personnel (e.g. PhDs awarded) and on flows of PhD holders display patterns of mobility and of brain drain. If they can be matched with FP-funded activities (e.g. PhDs from FP-funded research, Marie Curie actions), the FPs' contributions can be assessed.

Table 2 on the following page summarises the proposed potential priority indicators.

### **2.3 State-of-the-art and recommendations for further development of the methodological framework**

There is only a very limited potential for standardising evaluation-oriented indicators, because their choice and definition are closely linked to the individual questions and conditions of each evaluation. Therefore, an enhanced planning of indicator use is crucial for the success of future evaluations. In their preparation, necessary indicators should be identified early so that their production can be launched in time.

A further important issue is the improvement of preconditions for indicator use in FP planning and execution. This includes the availability of better baselines, against which achievements can be measured. This requires a conscious effort in the formulation of programme objectives to define them in a way which enables evaluators to measure their achievement (against spelled-out targets, competing regions or other relevant baselines).

The detailed discussion in Volume 2 shows that indicators and data to produce them are available in many, but not in all indicator areas. The analysis points to a number of gaps where substantial methodological advances are still necessary. For example, the methodological base of the indicator domain 'structuring effects', which is still in its infancy, should be developed further.

However, the resulting list of improvement areas should not simply be translated into a list of studies to be launched. In certain cases synergies can be used.<sup>32</sup> Therefore good links between the Commission services in charge of FP evaluation, other institutions active in this field and specialists for the methodological development of indicator definition, production and application<sup>33</sup> are crucial for an efficient advancement.

Methodologies and guidelines for their application should also be documented and disseminated consequently to ensure their consistent use and availability for future evaluations.

---

<sup>32</sup> e.g. in FP7's Social Sciences and Humanities Areas (which contains a specific action line on indicators), other direct contracts of the European Commission (like in bibliometrics or higher education) or specific calls (like ERC support actions). Some issues are also subject to other specialist work funded on a national or other base.

<sup>33</sup> for example the PRIME network and its spin-off, the European Network of Indicators Designers.

Indicator	Measure	Use	Feasibility <sup>34</sup>
<i>Programme management indicators</i>			
Proposals demography	No. of participation in calls per institutions and ratio of new entrants per proposal	Structuring effects on institutions (laboratory level) taking part in FP proposals	2
Success rates	Success rates; by FP activity, country and type of organisation	To look for success rates and to identify domains/countries where success rates are higher or lower than the average	2 - 3
Resources for proposals	Cost per proposal submitted	Analysing the actual workload of submitting a proposal	4
Management efficiency	Deviations from announced schedules, Actual proposal/contracts/payment times	Evidence for management evaluation. Possibilities to improve efficiency	1 - 2
Administrative bottlenecks	Index based on survey scales	Analysing management and other bottlenecks for project implementation	4
<i>Indicators for participation and funding</i>			
FP participation	Number of participations; breakdown by country, region, performer, FP activity	General information about the repartition of participation; achievement of FP objectives	1 - 2
Funding from FP	Repartition of funding; breakdown by country, region, performer, FP activity	General information about the repartition of funding; achievement of FP objectives.	1 - 2
Participants demography	Repartition of participants by number of FP participations; Breakdown by activity/country	Identify the core of repeating participants to FP activities and their degree of concentration. Look for continuity in participation between FP6 and FP7	2
<i>Indicators for scientific products and results</i>			
FP participants publication volume	Number of scientific peer-reviewed research publications of FP participants, World/EU market shares (per important research area, in Top10 Journals, per country or participant)	Assessment of the bibliometric "weight" and overall visibility of participants, for instance under the aspect of catching-up, critical mass, etc. Assessment of attractiveness of lead scientists	2 - 3
Impact factors of participants	Impact factors of authors or research teams involved in FP7 vs. impact factors of EU as a whole in the same themes/areas	Assessment of the overall visibility of actors, for instance under the aspect of catching-up, critical mass, etc. Assessment of attractiveness of lead scientists	2 - 3
Networking	Correlation indices between international co-publications of participants and impact factors, per various instruments, country and theme	Tracking the evolution of collaborative research through identification of networks generated or enhanced by the FPs and assessment of impact on the thematic scientific landscape; assessment of durability of effects	2 - 3
	Indicators for bibliometric networks (co-publications and citations), by country and by theme (Collaboration: volume/ratios of co-signed articles, collaboration patterns)	Zoom on patterns of citation between participants, position, role (centrality betweenness...) and particularly for the case of the lead scientists/main stream in the various instruments of the FPs	2 - 3
Excellence	Size-corrected indexes and 'productivity' measures combining bibliometric impact indexes and input data	Assessment of effectiveness / average visibility or prestige of outputs attributed to FP projects or programmes of the FPs; assessment of the effectiveness for a specific programme (output related to input)	2 - 3

<sup>34</sup> 1 = Indicator and data fully available; 2 = indicator and raw data available, processing and cleaning necessary; 3 = indicator and data partially available; 4 = indicator and data not available today. For details see Volume 2, Chapters 3 - 9.

Indicator	Measure	Use	Feasibility <sup>34</sup>
<i>Indicators for technological products and results</i>			
FP participants' technological output	Counts of patents	Assessment of the innovative performance and technology progress	2
	Counts of licensed patents	Assessment of FPs innovative technology outputs' commercialisation value and market potential	2 - 3
Technological output producers	New human resources generated by FPs	Capability of FP-funded activities to produce new human knowledge resources	2 - 3
	Innovation activities at firm level (Different indicators)	Measurement of the impact of FP participation on the innovation behaviour of firms	2
Macro-level indicators	European and national innovation	Measurement of innovation capabilities of European and national economies, if possible correlated with FP participation	2
	Technology products exchanges (Technology Balance of Payments)	Assessment of national payments and transaction on technology goods and services, contract forms and trading rules	1
<i>Indicators for economical, social and environmental impact (focus on indicators feasible short-term; see Volume2, Chapter 7 for long-term prospects)</i>			
FP participation patterns	Country, region and sector participation intensities	Identification of overall FP involvement and of participation patterns as a basis for comparison with economic performance, etc.	1 - 2
Relationship between firms' FP participation and their economic performance	Labour productivity, employment (growth), economic growth, value added (for sectors), exports	Assessment of overall performance of national economies	1 - 2
	Time lags between evolution of economic performance data and FP participation	Identification of correlations between economic performance and FP participation, assessment if the FP stimulates a dynamisation	2 - 4
<i>Indicators for structuring and integrating the ERA</i>			
<i>Short- to medium term (based on available data and concepts for upcoming evaluations)</i>			
Excellence indicators	Relative position of European research	Bibliometric data describing European positions in the high tail of scientific quality by domain, emergence of excellent players	2 - 4
Scientific collaboration	Co-authorships	Mapping of scientific collaboration relationships in the ERA. If FP participant identification is possible, these can be related to FP projects	3 - 4
	Repeated collaborations	Identification of sustainable research collaboration relationships, if possible correlated with FP participation	3 - 4
Production of scientific personnel	e.g. PhDs awarded, flows of PhD holders	Identification of patterns of mobility and brain drain; assessment of FP-stimulated mobility if possible to correlate with FP participation	2 - 4
<i>Longer-term (methodological base still to be developed) ◀ systematic ERA health check desirable, consisting of indicators for:</i>			
ERA effects and rationale	Enhancement of ERA knowledge generation and exploitation, development of critical masses, strengthening of European innovation capabilities, research contributions to solving Europe's grand challenges		
Progress in ERA implementation	Development of a European internal market for knowledge, research and innovation, building valuable initiatives implemented and funded at a European level, coordination/alignment of European research policies		
FP contributions to realising ERA	Identification of structuring effects which can be attributed to the FP and of its core elements' performance in creating such structuring effects		

Table 2: Summary of proposed priority indicators

### 3 Application of the indicator-based framework in evaluations

#### 3.1 Integration of indicators in evaluation logics and workflow

To exploit the full potential of indicators, evaluators must identify early in the evaluation process those indicators (as well as other types of evidence) which are necessary to answer their core questions. This may go beyond *performance-oriented indicators* (which provide evidence for assessing the programme's outcomes and fulfilment of objectives, appropriateness, effectiveness and efficiency) and include also *context indicators* (which describe the relevant 'landscape', e.g. programme budget, participants, projects, project sizes, geographical/sectoral distribution of participants, etc.).

Such an assessment of where indicators add most value to a specific evaluation should be an integrated part of evaluation logics and workflow (see Figure 3).

