

PLATTFORM ■ TECHNOLOGIE ■ EVALUIERUNG

Nr. 12 April 2001

■ THEMA

**Bibliometric Analysis as an Instrument for
Research Evaluation**

Anthony F. J. van Raan

**Bibliometric Visualisation of Research Networks -
Examples of the 4th Framework Program of the EC**

Edgar Schiebel et al.

■ BEITRAG

**Bewertungen von Programmen der EU-
Regionalpolitik - Zwischen Anspruch und Wirklichkeit**

Markus Gruber

■ ANNOUNCEMENT

Annual Conference DeGEval

Eine Initiative von
BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND
TECHNOLOGIE
JOANNEUM RESEARCH FORSCHUNGSGES.M.B.H.
ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
ÖSTERREICHISCHES FORSCHUNGSZENTRUM SEIBERSDORF
TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H.
TECHNOPOLIS AUSTRIA
FFF FORSCHUNGSFÖRDERUNGSFOND FÜR DIE GEWERBLICHE
WIRTSCHAFT
FWF FONDS ZUR FÖRDERUNG DER WISSENSCHAFTLICHEN
FORSCHUNG

Vorwort

In der letzten Veranstaltung PLATTFORM TECHNOLOGIE EVALUIERUNG am 6. März 2001 beschäftigten wir uns mit dem Themenschwerpunkt Bibliometrie in der Forschungs- und Technologieevaluierung. Wir konnten dazu als Referenten Anthony van Raan vom Zentrum für Wissenschafts- und Technikforschung der Universität Leiden (NL) sowie Edgar Schiebel vom Forschungszentrum Seibersdorf gewinnen.

Professor Anthony van Raan ist zweifellos einer der weltweit führenden Forscher auf dem Gebiet der Bibliometrie. Seit mehr als 15 Jahren beschäftigt er sich intensiv mit Fragen der Evaluierung mit Hilfe fortgeschrittener bibliometrischer Methoden. In seinem Beitrag beschreibt Anthony van Raan die Einsatzmöglichkeiten der Bibliometrie zur Unterstützung forschungs- und technologiepolitischer Entscheidungen.

In den letzten Jahren ist in der Bibliometrie ein starker Trend Richtung Visualisierung von wissenschaftlichen Leistungen zu beobachten. Edgar Schiebel stellt dazu ein am Forschungszentrum Seibersdorf entwickeltes Tool vor, das eine graphische Darstellung von Forschungsinhalten und Kooperationsmustern ermöglicht. Die Möglichkeiten dieser Methode werden anhand des Mappings von Forschungsk Kooperationen im 4. EU-Forschungsrahmenprogramm illustriert.

In einem Gastbeitrag beschreibt Markus Gruber vom Institut für Technologie- und Regionalpolitik (INTEREG) des Joanneum Research Erfahrungen und Perspektiven mit der Evaluierung von EU-Kohäsions- und Regionalprogrammen.

Viel Spaß beim Lesen!

Anton Geyer (für das Team der Plattform)

Preface

In our recent seminar PLATTFORM TECHNOLOGIE EVALUIERUNG on March 6, 2001 we intensively discussed the issue of bibliometrics as a tool for evaluating science and technology policies. The Plattform team was happy to welcome Anthony van Raan (Centre for Science and Technology Studies, Leiden University) and Edgar Schiebel (Austrian Research Centers Seibersdorf) as speakers.

Professor Anthony van Raan is undoubtedly among the leading researchers in the field of Bibliometrics world-wide. Since more than 15 years he has been working on advanced bibliometrics methods and their application for research evaluation. In his paper, Anthony van Raan illustrates the potential of bibliometrics-based methods to support decision making in science and technology policy.

In recent years bibliometrics has been increasingly concerned with the issue of visualisation of scientific excellence and research outcomes. Edgar Schiebel presents a tool for the graphic illustration of research contents and research networks. He highlights useful application areas for this mapping tool (which was developed by Edgar Schiebel and his team at the Austrian Research Centers Seibersdorf) presenting recent pieces of his work on research networks in the 4th RTD EU Framework Programme.

In an additional article to this issue, Markus Gruber of the Institute of Technology and Regional Policy (Joanneum Research) summarises the lessons learnt with respect to the evaluation of EU cohesion and regional policy programmes.

Enjoy reading this Newsletter!

Anton Geyer on behalf of the Plattform team

Bibliometric Analysis as an Instrument for Research Evaluation¹

Anthony F. J. van Raan

Abstract

Science is the driving force of our modern society. It is crucial to identify and to promote scientific quality by evaluation of research. Peer review is the central part of research evaluation but it may be influenced by subjective elements. We present a method to support peer review with an objective instrument. This method provides an indispensable tool for decision-making in science policy, particularly in priority setting at the level of research programs. A specific part of our method, science mapping, even provides the possibility of foresight.

Qualitative and Quantitative Evaluation

Peer review undoubtedly has to remain the principal procedure of quality judgement in research evaluation. But peer review has serious shortcomings and disadvantages. The major problem is subjectivity, i.e., dependence of the outcomes on the choice of individual committee members. This dependence may result in conflicts of interests, unawareness of the quality or a negative bias against younger people, newcomers to the field, particularly in the case of emerging, new interdisciplinary fields. Bibliometric assessment is not infected by such crucial disadvantages.

We do not plead for a replacement of peer review by bibliometric analysis. Subjective aspects are not

merely negative elements. In any judgement there must be room for the intuitive insights of experts. It is claimed however that for a substantial improvement of decision-making our bibliometric method has to be used in parallel with peer-based evaluation procedures.

The most crucial parameter in the assessment of research performance is *international scientific influence*. This paper addresses standardized procedures based on advanced bibliometric methods to detect scientific quality. Undoubtedly, the bibliometric approach is not an ideal instrument, working perfectly in all fields under all circumstances. But our approach works well in the large majority of the natural, medical, applied and behavioral sciences. These fields of science are the most cost-intensive and the ones with the strongest socio-economic impact.

Bibliometric assessment of research performance is based on one central assumption: scientists who have to say something important do publish vigorously in the open, international journal ('serial') literature. Our central statement, based on a long-standing experience, is that research evaluation should be performed systematically on the 'meso'-level of larger institutions, such as universities or major parts of universities, such as large faculties or institutes. After an overall assessment of these larger institutions, the evaluation can be narrowed down to research groups and programs within these institutions.

Outline of the Method

Publications are not the only, but certainly very important elements in the process of scientific communication. Work of high quality provokes reactions of colleague-scientists. They are the

¹ Revised version of an earlier paper of the author (van Raan 1999).

international forum, the 'invisible college', by which research results are discussed. In most cases, these colleague-scientists play their role as a member of the invisible college by referring in their own work to earlier work of other scientists. This process of citation is a complex one, and it does not provide the 'ideal' monitor on scientific performance. This is particularly the case on a statistically low aggregation level, e.g., just one publication. But the application of citation-analysis to the work, the 'oeuvre' of a group as a whole over a longer period of time, does yield in many situations a strong indicator of scientific performance.

An important condition is that citation-analysis is part of an advanced, technically highly developed bibliometric method. For a detailed discussion of this methodology we refer to recent overview papers of our group (Van Raan 1996, and see also our website²). We now focus on the *internationally standardized impact indicator CPP/FCSm*³, our 'crown' indicator that relates the measured impact of a research group or institute to a worldwide, field-specific reference value. This indicator shows whether the performance of a research group or institute is far below (indicator value < 0.5), below (indicator value 0.5 - 0.8), around (0.8 - 1.2), above (1.2 - 2.0), or far above (>2.0) the international (western world) dominated impact standard of the field. In the latter case, a measured impact far above the international standard, provides a very strong indication of high quality. In other words, by searching for research

groups with an impact above a specific threshold value, scientific excellence can be detected.

Example of Results

Table 1 shows a performance trend analysis of the medical faculty and university hospital of an established and renown university in a European Union member state. As a good indication of size the number of publications (in international, refereed journals, as defined in Van Raan 1996) is given. This number is about 1,000 per year.

Table 1: Institutional overall survey: Trend analysis of size in terms in publication output (P) and of the 'crown indicator' CPP/FCSm

Time period	P	CPP/FCSm
1988-1991	3,637	1.28
1989-1992	3,891	1.25
1990-1993	3,988	1.22
1991-1994	4,209	1.16
1992-1995	4,433	1.19
1993-1996	4,559	1.26
1994-1997	4,665	1.34

The higher the aggregation level, the larger the volume in publications and the more difficult it is to have an impact significantly above the international level. At the 'meso-level', *CPP/FCSm* above 1.2, such as in this case, means that the institution can be considered as a scientifically strong organization, with a high probability to find very good to excellent groups. Thus, the next step is the breakdown of the institution into smaller units, i.e., research groups.

The medical faculty provided us with the names of all senior researchers as far as they were still employed in 1998, and the allocation of these researchers over all research groups and programs (about 100). The bibliometric algorithms can now be repeated efficiently at the research group level. In most cases

2 www.cwts.leidenuniv.nl, go to 'Publications', e.g., "Assessment of the Scientific Basis of Interdisciplinary, Applied Research: Application of Bibliometric Methods in Nutrition and Food Research".

3 Number of 'external' citations received by the group or institute per publication over the given period of time (CPP), normalized to the same parameter worldwide (FCSm). See Van Raan 1996.

the publication volume at this level is between 10 and 20 per year. Groups with $CPP/FCSm > 3.0$ can compete easily *with top-groups at top US universities*. Thus, if the threshold value for the $CPP/FCSm$ indicator is set at 3.0, we filter out the excellent groups with high probability, see Table 2.

