



Fraunhofer Institute
Systems and
Innovation Research

Fraunhofer ISI Discussion Papers *Innovation System and Policy Analysis*, No 3/2003

ISSN 1612-1430

Karlsruhe, Oktober 2003

Informationen zur Forschungsevaluation in Deutschland – Erzeuger und Bedarf

Gutachten für die Geschäftsstelle der
Deutschen Forschungsgemeinschaft

Stefan Kuhlmann

Fraunhofer Institute for Systems and Innovation Research
and Utrecht University, Copernicus Institute

Thomas Heinze

Fraunhofer Institute for Systems and Innovation Research

Contents

ZUSAMMENFASSUNG:	3
1 ZIELSETZUNGEN UND GRENZEN DES GUTACHTENS	4
2 PRAXIS UND GRENZEN DER EVALUATION VON FORSCHUNG	6
2.1 Evaluation von Forschung in Deutschland.....	6
2.2 Methoden und Grenzen der Evaluation.....	8
3 ZWECKE, TYPEN UND AGGREGATIONSEBENEN EVALUATIVER INFORMATION	11
3.1 Zwecke	11
3.2 Typen.....	12
3.3 Aggregationsebenen.....	15
4 ERZEUGER UND ANGEBOTE EVALUATIVER INFORMATIONEN.....	17
4.1 Forschungsinput.....	17
4.2 Forschungsleistungen (Output und Wirkung)	21
4.3 Anbieter evaluativer Informationen	25
5 INFORMATIONSNUTZER UND DESIDERATA DER INFORMATIONSBEREITSTELLUNG	31
5.1 Nutzerinteressen	31
5.2 Desiderata.....	32
6 MÖGLICHE ORGANISATORISCHE ANTWORTEN	35
7 VERWENDETE LITERATUR	40

Zusammenfassung:

Informationen zur Forschungsevaluation in Deutschland – Erzeuger und Bedarf

(1) Der Ruf nach Rechenschaft über Leistungen, Qualität und Nutzen staatlich geförderter Wissenschaft und Forschung wird lauter. Wachsende Autonomiespielräume der Hochschulen durch Einführung von Globalhaushalten gehen Hand in Hand mit neuen, leistungsorientierten Modi der Verteilung von Forschungsmitteln. Politisch gewollter Wettbewerb zwischen Forschungseinrichtungen erfordert eine verstärkte Profilbildung und Wissenschaftsmanagement. Mit Blick auf den künftigen "europäischen Forschungsraum" versucht auch die Europäische Kommission Leistungsindikatoren zu identifizieren und in ein Berichtssystem zu implementieren. In der Summe bewirken diese Entwicklungen bei den Verantwortlichen in Forschungsplanung, -management, und -politik einen *wachsenden Bedarf an Daten und Informationen zu Forschungsleistungen und zu Grunde liegenden Strukturvariablen*.

(2) Das Gutachten beruht auf der Annahme, dass dieser wachsende Bedarf bislang nicht hinreichend gedeckt wird. In diesem Zusammenhang interessiert die Frage, ob es sinnvoll ist eine *Dienstleistungseinrichtung* einzurichten, die evaluative und forschungsstrukturbezogene Informationen bereit hält. Ein solches Angebot könnte sich an alle Akteure des Forschungssystems richten. Das Gutachten prüft diese Annahmen und Überlegungen – allerdings nur *explorativ und exemplarisch!*

(3) Das Gutachten präsentiert eine Auswahl von anerkannten Indikatoren und Verfahren, mithilfe welcher evaluative Informationen generiert werden können. Diese werden nach *Evaluationszwecken, Informationstypen und Aggregationsebenen* geordnet. Dabei ist bewusst zu halten, dass die Zusammenhänge zwischen "Investition" in Forschung auf der einen und Innovation und möglichen wissenschaftlichen, wirtschaftlichen, sozialen oder ökologischen "Erträgen" auf der anderen Seite komplex sind und *keinesfalls als schlichtes Input/Output-Modell* konstruiert werden können.

(4) Im stark fragmentierten, nicht zentralistisch organisierten deutschen Forschungssystem sind Angebot und Nachfrage evaluativer Information sehr unübersichtlich. Zur besseren Strukturierung wird dieser "Markt" skizziert, zum einen nach "Anbietern" evaluativer Informationen und zum anderen nach deren Nachfragern. Hierbei werden Desiderata mit Blick auf nachfragende Nutzer sichtbar. Der Bedarf ist um so größer, *je höher aggregiert die Betrachtungs- und Entscheidungsebene* der Informationsnutzer ist. Es fehlen bisher vor allem *kompatible Formate und Rhythmen* der Erfassung von Daten quer zu den Forschungseinrichtungen, mit Blick auch auf die internationale Anschlussfähigkeit. Solche Formate bilden die Voraussetzung für Vergleiche (institutionelle, sektorale, ...) auf aggregierten Niveaus. Dabei müssen die *unterschiedlichen Aufgabenstellungen* und fachlichen Orientierungen heterogener Forschungseinrichtungen unbedingt berücksichtigt werden.

(5) Das Gutachten skizziert die Aufgaben und die Konstruktion einer möglichen *Gemeinschaftsstelle "Clearinghouse"* – konzipiert als "intelligenter Netzknoten" –, getragen und gesteuert von einem *Konsortium* (d.h. einer freiwilligen Kooperation) aller relevanten deutschen Einrichtungen der Forschungspolitik, der Forschungsförderung, der Forschungsträgerinstitutionen sowie ausgewählten Einrichtungen der Wissenschafts- und Innovationsforschung.

1 Zielsetzungen und Grenzen des Gutachtens

Über ein Drittel aller in Deutschland betriebenen Wissenschaft und Forschung wird von der öffentlichen Hand finanziert, knapp zwei Drittel von Unternehmen der Wirtschaft. Die staatlichen Forschungsmittel fließen überwiegend an Universitäten und außeruniversitäre öffentliche Forschungseinrichtungen. Allein die Deutsche Forschungsgemeinschaft (DFG) finanziert etwa 40% aller Drittmittel an Hochschulen.

Traditionell finanziert die DFG Forschungsprojekte unter der Bedingung einer *ex ante*-Evaluation der wissenschaftlichen Qualität von Projektvorschlägen durch Fachkollegen (*peer review*). Dieses Verfahren genießt international hohes Ansehen. Eine *ex post*-Bewertung der Ergebnisse und Auswirkungen geförderter Forschung, etwa Projekte oder Programme der DFG und anderer Forschungsförderungseinrichtungen erfolgt jenseits einzelner Ad-hoc-Evaluationen bis heute jedoch nicht systematisch und regelmäßig. Das liegt nicht zuletzt daran, dass es kein nach einheitlichen Kriterien organisiertes Verfahren zur Erhebung von Input- und Output-Parametern des bekanntermaßen heteronomen, stark "versäulten" deutschen Forschungssystems auf unterschiedlichen Aggregatebenen gibt.¹ Damit ist es nicht nur schwierig, einen deskriptiven Überblick über die Forschungslandschaft zu erhalten; es ist auch kaum möglich, relevante Leistungsparameter - sofern sie gemessen werden - in Relation zu aussagefähigen, gleich aggregierten Input-Größen zu setzen, um Analysen der Dynamik sowie der Stärken und Schwächen des Forschungssystems zu betreiben. Strukturbildende Effekte von Forschungsförderung – etwa die Herausbildung international anerkannter Forschungsschwerpunkte ("Centres of Excellence") oder Auswirkungen auf das Bildungsangebot der Hochschulen sind ebenfalls bisher nicht systematisch erfassbar.