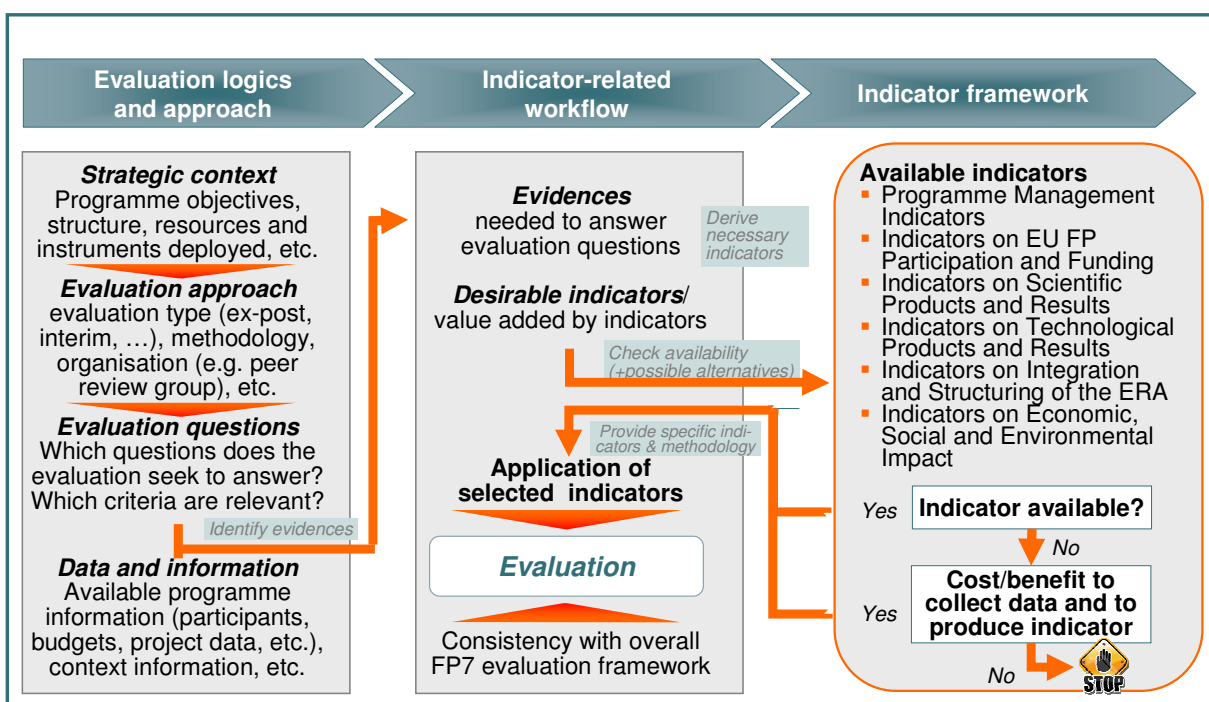


Figure 3: Integration of the proposed indicator-based framework in the evaluation process

Identification, production and application of indicators are part of the evaluators' responsibility. To do this in an efficient way, we propose to proceed in the following four steps:<sup>35</sup>

**Step 1: Identify programme rationale and establish links between objectives and outcomes**

A profound understanding of programme rationale and intervention logic is the necessary starting point for the design of the evaluation approach. At this stage, the translation of objectives into baselines against which achievements can be assessed prepares the ground for efficient choice and application of indicators.

**Step 2: Formulate evaluation questions**

Based on the programme rationale, the main evaluation questions must be formulated to focus the evaluation on key issues and to break down overall policy ques-

<sup>35</sup> Overview, for detailed examples see the following sections.

tions – which are typically too general to be answered directly – into detailed questions for which objective evidence can be provided.<sup>36</sup>

Step 3: *Identify necessary evidence and select desirable indicators*

To answer the evaluation questions, appropriate indicators and other evidence must be selected. The selection of the most appropriate type of evidence is pivotal for the appropriateness and accuracy of answers. On this basis, a limited number of indicators should be chosen which are most valuable for the evaluation.<sup>37</sup>

Step 4: *Produce and apply selected indicators*

In real life, not all indicators which would be desirable from the evaluators' viewpoint are immediately available or can be produced with a reasonable effort. For example, necessary data or methodological prerequisites for their construction may not be available. Therefore, 'ideal' and 'feasible' indicators must be reconciled to specify the 'best possible' set of indicators which can be produced and applied in the context of each evaluation.

The following sections describe the conceptual key elements of this approach:<sup>38</sup>

### **3.2 Identification of programme rationale and establishment of the link between policy/programme objectives and achievements**

To assess the relevance of formulated objectives, the appropriateness of their translation into programme design and thematic priorities, instruments and other programme elements and the efficiency of programme execution, evaluations should start with an analysis of objectives, differentiating between three levels:

(1) *Overarching policy objectives*

The *identification of relevant policy objectives and criteria for their achievement* should convert (as far as possible) broad policy objectives, which are often only formulated in general terms, into measurable performance measures and related baselines, against which progress can be measured (for example status quo vs. target level). Such measures can be constructed for a number of policy objectives.<sup>39</sup> But in other cases, elusive formulations force evaluators to define and validate workable proxies themselves to cope with objectives like making "the EU the most dynamic, competitive, sustainable knowledge-based economy..."<sup>40</sup>

(2) *Objectives of the FP and of its core elements*

A key assumption for pursuing broader policy objectives under the FPs is that research and development are major contributors to achieving economic growth, improved quality of European citizens' life, etc. This leads to the overarching aim of FP7 to contribute to "the Union becoming the world's leading research area". But such a formulation is far too general to measure its achievement directly. Therefore evaluators must break it down

---

<sup>36</sup> For example, the question „Has the FP been a success?“ is by far too general to be answered in a meaningful way without providing detailed reasons. For this purpose, it must be broken down into more precise questions, for which evidence can be provided – like “Has the FP achieved the desired contributions to mobilising the research potential of ...?”.

<sup>37</sup> The cost/benefit of creating indicators is discussed in Volume 2. This suggests to limit the production of indicators to those which have a high value added and are feasible with a reasonable effort.

<sup>38</sup> For further details, see the in-depth background descriptions in Volume 2, Chapters 3 – 8.

<sup>39</sup> For example: GDP development as a measure of economic growth, employment rate, technology balance of payments or environmental effects (e.g. measures for air, soil and water pollution).

<sup>40</sup> This objective contains a variety of components, including for example the level of research expenditure, European position in important research fields, evolution of private sector R&D investment, etc. Defining a proxy for the overarching objective requires composite indicators which are able to combine the elements to a 'big picture'. These do not yet exist to the desired extent.

into workable elements which can be answered on the basis of specific evidence. This takes place on two levels:

- *Identification of overall FP objectives and criteria for their achievement*

Formulated FP7 objectives include the promotion of world-class research, the stimulation of trans-national European research cooperation, enhanced dynamism, creativity and excellence of European research, strengthening the human potential of European research and structuring effects.<sup>41</sup> Baselines may be derived from relevant policy documents.

- *Identification of objectives of relevant FP elements and criteria for their achievement*

An important change in FP7 was that more room was given to the Specific Programmes and thematic priorities to respond autonomously to their specific needs. Therefore, intervention logic and success must be analysed also at this level, similar to the analysis at the overall FP level.<sup>42</sup>

### (3) Operational objectives

To assess if the FPs are implemented efficiently, evaluators must identify operational objectives and criteria for their achievement. At this level, objectives are usually very tangible. For criteria like the efficiency of programme execution performance measures can be formulated more easily (e.g. processing time of proposals and funding decisions, cost/effort allocated to administration, etc.).

The identification of objectives and criteria at these levels must build on a good understanding of how the FP elements interact within the overall FP as well as with other policy domains. Overarching policy objectives, the FP's objectives and those of the Specific Programmes, individual action lines, etc. cannot be treated as completely independent variables. They are interconnected in various ways.<sup>43</sup> It would be naive to assume that contributions of different FP elements can simply be 'added up' to obtain a measure for the overall FP's success.

Now, the next step is to analyse how actions which are initiated in response to the programme's objectives are linked to results. These causal links should be established in the design of the programme (e.g. in ex-ante evaluations) and justify the allocation of resources. Interim and ex-post evaluations validate these assumptions.

An established tool for representing these links is the 'Logic Chart' approach. It seeks to establish the relationships between broader policy objectives, programme-specific objectives, (expected) outputs and longer-term effects in a diagrammatic representation. Displaying these hierarchical relationships highlights logical links and interdependencies and permits the assessment of programme consistency.<sup>44</sup> Individual activities are performed within a limited time and lead to immediate outcomes, which can usually be observed with a relatively short delay. Programme objectives are typically closely related with intermediate impacts which occur often only several years after the end of the FP and are more difficult to observe. Ultimate impacts are correlated with broader policy objectives. But their time horizon may range

---

<sup>41</sup> see European Parliament (2006).

<sup>42</sup> This has to be done for each Specific Programme, action line or instrument separately to account for their high diversity. For example in the 'Collaboration' programme, analysis of objectives might focus on sector-specific scientific or technological targets. To some extent, it is also possible to assess structuring effects, see for example the evaluation approach in the ICT thematic area by DG INFSO ([http://ec.europa.eu/dgs/information\\_society/evaluation/index\\_en.htm](http://ec.europa.eu/dgs/information_society/evaluation/index_en.htm)).

<sup>43</sup> For example, research carried out under the 'Collaboration' programme has also a structuring effect, as well as measures taken under the 'People' (e.g. to enhance mobility of researchers), the 'Ideas' (e.g. through fostering trans-European excellence-oriented research collaboration) or the 'Capacities' programmes (e.g. through the development of trans-European infrastructures) do.

<sup>44</sup> For details see Volume 2, Chapter 2.

up to decades. For this reason and because of the described interference with other policy initiatives, they are most difficult to assess.

Figure 4 shows an illustrative example for a possible FP Logic Chart. Once drawn, it serves as the basis for the formulation of evaluation questions and for the identification of necessary evidence, including indicators, for the evaluation.