Table 2: Search for scientific excellence

Research group	P	CPP/FCSm > 3.0
Blood research: thrombosis, genetic coagulation factors	111	4.32
Immuno-genetics: cell membranes, HLA	30	3.12
Medical imaging	56	4.27
Clinical Epidemiology	169	4.23
Molecular and Genetic, Tumor-pathology	144	3.49

Five of the about 100 research groups and programs that can be regarded as excellent are identified. We also found ten groups with $CPP/FCSm$ between 2.0 and 3.0, which classifies them as very good. In total, there are 29 groups with a $CPP/FCSm$ above 1.5.

A further illustration of our method is the evaluation of the physics research groups of all universities in a country (van Leeuwen et al 1996). This evaluation is carried out a few years ago, covering the ten-year period 1985-1994 which is presented here. An update for 1995-2000 is in preparation.

Both the volume of publications and the $CPP/FCSm$ values for all 200 groups are displayed in Fig. 1. We observe around 10 excellent research groups, i.e., with $CPP/FCSm > 3.0$. The figure illustrates an earlier discussed general finding: the larger the unit, the more the average impact of the unit tends to lower (more 'average') values. Therefore it is important to split up larger groups into new, smaller groups to allow a more precise impact assessment. Otherwise,

excellent work will be 'hidden' within the bulk of a large group.

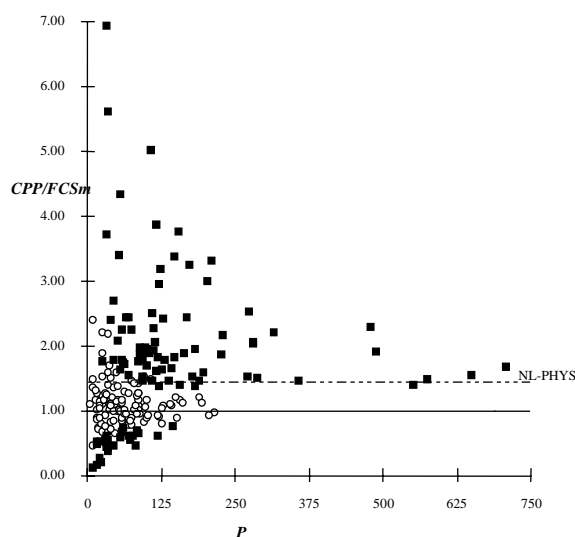


Figure 1: Evaluation of all physics research groups in a country

Abscissa: number of papers; ordinate: CPP/FCSm value. Open circles: groups with impact not significantly different from the international standard (indicated by the horizontal solid line $CPP/FCSm=1$). Black squares: groups above (below) the international standard. The dotted line represents the average ($CPP/FCSm = 1.45$) of the total of all 200 physics groups.

Bibliometric Mapping: Landscapes of Science

Finally, this paper will shortly address a second major line of advanced bibliometric methodology: bibliometric mapping. The basic idea is the following. Each year about a million scientific articles are published. For just one research field, such as Mathematics and Computer Science (MCS), the amount of papers is already about fifty thousand per year. How to keep track of all these developments? Are there cognitive structures and patterns 'hidden' in this mass of published knowledge, at a 'meta-level'?

Suppose each research field can be characterized by a list of most important, say 200, 'concepts', i.e., keywords or, in most cases, keyword-combinations (Noyons 1999). For MCS research such a list will cover words like *differential equation*, *optimization*, *chaos*, *fuzzy set*, *parallel computer*, *Monte Carlo simulation*, and so on. Each publication can be characterized by a subset from the total list of keywords. It is, as it were, a DNA fingerprint of a publication. For all publications, keyword-lists are compared pair-wise. In other words, these many thousand publications constitute a gigantic network in which all publications are linked together by one or more common keywords. The more keywords two publications have in common, the more these publications are related (keyword-similarity) and thus belong to the same research area or research speciality. In the biological metaphor: the more DNA two objects have in common, the more they are related. Above a certain similarity threshold, they will belong to a specific species.

We developed special mathematical techniques to unravel these publication networks, by word-similarity measurements, clustering of related publications, and mapping the ensemble of these clusters in a two-dimensional space. This procedure generates an underlying structure. The fascinating point is that these structures can be regarded as the cognitive, or intellectual structure of the scientific field. Clusters can be identified as sub-fields and research themes. As discussed above, the procedure is entirely based on the total of relations between all publications. Thus, the structures that are discovered are not the result of any pre-arranged classification system or whatsoever. Nobody prescribes these structures. The structures emerge solely from the internal relations of the whole universe of publications together. In other words, what is made visible by our mapping method, is the self-organized structure of science.

Our mapping procedure depends partly on expert input. Special internet-facilities enable experts to comment on the keywords (mostly word combinations, 'concepts') used to generate the structure of the field. In Figure 2 we present the structure of MCS by a 2D map of the structural relations between sub-fields. The electronic internet version (see our website) enables the user to move the mouse over sub-field circles. The name of the sub-field appears in the box at the top of the figure.

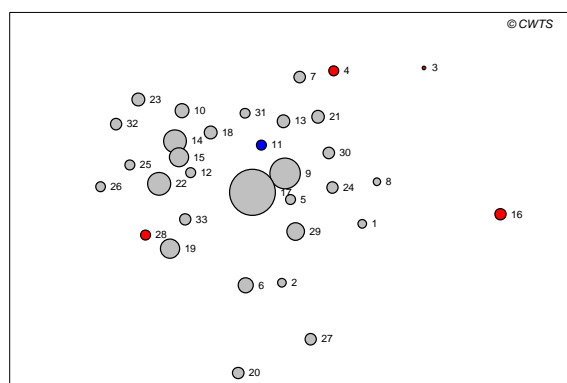


Figure 2: Map of Mathematics & Computer Science (1997-1998)

The size of the sub-fields represents the number of publications. For the list of sub-fields with corresponding number we refer to the appendix. The shades indicate a significant change of publication activity. Dark grey (4, 16, 28): increase of activity; black (11): decrease of activity. Badness-of-fit 0.22; distance correlation 0.88.

Concluding Remarks

In this paper we presented an advanced method for the clear, objective evaluation of research activities. These bibliometric performance indicators allow a crucial improvement of the peer review based evaluation. We discussed bibliometric mapping as a powerful tool for visualizing the cognitive landscape of a research field and its surrounding environment. Maps made over a series of years reveal trends and changes in structure, and extrapolation of a map

series can act as a 'foresight' system for near-future developments.

Furthermore, the position of actors can be put on the map. Thus a strategic map is created: who is where in science? With help of the performance indicators we are able to identify the strongest players. The combination of both bibliometric methods, performance assessment and mapping, appears to be a very powerful tool in the evaluation of research activities. Changes in maps over time (field structure, position of actors) may indicate the impact of research programs, particularly with respect to sub-fields characterized by research around social and economic problems. Our mapping methodology is also applicable in the study of the socio-economic impact of R&D.

Appendix: MCS Sub-fields (figure 2)

-
- 1 ATM/ performance evaluation
 - 2 Linear model/ bayesian approach/ monte carlo method
 - 3 case based reasoning
 - 4 Fuzzy set/ membership function
 - 5 large number/ strong law
 - 6 simulation study/ maximum likelihood/ asymptotic distribution/ distribution function
 - 7 Robotics
 - 8 scheduling problem/ single machine/ travelling salesman problem/ job shop
 - 9 artificial neural network/ computer simulation/ computational complexity/ Expert system
 - 10 Stability/travelling wave
 - 11 parallel computer/ parallel algorithm
 - 12 necessary condition/ optimal control
 - 13 Optimization
 - 14 Boundary Condition/ numerical simulation/ numerical experiment/ partial differential equation
 - 15 Dynamical system/ Chaos/ time series/ initial condition
 - 16 Internet/ WWW/ web site
 - 17 n 1/ Finite Group/ monte carlo simulation
 - 18 discrete time/ continuous time/ frequency domain/ non-linear systems
 - 19 necessary and sufficient condition/ Banach Space/ Hilbert space/ l 2
 - 20 Vertex/ regular graph/ chordal graph/ distance regular graph
 - 21 genetic algorithm/ objective function/ simulated annealing/ tabu search
 - 22 asymptotic behavior/ Boundary Value Problem/ approximate solution/ Exact solution
-

-
- 23 Finite Element
 - 24 real time/ petri net/ formal method
 - 25 Differential Equation/ 2nd order/ runge kutta method/ first order
 - 26 initial data/ cauchy problem/ global existence/ initial boundary value problem
 - 27 polynomial time/ linear time/ approximation algorithm/ bipartite graph
 - 28 r n/ positive solution/ bounded domain/ semilinear elliptic equations
 - 29 Complexity/ Lower Bound/ Upper Bound/ branch and bound algorithm
 - 30 Classification/ feature extraction/ discriminant analysis/ texture classification
 - 31 Robustness/ disturbance rejection
 - 32 Numerical method/ finite difference method
 - 33 sufficient condition/ asymptotic stability/ lyapunov function/ global stability
-