Angesichts des seit den 1990er Jahren verstärkten Rufes nach Rechenschaft über Leistungen, Qualität und Nutzen von staatlich geförderter Wissenschaft und Forschung wird es in steigendem Maße erforderlich, die "effektive" und "zielführende" Verwendung staatlicher Mittel zu dokumentieren. Darüber hinaus bedingen wachsende Autonomiespielräume der Hochschulen durch Einführung von Globalhaushalten neue, leistungsorientierte Verteilungsmodi der Ressourcenallokation. Weiterhin wird das in Gang gesetzte Prinzip des Wettbewerbs zwischen Forschungseinrichtungen eine verstärkte Profilbildung mithilfe gezielter planerischer Entscheidungen im Rahmen von forschungspolitischem Wissenschaftsmanagement erfordern. Mit Blick auf den künftigen "europäischen Forschungsraum" ist derzeit auch die Europäische Kommission bemüht, diesbezügliche Indikatorenbündel zu identifizieren und in ein Berichtssystem zu implementieren (vgl. Ragnarson 2001; European Commission 2000). In der Summe bewirken diese Entwicklungen bei den Verantwortlichen in Forschungsplanung, -management, und -politik einen *wachsenden Bedarf an Daten und Informationen zu Forschungsleistungen und zu Grunde liegenden Strukturvariablen*.

Das vorliegende Gutachten, angefertigt für die Geschäftsstelle der Deutschen Forschungsgemeinschaft, beruht auf der Annahme, dass dieser wachsende Bedarf an Daten und In-

¹ Siehe z.B. die "Systemevaluationen" der großen Forschungseinrichtungen in Deutschland; Evaluierungskommission 1998; Internationale Kommission 1999; Wissenschaftsrat 2000a, 2001. Vgl. auch Schmoch 2001.

formationen zu Forschungsleistungen bislang von vorhandenen Einrichtungen – z.B. den Verwaltungen der großen Forschungsträgerorganisationen; den Hochschulen und ihren Verbänden; dem Wissenschaftsrat – nicht hinreichend gedeckt wird. In diesem Zusammenhang wurde in Fachkreisen die Frage aufgeworfen, ob es sinnvoll ist eine *Dienstleistungseinrichtung* des deutschen Forschungssystems einzurichten, die evaluative und forschungsstrukturbezogene Informationen der verschiedenen Einrichtungen des Forschungssystem sammelt, strukturiert, aggregiert und für Zwecke der Analyse und Strategieentwicklung bereit hält. Ein solches Angebot könnte sich an alle Akteure des Forschungssystems richten.

Das Gutachten kann diese Annahmen und Überlegungen – im Rahmen knapper Zeit und Personalressourcen – *nur explorativ und exemplarisch* prüfen, sie also nicht systematisch und empirisch erschöpfend behandeln! Als Informationsquellen wurden einschlägige Literatur und Dokumente genutzt, außerdem führten die Verfasser zehn Hintergrundgespräche mit Experten aus Forschungspolitik, -verwaltung und Wissenschaft.

Zum Aufbau des Papiers: Bevor wir in Kap. 5 Desiderate der Informationsbereitstellung und in Kap. 6 mögliche organisatorische Antworten darauf diskutieren, wollen wir in Kap. 2 knapp Praxis und Grenzen der Evaluation von Forschung skizzieren, in Kap. 3 Zwecke, Typen und Aggregationsebenen evaluativer Information strukturieren sowie in Kap. 3 Erzeuger und Angebote evaluativer Information explorativ darstellen.

2 Praxis und Grenzen der Evaluation von Forschung

2.1 Evaluation von Forschung in Deutschland

In historischer Perspektive kann man in Deutschland wie auch in anderen industrialisierten Ländern wissenschaftsinterne und –externe Ansätze als zwei heterogene Entwicklungslinien von Evaluationsverfahren im Bereich von Wissenschaft, Forschung und Technologieförderung ausmachen; sie lassen sich als ein *Drei-Schalenmodell* darstellen²:

Erste Schale - Individuelle Forschungsleistungen: Den "Kern" bilden peer review-Verfahren und später zusätzlich Verfahren zur Messung der Forschungsleistung einzelner Forscher und Gruppen (Bibliometrie etc.) als *wissenschaftsinterne* Instrumente für die Entscheidung über die Allokation von Fördermitteln in Forschungseinrichtungen (vgl. hierzu Daniel 1993; Hornbostel 1997; van Raan 1988). Peer Review-Verfahren kommen im deutschen Forschungs- und Innovationssystem weithin zur Anwendung, insbesondere bei der ex ante-Bewertung von Projekten der grundlagen- und der langfristig anwendungsorientierten Forschung (Campbell/Felderer 1997). Dieses Verfahren ist das vorherrschende Evaluationsinstrument der Deutschen Forschungsgemeinschaft (DFG). Förderanträge werden von Peers bewertet, die im Vier-Jahres-Rhythmus von der gesamten wissenschaftlichen Gemeinschaft gewählt werden (Neidhardt 1988). Jeder Gutachter ist angehalten, den Antrag allein auf der Grundlage seiner wissenschaftlichen Qualität zu beurteilen. Im Frühjahr 2002 haben Senat und Hauptausschuss der DFG beschlossen, dieses bewährte System ("bottom-up") um strategie-orientierte, interdisziplinäre und internationale Komponenten zu ergänzen (DFG 2002a).

Zweite Schale - Programme: Um diesen Kern der Peer Review-Verfahren herum legt sich eine Schale, die aus Evaluationsstudien im Sinne von impact analyses forschungs- und innovationspolitischer Programme besteht. Die in Deutschland bisher durchgeführten Studien schließen konzeptionell überwiegend an die Tradition der Wirkungsforschung an: Grundlage bilden die vielfältigen und traditionsreichen Arbeiten vor allem der amerikanischen Wirkungsforschung (impact analysis), in Verbindung mit Ansätzen der policy analysis, die sich zu einem in vielen Politikfeldern (insbesondere der Sozialpolitik) angewendeten Instrument der Politikberatung ("policy-analytic movement"; vgl. Schön/Rein 1994: 11) entwickelt hat (Übersichten bei Widmer 1996; Patton 1997; Kuhlmann 1998: 86-111; Wollmann 1998). Dieses Evaluationskonzept kann als *wissenschaftsextern* bezeichnet werden, denn es wird überwiegend von Akteuren des politisch-administrativen Systems in Gang gesetzt in der Absicht, die Erreichung politisch gesetzter wissenschaftlicher, technologischer, ökonomischer oder gesellschaftlicher Ziele zu prüfen. Wirkungsanalysen haben sich seit den 1970er Jahren in Deutschland mit der Verbreitung von Programm-Politik in vielen Politikfeldern durchgesetzt (vgl. Derlien 1976; Mayntz 1980 und 1983; Hellstern/Wollmann 1984). Seit dieser Zeit wurde etwa die Hälfte des Forschungshaushalts in die Projekt- bzw. Programmförderung gelenkt, die andere Hälfte in die institutionelle Förderung: Im Unterschied zur institutionellen Förderung verfolgen Programme gezielt politische Steuerungs- und Gestaltungsabsichten, etwa im Bereich der Förderung von "Schlüs-

² Vgl. zum Folgenden Kuhlmann 2000: 303ff.

selstechnologien" oder der Stimulierung von Innovationstätigkeit in der mittelständischen Wirtschaft. Weitreichende Steuerungsansprüche verlangten nach Erfolgskontrolle: Die Programmevaluation und Wirkungsforschung erlebten seither einen nachhaltigen Aufschwung (Meyer-Krahmer 1989). Die Ausbreitung von Evaluationsverfahren ist außerdem eng mit der wachsenden Zahl von strategischen Programmen (der EU-Kommission) zur Förderung von Wissenschaft und Technologie verbunden. In Deutschland zeigte sich insbesondere das Bundesforschungsministerium aktiv; zwischen 1985 und 1993 ließ man etwa 50 größere Evaluationsstudien anfertigen (Kuhlmann/Holland 1995a). Als Evaluatoren treten dabei üblicherweise unabhängige Forschungsinstitute im Auftrag forschungspolitisch-administrativer Akteure auf; sie wenden ein relativ breites Spektrum von Evaluationskonzepten, -methoden und -instrumenten an (vgl. Becher/Kuhlmann 1995; mittlerweile eingebunden in internationale Netzwerke, vgl. Shapira/Kuhlmann 2002) und sind seit den 1990er Jahren auch professionell in einer "Deutschen Gesellschaft für Evaluation (DeGEval)" organisiert.