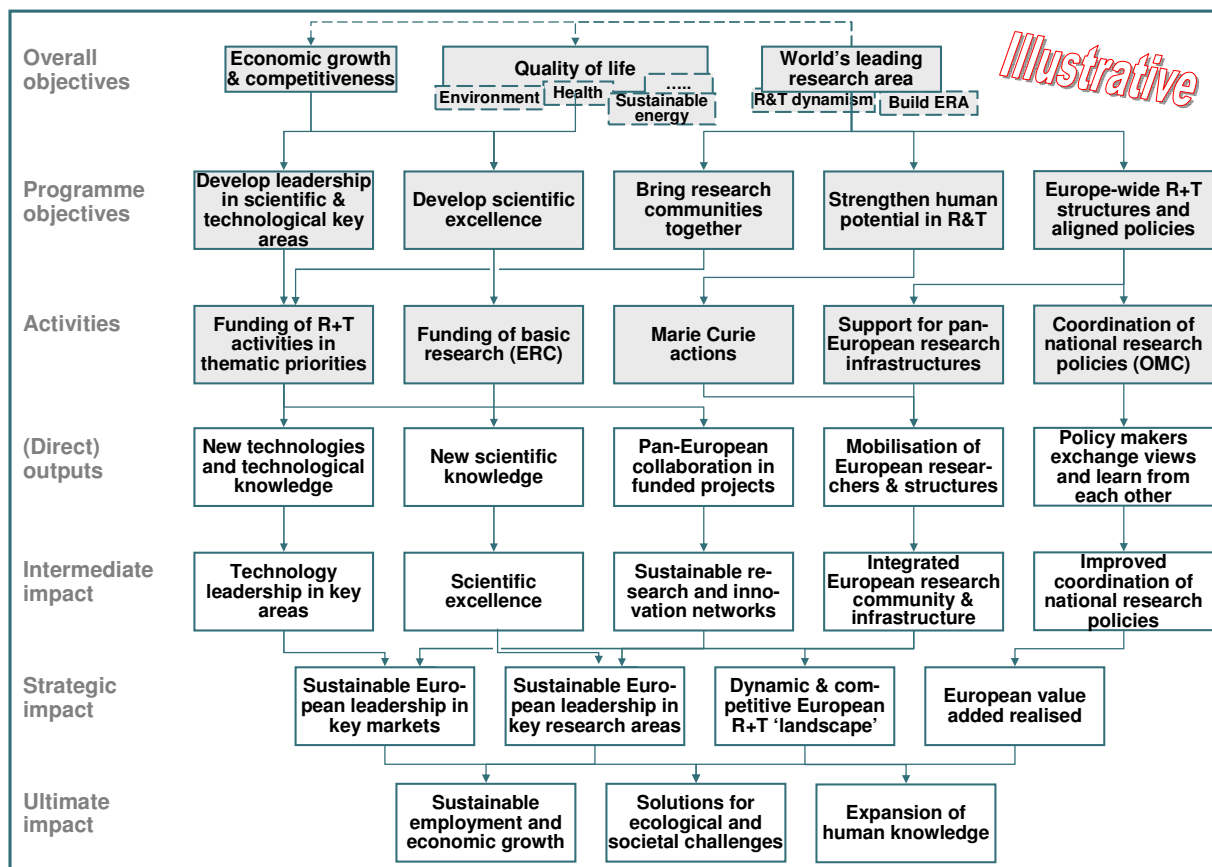


Figure 4: Illustrative attempt to construct a Logic Chart for the FP

### 3.3 Formulation of evaluation questions

The formulation of evaluation questions may start with a single top-level question which expresses the overall goal of the evaluation in a very condensed form.<sup>45</sup> Typically, this question asks about the overall success of the programme. Such a top-level question is usually far too general to be answered in a sufficiently precise and objective way on the basis of facts instead of unspecific convictions. Therefore, it must be broken down into more specific sub-questions, which can be answered in a reliable and objective way, based on facts.

In such a hierarchy, evaluation questions become increasingly specific until a level of detail is reached where they can be answered in an objective way using facts. Questions may be structured for example in a way which differentiates between the programme's appropriateness ("Are we pursuing the right objectives?"), its effectiveness ("Are we doing the right things?") and its efficiency ("Are we doing things right?"). This leads to the following structure:

<sup>45</sup> This approach is partially based on Minto (1987).

**Top-level question:**

- 1 Has the Framework Programme been a success?

**Detail-level questions:**

- 1.1 Have the FP's initial objectives proven to be valid?
- 1.2 Has the FP achieved its objectives and created the targeted outcomes and impacts?
- 1.3 Has the programme been executed efficiently?

As far as needed, these questions can be broken down further until a level is reached, where they are sufficiently specific to be answered on the basis of appropriate evidence. For question 1.2, this could be for example in the following way:

- 1.2.1 Have the specified immediate outcomes been produced?
- 1.2.2 Have the specified intermediate impacts been created?
- 1.2.3 Has the FP made desired contributions to achieving overarching policy objectives?

If this approach is applied consequently, a comprehensive and structured set of evaluation questions emerge, as shown in Figure 5.

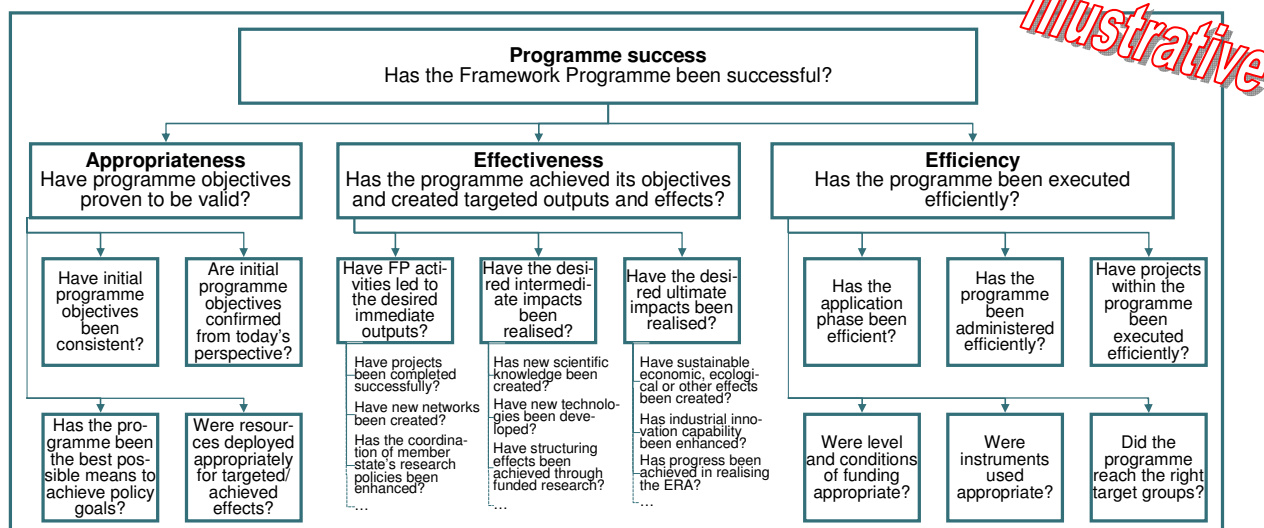


Figure 5: A possible hierarchy of questions for a hypothetical FP evaluation situation

In a real-life evaluation situation, these questions and evidence must be adapted and refined as a function of the type of evaluation and its key issues (overall FP ↔ Specific Programme ↔ ...), its timing (ex-post ↔ interim ↔ ex-ante), the availability of data, etc.

### 3.4 From evaluation questions to evidence and indicators

Once the evaluation questions have been identified, the necessary evidence must be specified in the next step. Here, indicators compete with other types of evidence (see Table 3). This choice cannot be undertaken in a mechanistic way. It requires a deep understanding of the evaluation questions and of the characteristics of different types of evidence.<sup>47</sup> In this process, the number of indicators produced and applied is important. Typically, a limited number of indicators provide the bulk of the desired

Method	Description
Peer review/ expert judgment	Opinions and recommendations are sought from experts specific to the field
Survey	Asks multiple parties a series of questions, generates both quantitative and qualitative data
Bibliometrics	Analysis of publications and citations data, using it as an indicator of science quality
International benchmarking	Comparison of FP research with research quality in other countries, usually by peer review
Case study	In-depth analysis of a sample of a specific issue(s)
Historical tracking	Tracing backward from a research outcome to identify precursor developments
Social analysis	Identifying and studying the contribution of research to quality of life
Economic analysis	Identifying the economic benefits and effects of research activities

Table 3: Typical types of evidence used in evaluations<sup>46</sup>

evidence. Overextending this number creates extra effort without adding significant value. In addition, their feasibility must be considered, bearing in mind that the 'ideal' indicator (which would be desirable from the viewpoint of the evaluation questions) may not always be available. Therefore it must be reconciled with the best 'feasible' indicator (which is defined by the availability of reliable data and a sound methodological base).

Depending on the type of evaluation questions, indicators may be taken from different indicator categories (as defined in Chapter 2). Table 4 proposes a first general categorisation in which domain appropriate indicators may be found.

Category of evaluation questions	Programme Management indicators	Indicators for FP participation & funding	Scientific output indicators	Technological output indicators	Longer-term impact indicators	ERA structuring indicators
Appropriateness of objectives		✓	✓	✓	✓	✓
Achievement of immediate outputs			✓	✓		
Achievement of intermediate impacts			✓	✓	✓	✓
Achievement of ultimate impacts			✓	✓	✓	✓
Achievement of structuring effects		✓	✓	✓	✓	✓
Efficiency of programme execution	✓	✓	✓	✓		

✓ = strong correlation; ✓ = less strong correlation

Table 4: Link between categories of evaluation questions and indicator areas

<sup>46</sup> Source: BBSRC (2005).

<sup>47</sup> For example, ex-ante, interim and ex-post evaluations have very different purposes, ranging from the anticipation of expected results and effects (in order to orient resources towards the best research fields) to the assessment of achievements. In these different situations, data availability and resulting possibilities to construct indicators vary also considerably.

The value of the different indicator categories for assessing the achievement of objectives depends also on the type of objective whose achievement shall be assessed (see Table 5).

On this basis, evaluators can now select those indicators which are most suitable for answering their specific key questions. They may draw on already existing indicators, if one of those answers the question. In other cases, available data must be checked for possibilities to construct complementary indicators. The in-depth description of the proposed indicator domains in Volume 2 provides the necessary conceptual background for this task. If no data are available off the shelf for certain questions, the evaluation team will have to collect data itself, for example through questionnaires, etc.