References

- Leeuwen, Th.N. van, E.J. Rinia and A.F.J. van Raan (1996). Bibliometric Profiles of Academic Physics Research in the Netherlands. In: Quality Assessment of Research: An Analysis of Physics in the Dutch (Netherlands) Universities in the Nineties. Utrecht: VSNU, 1996 (ISBN: 90-5588-041-8).*
- Noyons, E.C.M. (1999), Bibliometric mapping as a science policy and research management tool. Thesis Leiden University. Leiden: DSWO Press.*
- Raan, A.F.J. van (1996). Advanced Bibliometric Methods as Quantitative Core of Peer Review Based Evaluation and Foresight Exercises. Scientometrics 36, 397-420.*
- Raan, A.F.J. van (1999). Evaluating the scientific excellence of research programmes: a pivot of decision-making. The IPTS Report, No. 40, December 1999, 30-37 (Published by the Joint Research Centre of the European Commission).*

Author

Professor Dr. Anthony F.J. van Raan
 Centre for Science and Technology Studies (CWTS)
 Universiteit Leiden, Wassenaarseweg 52
 NL - 2300 RB Leiden, Netherlands
 P.O. Box: PB 9555
 Phone: +31 71 5273909 (Secretary) or 31 71 5273942
 Fax: +31 71 5273911
 E-mail: vanraan@cwts.leidenuniv.nl

<http://www.cwts.leidenuniv.nl>

Bibliometric Visualisation of Research Networks – examples of the 4th Framework Program of the European Community¹

Edgar Schiebel, Clemens Widhalm.,
Michaela Topolnik, Alexander Kopcsa,
Matthias Weber

Abstract

In this paper we focus on the analysis of networks of collaboration that were set up within the European Union's Fourth Framework Programme (FP4) by applying co-occurrence analysis and visualisation in 2-dimensional knowledge maps. The ability to generate research networks was one of the key objectives of the Framework Programme in order to foster the long-term competitiveness of the European economy. Therefore the focus of the analysis is on the patterns of co-operations and collaboration in research networks within FP4, involving the industrial, research and educational sectors of all EU-member states. It will look both at the aggregate situation of FP4 as a whole, but also at the situation in one of the specific programmes, namely the Transport Research Programme. The objective is to identify specific features within the patterns of co-operation which are of particular relevance to European research, such as the respective roles of universities, industry and non-university research institutions and the role of centres of excellence. The article highlights and assesses the potential of the presented methods and tools for evaluation purposes in connection with the creation of the European Research Area.

Introduction

Looking at EU activities in the field of technology policy and research policy, the *Framework Programme for Research, Technological Development and Demonstration* (FP) are the central instrument to foster and to strengthen RTD at a European level. Since the early Eighties when the first Framework Programme was designed, they have become more important, as expressed in the financial resources that were attributed to the FPs. The total budget of the FP4 is more than three times that of the first in nominal value, with a doubling of resources after the FP3. The currently running FP5 has further grown by about 10%. It was also with the FP4 that for the first time all 15 current EU-Member States fully participated, and it also covered a great variety of different scientific and technological areas by thematic and horizontal programmes.

Our analysis of the participation in the EU Framework Programmes concentrates on identifying specific features within the patterns of co-operation and collaborative networking within FP4, which are of particular relevance to the ongoing European Research Area debate. It will focus on the respective roles of industrial, university and non-university research, and on the participation of different Member States.

The analytical tool that is used to produce the collaboration maps in this paper goes far beyond conventional statistics. On a descriptive level the visualised networks allow to identify different types of collaborative R&D patterns between countries, sectors, organisations etc. and to use these observations as a basis for further investigations.

1 Revised version of a paper by Widhalm, C., Topolnik, M., Kopcsa, A., Schiebel, E. and Weber, M. (2001).

Data basis

For this analysis we used data of the CORDIS database of the European Commission. Due to the high amount of projects we restricted the dataset to projects financed by the EU within FP4 (1995-1998). At that time (December 1998) more than 10.000 projects with over 47.000 participations across all thematic and horizontal programmes of FP4 were listed in the CORDIS database. For the network analysis we selected those organisations, that participated in three or more projects (about three quarters of the partners participated in only one project and were not considered to be a part of a network). This selection led to 2700 partners, who participated 30.000 times in 8.800 projects. The network showed about 53.000 collaborative links. Partners from the 15 EU states across all sectors (industry, universities, research institutes) were included.

Bibliometrics

For the analysis of networks we suggest the co-word analysis. The co-word analysis is a relational approach and well known in the research field of bibliometrics. Relations between technological developments, different fields of application and leading experts can be determined using bibliometrics (Grupp et al 1990). Due to the manifold possibilities of analysing large amounts of information and documents the importance of bibliometric analysis increased also for other applications, like competitive analysis of companies and their products, structuring of internal documents for knowledge management, patent analysis and management or analysis of co-operation behaviour of institutions or persons (Noll and Schiebel 2000).

Bibliometric methods go for structuring of electronically stored information in internal or external databases.

The structuring is based on the calculation and often visualisation of relations between objects, such as documents, keywords, authors or institutions. The relations are derived from indicators that can be defined through different models (van Raan 1992).

Evaluating patterns of collaboration in the Fourth Framework Programme

In this section, we will look at aggregate patterns of co-operation that can be discerned for the FP as a whole.

Industries, universities and non-academic research organisations

In Figures 1 educational, industrial and non-academic institutions are highlighted. A clear division can be noticed between industrial institutions on the one hand, and universities/higher education on the other. This indicates a preference of institutions of these two types to co-operate with partners of the same organisational type and to form more steady networks.²

When looking at the non-academic research organisations (ror), it can easily be observed, that quite often they do co-operate with partners from other organisational types because they are evenly spread across the figure and to not form a sector by their own. To all appearances non-academic organisations fulfil an intermediary function between educational and industrial institutions. This is a quite trivial result but demonstrates the kind of visualisation of this relationship.

² That does not mean that they do not co-operate with other partners at all but just that their collaborative links are more intense with partners of their own kind. If we would look just at the sheer number of participations we would not be able to see this specific pattern, on the contrary we would presume that the co-operation is quite balanced.

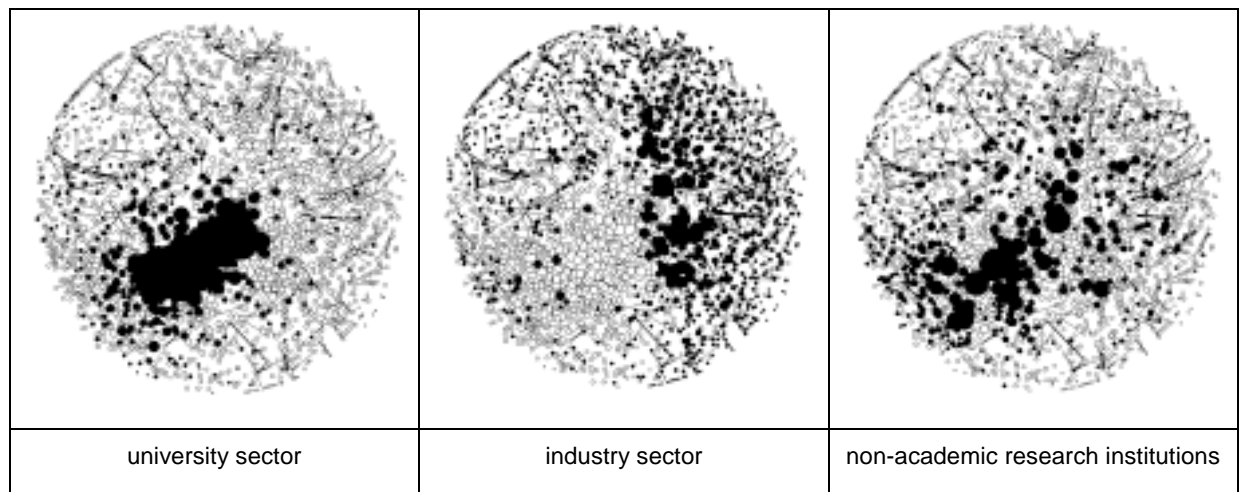


Figure 1: Co-operation network in the 4th European FP. University, industry and non-academic research institutions are marked in black respectively.

Evaluating patterns of collaboration in the Transport Research Programme

More detailed observations can be made when moving deeper to the level of individual programmes. In this section, we will analyse some of the patterns that result from applying the bibliometric tool to the Transport Research Programme.¹

The main thematic clusters

Figure 2 shows the result of the bibliometric analysis of the Transport Research Programme. The graphical representation shows eight well-delimited clusters of dense collaboration. Their titles have been chosen on the basis of a content analysis of the underlying project descriptions that make up the database. It is possible to identify the eight clusters with key thematic areas or priority themes of the Transport Research Programme: Strategic modelling of European transport, Road and urban transport, Rail transport, Intermodal transport, Inland navigation, Air

traffic, Maritime shipping, Sustainable mobility for Europe.

This is a quite interesting result because it allows to show that the Transport Research Programme has indeed led to a grouping of organisations around its priority themes. Whether these grouping also develop into more stable networks of collaboration is an issue that should be analysed by looking at the further evolution under FP5.

Another interesting feature is the respective density of collaboration within the individual clusters. Especially in maritime waterborne transport and to some extent also in integrated transport the collaborating institutions are only loosely clustered, pointing to rather occasional collaborations in these areas. For integrated transport this can be explained by the fact that we are often dealing with partners coming from the individual modes, brought together by an intermodal “integrator” or service provider. In maritime waterborne transport, the small number of incidences may be the reason for the scattered character of this cluster.

¹ Some transport-related research projects of the 4th FP were not conducted within the Transport Programme, but within other programmes of FP4: ICT, Energy, Environment and Industrial Materials and Technologies.