Dritte Schale - Institutionen: Hier geht es um die Leistungsfähigkeit ganzer Forschungsinstitutionen. In Deutschland spielen seit langem die Begutachtungen definierter Wissenschaftsgebiete durch den Wissenschaftsrat eine wichtige Rolle; bei der Neustrukturierung der Forschungslandschaft Ostdeutschlands nach der Vereinigung übernahmen sie sogar eine gestaltende Funktion (Block/Krull 1990). Seit den 1990er Jahren wurden institutionelle Evaluationen immer häufiger durchgeführt (Kuhlmann/Holland 1995b). Im Frühjahr 1999 schloss eine internationale Kommission eine "Systemevaluation" der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft ab (im Auftrag der Regierungschefs von Bund und Ländern; vgl. Internationale Kommission 1999). Zeitlich parallel wurde auch eine Systemevaluation der Fraunhofer-Gesellschaft durchgeführt (vgl. Evaluierungskommission 1998). Eine aufwändige Evaluation der Einrichtungen der "Wissenschaftsgemeinschaft G.W. Leibniz (WGL)" wurde 2000 abgeschlossen (Wissenschaftsrat 2000; Röbbcke/Simon 2001) sowie 2001 Systemevaluationen der Großforschungseinrichtungen der "Helmholtz-Gesellschaft" (Wissenschaftsrat 2001) und der Institutionen der "industriellen Gemeinschaftsforschung" (Kommission Systemevaluation 2001). Bei der Evaluation von Forschungsinstitutionen werden zunehmend *wissenschaftsinterne und -externe* Fragestellungen und Informationsquellen *kombiniert*.

Parallel zum gewachsenen Interesse an der Leistungsfähigkeit öffentlich geförderter Forschungsprogramme und -institutionen widmete auch die *Privatwirtschaft* der Effizienz, Effektivität und strategischen Ausrichtung ihrer Laboratorien erhöhte Aufmerksamkeit (z.B. Brockhoff 1999; Gerpott 1999; Bürgel et al. 1996).

Zusammenfassend lässt sich die deutsche Evaluationspraxis im Bereich der Leistungen von Forschung und Forschungsinstitutionen einerseits als *entwickelt* und *dynamisch* bezeichnen, andererseits aber auch als *zersplittert* charakterisieren: Sie ist entwickelt hinsichtlich des hohen Maßes an Selbstorganisation bei wissenschaftsinternen Verfahren, die Konsens und Verpflichtung zwischen den Forschern fördern, sowie dynamisch wegen vielfältiger neuer Evaluationsansätze und Strategieentwicklungen in allen Institutionen und auf allen Ebenen, jedoch auch unsystematisch und zersplittert, weil die verschiedenen institutionell orientierten Evaluationsbemühungen bisher nur wenig aufeinander abgestimmt werden konnten.

Weiterhin ist festzustellen, dass sich viele Evaluationsinitiativen (interne und externe Ansätze) nicht mehr allein auf die Feststellung der Qualität einzelner Forschungsleistungen

beschränken. Zusehends soll evaluative Information auch zur Erhellung *strukturbildender* sowie *sozio-ökonomischer Effekte* von Forschung und Forschungsförderung beitragen (Airaghi et al. 1999; European Commission 2002a). Ein gemeinsames Raster von Indikatoren zur Erfassung und Analyse solcher Effekte wurde bisher jedoch noch nicht verabredet.

2.2 Methoden und Grenzen der Evaluation

Wir kennen heute vielfältige Methoden zur Feststellung erzielter oder potentiell erzielbarer Leistungen und Wirkungen von Forschung (vgl. Fisch/Daniel 1986, Daniel/Fisch 1988, van Raan 1988; Daniel 1993; Bozeman/Melkers 1993; Grupp/Kuntze/Schmoch 1995; Hornbostel 1997, 2000; Grupp 1998). Neben der "klassischen" Bewertung von Forschungsleistungen durch Fachkollegen (peer review) sind die wichtigsten der Vorher-/Nachher-Vergleich, der Kontroll- oder Vergleichsgruppenansatz, sowie qualitative Analysen (u.a. Plausibilitätsüberprüfungen, Schätzurteile). Sie können mit unterschiedlichen Indikatoren (finanzieller Aufwand für Forschung und Entwicklung, Veröffentlichungen, Zitate, Patente, ökonomische, soziale, technische Kenngrößen, etc.), Datensammlungsverfahren (Statistiken, Fragebögen, Interviews, Fallstudien, Panel etc.) und Datenanalyseverfahren (ökonomische Modelle, Cost-/Benefit-Analysen, andere statistische Verfahren, Technometrie, Bibliometrie, Peer Review) einzeln oder kombiniert verwendet werden. Bei allen notwendigen Bemühungen um objektivierende Verfahren und geeignete Indikatoren muss davor gewarnt werden, quantitative Indikatoren allein als hinreichend für die Evaluation von Forschungsleistungen zu betrachten (Barré 2001). Der verständliche Wunsch nach einem standardisiert anwendbaren "Indikatoren-Werkzeugkasten" ist bei der Tendenz zur Verfolgung komplexer politischer Zielsetzungen nicht erfüllbar (Airaghi et al. 1999), auch wenn immer wieder Versuche in dieser Hinsicht unternommen werden (z.B. European Commission 2002b).

Ein aus der Perspektive der Evaluationsforschung umfassendes Konzept der Evaluation forschungspolitischer Initiativen erstreckt sich über alle in *Tabelle 1* aufgeführten Fragestellungen (Arnold/Guy 1997; siehe auch Kuhlmann 2000; Kuhlmann/Holland 1995a; Callon et al. 1995; Meyer-Krahmer 1989: 50-60; Rip 1990):

Tabelle 1: Typische Evaluationsanliegen und Fragen (nach Arnold/Guy 1997)

Appropriateness:	Was it the right thing to do?
Economy:	Has it worked out cheaper than we expected?
Effectiveness:	Has it lived up to the expectations?
Efficiency:	What's the return on investment (ROI)?
Efficacy:	How does the ROI compare with expectations?
Process efficiency:	Is it working well?
Quality:	How good are the outputs?
Impact:	What has happened as a result of it?
Additionality:	What has happened above what would have happened anyway?
Displacement:	What hasn't happened which would have happened in its absence?
Process improvement:	How can we do it better?
Strategy:	What should we do next?