Indicator type	Broader policy objectives					Specific programme element objectives						
	Econ. growth/employment	Competitiveness of European industry	Improved sustainability/quality of life	Performance of Europ. Research system	European added value	Development of the ERA	Scientific/technol. leadership in key areas	Enhanced pan-European research collab.	Enhanced mobility of researchers	Enhanced involvement of SMEs	Efficiency of instruments deployed	Efficiency of FP programme management
Programme management indicators					✓	✓				✓	✓	✓
Indicators for FP participation and funding		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
Indicators for scientific products and results		✓	✓	✓	✓	✓	✓	✓			✓	✓
Indicators for technological products and results	✓	✓	✓	✓	✓		✓	✓			✓	
Indicators for structuring of the ERA		✓		✓	✓	✓	✓	✓	✓		✓	
Indicators for economical, social & environm. impact	✓	✓	✓		✓							

Legend: ✓ = high relevance; ✓ = existing, but limited relevance

Table 5: Generic allocation of indicator types to policy/programme objectives (examples)

Table 6 on the following page shows some illustrative examples for resulting links between evaluation questions and possible indicators to answer them.

*Illustrative examples*

Evaluation question	Performance target	Evidence/indicator
<i>Effectiveness-related questions</i>		
Has new scientific knowledge been created?	<ul style="list-style-type: none"> <li>▪ Scientific leadership in funded target areas</li> </ul>	<ul style="list-style-type: none"> <li>▪ Scientific publications/ bibliometric indicators</li> </ul>
Have new technologies been created?	<ul style="list-style-type: none"> <li>▪ Technological innovation, based on funded research</li> </ul>	<ul style="list-style-type: none"> <li>▪ Patents filed/granted</li> <li>▪ Data on new product development based on fund. research</li> </ul>
Has industrial innovation capability been enhanced?	<ul style="list-style-type: none"> <li>▪ Enhanced industrial research activities?</li> <li>▪ Efficient knowledge transfer</li> </ul>	<ul style="list-style-type: none"> <li>▪ Industrial R&amp;D expenditure in specific area</li> <li>▪ Participant feedback on use of research results</li> </ul>
Have sustainable economic, ecological or other effects been achieved?	<ul style="list-style-type: none"> <li>▪ Superior properties of new products</li> <li>▪ Sustainable competitive advantage</li> <li>▪ Contributions to economic growth and employment</li> </ul>	<ul style="list-style-type: none"> <li>▪ Progress in key technology/ product characteristics</li> <li>▪ Revenue growth/employment in specific area, based on stimulated innovation</li> <li>▪ New product rate</li> </ul>
Have structuring effects been achieved through funded research?	<ul style="list-style-type: none"> <li>▪ Enhanced trans-European collaboration</li> <li>▪ European value added to research in the ERA</li> </ul>	<ul style="list-style-type: none"> <li>▪ Number and lifetime of created collaborations and networks</li> </ul>
Have scientific and operational programme objectives been achieved?	<ul style="list-style-type: none"> <li>▪ Creation of targeted new knowledge and scientific progress</li> </ul>	<ul style="list-style-type: none"> <li>▪ Achievement of specific project/programme objectives</li> </ul>
<i>Efficiency-related questions</i>		
Has the application phase been efficient?	<ul style="list-style-type: none"> <li>▪ Efficient treatment of calls and proposals</li> <li>▪ Motivation to participate<sup>48</sup></li> </ul>	<ul style="list-style-type: none"> <li>▪ Processing time of proposals</li> <li>▪ Applicant feedback (also on clarity of calls, transparency of selection, etc.)</li> </ul>
Has the programme been administered efficiently?	<ul style="list-style-type: none"> <li>▪ Efficient programme administration</li> </ul>	<ul style="list-style-type: none"> <li>▪ Administrative cost</li> <li>▪ Project time spent on administrative issues</li> <li>▪ Participant feedback</li> </ul>
Have funded projects been executed efficiently?	<ul style="list-style-type: none"> <li>▪ High project efficiency</li> </ul>	<ul style="list-style-type: none"> <li>▪ Number of budget overruns/ unmet deadlines</li> </ul>
Were levels and conditions of funding appropriate?	<ul style="list-style-type: none"> <li>▪ Avoid over- or underspending on specific programmes</li> <li>▪ Ensure appropriate project size and duration</li> </ul>	<ul style="list-style-type: none"> <li>▪ Ratio applied/granted budgets</li> <li>▪ Participation levels of specific target groups<sup>49</sup></li> </ul>
Were instruments used appropriate?	<ul style="list-style-type: none"> <li>▪ Ensure best possible project conditions through choice of best suited instrument</li> </ul>	<ul style="list-style-type: none"> <li>▪ Ratio applied/granted budgets</li> <li>▪ Applicant/participant feedback</li> </ul>
Did the programme reach the right target groups?	<ul style="list-style-type: none"> <li>▪ Ensure involvement of target groups where research funding achieves high leverage</li> </ul>	<ul style="list-style-type: none"> <li>▪ Share of overall programme budget by target group (target vs. actual)</li> </ul>

Table 6: Relation between evaluation questions and indicators (illustrative examples)

<sup>48</sup> For example, certain potential participants are difficult to motivate because they fear an overly high administrative burden already in the proposal phase.

<sup>49</sup> For example, SMEs tend to refrain from projects whose duration, complexity or resource absorption go beyond their more short term-oriented strategic scope and/or limited resource base.

### 3.5 Integrated, life cycle-oriented framework for FP evaluations and management

A shortcoming of many programme designs is their lack of provisions for comparing anticipated and real impacts.<sup>50</sup> Evaluations must be able to build on a validated baseline against which results and success can be measured. This would require that criteria for assessing achievement of objectives are used consistently throughout the whole programme life cycle. In such an integrated system the formulation of objectives would already include (measurable) criteria for target achievement. These form the basis for programme management and for the assessment of relevance, outcomes and efficiency. If combined with a timely monitoring this would provide policy decision makers also with an ‘early warning’ system for possible deviations. Figure 6 sketches such an *integrated, FP life cycle-oriented indicator framework* which is based on a consequent use of a set of core indicators.

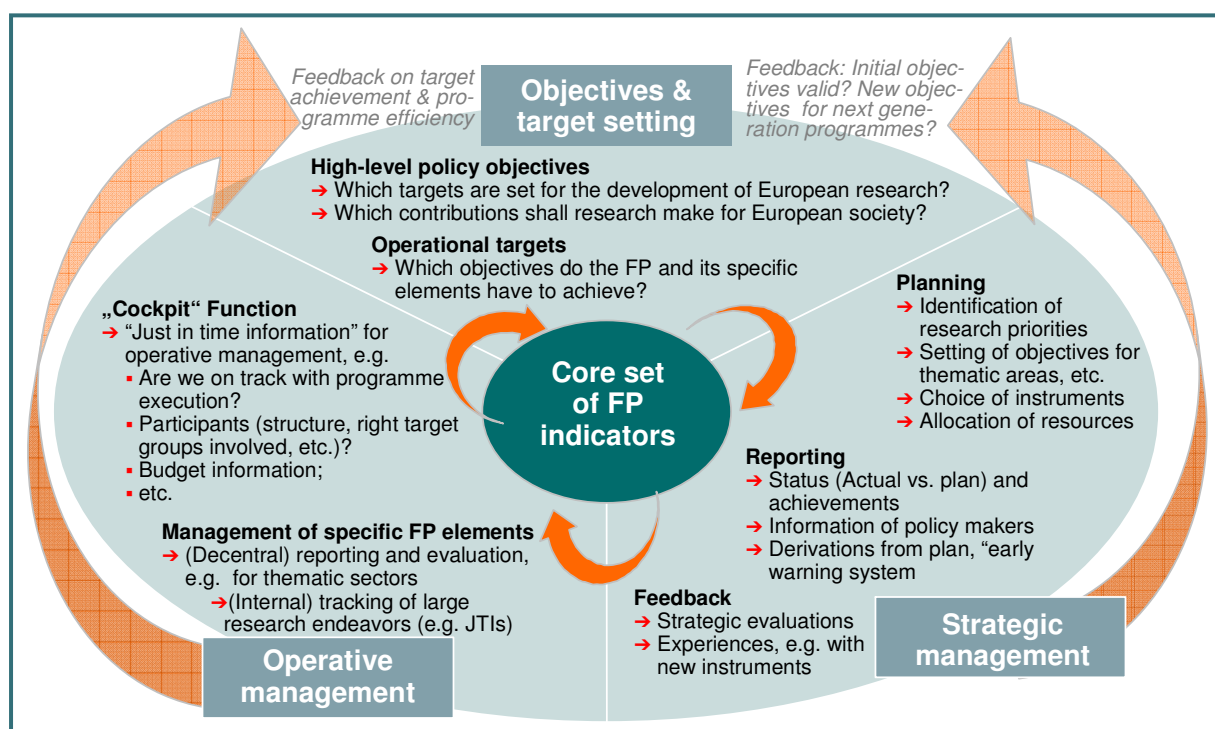


Figure 6: Use of indicators as a "going concern" throughout all phases of the Framework Programme life cycle

In this concept, the defined core indicators build the bridge between objectives, programme planning, programme execution and assessment of achievements in the following steps:

- *Planning and launch phase*

Already in the programme design phase, objectives should be linked to clear, measurable criteria – on the overall policy level as well as on the programme level. Programme objectives should be spelled out sufficiently specifically so that their achievement can be

<sup>50</sup> A recent screening of >150 impact assessments carried out by Commission services (Watson et al. (2007)) showed that ~88% of all Impact Assessments outlined some form of monitoring system for the proposed policy initiative; however, only 25% provided concrete indicators. The monitoring tools and provisions for ex-post assessments varied from DG to DG. In some cases there were specific detailed monitoring systems of quantitative economic or environmental indicators (e.g. the Thematic Strategy on Air Pollution has a scheduled review in 2012, which will make use of the data included in the Impact Assessment). In other areas there was a broader evaluation against a variety of criteria (e.g. enlargement process which was based on an impact evaluation loop, where a country's progress was assessed against the Copenhagen criteria via annual reports).

tracked.<sup>51</sup> Due to the high diversity of FP elements, such criteria may vary considerably. However, for each programme element, it should become visible which criteria define its success and how it contributes to the achievement of overarching FP objectives.