Zooming into the centres of excellence in European transport research

A final interesting feature of the transport research programme consists of the outstanding role of a rather limited number of research institutions that are located right in the centre of the overview picture of the Transport Research Programme (Figure 2). The extent of their involvement and the cross-cutting role they play in many fields of transport research indicates that they are very well positioned to develop into "centres of excellence" or at least into key network nodes in the envisaged European Research Area in the transport field. In other words, their central position does not mean that they are just core research actors in the horizontal area of strategic modelling, but that they are involved in many research fields in transport. This can be shown by searching deeper in the underlying database and to perform a content analysis of their research portfolios. In other words, their central position on the figure is thus the consequence of the clustering algorithm. The six key organisations identified are: Technical Research Centre of Finland (VTT), Institut National de Recherche sur les Transports et leur Sécurité (INRETS), Netherlands Organizations for Applied Scientific Research (TNO), University of Leeds, Transport Research Laboratory (TRL), and National Technical University of Athens.

Conclusions

The 4th FP involved more than 10.000 RTD-projects where several partners took part in each of them. To gain insight into the typical co-operational behaviour

within 4th FP and therefore into a part of the European research area just hierarchical and one-dimensional structuring of the respective data is not sufficient. The complexity of this research network is multidimensional and therefore multidimensional indicators like co-occurrences are to be used for an adequate characterisation. Unfortunately, the interpretation of multidimensional indicators is extremely difficult, and this is where the main advantage of the methodology and tool used in this article lies. As a reasonable compromise a graphical mapping methodology based on a spring model is presented where the multidimensional problem is reduced to a 2-dimensional knowledge-map with minimum aberrations to the original system.

The analysis as well as the graphical display of networks of partners provides an information base to analyse the structure of and to characterise Europe-wide collaboration. Visualisation of networks indicates specific co-operation patterns beyond quantitative one-dimensional analysis and thus is a valuable method for the characterisation of co-operational networks.

It could be shown that the tool is able to visualise features of European RTD which are becoming increasingly important. Both the patterns of networking and collaboration, but also the role and impact of individual research institutions are important pieces of information for the ongoing debate on the European Research Area. The usefulness of this method was illustrated by example of patterns of co-operation behaviour with respect to the organisational types of partners involved, the countries of origin and the thematic cluster they form.

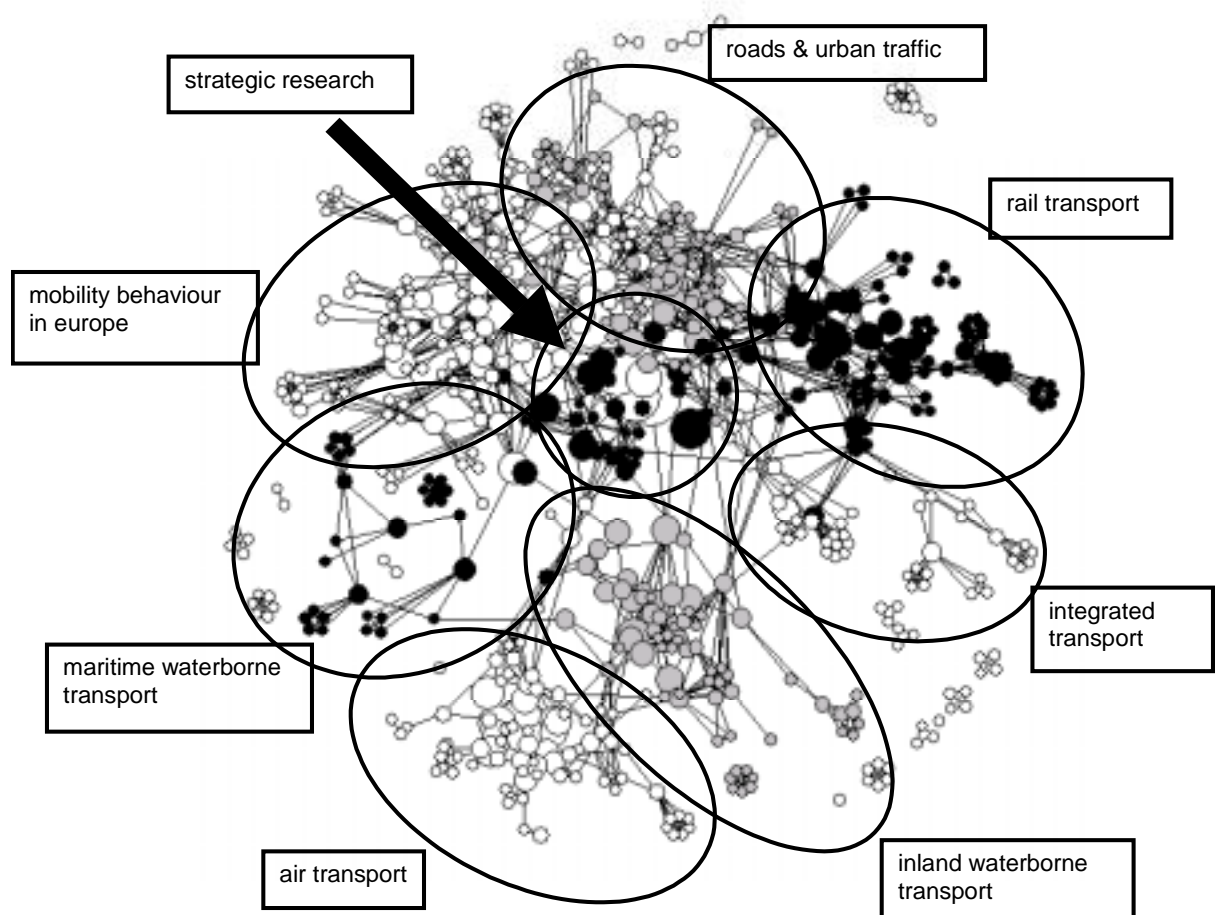


Figure 2: The main clusters of collaboration within the Transport Research Programme and their thematic orientation

Research in the future will concentrate on improving the possibilities to follow transformation patterns of the collaboration networks. In the current version of the tool, these are difficult to realise due to the clustering algorithm that is applied. This is going to be the next research challenge in methodological terms, which would make the tool even more versatile for evaluation purposes.

The evaluation results of the FP4 and the Transport Research Programmes give evidence of the extent to which the objectives of EU FPs could be met. The stimulation of co-operation between partners from different countries and organisational types is one of

the objectives pursued by the EU (co-operation principle), the bridging between technological fields and scientific disciplines is another one (principle of horizontality) (Grande 1996). The analysis has shown that educational and industrial partners tend to co-operate with partners of their own kind rather than taking part in cross-organisational networks, especially when looking at the overall Framework Programme. However, selected research segments can be identified, where a core network of industrial partners is well connected to collaborators from the educational sector. Overall, the intended networking effect of FP4 could not be fully met, but the even

distribution of non-academic research organisations across all fields of research indicates that they are well positioned to fulfil a bridging function in the future.

In the Transport Research Programme the different organisational types are better distributed, though with differences between individual segments of transport research, but it would be necessary to carry out a deeper analysis of individual networks to confirm this observation. In future programme and proposal evaluations more attention should be paid to that specific function of non-academic research organisations.

The analysis of collaboration patterns with our methodology at the level of the Transport Research Programme lead to some very interesting findings. It was possible to identify the main thematic clusters and show that the clusters of dense collaboration are identical to the main thematic areas of the programme. Apparently, the programme has indeed led to the creation of networks around these themes, even if their sustainability should be checked once data on the transport actions of the Fifth Framework Programme becomes available. This observation points to two further research questions to be addressed. First, it would be interesting to look for similar thematic clusters in other research programmes within FP4. Second, also at the aggregate level of FP4 the same question could be raised, especially with respect to research of a cross-cutting nature.

Finally, within the Transport Research Programme "centres of excellence" could be identified. These institutions are active in many segments of transport research and are therefore very well positioned to become the future key nodes of the European research area in the transport field in the future.

References

- Callon, M., J. P. Courtial, W. A. Turner and S. Bauin (1983). *From Translations to Problematic Networks: an Introduction to Co-Word Analysis*. *Social Sciences Information*, 22, 191-235.
- Capron, H. et al. (no date). *National Innovation Systems. Pilot Study of the Belgian Innovation System. Study carried out for the Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC) in the context of the OECD Working group of Innovation and Technology Policy*.
- Dumont, M. and Meussen, W. (1997). *Some results on the graph-theoretical identification of micro-clusters in the Belgian Innovation System, presented at the OECD/DSTI-workshop on Cluster-analysis and cluster-based policies, Amsterdam, October 1997*.
- European Commission (1997). *Second Report on Science and technology Indicators 1997, EUR 17639, European Commission, Luxembourg, DG XII*.
- European Commission (1999). *Research and Technological Development Activities of the European Union – 1999, European Commission COM (1999) 284, Brussels*.
- European Commission (2000). *Towards a European Research Area, European Commission, COM (2000) 6, Brussels*.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994). *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies, Sage: London*.
- Grande, Edgar (1996). *Forschungspolitik in der Politikverflechtungsfalle? Institutionelle Strukturen, Konfliktdimensionen und Verhandlungslogiken europäischer Forschungs- und Technologiepolitik; in: Politische Vierteljahresschrift 36, 460-483*.
- Grupp, H., T. Reiss, U. Schmoch (1990). *Knowledge interface of technology and science – developing tools for strategic R&D-management. FhG – ISI report to the Volkswagen Foundation, Karlsruhe*.
- Gusmao, Regina (2000). *Developing and using indicators of multilateral S&T co-operation for policy making: the example of European research programmes. Scientometrics, Vol 47, no. 3 (2000), pp. 493-514*.
- Kopcsa, A. and E. Schiebel (1998). *Science and Technology Mapping: A New Iteration Model for Representing Multidimensional Relationships; Journal of the American Society for Information Science JASIS, Jan, 1998, No. 1/ Vol49, p.7ff*.
- Kostoff R. (1993). *Co-Word Analysis; Kluwer Academic Publishers, 63-78*.