Auch wenn Evaluationsverfahren heute in weiten Bereichen staatlicher Forschungspolitik zunehmend Anwendung finden und Akzeptanz genießen, so lässt sich doch die "Richtigkeit" und "Objektivität" ihrer Ergebnisse und der daraus abgeleiteten Empfehlungen mit dem Hinweis auf unzureichende Evaluationsmethodik immer wieder in Zweifel ziehen. Bezweifelt wird die methodische Seriosität von Evaluationsstudien sowohl von Wissenschaftlern wie von Politikern. Schön/Rein (1994: 12/13) identifizieren die folgenden Probleme:

- Die verwendeten Evaluationskriterien vernachlässigen die Tatsache, dass die meisten Fördermaßnahmen "multiple, conflicting, and evolving purposes" verfolgen.
- Fördermaßnahmen werden häufig evaluiert, ohne dass ihr Entstehungskontext hinreichend verstanden worden ist ("black box problem").
- Evaluation erwarb vielfach den Ruf eines "Killers", wenn sie den (teilweisen) Misserfolg einer Fördermaßnahme nachwies, ohne sich mit den möglichen Ursachen zu beschäftigen.
- Häufig wird vor allem die Perspektive der politisch-administrativen Förderer (oder staatlicher Aufsichtsorgane) eingenommen, jedoch die Sicht der "Betroffenen" vernachlässigt.

Zu diesen generellen Schwierigkeiten treten im Anwendungsbereich "Forschung" spezifische Probleme (Airaghi et al. 1999; Jasanoff 1997, xiv): (1) Forschung hat vielfältige Effekte. Zu den typischen kurzfristigen wünschbaren Wirkungen geförderter Forschung gehören Erkenntnisgewinne, wissenschaftliche Veröffentlichungen, Beiträge zum Bildungsangebot, technologische Fortschritte, Umsatzsteigerungen und vergrößerte Marktanteile, Verbesserungen des Know-how sowie neuartige Kontakte. Die Effekte von Forschung gehen aber weit darüber hinaus: Forschung übt auch vielfältige Wirkungen auf Akteure aus, die nicht unmittelbar daran teilgenommen haben – solche Einflüsse sind allerdings quantitativ schwer zu messen. (2) Evaluationsergebnisse werden häufig frühzeitig für politische Entscheidungen benötigt, manchmal sogar bevor die Forschungsarbeiten abgeschlossen sind und normalerweise bevor sich die gesamte Breite möglicher sozio-ökonomischer Effekte entfaltet haben kann. (3) Viele Leistungen und Effekte lassen sich nicht eindeutig einem bestimmten Projekt oder Programm zuordnen. Sie können die Folge einer Kombination von Einflüssen sein, wozu auch die Verfahren der praktischen Umsetzung von Forschungsergebnissen zählen.

Eine weitere Herausforderung der Leistungsfähigkeit von Evaluationsverfahren bildet der anhaltende Wandel der kognitiven Strukturen moderner Wissensproduktion, der gegenwärtig unter einer Vielzahl von Aspekten diskutiert wird. Als prägnanter Beitrag ist die These eines "neuen Modus der Wissensproduktion (mode 2)" von Gibbons et al. (1994) in diesem Diskurs hervorgetreten, die ähnlich auch bei Callon (1992), Funtowicz/Ravetz (1993), Funtowicz (2001), Rip/van der Meulen (1996), Etzkowitz/Leydesdorff (2000), Rip (2002), Ziman (1995; 2001) und anderen zu finden, aber auch kritisiert worden ist (z.B. Weingart 1997). Betont wird nicht nur, dass der vormalige institutionelle Primat der Universitäten als zentrale gesellschaftliche Agenturen der Wissensproduktion pluralisiert wird, und zwar in Form neuer und alternativer wissensproduzierender Einrichtungen (Gibbons et al. 1994: 6). Argumentiert wird vor allem, dass die neue Form der Wissensproduktion durch Problemorientierung statt durch Erkenntnisinteresse geprägt ist, mithin epistemisch pluralisiert wird und zunehmend unscharfe Grenzen zwischen Grundlagen- und Anwen-

dungsforschung sowie zwischen den wissenschaftlichen Disziplinen erzeugt. Disziplinär verengte oder mechanistische, oberflächliche Evaluationsindikatoren und –verfahren greifen hier zu kurz.

Die offensichtlichen Grenzen der Leistungsfähigkeit von Evaluationsverfahren lassen sich nur kompensieren, wenn ihr Verwendungskontext bewusst gehalten und ihre Ergebnisse mit Vorsicht bewertet werden (vgl. Kuhlmann 1998: 97-106):

- Evaluatoren und Evaluationsnutzer müssen sich klar darüber werden, was sie eigentlich wissen wollen, welche Fragen in welcher Breite, Tiefe und Radikalität zu stellen sind. Dabei bleibt die analytische Reichweite von Evaluationsstudien immer beschränkt: Wenn man einen prinzipiell unbegrenzten Wirkungsraum politischer Maßnahmen annimmt, in den zugleich immer auch andere Wirkkräfte strahlen, dann bereitet die Rekonstruktion von Wirkungszusammenhängen umso mehr Probleme, desto umfassender und allgemeiner die Input- und Outputgrößen definiert werden.
- Die Evaluationsforschung im Bereich der Forschungs- und Innovationspolitik geht heute davon aus, dass "realistische" Informationen über ein Evaluationsobjekt nur durch die kombinierte Verwendung verschiedener sozialwissenschaftlicher Methoden und Indikatoren erarbeitet werden können.
- Dennoch sind der "Objektivität" von Evaluationsergebnissen enge Grenzen gesetzt, die durch eine gezielte Berücksichtigung verschiedenartiger Akteurperspektiven allerdings "überwunden" werden können.

Vorsicht und Sorgfalt sind also angebracht. Wissenschafts- und Innovationsforschung wissen heute, dass die Zusammenhänge zwischen "Investition" in Forschung und Innovation und möglichen wissenschaftlichen, wirtschaftlichen, sozialen oder ökologischen "Erträgen" komplex sind und keinesfalls als schlichtes Input/Output-Modell konstruiert werden können (z.B. Grupp 1998, 324-336). Nichtwissen hat wissenssoziologisch gesehen durchaus auch eine "Funktion". In gewissem Umfang muss die Gesellschaft "blinde Flecken" hinsichtlich der Leistungen der Forschung und ihrer Einrichtungen akzeptieren und institutionalisieren, damit sie diese und sich selbst nicht überfordert.

3 Zwecke, Typen und Aggregationsebenen evaluativer Information

Dieses Gutachten will explorativ die Erzeugung von und den Bedarf an evaluativen Informationen mit Blick auf eine mögliche Dienstleistungseinrichtung skizzieren. Im stark fragmentierten, nicht zentralistisch organisierten deutschen Forschungssystem sind Angebot und Nachfrage evaluativer Information sehr unübersichtlich. Zur besseren Strukturierung dieses "Marktes" sollen im Folgenden verschiedene Zwecke, Typen und Ebenen der Aggregation evaluativer Information unterschieden werden. Die Kapitel 4 und 5 werden anschließend die "Anbieter" und "Nachfrager" in einem solchen Markt skizzieren.

3.1 Zwecke

In der Literatur besteht Einigkeit darüber, dass die Evaluation von Forschungsleistungen mit erheblichem, derzeit eher wachsenden Zeit- und Geldaufwand verbunden und schon deshalb kein bloßer Selbstzweck ist, sondern einer legitimen Begründung bedarf. Über diese Begründung müssen sich Evaluatoren im Klaren sein und sie auch gegenüber den zu Evaluierenden transparent machen (Daniel 2001: 49, auch: Fisch/Daniel 1986: 11). Welche Zwecke können dies sein?