- *Programme execution*

During programme implementation, indicators which trace the formulated criteria provide continuous information about the 'health' of FP and ERA and about actual vs. expected achievement of objectives. Operative indicators inform about the efficiency of programme implementation. Thus, decision makers can also intervene if corrective action is necessary, for example if programme elements fail to achieve desired results.

- *End of programme*

If the same set of criteria and indicators is used in evaluations at defined milestones and towards the end of the programme, this makes it easier to assess if initial objectives have been achieved, if these initial objectives have proven to be valid in the light of new developments and/or if the best possible use has been made of available resources.

### 3.6 Recommendations for implementation

The detailed analysis of indicator domains in Volume 2 confirms that there is no lack of possible indicators which could contribute to FP evaluations. Their wider and systematic use would certainly contribute to transparency, efficiency and credibility of FP evaluations.

However, the analysis shows also that there is a major difference between typical traditional indicators, which are of a generic nature and produced according to standardised methodologies, and the type of indicators needed for research programme evaluation. Evaluation-oriented indicators are mostly ad-hoc indicators which have to be customised, depending on evaluation context and available data. Furthermore, instead of providing firm quantitative answers, this type of indicators tends also to provide often only partial evidence, which needs to be combined with other types of evidence or with (subjective) expert assessment to lead to meaningful answers.

Therefore, a profound change in our understanding of the role of indicators in evaluations is necessary. Even indicators which are incomplete or methodologically not fully developed can be useful if they add constructive elements to the evaluation debate. But this means that an enhanced indicator use will not make evaluations automatically simpler, faster or cheaper. Even an additional resource investment may become necessary. Evaluators must weigh carefully the new insights obtained against the additional effort of creating new indicators.

Implementation of such a framework requires considerable time and effort. For evaluations which are due in the next years, there is not enough remaining time to generate new data or indicators, whose collection and production has not yet been instigated. Therefore, within this timeframe the focus should be on providing optimal indicator-based support on the basis of existing indicators and available data.

Denominated evaluators should specify as soon as possible evaluation questions and to derive from these their desired 'ideal' indicators. The proposed generic structure of evaluation questions (see Figure 5, page 26) can serve as a starting point.<sup>52</sup> To provide the necessary evaluation inputs, the generic set of data which was proposed in Table 2 can also be used as a starting point. Depending on the evaluators' expressed preferences, further indicators can be supplied using the sources which are discussed in the relevant chapters of Volume 2. The

---

<sup>51</sup> A typical example is the 3% target. Since it has been introduced, both the EC as a whole and its member states can follow regularly their progress and take appropriate measures.

<sup>52</sup> Necessary adaptations can be taken from Chapters 3–8 of Volume 2, depending on the evaluators' priorities and needs.

'best feasible' set of indicators results from a comparison of 'demand side' and 'supply side' (See Table 7).

Evaluation question	Desirable evidence		Indicator availability/feasibility	Data source	Alternative?
	Indicator	Other			

*Illustrative*

*Table 7: Proposed scheme for selecting indicators for upcoming evaluations*

Indicators which are already available can be used immediately on an 'as is' basis. Indicators, for which the necessary data are available and which can be constructed with a reasonable effort should be produced – if possible – already in advance. Indicators for which the necessary data are not (fully) available or which cannot be constructed with a reasonable effort within the given time frame, should be replaced by the best available proxies (or, if this is not possible, disregarded) for this specific evaluation. But they should be considered as important candidates for the longer-term development of a consistent indicator framework, which is described in the following section.

Beyond this short-term stream of action, a systematic longer-term initiative to implement the proposed amended framework consequently should be pursued to ensure optimal conditions for indicator use in future FP monitoring and evaluations. Its priorities should be on a consequent implementation of the proposed indicator-based methodological framework and further development of the methodological foundations as described in Chapter 2.

## 4 Application of indicators in FP monitoring

### 4.1 Development options for the FP reporting ‘landscape’

The discussion of current practices in Chapter 1.2 has confirmed that there is a rich, but heterogeneous FP reporting ‘landscape’. Its elements fulfil different purposes to a varying extent:

- *Basic information function*

The most basic type of monitoring, limited to a regular representation of relevant facts (usually in the form of statistical data) is provided by the *Key figures of Science, Technology and Innovation* which are published by the EC (European Commission (2007g)<sup>53</sup>), by the statistics published by Eurostat.<sup>54</sup> the CORDIS database and other sources.<sup>55</sup> But not all necessary raw data which are necessary to produce the proposed indicators are available from these sources.

- *Analytical function*

A more profound analysis of the situation of European research, FP implementation status and underlying causalities requires an interpretation of facts, based on in-depth analyses and expert judgement. This is achieved today especially through the *Annual Report* and *FP Monitoring Report* exercises.<sup>56</sup> The *Annual Management Plan* deepens this analysis further for Commission-internal planning purposes (see Chapter 1.3). But such profound analyses at this level of detail require considerable effort and time. Depending on applied practices there may also be a danger of limited consistency between documents or of limited continuity in the observation of important parameters, as the focus of expert opinions may change over time.<sup>57</sup>

- *‘Health check’ function*

Both the *Annual Report* and the *Annual Management Plan* highlight shortcomings and opportunities and provide – to a certain extent – an insight into the ‘health’ of European research with the help of defined indicators. Under the new FP monitoring principles, the *FP7 Monitoring Report* uses also a range of indicators to make the programme’s implementation status transparent.<sup>58</sup>

However, this information is embedded in relatively large, extensive reports which are generated in resource-consuming exercises. Therefore, such information about the ‘health’ of ERA and FP is only available with some delay and addresses primarily experts. It is less easily ‘digestible’ for high-level decision makers and/or a larger audience.<sup>58</sup> Currently, there is no complementary ‘rapid snapshot’ format, which provides this type of information in a rapid, concise format which can be communicated easily.

The current reporting ‘landscape’ supports especially the basic information and analytical functions (see Table 8). Efficient FP governance would benefit from further enhancements of this reporting ‘landscape’. Even if publicly funded research enjoys – for valid reasons – a

---

<sup>53</sup> See also [http://ec.europa.eu/invest-in-research/monitoring/statistical01\\_en.htm](http://ec.europa.eu/invest-in-research/monitoring/statistical01_en.htm).

<sup>54</sup> A typical example on a national level are the databases on the US national science and innovation base, provided by the NSF’s Division of Science Resources Statistics, see <http://nsf.gov/statistics/>. For details, see Volume 2, Chapter 1.

<sup>55</sup> For details, see Volume 2, Chapter 1.

<sup>56</sup> See Chapter 1.2, page 7 and Volume 2, Chapter 1.3. Similar approaches are pursued on a national level in several countries. For example in Germany, after a recent reorganisation the *Bundesbericht Forschung und Innovation* (which has a basic information function) is complemented by an independent high-level expert group’s assessment. Both shall be published biannually in a coordinated way (Source: BMBF (2008)).

<sup>57</sup> Recent changes in the systems address the latter issue. For example, the new *FP7 Monitoring Report* introduces a set of indicators which shall be used continuously (See Chapter 1.2, page 8).

<sup>58</sup> This is the case in certain other areas. For an example, see the European Sustainable Development Indicators (Eurostat (2007) and literature quoted therein).

high degree of autonomy,<sup>59</sup> actual information about achievement of objectives, generated/anticipated outputs, impacts, etc. supports better-informed decision making. Actual information, generated with a justifiable effort and presented in a concise, action-oriented format, enhances decision makers' timely understanding of the FP's likelihood to achieve its objectives, points to possible areas where corrective action must be taken and creates first insights into underlying causes.

Source	Function			Target Group			Remarks
	Basic inform.	Analytical	'Health check'	Policy makers	Progr. mgmt.	Stakeholders	
Statistics (Key Fig., Eurostat)	✓					✓	
CORDIS	✓					✓	
Online datab. (eCorda, etc.)	✓				✓		Not publicly available
Annual report		✓	✓	✓	✓	✓	
FP7 Monitoring Report		✓	✓	✓	✓	✓	
Ann. Management plan		✓	✓		✓		Not publicly available

Table 8: Overview over main elements of the current FP-related reporting 'landscape'

Therefore we suggest considering two possible areas for amendments: A *review and extension of the set of indicators used for FP monitoring* could extend the information provided by monitoring beyond the discussed limits of currently used indicators (See the discussion in the evaluation context in Chapter 3). As methodological progress will make for example new composite indicators available, these should also be incorporated in reporting systems. A *review of the FP-oriented reporting 'landscape'* could combine the merits of the three functions described on page 33 for example though increased use of synergies between monitoring activities or the introduction of a complementary pragmatic 'health check' function.

## 4.2 Indicators for FP monitoring

The indicator categories proposed in Chapter 2 are in principle also valid for monitoring. Their use as an extension of the current indicator sets applied in *Annual Report*, *FP7 Monitoring Report* and *Annual Management Plan* might be helpful to enhance coverage of relevant issues in these key documents.