- Leyersdorf, K. (1989). *Words and Co-Words as Indicators of Intellectual Organisations*; *Research Policy*, 18, 209-223.
- Leydesdorff, L. (2000). *The Triple Helix: An evolutionary model of innovations*; *Research Policy*, 29, No. 2, 243-255.
- Lukkonen, Terttu et al. (1999). *Finnish Participation in the Fourth Framework Programme*, Tekes, Helsinki.
- Noll, M. and E. Schiebel (2000). *Bibliometric Analysis for Knowledge Monitoring*, OEFZS-Report S-0083.
- Peterson, J. / Sharp, M. (1998). *Technology Policy in the European Union*, Macmillan
- Report of the Independent Expert Panel (2000). Five-Year Assessment of the European Union Research and Technological Development Programmes, 1995-1999*, European Commission.
- Van Raan, A.F.J. (1992). *Advanced bibliometric methods to assess research performance and scientific development: basic principles and recent practical applications*. *Research Evaluation*, 3, 3, 151-166.
- Van Raan, A.F.J. (2000). *R&D evaluation at the beginning of the new century*; *Research Evaluation*, 9, No. 2, 81-96.
- Widhalm, C., Topolnik, M., Kopcsa, A., Schiebel, E. and Weber, M. (2001). *Evaluating Patterns of Co-operation: Application of a bibliometric visualisation tool to the 4th Framework Programme and the Transport Research Programme*. Forthcoming in: *Research Evaluation*.
- Removille, J. and B. Clarysse (1999). *Intra-European scientific co-operation: measuring policy impact*; *Research Evaluation*, 8, No.2, 99-109.

Author

Edgar Schiebel, Clemens Widhalm,

Alexander Kopcsa:

Austrian Research Centers Seibersdorf,
Geschäftsfeld Technologiemanagement

Matthias Weber:

Austrian Research Centers Seibersdorf,
Geschäftsfeld Technologiepolitik

Michaela Topolnik:

PROVISO Project Team, Austrian Federal Ministry of
Transport, Innovation and Technology

Contact:

Dr. Edgar Schiebel

ARCS-Austrian Research Centers Seibersdorf

Department Technology Management

2444 Seibersdorf, Austria

Phone: + 43 50550 3830

Fax: +43 50550 3888

E-mail: edgar.schiebel@arcs.ac.at

<http://www.arcs.ac.at/gesch/S/ST>

Bewertungen von Programmen der EU-Regionalpolitik – Zwischen Anspruch und Wirklichkeit

Markus Gruber

Abstract

The European Union allocates a considerable amount of money – 35% of its budget – to regional policy. In consequence comprehensive evaluation activities are carried out appraising the structural funds interventions. The European Union makes great efforts to develop evaluation standards and to structure methodological discussions. In practice however, the European Commission focuses too much on quantification of employment effects of their interventions. This can be traced back to high political pressure to substantiate impacts and success of structural funds interventions. The efforts undertaken to quantify employment effects on the one hand and the complexity of regional development programmes on the other hand often leads to a „stuck in the middle“ evaluation exercises. Impact analysis – in the sense of appraising net effects – are barely carried out, and if so they are driven by very pragmatic – often inadequate – approaches. Little scope remains for innovative, qualitative evaluation approaches which are often more suitable for complex regional development programmes in particular with regard to the stimulation of a learning process within the programme partnership.

Einleitung

Die Europäische Union gibt heute 35% ihres Budgets für die Kohäsions- und Regionalpolitik aus und finanziert über die Strukturfonds europaweit eine große Zahl an Entwicklungsprogrammen mit. Alleine in Österreich wurden in der vergangenen Programmperiode 1995-1999 zwölf Regionalprogramme (Ziel 1, 2, 5b) und 23 Programme aus Gemeinschaftsinitiativen ko-finanziert. Im letzten Newsletter (Nr. 11 vom November 2000) hat Rudolf Schicker von der Österreichischen Raumordnungskonferenz über die Ergebnisse und Verwertbarkeit der in der letzten Programmperiode durchgeführten Evaluierungen in Österreich berichtet. Der gegenständliche Artikel wirft einen generellen Blick auf die im Rahmen der EU-Strukturfonds zur Anwendung kommende Evaluierungssystematik, den Ansprüchen der Kommission und die „empirisch“ beobachtbare Praxis der Evaluation. Dabei wird sich zeigen, dass eine erhebliche Diskrepanz zwischen Ansprüchen der Kommission, die ihrerseits in vielen Bereichen nicht immer adäquat sind, und der Wirklichkeit der Evaluierungs-Praxis besteht.

Strukturfonds-Reform 1988 ist dem Grunde nach heute noch gültig

Bis Ende der 80er Jahre war die EU-Regionalpolitik durch fehlende Koordination und ineffiziente Förderung von Einzelprojekten gekennzeichnet. Dementsprechend war auch die Evaluationsmethodik wenig entwickelt. Mit der Reform der Strukturfonds 1988 wurden die Prinzipien (1) der Fokussierung des Ressourceneinsatzes auf prioritäre Ziele, (2) der Vergabe der Mittel auf Basis einer mehrjährigen integrierten Programmplanung, (3) einer partnerschaftlichen Planung und Umsetzung und (4) der

Additionalität der EU-Mittel¹ eingeführt. Darüber hinaus müssen die Strukturpolitik und die darauf aufbauenden Entwicklungsprogramme mit anderen EU-Politiken (Umwelt, Wettbewerb) kompatibel sein. In mehreren Reformschritten wurde dieses System in den 90er Jahren angepasst. Dem Grunde nach wurde es aber bis heute beibehalten.

Begleitung und Bewertung im Rahmen der Strukturfonds-Programme

Gerade in Anbetracht der eingesetzten umfangreichen finanziellen Mittel ist es verständlich, dass es sowohl für die Europäische Union insgesamt als auch für die Regionen von Interesse ist, wie effizient und zielgerichtet die Mittel eingesetzt werden. Entsprechend werden umfassende Evaluierungsaktivitäten verfolgt, die den Einsatz der Mittel bewerten sowie Vorschläge zur Verbesserung des Mitteleinsatzes entwickeln.

Durch die Reform 1988 wurde daher ein umfangreiches System zur Begleitung der Programme eingeführt: Dies umfasst die Implementierung von Monitoring-Systemen, von partnerschaftlich (zwischen EU- und Mitgliedsstaaten) besetzten Begleitausschüssen sowie die Durchführung einer Ex-ante und Ex-post Evaluierung.

Während der Förderungsperiode 1989-1993 wurden diese Vorschriften erstmals umgesetzt. Allerdings waren sowohl Monitoring als auch Evaluierung jene Aspekte der Strukturfonds-Reform, denen in der tatsächlichen Umsetzung am wenigsten Beachtung geschenkt wurde.

Erst zu Beginn der 90er Jahre war ein Wendepunkt erkennbar. 1992 „situated at a turning point between

the end of the first actions lasting several years and the preparation of the ensuing Community interventions, evaluation work received fresh impetus“ (MEANS 1993 zit. nach Bachtler/Taylor 1999). Im Zuge einer weiteren Ausweitung der Strukturfondsmittel und der Diskussionen über die Wirtschafts- und Währungsunion wurde sowohl von Seiten der Kommission als auch weiterer Institutionen der Gemeinschaft (z.B. Europa-Parlament) ein Bedarf nach umfangreicheren Informationen über die Implementierung der Programme und ihrer Wirkungen erkannt.

Von Seiten der Kommission wurden daher von 1994 an Maßnahmen zur Verbesserung der Effektivität des Evaluierungsprozederes getroffen. Als wesentliche Schritte sind zu nennen:

- die Schaffung spezieller Evaluierungseinheiten innerhalb der Kommission zur Koordination der Evaluierungsanstrengungen
- die Schaffung des MEANS-Programmes „Evaluation Methods for Actions of a Structural Nature“, um eine europäische Evaluierungskultur durch die Verbesserung der Qualität und Nutzbarkeit der Evaluierungen zu entwickeln.
- die Entwicklung von Leitlinien zur Evaluation sowie ein verstärkter Informations- und Erfahrungsaustausch zwischen Programm-Managern und Evaluatoren.

Aus heutiger Sicht gelten daher folgende Rahmenbedingungen:

- Hohe **Verbindlichkeit** und Stellenwert durch Verankerung in den Verordnungen;
- Umfassendes **Design** von Konzeptebene, operativer Ebene (Organisation der Umsetzung), Programmvollzug, Wirkungsanalyse einschließlich der Möglichkeit internationaler Vergleiche

¹ Additionalität bedeutet in diesem Zusammenhang, dass EU-Mittel keine nationalen Budgetansätze ersetzen dürfen, also zusätzlich sein müssen.

(siehe hinsichtlich der unterschiedlichen Ebenen Übersicht 1);

- Orientierungsrahmen für die Programm-Entwicklung und Evaluation durch Systematisierungsversuche der Kommission und einen organisierten Erfahrungsaustausch.
- Verbindliche **Festlegung** des gesamten Evaluationszyklus von Ex-ante-, Zwischen- und Ex-post-Evaluierung (seit 1994).