Qualitätssicherung: Zunächst liegt es im *Interesse der Wissenschaft selbst*, die Produktion wissenschaftlicher Arbeiten von hoher Qualität sicherzustellen und Bewertungsverfahren für Forschungsleistung durchzuführen. Um dieses operative Ziel zu erreichen, werden eingereichte Forschungsanträge seit langem vorwiegend ex ante durch ausgewiesene Fachkollegen begutachtet. Dieses Verfahren des ex ante Peer Review, das nicht die realiter erbrachte Forschungsleistung, sondern die fachwissenschaftliche Qualität der Fragestellung sowie der verwendeten Theorien und Methoden eingereicherter Projektanträge bewertet, hat sich in historischer Perspektive etabliert und auch bewährt. Er ist zu einem Mittel wissenschaftlicher Selbststeuerung geworden (vgl. Neidhardt 1988, Daniel 1993).

Nutzenoptimierung: Forschung betreibende Einrichtungen stehen einer vielgestaltigen *gesellschaftlichen Umwelt* gegenüber, die "Forschungsleistung" in unterschiedlicher Weise beurteilt. Unternehmen etwa werden sich vorzugsweise für diejenigen Forschungsleistungen interessieren, die im wirtschaftlichen Sinne verwertbar sind. Eine unter dieser Prämisse durchgeführte Evaluation wird vorzugsweise darauf abstellen, wie hoch der Anteil wirtschaftlich anschlussfähigen Wissens an der gesamten Wissensproduktion einer Forschungseinrichtung ist und wie gut es sich für wirtschaftliche Verwertungsprozesse verwenden lässt.

Effizienz- und Effektivitätskontrolle: Akteure der Forschungs- und Technologiepolitik werden fragen, *wie effizient und effektiv aus Steuergeldern finanzierte Forschungstätigkeit* durchgeführt wird. Damit werden Forschungserträge zum einen in Relation zu den eingesetzten Ressourcen wie z.B. Personal, Infrastruktur, Fördermittel (Effizienz) und zum anderen in Relation zu vereinbarten Zielvorstellungen (Effektivität) gesetzt. Auf diese Weise werden politische Akteure Forschungsförderung auf ihren Zielerreichungsgrad hin über-

prüfen wollen, um den Erfolg der Ressourcenallokation zu dokumentieren, das heißt politisches Handeln zu legitimieren.

Selbststeuerung: Neben solchen von "außen" an das Wissenschaftssystem gerichteten Evaluationszwecken ist die Beschreibung und Bewertung von Forschungsleistung aber auch für die *Forschungseinrichtungen (als Organisation) selbst* von Interesse. Je nach der *spezifischen "Mission"* einer Einrichtung (z.B. Grundlagenforschung; angewandte Forschung und Innovation; problemorientierte Forschung) hilft evaluative Information über Produktivität, Arbeitsweise, und wissenschaftliche Reputation dabei, Stärken und Schwächen im internen, interinstitutionellen, internationalen etc. Vergleich zu identifizieren und ggf. zu steuernd reagieren. In diesem Sinne wird Evaluation von "innen" her begründet und kann als "Wissenschaftsmanagement" professionalisiert werden.

Strukturverbesserung: Die Perspektive der empirischen Wissenschaftsforschung wird Forschungsleistung konzipieren als kreatives, neues Wissen, mit dem innerwissenschaftlich definierte Probleme besser als bisher behandelt bzw. gelöst werden können. Der Zweck von Evaluation wird folglich darin bestehen, nach den *Bedingungen produktiver Forschung* zu fragen, d.h. Forschungshypothesen aufzustellen und mit empirischem Material zu überprüfen, warum bestimmte Personen oder Einrichtungen produktiver sind als andere.

3.2 Typen

Um explorativ zu prüfen, in welchem Umfang, in welchem Format und über welche Kanäle Informationen zu Forschungsfinanzierung und -leistungen erzeugt und verbreitet werden, benötigt man eine überblicksartige Darstellung möglicher Informationstypen, die für die genannten Zwecke verwendet werden können. In den letzten Dekaden ist diesbezüglich eine relativ laborierte Wissenschafts- und innovationsindikatorik entwickelt worden, die als Grundlage für eine *Matrix von Informationstypen* dienen kann. Die Tabellen 2, 3 und 4 stellen ein Set an zentralen Indikatoren vor, die sich zum einen auf *Input*-Größen in das Forschungssystem beziehen, und zum anderen auf *Output*- bzw. *Wirkungsindikatoren*, welche einen Bewertungskorridor für Forschungsleistung eröffnen.³

Die Tabellen sind nicht als vollständig zu verstehen, sondern zeigen vor allem standardisierte und mit der deskriptiven sowie schließenden Statistik auswertbare Kennziffern. Eine vollständige Matrix müsste zudem auch Informationen enthalten, die nur beschränkt standardisierbar sind, wie sie z.B. bei Begehungen erhoben oder durch problemorientierte empirische Untersuchungen gewonnen werden können. Unterteilt sind die Tabellen erstens in *quantitative Maße*, worunter im weitesten Sinne abzählbare In- und Output-Einheiten verstanden werden können, und *qualitative Maße*, durch welche quantitative

³ Die Tabellen 2, 3 und 4 wurden mithilfe mehrerer einschlägiger Publikationen zur Wissenschaftsindikatorik erstellt. Vgl. insbesondere Bolsenkötter (1986), Fisch/Daniel (1986), Alewell (1988), Giese (1988), Daniel/Fisch (1988), Weingart et al. (1991), Stock (1994), Hornborstel (1997), Daniel (2001), Liefner (2001). Das in diesen Quellen vorfindbare Material ist weit reichhaltiger, als hier dargestellt werden kann. Insoweit handelt es sich um eine Auswahl.

Relative inhaltlich näher spezifiziert und auch strukturelle Informationen angezeigt werden können.

Beispiel. Eine abzählbare, *quantitative* Input-Größe stellt der für FuE investierte Ressourcenbetrag in Humankapital dar, der in einer bestimmten Zeitperiode für einen sachlichen Zweck verwendet wird. Dieser Betrag ist zwar als solcher bereits informationshaltig; allerdings steigt seine Aussagekraft, wenn er inhaltlich näher bestimmt, d.h. *qualitativ* erweitert wird, z.B. hinsichtlich seiner Herkunft (Grundmittel, Drittmittel), oder seiner Verwendung auf unterschiedlichen Hierarchie-Ebenen (Mittelbau, professorale Ebene) oder Altersstufen. Wird ein solches Wechselverhältnis von quantitativen und qualitativen Aspekten systematisch betrieben, erhält man Informationen mit *Strukturwert*, mit welchen dann nicht allein Vergleichbarkeit mit ähnlichen Einheiten hergestellt werden kann, sondern auch Defizite (z.B. zu geringe Investition in junges wissenschaftliches Humankapital) und Stärken (dynamisches Forschungsprofil) identifiziert werden.