However, the choice of indicators and their application are different:

### (1) Input and participation dimension

To ensure that resource deployment and participation are in line with the FP's plans and objectives, programme responsables need continuous information about invested resources, about their allocation to programme elements and about participation. Such actual information can be obtained in particular from FP documentation. Currently, efforts are undertaken to improve quality, consistency and availability of data. To support an enhanced reporting (e.g. timely observation of participation patterns), these should be combined with more reporting-oriented data structures.

<sup>59</sup> State-of-the-art methodologies for monitoring large research project portfolios are current in the private sector. However, such concepts cannot be transferred 1:1 to the management of a large public research programme like the FP, because the circumstances and stakeholder needs are completely different. For example, in private sector R&D, corporate management can modify or stop projects usually at any time if it does not fulfil its targets or becomes strategically obsolete.

(2) *Direct output dimension*

In this area, monitoring should create transparency about the success of funded research in terms of scientific and technological outputs. As most funded projects are still in progress during reporting periods, typically their scientific and technological results cannot (yet) be assessed accurately. However, for programme management it is important to gain at least a feeling for their likelihood to deliver expected outputs. Therefore, appropriate proxies must be sought, for example on the basis of extrapolations from available project/programme information or from participants' self-assessment in mid-term project reporting, etc.<sup>60</sup> However, this requires that reporting formats provide more codifiable information about progress instead of open questions.

(3) *Mid- to long-term impact dimension*

Mid- to long-term effects of FP-funded research occur in general only with a considerable delay after the end of funded research. Therefore, appropriate proxies must be sought also here. Ideally, these should indicate if the FP has the desired effects on European innovation performance, competitiveness, etc. and creates the targeted socio-economic and other benefits. Certain measures like tracking of Europe's position in high-tech sectors are available. But as discussed in more depth in Chapter 2.2, time lags and attribution problems make it difficult to establish causalities.

(4) *Structuring effects dimension*

Today, monitoring elements which inform about progress in ERA implementation and the FP's contributions to structuring European research exist only in a rudimentary form. Because this is a relatively new area of interest and due to the complex nature of underlying causalities, appropriate proxies still have to be developed. Developing the methodological base for this and its inclusion in reporting systems should be pursued with high priority.

(5) *Efficiency dimension*

The monitoring objective in this area is to inform decision makers about the efficiency of FP implementation. This includes for example administrative efforts, throughput times for proposals, etc. For this purpose, data are available from FP reporting, e.g. the number of proposals received per call, average processing time of proposals, administrative share of overall budget, etc.<sup>61</sup> These help to trace the efficiency of programme implementation.

Table 9 on the following page shows a first attempt to identify possible relevant monitoring indicators on this basis.

---

<sup>60</sup> Since the Specific Programmes pursue different objectives and approaches, different output types must be considered. The 'Collaboration' Programme yields project-specific research results which can be specified rather precisely. The outputs of the 'Ideas' programme cannot be measured on the same basis. Here, measures should be directed towards achievement of the excellence target. To trace the direct outputs of the 'People' Programme, indicators are needed which provide evidence for impacts on the relative strength and development of the human resource base of European research, mobility, etc. For the 'Capacities' programme, which is mostly devoted fostering the integration of European research, relevant indicators still need to be defined.

<sup>61</sup> For an example, see the working paper Performance indicators in the frame of Networks of Excellence - Methodological guidance note, undated, available from [ftp://ftp.cordis.europa.eu/pub/fp6/docs/performance\\_indicators\\_noes.pdf](ftp://ftp.cordis.europa.eu/pub/fp6/docs/performance_indicators_noes.pdf). This document identifies a range of potential performance indicators for Networks of Excellence but does not specify or quantify them further.

Key issues	Possible indicators
<p>Input/particip. Are FP resources spent appropriately?  Member states participating on an even level? Target groups reached to the desired extent? ...</p>	<ul style="list-style-type: none"> <li>▪ Actual funding allocations vs. importance of research areas</li> <li>▪ Distribution of allocated funds per Specific Programme, thematic priority, etc.</li> <li>▪ Funding share/member state</li> <li>▪ Funding share/target group</li> <li>▪ ...</li> </ul>
<p>Direct output Do funded projects yield the targeted new scientific knowledge? Do funded projects yield the targeted new technologies? Will projects be completed successfully? Does FP funding attract excellent European researchers? Does the level of European research excellence increase?  Does the FP contribute to the emergence of European research excellence? ...</p>	<ul style="list-style-type: none"> <li>▪ Bibliometric measures (availability?)</li> <li>▪ Interim progress/expectations of participants</li> <li>▪ New patents (availability?)</li> <li>▪ Interim progress/expectations of participants</li> <li>▪ Actual vs. plan (participant assessment)</li> <li>▪ Degree of participation of European top institutions</li> <li>▪ Timeline; compared with other leading regions, ... → timeline can show trends, but impact of actual research only visible later</li> <li>▪ Measures for relative research position (e.g. weighted share of publications) → timeline can show trends</li> <li>▪ ...</li> </ul>
<p>Mid-/long-term effects Does the EC make the targeted progress towards policy objectives, e.g. the 3% target? Does Europe's position in key technology areas improve? Does Europe's competitiveness, position and employment in technology-intensive sectors improve? Does the FP make the desired contributions to achievement of overall policy objectives? ...</p>	<ul style="list-style-type: none"> <li>▪ Actual vs. target</li> <li>▪ Measures for relative technology position (e.g. patent analysis)</li> <li>▪ Measures for competitive position (e.g. market share in high-tech goods)</li> <li>▪ Employment rate (espec. high-tech sectors)</li> <li>▪ (no good measures available today)</li> <li>▪ ...</li> </ul>
<p>Structuring effects Is progress in the realisation of the ERA achieved, as expressed by achievement of key objectives?  Does the FP make the desired contributions (in comparison with other policy measures)? ...</p>	<ul style="list-style-type: none"> <li>▪ Level of researchers' mobility between institutions, disciplines, sectors and countries</li> <li>▪ Degree to which world-class research infrastructures emerge</li> <li>▪ Emergence of top clusters and research communities</li> <li>▪ level of effective knowledge-sharing, notably between public research and industry</li> <li>▪ Level of coordination of research programmes and priorities across Europe?</li> <li>▪ Level of cooperation and research partnerships between Europe and its global partners</li> <li>▪ (no good measures available today)</li> <li>▪ ...</li> </ul>
<p>Efficiency Do the applied instruments achieve the desired effects? Are they accepted and relevant from the participants' perspective? Are administrative processes (e.g. processing of proposals, etc.) executed efficiently? Is the overall administrative effort minimised? ...</p>	<ul style="list-style-type: none"> <li>▪ (measurable by indicators – or studies required?)</li> <li>▪ Efficiency measures (e.g. average processing time of tenders, funding decisions, etc.)</li> <li>▪ Measures for participants efforts (e.g. time spent on administrative requests)</li> <li>▪ Administrative share of total budget</li> <li>▪ ...</li> </ul>

Table 9: Possible key issues and indicators for the relevant dimensions of FP monitoring

### 4.3 Complementary 'health check' function

A complementary 'health check' format could provide decision makers and stakeholders with rapid information about the FP's status and point to possibly critical areas, where a more profound second view analysis and/or corrective action might be necessary. This approach focuses the attention on a limited number of priority areas, whose status is communicated in a straightforward, user-friendly and convincing format. This could be realised in a three layer pyramid indicator structure, as depicted in Figure 7.<sup>62</sup>

- At the top of the pyramid, *Lead* (or level-1) indicators provide critical strategic information if FP or ERA are 'on track'. These must be well-established indicators, which are robust, continuously available for all EU Member States and have a high communication value.
- At the second level, *FP Priority objective* indicators focus on the achievement of strategic FP objectives which stand for the overall FP and ERA 'health'. As the lead indicators in their respective sub-themes, they provide first indications for possible reasons of deviations. They must also be robust and available for EU Member States.
- The third level of *explanatory variables* provides complementary information, for example to gain a better understanding of the reasons for achievement (or non achievement) of higher-level objectives. This level consists of indicators related to actions outlined in the FP strategy or to other issues which are useful to analyse progress towards the FP objectives.
- *Contextual indicators* can be added where necessary. They are not intended to monitor directly any of the FP's objectives. Instead, they provide valuable background information on issues having direct relevance for the FP.

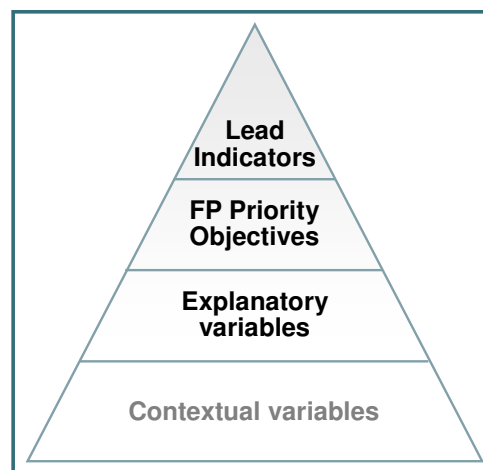


Figure 7: Possible indicator hierarchy for FP monitoring

Such a reporting tool could be worked out in collaboration with EC units involved in FP management and with other important stakeholders to ensure its acceptance and focus on those issues which its users perceive as most relevant. Table 10 shows representative examples for such a possible indicator structure. This type of reporting format might also include timelines, which show in a graphical form the evolution of critical parameters. This has the advantage that progress and remaining gaps to fulfil targets are immediately visible. To draw the readers' attention to critical areas, it might also use elements like for example the 'traffic lights' format, where 'green' indicates that the parameter is in good condition and targets are likely to be achieved, 'yellow' that development should be observed thoroughly and 'red' that target achievement is clearly in danger.