Zur Verknüpfung von Zielen und Indikatoren

Eine weitere Stärke ist der Versuch, Zielhierarchien bei der Planung aufzubauen und diese mit Indikatoren-Ketten zu verknüpfen. Formal bauen die Überlegungen auf das Logical framework² auf, wodurch in der Planung ausgehend vom Globalziel auf unterschiedlichen Ebenen die dazugehörigen Indikatoren sozusagen von oben nach unten festgelegt werden. Bei der Begleitung wird in umgekehrter Reihenfolge – also praktisch von unten nach oben – die Zielerreichung überprüft.

Alle Interventionen der Programme werden in einem spezifischen Kontext in Bezug auf ein bestimmtes globales Ziel geplant. Dieses Ziel ist wiederum für die Interventionsstrategie ausschlaggebend und führt zu einer Reihe spezifischer, im Wesentlichen auf die Förderschwerpunkte ausgerichteter Ziele. Jedes spezifische Ziel wird zur Erreichung der operationellen Ziele mit Hilfe konkreter Maßnahmen umgesetzt. Die entsprechenden Zielsetzungen werden – soweit dies möglich ist – auf unterschiedlichen Ebenen quantifiziert. Durch diese Vorgangsweise soll ein Zusammenhang zwischen Zielen und Indikatoren-Ketten entwickelt werden. In der Begleitung bzw.

Bewertung erfolgt die Betrachtung in umgekehrter Reihenfolge. Demzufolge werden die:

- **operationellen Ziele** anhand des Outputs ausgedrückt (z.B. Bereitstellung von Schulungsmaßnahmen für Langzeitarbeitslose);
- **spezifischen Ziele** anhand der Leistung ausgedrückt (z.B. Verbesserung der Beschäftigungsfähigkeit von Langzeitarbeitslosen durch Schulungsmaßnahmen);
- **globalen Ziele** anhand der Wirkung ausgedrückt (z.B. Verringerung der Zahl der Langzeitarbeitslosen).

Es wird damit einerseits eine systematische Verknüpfung von Zielen und Indikatoren versucht und andererseits werden „Wirkungsketten“ innerhalb des Indikatorensystems dargestellt. Demnach unterscheidet man folgende Indikatoren:

- **Input-Indikatoren** geben an, wie viele finanzielle Mittel eingesetzt wurden
- **Output-Indikatoren** beziehen sich auf die unmittelbare Fördertätigkeit. Sie werden häufig in physikalischen oder finanziellen Einheiten gemessen (z.B. Kilometer gebauter Straße, Zahl der finanziell unterstützten Unternehmen, Zahl der bereitgestellten Ausbildungsplätze usw.)
- **Ergebnisindikatoren** beziehen sich auf die direkten **und** unmittelbaren Auswirkungen eines Programms. Die Indikatoren können physischer (Verringerung der Fahrtzeiten, Zahl der Auszubildenden mit erfolgreichem Abschluss) oder finanzieller Art sein (Investitionen des Privatsektors, Verringerung der Transportkosten usw.)
- **Wirkungsindikatoren** beziehen sich auf die Wirkungen des Programms über die unmittelbaren Auswirkungen auf die Endbegünstigten

2 Vgl. Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung.

hinaus, die zum Teil erst mit einem time-lag beobachtbar werden.

Die Kommission hat in ihren Arbeitspapieren diesbezüglich auch Listen von Indikatoren vorgeschlagen, die in den Programmen 2000-2006 zur Anwendung kommen sollen. Um eine internationale Vergleich- und Aggregierbarkeit zu gewährleisten wurden für unterschiedliche Themenbereiche sogenannte Core-Indikatoren definiert, die einer einheitlichen Definition folgen sollen.

Übermäßige Orientierung an der Quantifizierung von Beschäftigungseffekten

Das Regelwerk der EU-Strukturfonds lässt dem Evaluatoren prinzipiell einen breiten Spielraum, sowohl statistisch fundierte als auch beschreibende, quantitative als auch qualitativ ausgerichtete Analysen vorzunehmen, die zudem in den sozio-ökonomischen Entwicklungs-Kontext zu stellen sind. Insofern bestehen keine grundsätzlichen rechtlichen Einschränkungen, die Bewertung und Beurteilung von EU-Strukturfonds-Programmen gegebenenfalls auch so vorzunehmen, dass sowohl stärker qualitative als auch kontext-bezogene Analyse-Elemente berücksichtigt werden. Ein Zwang zur Quantifizierung, wenn diese sich als nicht sinnvoll und ergebnisbringend erweist, kann aus den Strukturfonds-Verordnungen in keiner Weise abgeleitet werden.

Dagegen scheint die Praxis so zu sein, dass die Kommission diese Grundsätze bei Programm-Konsultations- und Programm-Genehmigungsverfahren keineswegs strikt befolgt und vielmehr eine nahezu ausschließliche oder doch zumindest dominante quantitative Orientierung verlangt, insbesondere bezogen auf die Berechnung von Beschäftigungseffekten (vgl. BAW 1997). Die Begründung liegt hier wohl im hohen politischen Legitimationsdruck dem die EU-Regionalpolitik ausgesetzt ist.

Von Seiten der Kommission wird auf die Erfassung der Netto-Beschäftigungseffekte Wert gelegt. Bezogen auf die Ermittlung von Netto-Beschäftigungseffekten kann die Systematik der Kommission in drei Schritte zerlegt werden.

- Im ersten Schritt werden zusätzliche Arbeitsplätze als direktes Ergebnis, als direkte Folge oder als indirekte Folge von regionalpolitischen Aktivitäten unterschieden. Dabei handelt es sich also um Brutto-Effekte.
- Im zweiten Schritt wird differenziert zwischen befristeten / temporären Arbeitsplatzeffekten, Dauerarbeitsplätzen und gesicherten Arbeitsplätzen.
- Schließlich werden in einem dritten Schritt die Brutto-Größen auf Netto-Größen reduziert. Das heißt: Zum einen sollen die Mitnahme-Effekte abgezogen werden, zum anderen soll das Ergebnis um Verlagerungs- und Substitutionseffekte bereinigt werden.

Der Kommission ist es zwar gelungen, durch Vorlage entsprechender Definitionen und Leitlinien einen Orientierungsrahmen zu schaffen – und dies ist auch zu begrüßen – allerdings werden dadurch die eigentlichen Probleme der Evaluation (vor allem Ermittlung von Mitnahme-, Substitutions- und Verlagerungseffekten) keineswegs einer Lösung näher gebracht.

Praxis der Strukturfonds-Evaluationen

Die „empirischen“ Erfahrungen im Rahmen der Entwicklung, Umsetzung und Bewertung von Strukturfonds-Programmen zeigen eine deutliche Relativierung dieser von Kommissionsseite gestellten Erwartungen.

(1) Das „logical framework“ wird in der Planung und Evaluierung der Programme kaum in die Praxis umgesetzt. Die Ziele sind – wenn auch deutliche

Verbesserungen in den letzten 10 Jahren zu beobachten sind (Bachtler/Taylor 1999) – häufig unspezifisch und wenig operationalisiert. Dies hat auch systemimmanente Ursachen, versuchen sich die Regionen doch so wenig wie möglich in der Programm-Planung zu binden und durch unspezifische Zielsetzungen eine größtmögliche Flexibilität für die Umsetzungsphase zu erhalten.

(2) Daraus resultieren jedoch zahlreiche Probleme der Evaluationspraxis:

- in den Programmen ist vielfach ein häufig unsystematisches Nebeneinanderstellen von Indikatoren beobachtbar;
- damit ist die Verknüpfung der strategischen Zielsetzungen mit den Indikatoren kaum beobachtbar.³

In der Praxis steht alleinig die Zählung von Arbeitsplätzen im Vordergrund, auch für Maßnahmen wo dies kaum adäquat erscheint. Darüber hinaus sind häufig auch Mehrfachzählungen von geschaffenen Arbeitsplätzen zu beobachten (beispielsweise zwischen Infrastruktur- und betrieblichen Förderungen).

Die seit der Programmperiode 2000-2006 verfolgte Gliederung in Output, Ergebnis und Wirkungsindikatoren brachte deutliche Verbesserungen in der Systematisierung der Indikatoren in der Programm-Planung. Allerdings hat sie die Schwäche, dass hier Kausalketten unterstellt werden, die nicht immer nachvollziehbar sind.

3 Ein als good practice zu bezeichnendes Beispiel ist hier jedoch der von Bremen verfolgte Ansatz der Verknüpfung von Zielhierarchien mit Indikatoren. Eine ähnliche Vorgangsweise wurde für Österreich im Rahmen der Ziel-2-Zwischenevaluierung 1995-1999 vorgeschlagen (vgl. Gruber et al. 1998, publiziert in der ÖROK-Schriftenreihe Nr. 140).

(3) Zurückzuführen ist ein Gutteil der ungelösten Probleme vor allem auch darauf, dass die Systematik der Kommission nach wie vor sehr stark auf Strukturfonds-Programme aufbaut, die auf traditionelle Infrastrukturförderungen sowie einfache Investitionen ausgerichtet waren. Spätestens seit Mitte der 90er Jahre ist jedoch eine starke Ausdifferenzierung der Instrumente und ein Übergang auf eher „weiche“, nachfrageorientierte Maßnahmen wie Beratung, Kooperationsförderung, Technologietransfer, F&E etc. zu beobachten. Sowohl methodische Anforderungen als auch die Indikatoren sind jedoch noch zu sehr an den traditionellen Interventionen orientiert. Wiewohl anzumerken ist, dass hier natürlich nicht nur die EU-Kommission, sondern die Evaluations-Forschung insgesamt noch vor großen Herausforderungen steht.