Tabelle 2: Typen evaluativer Input-Information (Auswahl)

	quantitativ	qualitativ
Personal	Anzahl Wissenschaftler Anzahl Nicht-Wissenschaftler	Anteil der promovierten Wissenschaftler Relation Wiss./Nicht-Wiss. Relation Mittelbau/C3-C4 Stellen Altersstruktur Disziplinenportfolio
Finanzierung	Volumen Veränderung des Volumens	Herkunft (Grundfinanzierung, Drittmittel) Verwendbarkeit (Haushaltsrecht)
Infrastruktur	Quadratmeter Fläche	Alter, Baulicher Zustand
	Anzahl Personalcomputer	Leistungsfähigkeit

Tabelle 3: Typen evaluativer Output-Information (Auswahl)

	quantitativ	qualitativ
Forschungshandeln	Projekte (Drittmittel)	Laufzeit Finanzierung (DFG, Ministerien) Beitrag zur Profilbildung
	Forschungsaufenthalte	Regelmäßigkeit Finanzierung (Stipendien, Kooperationsverträge)
Wissensprodukte	Publikationen	Monographie/Zeitschriftenaufsatz referiert/nicht-referiert national/international
	Patente	Technologie (Spitzentechn./Hochtechnologie)
Qualifikation wissenschaftlichen Nachwuchses	Promotionen	Benotung Promotion Dauer Promotion

Tabelle 4: Typen evaluativer Wirkungsinformation (Auswahl)

	quantitativ	qualitativ
Sichtbarkeit (in der Wissenschaft)	Zitation	Nachhaltigkeit/Halbwertszeit der Zitationen
	Rezensionen, Repliken	in referierten/internationalen Fachzeitschriften
Reputation (in der Wissenschaft)	Berufungen	Reputation der rufenden Universität
	Forschungspreise, Stipendien	Art und Status des Preises, der Stipendien
Ausstrahlung (in die gesellschaftliche Umwelten)	Lizenzen, Patentzitate	Art der Erlösverwendung Technologiegebiet
	Gutachten, Beratungstätigkeiten	Art der auftraggebenden Einrichtung (Stiftungen, Ministerien)
	Vorträge	Art des Veranstalters (z.B. Industriekammer, Akademie)

3.3 Aggregationsebenen

Tabelle 5 dokumentiert verschiedene Ebenen der Aggregation evaluativer Informationen; dabei wird unterschieden zwischen *vertikaler* Aggregation, welche die Tiefenschärfe angibt, mit der Leistungsindikatoren in Forschungseinrichtungen erhoben werden können, und *horizontaler* Aggregation, welche die Ebenen der vertikalen Stufung übergreift. Horizontale und vertikale Ebene sind in der Evaluationspraxis häufig nicht zu trennen. Beispielsweise untersuchte der Wissenschaftsrat (1999) das Feld Energieforschung auf disziplinärer und organisationaler Ebene; die Evaluation der Fächer Wirtschaftswissenschaften und Geschichte durch die Wissenschaftliche Kommission Niedersachsen (1999, 2001) wurde auf Fakultäts- und Forschergruppen-Ebene durchgeführt.

Die Ebene des Einzelwissenschaftlers – in den Tabellen der Vollständigkeit halber aufgeführt, denn sie stellt für manche Aggregatinformationen die Basisinformation dar – wird im weiteren Gutachten nicht mehr explizit berücksichtigt.

Neben den für die universitäre Forschung wichtigen Differenzierungen zwischen Fächern und Lehrstühlen/Instituten, findet sich auch die vor allem für Forschungsinstitute relevante Unterscheidung von Abteilungen bzw. Instituten und Gesamtorganisation.

Tabelle 5: Zwecke und Aggregationsebenen der Verwendung evaluativer Information

Aggregationsebenen evaluativer Information	Evaluationszwecke				
	Qualitätssicherung	Nutzenoptimierung	Selbststeuerung	Effizienzkontrolle	Strukturverbesserung
<i>Vertikal</i>					
Einzelwissenschaftler	?	?	?	?	?
Forschungsgruppe; Lehrstuhl	?	?	?	?	?
Fakultät; Institut; Abteilung	?	?	?	?	?
Universität; Forschungsträgerorganisation	?	?	?	?	?
<i>Horizontal</i>					
Disziplin; Themenfeld	?	?	?	?	?
Programm	?	?	?	?	?
Institutionen- "Landschaft"	?	?	?	?	?

Explizit berücksichtigt ist die Forschungsprogrammebene, etwa im Sinne von Sonderforschungsbereichen oder sonstigen interdisziplinären Forschungs-Clustern sowie forschungspolitischen Förderprogrammen.

Die organisationsübergreifende strategische Analyse der nationalen, internationalen oder sektoralen Institutionen-"Landschaft" des Forschungssystems bildet die oberste Aggregationsstufe evaluativer Information.

4 Erzeuger und Angebote evaluativer Informationen

4.1 Forschungsinput

Generell ist Hornbostels (2000) Feststellung zuzustimmen, dass dauerhafte Berichtsstrukturen für Input-Informationen allenfalls für die Berichtssysteme der OECD, der nationalen und europäischen amtlichen Statistik auf hohem aggregierten Niveau vorhanden sind. Tiefere Aggregationsebenen werden in Deutschland nicht systematisch erfasst. Für die einzelnen Ebenen ergibt sich folgendes Bild.

Für die *Aggregatebenen Programm und Institutionen-"Landschaft"* lassen sich eine Vielzahl relevanter Publikationen feststellen, z.B. der Bundesbericht Forschung (BMBF 2002), der Grund- und Strukturdaten-Bericht des BMBF (BMBF 2001a), die Fachserien 11 und 14 des Statistischen Bundesamtes (StBA 2001a, 2001b), die Abschlussberichte zu den Systemevaluationen (WR 2000a, WR 2001, Internationale Kommission 1999, Evaluierungskommission 1998), und nicht zuletzt Jahresberichte von Forschungseinrichtungen und übergeordneten Berichtsstellen der großen Forschungseinrichtungen wie z.B. der MPG und FhG (MPG 2001, FhG 2001). Darüber hinaus sind datenbankgestützte Informationssysteme des BMBF sowie der DFG im Aufbau, die Informationen zu Forschungsausgaben nach ihrer Aufgabenstruktur und -entwicklung in Deutschland (DASTAT), hochschulspezifische Datenmengen (ICEbmbf) sowie Angaben zu DFG-geförderten Programmen bereithalten (GEPRIS).

Insbesondere die ministeriellen und der amtlichen Statistik zugehörenden Berichtsstrukturen stellen ab auf hochaggregierte Beschreibungen der finanziellen und personellen Ausstattung der institutionellen Teilbereiche der deutschen Forschungslandschaft (Hochschulen, außeruniversitäre Forschung, Wirtschaftsforschung), aufgeschlüsselt nach regionalen (Bundesländer), disziplinären und sektoralen Kriterien. Im Sinne von Tabelle 1 sind dabei vor allem Personal- und Finanzdaten abgedeckt, in geringerem Umfang bzw. gar nicht erscheinen Daten zur Gebäude- und sonstigen Ausstattung der Forschungseinrichtungen.

Generell erscheint bei diesen Daten problematisch, dass hochschulbezogene Statistiken spezifische Abgrenzungsprobleme bei Forschung und universitärer Lehre aufweisen, so dass nicht klar abgrenzbar ist, welche Ressourcen realiter für Forschung aufgewendet werden. Zudem ist selbst auf dieser hoch aggregierten Ebene ein nur in Ansätzen entwickeltes Struktur-Berichtssystem vorhanden. Zwar sind beim jüngsten Bericht des Statistischen Bundesamtes relativ ausführlich wichtige Strukturkennziffern zur Ausstattung der Hochschulen, beispielsweise Grundmittel je Studierender, Drittmittel je Professorenstelle oder Ausgaben der Hochschulen nach Fächergruppen dokumentiert (StBA 2001b). Allerdings erfährt man nichts über den Anteil promovierter Wissenschaftler in Hochschulen, an außeruniversitären Einrichtungen oder Forschungseinrichtungen der Wirtschaft, ebenso wenig über deren Alterstruktur insgesamt oder aufgeschlüsselt nach Sektionen⁴ bzw. Fachdisziplinen. Die bisherige Aufschlüsselung der Personalstruktur nach Besoldungs- und

⁴ Die MPG organisiert ihre Institute in eine chemisch-physikalisch-technische, biologisch-medizinische und geisteswissenschaftliche Sektion; die FhG bindet einen Teil ihrer Institute in thematische Forschungscluster ein. Ähnliches gilt für WGL und die HGF.

Statusgruppen trägt die Handschrift einer für ministerielle Bedarfe zugeschnittenen Datenerhebungspraxis.