<sup>62</sup> This concept draws to some extent on the monitoring approach which is used for the European Commission's monitoring framework for its Sustainable Development Strategy (see European Commission (2007h), Eurostat (2007) and literature quoted therein for details).

Level 1	Level 2	Level 3
<p><i>Objective 1: Contribute to intensification of European Research</i></p> <p>Lead Indicator: Research intensity</p>	<b>Sub-theme 1.1: Overall research investment</b>	
	Research investment as % of GDP	Dispersion of research investment per member state Dispersion of research investment per sector, public vs. private sector
	<b>Sub-theme 1.2: Development of outputs of European research</b>	
	Measures for research output, e.g. bibliographic data number of patents	Overall measures for scientific activity (e.g. number of publications), relative to other leading global regions Measures for stimulated innovation activities (e.g. number of patents) Breakdown by member country, sector, etc.
	<b>Sub-theme 1.3: Effect of Framework Programme</b>	
	Measure for impact of Framework Programme (Additionality)	
	<b>Sub-theme (n).1: ...</b>	
	....	....
	<b>Sub-theme (n).2: ....</b>	
	....	....
	<b>Sub-theme (n).3: ...</b>	
	....	....

For illustration only

Table 10: Possible detailed hierarchy of indicators (schematic)

## 5 Necessary prerequisites

### 5.1 Priorities for further development of the data base for evaluation and monitoring

The indicator and data 'landscape' described in Chapter 1 offers access to a multitude of detailed information. But at the same time, it is very heterogeneous in terms of sources, formats, data quality and reliability. And the detailed discussion of data requirements and actual sources in Volume 2, Part B (which takes the perspective of *the requirements for indicator production*) reveals also that there are considerable data gaps in certain indicator domains. These are partially due to the absence of raw data, but in many cases also to data quality, formats and structures. FP data are collected essentially under the aspect of the requirements of administrative handling and technical follow-up of funded projects. This leads to data structures which are very difficult to use for indicator production in most cases. As a consequence, the wealth of FP-internal data is currently not fully exploited.

Among other, project reporting formats do not support the collection of usable data about project status and progress, expected/achieved outcomes and effects (e.g. development of pan-European research collaborations) beyond some very limited – and mostly qualitative – information.<sup>63</sup> Participants' inputs are delivered to a large extent in an unspecific generalised text format. This leads to heterogeneous responses which are also often too 'soft' to permit objective cumulative assessments. In addition, no reliable data can be obtained about the real achieved (vs. expected) longer-term impacts, since reporting ends with project termination.

But this is not an issue of collecting more data! One might even argue that less data should be collected from projects. In most cases, the problem is more related to the concepts behind data collection and to different needs of programme administration and evaluation. There is a need for enhanced data standardisation and quality, especially across different FPs. Databases for indicator production must be based on sound conceptual and methodological definitions, including definitions of variables, taxonomies and exclusion/inclusion criteria.<sup>64</sup>

Improved participation data are particularly desirable. Existing databases like E-CORDA build mostly on legal entities (even if containing also information on participating research units). This is a viable option for managing contracts and legal requirements. But to be meaningful for evaluation purposes, a rapid and reliable identification of participants *at laboratory level* is important for many meaningful indicators (See the detailed discussion of this issue in Volume 2, Part B). A *common participation database* for the whole FP, possibly maintained on a cross-FP base in the future, would improve this situation considerably. Therefore, an important recommendation is to launch as soon as possible an initiative to design such a database (which should, of course, be closely linked to E-CORDA). Without such a database, it will be difficult to perform in-depth evaluations of FP participation patterns with the help of quantitative data, for example through matching participations with large bibliometrics databases or with survey results.<sup>65</sup>

There are two further improvement potentials: (1) A *common system of logging of administrative events* would be very useful to analyse programme management and administration.

---

<sup>63</sup> The reporting standards for project participants have recently been updated (See the *Guidance Notes on Project Reporting* for FP7 Collaborative Projects, NoEs, coordination and support actions and research for the benefit of specific groups, Version 07/03/2008) . In the new formats, the aspects mentioned in the text have received more attention. However, further amendments may be necessary to permit the development of an indicator-based tracking of project outcomes.

<sup>64</sup> For an in-depth discussion, see Volume 2, Chapter 9.6.

<sup>65</sup> This task should not be underestimated. It requires identification of individual research units, careful tracking of organisational changes and data validation. See Volume 2, Chapter 9.6 for details.

This system should be restricted to a core of key events relevant for monitoring and evaluation. (2) A joint system to survey participants and projects can track the scientific and technological production of projects, participants' opinions and feedbacks on programme management and project impacts. Even if this information is available in principle from project reports, implementation plans, ad hoc surveys, etc., there is no efficient standardised instrument which supports statistics and searches (for example using separate fields for distinct items, list fields instead of free text, etc.). For both of these, pilot projects should be launched as soon as possible. The target should be to have functioning systems in place for the next FP, even if FP7 will presumably not benefit from them because of their development time.

Despite these comments, existing data and databases have their value. To exploit their potential, a system of dedicated databases for indicator purposes should be set up which condensates, cleans and structures available information. The key to achieving this is *simplification*: indicator databases do not have to include all kind of data required for management purposes – from phone numbers to details on project milestones and reporting. They should concentrate on those data which are critical for indicator production and exclude others. These data should be based on taxonomies and closed lists, free field data are in most cases of little use, especially for large programmes, where hand-cleaning is too resource-consuming. Such a system should also include intermediate data layers, which combine and compile available data in a way which supports the production of composite indicators.

Such databases should be designed in collaboration with indicator experts. Available data and sources must be reviewed and complementary data structures be developed. Figure 8 sketches a possible generic model in support of indicator production and evaluation.

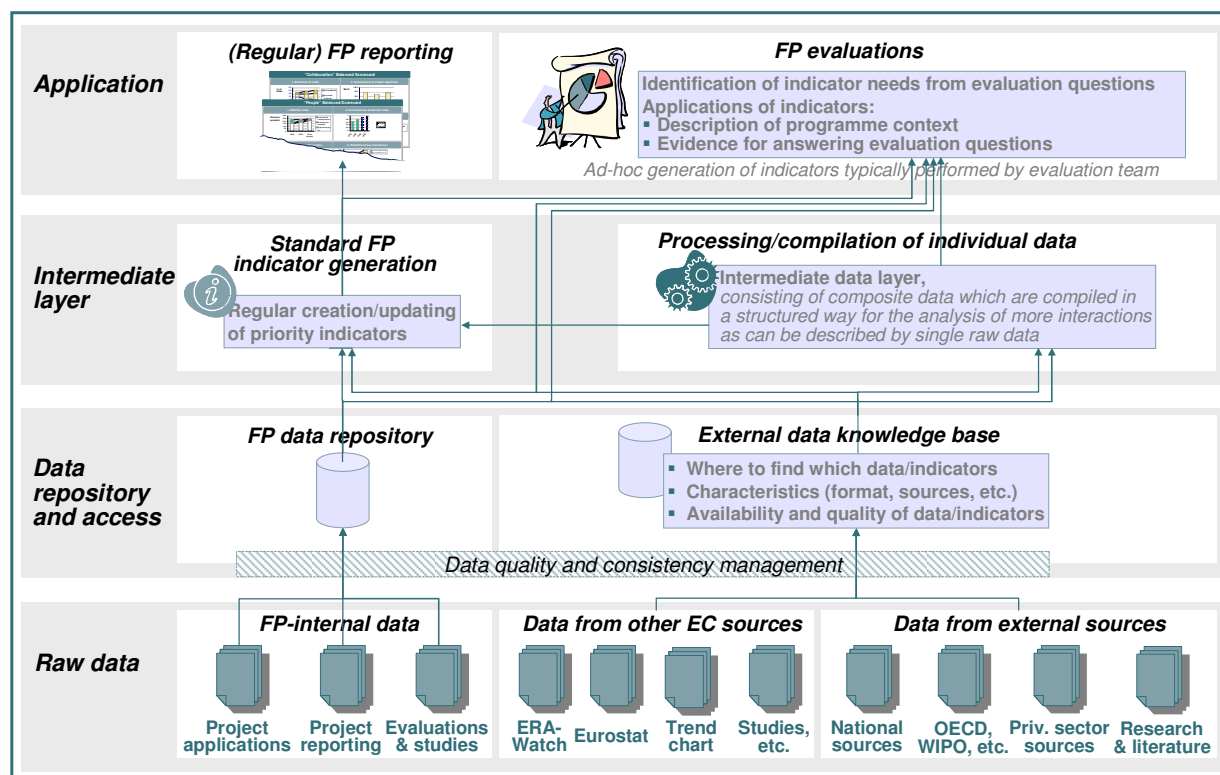


Figure 8: Possible data model to support future FP monitoring and evaluation activities

## 5.2 Process and structural prerequisites

The call for an enhanced use of indicators raises inevitably the question of the availability of the necessary competences and resources for their design, production and maintenance. This cannot be treated as a 'side activity'. Indicator design, development and production are highly specialised activities which require specific competencies, both in conceptual terms and concerning methodology and technical treatment of data. Continuity is also essential: the value of most indicators increases considerably after repeated production and use because of the value of time-series, but also because standards and application/interpretation competencies emerge. The field of indicators is no place for stop-and-go exercises.