(4) Wirkungsabschätzungen – auch in Ex-post-Analysen – sind kaum zu beobachten (EPRC/FAI 2000). Die Analyse der Evaluationsstudien, von „Meta-Studien“ sowie methodischen Konzepten der EU-Kommission⁴ zeigte:

- Aussagen über Mitnahme- und Verlagerungseffekte behelfen sich meist mit Surveys, die an die Förderungsempfänger adressiert sind. Die Frage lautet, ob das Projekt auch ohne Förderung durchgeführt, zu einem späteren Zeitpunkt durchgeführt oder in kleinerem Umfang durchgeführt worden wäre. Auf Basis der Surveys wird versucht, Koeffizienten für Mitnahme- und Verlagerungseffekte zu schätzen.
- Als Alternative werden häufig Koeffizienten und Benchmarks für Mitnahme- und Verlagerungseffekte praktisch „off the shelf“ verwendet, was

4 EPRC/FAI (2000); Ernst & Young (1999), MEANS Collection (1999), EK (1995), Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung

prinzipiell als problematisch einzuschätzen ist, wenn nicht berücksichtigt wird, dass diese nicht nur nach Sektoren stark variieren, sondern vom Design der Maßnahmen und dem Auswahlprozess der Projekte stark abhängig sind.⁵

- Kontrollgruppen-Ansätze werden kaum verwendet, wenn, dann meist nur im Rahmen der ESF-Förderung (Ziel 3). Lediglich im Rahmen einer Evaluierungsstudie im Ziel-1-Gebiet Irland wurde ein Kontrollgruppenansatz verwendet, um Netto-Effekte für Maßnahmen des EFRE zu ermitteln.

Die Evaluationen bleiben meist auf der Ebene der Bewertung des Programm-Designs, der Bewertung des Programm-Vollzugs sowie der Analyse der Umsetzungsstrukturen hängen, Wirkungsanalysen werden hingegen kaum vorgenommen.

Die Schwierigkeit der Strukturfonds-Evaluationen sind wohl die Folge des hohen Legitimationsdrucks der EU-Kommission, der wiederum zu einer unmäßig starken Betonung von Beschäftigungseffekten führt. Zu bewerten sind komplexe, multi-sektorielle Programme mit einer Vielzahl an Interventionsbereichen (von Technologie, F&E, Modernisierungsinvestitionen, Beratungs-, Gründungs-, Tourismusförderung etc.). Die Komplexität der Programme einerseits und der Erwartungshaltung der Kommission andererseits führt zu einem „stuck in the middle“ in der Begleitung und Bewertung. Die Notwendigkeit dieser Erwartungshaltung zumindest in Ansätzen zu gerecht zu werden, bindet einen Großteil der Ressourcen der Begleitung (Monitoring) und der Evaluation mit der Sammlung von Daten, die in der

Folge jedoch nur geringe Aussagekraft haben. Aus pragmatischen Gründen werden dann vielfach inadäquate Ansätze wie die Abschätzung von Mitnahmeeffekten oder die alleinige Abschätzung von Wirkungen von F&E-Förderungen durch die Befragung von Firmen vorgenommen. Gleichzeitig werden durch diese Bemühung zur Quantifizierung der Zahl der Arbeitsplätze auch die Ressourcen der Evaluatoren übermäßig beschränkt, sodass nicht mehr genügend Raum für alternative Evaluierungsansätze bleibt, die gerade in den vorliegenden Multi-Akteursprogrammen von Bedeutung sein können und einen Beitrag zum reflexiven Lernen leisten könnten.

Perspektiven

Die Perspektiven der Weiterentwicklung könnten in drei Bereichen zusammengefasst werden:

(1) Eine Reduktion der Erwartungshaltung hinsichtlich der Möglichkeiten zur Ermittlung von Netto-Effekten in komplexen regionalen Entwicklungsprogrammen und einer Erhöhung der Bereitschaft qualitative Ergebnisse anzuerkennen.

(2) Ergänzung bisheriger Ansätze durch die im MEANS-Programm Volume 4 dargestellten Ansätze wie concept mapping of impacts oder den Einsatz von scoring scales, matrix of cross impacts. Zum derzeitigen Zeitpunkt sind diese Methoden erst in einzelnen Programmen pilothaft erprobt worden. Allen Ansätzen ist gemein, dass neben bisherigen Analysen auch ergänzend qualitative Effekte erfasst sowie eine Gewichtung der Maßnahmen hinsichtlich ihres Beitrages zur Erreichung des Programmzieles vorgenommen werden. Das Wesen dieser Bewertungsmethoden liegt in der Einbeziehung der Stakeholders in den Evaluierungs- und Bewertungsprozess und der Unterstützung des reflexiven Lernens. Als Voraussetzung gilt wohl – insbesondere in der Konkurrenz

5 Allerdings wurden in einer hier als good practice zu bezeichnenden Studie 250 Evaluations-Studien analysiert, um daraus eine Liste von Koeffizienten zur Berechnung von Netto-Effekten bei Arbeitsplatzzahlen zu generieren. Diese Liste wurde in der Folge in einer Evaluierungsstudie eingesetzt (Yorkshire and Humberside Ziel-2-Evaluierung) (EPRC/FAI 2000).

um knappe Ressourcen – die dominant quantitative Orientierung der Kommission aufzuweichen.

(3) Der dritte Bereich liegt in einer kreativeren Herangehensweise von Seiten der Auftraggeber im Rahmen der Beauftragung von Evaluationen: Eine kontinuierlichere Einbeziehung von Evaluatoren in Form einer Begleitevaluation würde die Wahrung des laufenden Überblicks und gleichzeitig die sukzessive Bewertung einzelner Schwerpunkte ermöglichen; der verstärkte Einsatz von thematischen Evaluationen – programmübergreifend – gäbe den Spielraum, methodisch sauber Maßnahmenswerpunkt zu bewerten und Wirkungsanalysen durchzuführen.

Bachtler J./Polverari, L./Taylor S./Ashcroft B./Swales, K., Methodologies used in the Evaluation of Effectiveness of European Structural Funds: A Comparative Assessment, Study of the European Policies Research Centre and Fraser of Allander Institute, University of Strathclyde (EPRC/FAI), Glasgow 2000

Bachtler, J./Taylor S., Objective 2: Experiences, Lessons and Policy Implications, European Policies Research Centre, Glasgow, Juli 1999

BAW-Monatsberichte, Evaluierung Europäischer Strukturfonds-Programme, Zum Stand der Kunst – Teil II, Bremer Ausschuß für Wirtschaftsforschung, Bremen 4/1997

European Commission (EC) (2000a), The New Programming period 2000 – 2006: methodological working papers, Working paper 3, Indicators for Monitoring and Evaluation, An indicative methodology

European Commission (EC) (2000b), Annual Evaluation Review 1999, Highlights of results from evaluation of expenditure programmes, Working paper of the Commission departments;
http://europa.eu.int/comm/budget/evaluation/en/bilan/evaluationreview1999.html#_TOC475502965

European Commission (EC) (1999a), Evaluating socio-economic programmes, MEANS Collection Volume 1-6; Luxemburg, 1999

European Commission (EC) (1997), The New Regional Programmes 1997-99 under Objective 2 of the Community's Structural Policies – focusing on Job Creation, COM (97), 524 final, Commission of the European Communities, Brussels

Europäische Kommission (EK) (1999), Reform der Strukturfonds 2000-2006, Eine vergleichende Analyse, 1999

Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung. (DE/05/95/52070100.p00(FR).

European Policies Research Centre (EPRC 2000), At the Starting Block: Review of the new Objective 2 Programmes, Glasgow 2000

Österreichische Raumordnungskonferenz, Zwischenbewertung der Interventionen der Ziel-2-Programme, des RESIDER-II- und des RECHAR-II-Programmes in der Programmperiode 1995-99 in Österreich, Studie der JOANNEUM RESEARCH in Kooperation mit dem Bremer Ausschuß für Wirtschaftsforschung und dem European Policies Research Centre, Schriftenreihe Nr. 140, Wien 1998

Autoren-Hinweis

Markus Gruber

Joanneum Research, Institut für Technologie- und Regionalpolitik, Elisabethstraße 20
 A-8010 Graz
 Tel ++43 316 876 1843
 E-mail: markus.gruber@joanneum.at