Für die *Aggregationsebene Fachdisziplinen und Forschungsfelder* (auf organisationaler und departmentaler Ebene) liegen die Querschnittsbegutachtungen des Wissenschaftsrates in den Feldern Umwelt-, Material- und Energieforschung vor (WR 1994, 1996, 1999).⁵ In diesen Studien wurden in Form von Selbstevaluationen (schriftliche Befragung der Einrichtungen) und Fremdevaluationen (Begehungen) Informationen zur personellen und materiellen Ausstattung der untersuchten Forschungseinrichtungen – analog Tabelle 2 – erhoben. Weiterhin weist die DFG die in den 1990er Jahren bewilligten Mittelflüsse an Einrichtungen der universitären und außeruniversitären Forschung aus, gegliedert nach 16 Fachgebieten. Deutlich wird dadurch, an welchen Orten DFG-finanzierte Forschung fachlich ihre Schwerpunkte findet - sowohl mit Blick auf einzelne Einrichtungen, als auch mit Blick auf Förderregionen, die in der Dokumentation kartographisch dargestellt sind (vgl. DFG 1997, 2000, 2002b). Weiterhin hat Niedersachsen seit März 1999 mit der systematischen Evaluation der universitären Fachdisziplinen begonnen, die auf Fakultäts- bzw. Fachbereichsebene evaluative Input-Informationen erheben. Diese von der Wissenschaftlichen Kommission Niedersachsen durchgeführten Evaluationen liegen für die Fächer Chemie, Geschichte, Biologie, Bauingenieurwesen/Architektur, Elektrotechnik, Maschinenbau, Wirtschaftswissenschaften, Informatik und Berufswissenschaften bereits vor. Niedersachsen ist bislang das einzige Bundesland, das Input-Daten auf dieser tiefen Aggregationsebene vorhält (dabei bisher aber nicht öffentlich zugänglich macht); bemerkenswerte Ansätze in dieser Richtung werden mittlerweile auch in Baden-Württemberg und Bayern verfolgt. Eine bundesweite Übersicht der universitären Fachdisziplinen ist derzeit nicht verfügbar.⁶

Einen erster Schritt in dieser Richtung ist in der kürzlich erschienen Studie *Forschungslandkarte Hochschulmedizin* zu sehen, die vom Fraunhofer-Institut für Systemtechnik und Innovationsforschung im Auftrag des BMBF und des Medizinischen Fakultätentages durchgeführt wurde (BMBF/Medizinischer Fakultätentag 2002). Neben dem Schwerpunkt der Untersuchung – einem Überblick über die externen und internen Forschungsschwerpunkte der Medizinischen Fakultäten in Deutschland – wurden auch zahlreiche Daten zum Forschungsinput, z.B. Anzahl der Professuren, wissenschaftliche Vollzeitkräfte auf Planstellen, Höhe des Landeszuschusses für Forschung und Lehre erhoben. Damit bietet die Forschungslandkarte Hochschulmedizin einen umfassenden Überblick über *alle medizinischen Fakultäten Deutschlands*, mithin ein Tableau qualitativer und struktureller Informationstypen, die zur Bewertung von Forschungsleistungen relevant erscheinen.⁷

Für die *Aggregationsebene von Universitäten und ihren Untereinheiten* werden in der Regel von den Wissenschaftsministerien der Bundesländer erfasst. Für Baden-Württemberg stellt der FuE-Monitor des dortigen Statistischen Landesamtes Finanzdaten für jede baden-württembergische Universität zusammen (StLABW 2001). So sind die FuE-Ausgaben nach Herkunft (z.B. Grund- und Drittmittel) und Fachgebietsclustern ausgewiesen sowie auf die

⁵ Vgl. dazu die publizierten Ergebnisse des Wissenschaftsrates (1994, 1996, 1999).

⁶ Diese Berichte sowie Hinweise zur Rechtsgrundlage und Methodik der Evaluation von Forschungsleistungen an niedersächsischen Hochschulen sind auf folgender Website abrufbar:
<http://www.wk.niedersachsen.de/Materialien/material.htm>

⁷ Analyseebene sind die Medizinischen Fakultäten als Ganzes, nicht einzelne Abteilungen oder Kliniken.

Anzahl der Studierenden relationiert. Allerdings fehlen weitgehend Strukturmaße, die Profilunterschiede zwischen einzelnen Universitäten ausweisen würden. Ähnlich liegt es mit dem Bericht des Wissenschaftsrates zur Drittmittelsituation der Hochschulen (WR 2000b). Hier werden relativ ausführlich die eingeworbenen Drittmittel strukturell untersucht, etwa auf Lehrstuhl- oder An-Instituts-Ebene bzw. im regionalen Vergleich. In diesem Bericht wird zu Recht kritisiert, dass die für solche Analysen erforderlichen Strukturinformationen bisher nur schwer oder gar nicht erhältlich sind (vgl. WR 2000b: 79-95). Wiederum zu nennen ist das DFG-Ranking, das die bewilligten Forschungsmittel nicht nur nach Regionen und Fachgebieten, sondern auch nach Hochschulen differenziert ausweist (vgl. DFG 1997).⁸

Auf der Aggregationsebene einzelner Forschungsträgerorganisationen und ihrer Einzelinstitute bzw. Forschungszentren ist die Datenlage vergleichsweise heterogen. Relativ gut beschrieben sind die außeruniversitären Einzelinstitute, die einerseits regelmäßige, zumeist jährliche Selbstbeschreibungen in Form von Jahresberichten anfertigen, und von welchen andererseits im Rahmen der Systemevaluationen Fremdbeschreibungen angefertigt worden sind. Im Einzelnen ergibt sich folgendes Bild.

*Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz*⁹: Im Rahmen der Systemevaluation der WGL sind alle Einzelinstitute hinsichtlich ihrer *Personal- und Finanzausstattung* beschrieben worden. Die Angaben sind quantitativ gut, gehen jedoch nicht über eine zaghafte qualitative Beschreibung hinaus. So werden zwar Drittmittelquoten der Gesamtfinanzierung und das disziplinäre Portfolio ausgewiesen. Es fehlen jedoch relevante Strukturmaße wie Alterstruktur des Wissenschaftspersonals oder *Infrastrukturmaße*. Diese Situation spiegelt sich ebenso wider bei einer konkreten Recherche der Selbstbeschreibungen von fünf Instituten. Über Jahresberichte bzw. Internetpräsentationen ist im Regelfall sogar noch weniger Informationsmaterial zu beschaffen. Auffällig ist diesbezüglich die große Heterogenität, mit welcher einzelne Institute Input-Größen zur Verfügung stellen.

Allerdings deutet sich an, dass in Zukunft ausführlichere Input-Informationen zur Verfügung stehen werden. Der Senat der WGL hat am 11. März 2002 Grundsätze zu Aufgaben und Verfahren der künftigen Evaluation der WGL-Institute verabschiedet, welche zum einen eine interne Evaluation durch Beiräte – so wie seit langem bei der MPG und der FhG verwirklicht, und zum anderen eine externe Evaluation durch den Senatsausschuss Evaluation vorsieht.¹⁰

*Helmholtz-Gemeinschaft deutscher Forschungszentren*¹¹: Im Rahmen der Systemevaluation der HGF sind - wie bei der WGL - alle Zentren hinsichtlich ihrer *Personal- und Finanzausstattung* beschrieben worden. Die Angaben sind quantitativ und qualitativ deutlich präziser und zudem umfangreicher als bei der WGL. So werden insbesondere Strukturma-

⁸ Die DFG weist in der zweiten Ausgabe des Ranking auch die Mittelflüsse an die außeruniversitären Einrichtungen aus (vgl. DFG 2000).