As a consequence, appropriate resources and organisational structures are necessary for a consequent enhanced use of indicators in evaluations. To achieve this, it is useful to distinguish between different functions and their organisation:

- (1) *Basic methodological studies*, dealing with feasibility and design of specific indicators and their production can best be subcontracted to specialised institutions. But suitable follow-up by the Commission services has to be secured to integrate their results in FP evaluation practices. Such a strong supervision and coordination is also necessary to avoid overlaps in studies dealing with indicator aspects, including for example projects in the FP7 Social Sciences and Humanities Area (which contains a specific action line on indicators), other direct contracts of the European Commission (like in bibliometrics or higher education) or specific calls (like ERC support actions), which have been or will be launched in parallel.
- (2) The production of the specified *core indicators* and the maintenance of the *basic indicator databases* require continuity and sound maintenance. This should be done by specialists familiar with specific requirements of such databases. This may be done either by a dedicated unit in the Commission services or through long-term service contracts.
- (3) Indicator-oriented *support for evaluation panels* should not be outsourced, given its ad-hoc nature and the close linkages between evaluation activities, internal data structures and FP governance. This requires that specific competencies and capacities are set up in the Commission services.
- (4) Beyond these activities, the Commission services in charge of FP evaluation should maintain close linkages with experts in the indicator domain to be continuously aware of the state-of-the art, to avoid duplications and to identify critical areas for action. Such a stable network of experts might include for example closely collaboration with existing and newly created indicator-oriented networks, e.g. the European Network of Indicators Designers, a spin-off of the PRIME Excellence network.

To provide appropriate capacities for these tasks, there are different options. The question if this can best be done in existing structures or if it requires the creation of a new, dedicated unit depends on the exact balance between in-house and subcontracted activities and requires further discussion in the implementation of study results.

## Literature quoted and suggested readings

(For further readings, see the detailed list of literature in Volume 2 of this study)

Arnold, E. (2005), *What the Evaluation Record tells us about Framework Programme Performance*, published by Technopolis Group

BBSRC (Biotechnology and Biological Sciences Research Council; 2005), *Evaluation Strategy for Research Programmes and Schemes*, London, February

BMBF (2008), *Bundesbericht Forschung und Innovation 2008*, published by the German Federal Ministry of Education and Research, Berlin/Bonn

European Commission (2008), *Results of the Public Consultation on the Green Paper "The European Research Area: New Perspectives"*, Commission Staff Working Document, SEC(2008) 430, Brussels, April 2

European Commission (2009), *Evaluation of the Sixth Framework Programmes for Research and Technological Development*, Report of the Expert Group, Brussels, February 2009

European Commission (2009a), *First FP7 Monitoring Report*, published by the Research Directorate General, Brussels, February 13

European Commission (2008a), *Challenging Europe's Research: Rationales for the European Research Area (ERA)*, Report of the ERA Expert Group, Brussels

European Commission (2008b), *Second strategic review of Better Regulation in the European Union*, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, SEC(2008) 120, Brussels, January 30

European Commission (2008c), *Towards Joint Programming In Research: Working together to tackle common challenges more effectively*, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2008) 468 final, Brussels, July 15

European Commission (2007), *Inventing our future together; The European Research Area: New Perspectives*, Green Paper, COM(2007)161, Brussels, April 4

European Commission (2007a), Communication to the Commission from Ms. Grybauskaitė in agreement with the President, *Responding to strategic needs: reinforcing the use of evaluation*, SEC(2007) 213

European Commission (2007b), Report from the Commission, Annual Report on research and technological development activities of the European Union in 2006, COM(2007)519 final, Brussels, September 13

European Commission (2007b1), Brussels, 13.9.2007, Commission Staff Working Document., Accompanying the Report from the Commission, Annual Report on research and technological development activities of the European Union in 2006, COM(2007)519 final, Brussels, September 13

European Commission (2007b2), Statistical Annex to the Report from the Commission, Annual Report on research and technological development activities of the European Union in 2006, available from [http://ec.europa.eu/research/reports/2007/pdf/statistical\\_annex\\_en.pdf](http://ec.europa.eu/research/reports/2007/pdf/statistical_annex_en.pdf)

European Commission (2007c), *Green Paper: The European Research Area: New Perspectives*, COM(2007) 161 final

European Commission (2007d), *Monitoring Industrial Research: The 2007 EU industrial R&D Investment Scoreboard*, published jointly by the Directorate General Research and the Joint Research Centre, Luxembourg

European Commission (2007e), *Monitoring 2006; Implementation of Indirect Research Activities of the EC and Euratom Sixth Framework Programmes*; Brussels

European Commission (2007f), *Third Status Report on European Technology Platforms*, Report compiled by a Commission Inter-Service Group on European Technology Platforms, Brussels, March

European Commission (2007g), *Key Figures 2007 Science, Technology and Innovation; towards a European Research Area*, published by the European Commission, Directorate-General for Research, EUR 22572 EN, Brussels

European Commission (2007h), *Progress Report on the Sustainable Development Strategy 2007*, Communication from the Commission to the Council and the European Parliament, COM(2007) 642 final, Brussels, October 22

European Commission (2005), *Proposal for a Council Decision concerning the specific programme: "Ideas" implementing the 7th Framework Programme (2007-2013) of the European Community for research, technological development and demonstration activities*

European Commission (2005a), *Impact Assessment Guidelines* of June 15, 2005 with March 2006 update, SEC(2005) 791

European Commission (2005b), *Impact Assessment and ex ante Evaluation*, Annex to the Proposal for the Council and European Parliament decisions on the 7th Framework Programme (EC and Euratom), Commission Staff Working Paper, COM(2005) 119 final

European Commission (2005c), *Communication from the Commission on an EU Common Methodology for Assessing Administrative Costs imposed by Legislation*, SEC(2005)1329, Brussels, October 21

European Commission (2004), *Five-Year Assessment of the European Union Research Framework Programmes 1999-2003*, Report of the Five-Year-Assessment Panel, published by the European Commission, December 15

European Commission (2003), *Investing in research: an action plan for Europe*, Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions, COM(2003) 226 final

European Commission (2002), Communication for the Commission from the President and Mrs Schreyer, *Evaluation Standards and Good Practice*, C(2002) 5267, December 23

European Commission (2000), *Ex-ante evaluation – a practical guide for preparing proposals for expenditure programmes*, working paper, DG Budget

European Commission (2000a), *Towards a European Research Area*, Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions, COM (2000) 6, Brussels, January 18

European Court of Auditors (2007), *Evaluating the EU Research and Technological Development (RTD) framework programmes - could the Commission's approach be improved?*, Together with the Commission's replies, Special Report No 9/2007

European Parliament (2006), Decision No 1982/2006/EC of the European Parliament and of the Council of 18 December 2006, concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007-2013), published in the Official Journal of the European Union L 412/1, December 30

European Parliament (2002), *Decision No 1513/2002/EC of the European Parliament and of the Council of 27 June 2002. concerning the Sixth Framework Programme of the European Community for research, technological development and demonstration activities, contributing to the creation of the European Research Area and to innovation (2002 to 2006)*, published in the Official Journal of the European Communities, L232/1, August 29

European Parliament (1999), *Decision No 182/1999/EC of the European Parliament and of the Council of 22 December 1998 concerning the fifth framework programme of the European Community for research, technological development and demonstration activities (1998 to 2002)*, published in the Official Journal of the European Union L 026, February 01

Eurostat (2008), *Science, technology and innovation in Europe – 2008 Edition*, published by the Office for Official Publications of the European Communities, Luxembourg

- Eurostat (2007), *Measuring progress towards a more sustainable Europe*, 2007 monitoring report of the EU sustainable development strategy, published by the Office for Official Publications of the European Communities, Luxembourg
- Georghiou, L., and Polt, W. (2004), *EU Evaluation Practice and Toolkit; Applicability for Japan*, InTeReg Research Report Projekt Nr. RTW.2003.AF.022-01, published by Joanneum Research, Vienna, July
- Kellog (2004), *W.K. Kellogg Foundation Logic Model Development Guide*, Battle Creek, Michigan, January
- Leoprechting, R. von (2001), *Zur Reform des Planungs- und Budgetierungsprozesses der Europäischen Kommission: eine Fallstudie*, Beitrag zur Jahrestagung der Deutschen Gesellschaft für Evaluation in Speyer, October 4. and 5.
- McLaughlin, J.A, Jordan (1999), G.B., *Logic models: a tool for telling your programs performance story*, Evaluation and Program Planning, Volume 22, Issue 1, Spring 1999, Pages 65-72
- Minto, B. 1987, *The Pyramid Principle* (Third Edition), Minto International, London
- Nazar, B., *The Logic Model – past, present and future*, paper published by The Center for Applied Management Practices, Inc., Camp Hill, PA, available from <http://www.appliedmgt.com/Papers%20and%20stuff/The%20Logic%20Model%20Past%20Present%20and%20Future.pdf>
- Watson, J. Wolff, J, Kuehnemund, M, Ward, B., Burke, S. and Kitchener, M. (2007), *Evaluation of the Commission's Impact Assessment System*, Final Report submitted by The Evaluation partnership Limited, Richmond (UK), April
- Williams, K., de Laat, B., and Stern, E. (2002), *The Use of Evaluation in the Commission Services*, Final Report, published by Technopolis France, Paris, October
- Wholey, J.S., Hatry, H., Newcomer, K. (Editors; 2004), *Handbook of Practical Program Evaluation*, 2nd Edition, Jossey-Bass, May
- Wholey, J.S. (1987), *Evaluability assessment: developing program theory*, in Bickman, I. (Ed.), *Using program theory in evaluation*. New Directions for Program Evaluation, No. 33, Jossey-Bass, San Francisco
- Wholey, J.S. (1983), *Evaluation and effective public management*, Little Brown, Boston
- Wholey, J.S. (1979), *Evaluation: promise and performance*, Urban Institute, Washington, D.C.