Übersicht 1 Thematische Schwerpunkte der Strukturfonds-Evaluation

Schwerpunkt	Beurteilung	Ex-ante	Interim	Ex-post
Erfahrungen über Wirkungen bisheriger Programme	... der Erfahrungen über Effektivität und Effizienz der Programme aus vorangegangenen Programm-Perioden und deren Berücksichtigung im aktuell vorliegenden Programm.			
Analyse der Ausgangssituation	... der sozio-ökonomischen Ausgangssituation und Ableitung eines Stärken/Schwächen-Profiles. Speziell zu beachten sind folgende Politikschwerpunkte: Umweltsituation, Chancengleichheit, Beschäftigung, Innovation, KMU.			
Strategiekonzept	... der vorgelegten Strategien hinsichtlich ihrer Adäquanz und Kohärenz; ... des Strategiekonzeptes hinsichtlich der Politikbereiche: Umwelt, Chancengleichheit, Beschäftigung, Innovation, KMU.			
Zielsetzungen	... der gewählten Zielsetzungen sowohl hinsichtlich der Kohärenz mit der Analyse als auch der Frage ob die Zielsetzungen ausreichend quantifiziert wurden und die gewählten quantitativen Ziele als realistisch zu beurteilen sind.			
Maßnahmen	... der einzelnen Maßnahmen und des Maßnahmen-Mix auf Adäquanz und Kohärenz.			
Externe Kohärenz mit EU / nationaler Politik	... der Kohärenz des Programmes hinsichtlich der EU-Politik-Bereiche sowie der nationalen Politiken.			
Finanztableaus, Gewichtung der Maßnahmen	... der Adäquanz der gewählten finanziellen Gewichtung der einzelnen Programm-Teile und Maßnahmen.			
Indikatoren	... der Aussagekraft, Konsistenz und Adäquanz der bereits vorliegenden Indikatoren			
Partnerschaftsprinzip	..., ob im Rahmen der Vorbereitung / Umsetzung des Programmes die Regelungen hinsichtlich der Partnerschaft entsprechend eingehalten wurden.			
Operationelle Rahmenbedingungen	... der (geplanten) Verfahren der Programm-Umsetzung. - der Organisationsstrukturen hinsichtlich Klarheit, Transparenz, Informationsflüsse - der eingesetzten Mechanismen der Umsetzung (Ausschreibungen, Antragsverfahren, Publizität und Öffentlichkeitsarbeit) - der Projektselektions-Kriterien und -mechanismen - des Begleit(=Monitoring)systems.			
Programm-Vollzug - Inputfaktoren - Finanzielle Indikatoren	... des Standes der Mittelbindungen und Auszahlungen nach Programm / Prioritäten / Maßnahmen und Feststellen von Gründen für Abweichungen zu den geplanten Finanztabellen.			
Programm-Vollzug - Output-Faktoren	... des Programmfortschritts hinsichtlich der Erreichung der festgelegten Output-Ziele. Rückschlüsse auf die Zielerreichung; Spezielle Beachtung der horizontalen Schwerpunkte der Union Umwelt, Beschäftigung, Chancengleichheit, KMU, Innovation.			
Wirkungsanalyse	... der Programm-Wirkungen auf Basis der Programm-Ziele; Abschätzung von Brutto- bzw. Netto-Effekten.			

Quelle: eigene Zusammenstellung basierend auf VO 1260/99 sowie Arbeitspapieren der EU-Kommission

	Schwerpunkt der Evaluation
	Zu behandelndes Thema

Announcements

Annual Conference 2001 of the German Evaluation Society (DeGEval)

The next annual conference of the German Evaluation Society will be held at the German Post-Graduate School of Administrative Studies in Speyer on 4 and 5 October 2001 (for details on the conference venue, see: <http://www.hfv-speyer.de/>).

The conference motto will be **"Evaluation in Administration and Management"**. Up to 300 scientists, politicians, experts from administration and enterprises as well as practitioners from the fields of schools, university, companies, social work, health sector, environment, media, research and development policy, will discuss success controls, impact analyses and efficiency measurement in this interdisciplinary forum. The main focus of different working groups will also be on the improvement of evaluation by standards of good evaluation practice as well as the training and further training of evaluators and contractors.

Interested persons who want to discuss further topics pertaining to evaluation in own panel sessions, besides the DeGEval standing working group meetings, are cordially invited. The DeGEval board would especially welcome the organisation of sessions on the topics methods, utilisation of evaluations and evaluation of consumer protection policy. Please contact the DeGEval office with a short sketch of your theme.

You will find further details on the conference under www.degeval.de. You can put your name down for the conference under speyer@degeval.de as of now.

Contact address:

DeGEval Office, Mr. Berthold Schobert
P. O. Box 26 01 29, D-50514 Köln, Germany
Tel./Fax: +49-221/4600032
E-Mail: info@degeval.de; <http://www.degeval.de>

Jahrestagung 2001 der Deutschen Gesellschaft für Evaluation (DeGEval)

Am 4. und 5. Oktober 2001 wird die nächste Jahrestagung der Deutschen Gesellschaft für Evaluation an der Hochschule für öffentliche Verwaltung in Speyer stattfinden (Details zum Tagungsort <http://www.hfv-speyer.de/>).

Die Tagung steht unter der Überschrift **"Evaluation in Verwaltung und Management"**. Zu diesem interdisziplinären Forum werden bis zu 300 Wissenschaftler, Politiker, Experten aus Verwaltung und Unternehmen sowie Praktiker aus den Bereichen Schule, Hochschule, Betrieben, Soziale Arbeit, Gesundheit, Umwelt, Medien, Forschungs- und Entwicklungspolitik über Erfolgskontrollen, Wirkungsanalysen und Effizienzmessung diskutieren. Im Mittelpunkt der Arbeitsgruppen steht dabei die Verbesserung der Evaluation durch Standards guter Evaluationspraxis sowie die Aus- und Weiterbildung von Evaluatoren und Auftraggebern.

Interessenten, die neben den Arbeitskreisen weitere Themen der Evaluation in eigenen Arbeitsgruppen diskutieren wollen, sind herzlich eingeladen. Der Vorstand würde sich besonders über die Organisation von Arbeitsgruppen zu den Themen Methoden, Nutzung von Evaluationen und Evaluation von Verbraucherschutzpolitik freuen. Bitte wenden Sie sich mit einer kurzen Darstellung an die Geschäftsstelle der DeGEval.

Weitere Einzelheiten zur Tagung finden Sie unter [http://www.degeval.de/](http://www.degeval.de). Sie können sich ab sofort unter speyer@degeval.de vormerken lassen.

Kontakt:

Geschäftsstelle DeGEval, Herrr Berthold Schobert
Postfach 26 01 29, D-50514 Köln, Germany
Tel./Fax: +49-221/4600032
E-Mail: info@degeval.de; www.degeval.de

- Der Newsletter der PLATTFORM■TECHNOLOGIE■EVALUIERUNG ist in Zukunft auch im **pdf Format** erhältlich. Wenn Sie an einer elektronischen Zusendung interessiert sind, so richten Sie bitte ein kurzes email an plattform@joanneum.at oder faxen Sie diesen Abschnitt an 43 – 1 – 581 75 20 28 20.
- In the near future, the Newsletter PLATTFORM■TECHNOLOGIE■EVALUIERUNG will also be available in **pdf format**. If you are interested, please contact plattform@joanneum.at or send a fax to 43 – 1 – 581 75 20 28 20.

Name / Institution:

Adresse:

email:

Für den Inhalt dieser Ausgabe verantwortlich:

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION
UND TECHNOLOGIE; Renngasse 5, A-1010 Wien
Dr. Rupert Pichler (rupert.pichler@bmv.gv.at)

ÖSTERREICHISCHES FORSCHUNGSZENTRUM
SEIBERSDORF, Systemforschung Technik-Wirtschaft-
Umwelt, A-2444 Seibersdorf
DI Anton Geyer (anton.geyer@arcs.ac.at)

JOANNEUM RESEARCH, Institut für Technologie- und
Regionalpolitik, Wiedner Hauptstraße 76, A-1040 Wien und
Elisabethstraße 20, A-8010 Graz
Mag. Klaus Zinöcker (klaus.zinoecker@joanneum.ac.at)

Weitere Plattformmitglieder:

ÖSTERREICHISCHES INSTITUT FÜR
WIRTSCHAFTSFORSCHUNG, A-1103 Wien, PF 91
Mag. Gernot Hutschenreiter
(gernot.hutschenreiter@wifo.ac.at)

TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H.,
Grillparzerstr. 7/8, A-1010 Wien
Dr. Michael Stampfer (michael.stampfer@tig.or.at)

FFF-Forschungsförderungsfonds für die Gewerbliche
Wirtschaft, Kärntnerstraße 21-23, A-1015 Wien
Mag. Klaus Schitzer (klaus.schnitzer@fff.co.at)

TECHNOPOLIS, Prinz Eugen Straße 80/12, A-1040 Wien,
DI Fritz Ohler (fritz.ohler@technopolis-group.com)

FWF-Fonds zur Förderung der wissenschaftlichen
Forschung, Weyringergasse 35, A-1040 Wien
office@mails.ffw.univie.ac.at

Medienhinweis:

PLATTFORM■TECHNOLOGIE■EVALUIERUNG ist ein
unregelmäßig erscheinendes offenes Forum zur Diskussion
methodischer und inhaltlicher Evaluierungsfragen in der
Technologiepolitik. © Wien 2001

Herausgabe und Versand: Joanneum Research, Institut für
Technologie- und Regionalpolitik, A-1040 Wien, Wiedner
Hauptstraße 76.

PLATTFORM■TECHNOLOGIE■EVALUIERUNG

Die PLATTFORM■TECHNOLOGIE■EVALUIERUNG ist
eine Initiative des österreichischen
BUNDESMINISTERIUMS VERKEHR, INNOVATION
UND TECHNOLOGIE. In der Zusammenarbeit von
drei österreichischen Forschungseinrichtungen,
JOANNEUM RESEARCH, WIFO und
FORSCHUNGSZENTRUM SEIBERSDORF, der
TECHNOLOGIE IMPULSE GESELLSCHAFT,
TECHNOPOLIS sowie den beiden Fonds FFF und
FWF werden Themenstellungen zur
Technologieevaluierung erarbeitet und – z.T. mit
Einbeziehung namhafter ExpertInnen – in einem
Fachkreis diskutiert.

Der Newsletter beinhaltet Fachbeiträge zu Fragen der
technologepolitischen Evaluierung. Die Herausgabe
erfolgt in zeitlicher als auch inhaltlicher Abstimmung
mit Plattform-Veranstaltungen, um die
Synergiewirkungen eines breiten Austauschforums zu
nutzen.