⁹ Grundlage der folgenden Ausführungen sind Dokumentrecherchen der Jahresberichte von fünf WGL-Instituten sowie der Abschlussbericht des Wissenschaftsrates zur Systemevaluation der WGL (WR 2000).

¹⁰ Die WGL hat den Autoren des Gutachtens diese internen Dokumente freundlicherweise zur Verfügung gestellt.

¹¹ Grundlage der folgenden Ausführungen sind Dokumentrecherchen der Jahresberichte von fünf HGF-Zentren sowie der Abschlussbericht des Wissenschaftsrates zur Systemevaluation der HGF (WR 2001).

Be systematisch erhoben; Finanz- und Personaleinsatz werden zugleich zentren- und themenbezogen aufgeschlüsselt; die Personalstruktur wird ausführlich differenziert (Doktoranden, Diplomanden, Gastwissenschaftler, Frauenanteil) sowie deren Entwicklungsdynamik in einem Fünfjahresfenster dargestellt, insbesondere mit Blick auf Berufungs- und Qualifikationsentwicklungen. Insgesamt ist das Informationsangebot über HGF-Zentren somit deutlich besser als bei den anderen Forschungseinrichtungen.

Zu beachten ist allerdings, dass ein "Zentrum" der HGF nicht vergleichbar ist mit einzelnen "Instituten" der WGL, FhG oder MPG: HGF-Zentren sind selbst Institutsverbünde und z.T. bis zu 30mal größer als einzelne Institute der anderen Forschungsorganisationen. Deshalb wurde für fünf Zentren die Verfügbarkeit von Input-Daten auf Institutsebene im Internet recherchiert mit dem Ergebnis, dass hier die Informationslage ähnlich verbesserungsfähig erscheint wie bei der WGL. Zwar stehen im Regelfall Personaldaten in quantitativer Form zur Verfügung; allerdings sind Finanz-Größen sowie Infrastrukturdaten nur schwer zugänglich zu machen und wenn, dann nur in heterogener Qualität.

*Fraunhofer-Gesellschaft*¹²: Im Rahmen der Systemevaluation der FhG wurden von der Evaluierungskommission keine Informationen hinsichtlich der institutsbezogenen Personal- und Finanzausstattung im Abschlussbericht dokumentiert. Allerdings enthalten die von der FhG-Zentrale publizierten Gesamtjahres- und Finanzberichte nach Instituten aufgeschlüsselte *Finanzdaten*. Dieses Bild spiegelt sich auch in den Jahresberichten der einzelnen Institute wider, die in vergleichsweise homogener Form Finanzdeckdaten bereithalten. *Personaldaten* werden in den Jahresberichten der gesamten FhG nur überblickshaft und nicht institutsbezogen ausgewiesen. Solche Informationen sind dagegen in den einzelnen Institutsdokumentationen in quantitativer Hinsicht gut erhältlich. Ähnliches gilt auch für *Infrastrukturinformationen*. Insgesamt fällt auf, dass weiterführende qualitative und strukturelle Dimensionen fehlen. Am vergleichsweise besten dokumentiert sind diesbezüglich finanzielle Inputgrößen.

*Max-Planck-Gesellschaft*¹³: Im Rahmen der Systemevaluation der MPG wurden – ebenso wie bei der FhG - keine institutsbezogenen Input-Daten schriftlich zugänglich gemacht. Zwar ist davon auszugehen, dass solche Daten für den Evaluationszweck erhoben wurden; über deren Umfang und Qualität kann jedoch hier nichts gesagt werden. Auch sonst veröffentlicht die MPG keine Gesamtberichte, welche Input-Größen institutsbezogen ausweisen würden. Die allgemeinen Informationen über *Personal- und Finanzgrößen* der MPG, die sich aus den institutsübergreifenden Jahresberichten ableiten lassen, werden durch die Jahresberichte der Einzelinstitute in quantitativer Hinsicht nur bedingt disaggregiert. Allerdings finden sich – im Unterschied zur FhG – qualitative Informationen, z.B. der Anteil promovierter Wissenschaftler oder nach Disziplinen aufgeschlüsselte Beschäftigtenkategorien (Wissenschaftler, Nicht-Wissenschaftler, Studentische Hilfskräfte usw.). *Infrastrukturinformationen* sind dagegen nur in aggregierter Form erhältlich und werden nur zum Teil in den Jahresberichten der Einzelinstitute ausgewiesen. Der Dokumentation

¹² Grundlage der folgenden Ausführungen sind Dokumentrecherchen der Jahresberichte von fünf FhG-Instituten sowie der Abschlussbericht des Wissenschaftsrates zur Systemevaluation der FhG (Evaluierungskommission 1998).

¹³ Grundlage der folgenden Ausführungen sind Dokumentrecherchen der Jahresberichte von fünf MPG-Instituten sowie der Abschlussbericht der Internationalen Kommission zur Systemevaluation der DFG und der MPG (Internationale Kommission 1999).

weiterführender qualitativer und struktureller Datenbestände wird kein vorrangiges Gewicht zugemessen. Alles in allem ist die informative Situation mit der der FhG vergleichbar.

4.2 Forschungsleistungen (Output und Wirkung)

Neben evaluativen Informationen zur Ressourcenlage im Bereich der Forschung sind für eine Bewertung Struktur bildender Effekte vor allem Informationen zu Forschungsleistungen interessant. Tabelle 3 weist zum einen Forschungsoutputs im weitesten Sinne aus.¹⁴ Darunter sind zu verstehen *Output-Indikatoren des Forschungshandelns*, beispielsweise angezeigt durch Projekte; *Indikatoren der Wissensproduktion*, beispielsweise Publikationen und Patente; sowie *zur Qualifikation wissenschaftlichen Nachwuchses*, angezeigt etwa durch Promotionen. Tabelle 4 weist zum anderen Forschungswirkungen aus, d.h. mittelbar dem Forschungshandeln zurechenbare Resultate. Hierzu zählen die *Sichtbarkeit des generierten Wissens*, angezeigt durch Zitationen von Publikationen oder Lizenznahmen von Patenten; *Reputationswirkungen im Forschungssystem*, die sich durch Forschungspreise oder Rufe äußert; sowie die *Ausstrahlung in die gesellschaftliche Umwelt*, etwa in Form von Gutachten oder Vorträgen.

Beide, Output- wie Wirkungsindikatoren, lassen sich jeweils qualitativ und strukturell näher bestimmen. Dazu gibt es in der Forschungsliteratur eine intensive Diskussion über sinnvolle Operationalisierungen solcher weiterführenden Informationstypen (z.B. Buchholz 1995, Quoniam et al. 1995). Nicht zuletzt die Erfahrungen des Higher Education Funding Council for England bei der Evaluation der britischen Universitäten (Research Assessment Exercises) haben dazu geführt, Indikatoren nicht "theorielos" zu verwenden, sondern an im vorhinein zu bestimmende Ziele bzw. Zielkorridore zu binden, und in einem kontinuierlichen Abstimmungsprozess mit den zu evaluierenden Akteuren zu halten, um letztlich zu sinnvollen Aussagen zu gelangen und Steuerungspotenziale zu ermöglichen.

Wir haben wiederum *explorativ* geprüft, welche der exemplarisch aufgeführten quantitativ-qualitativ-strukturellen Informationstypen gegenwärtig in Deutschland auf unterschiedlichen Aggregationsebenen verfügbar sind. Dabei ergibt sich folgendes Bild.

Für die *Aggregatebene Institutionen-"Landschaft"* ist die Situation bei Leistungsindikatoren derjenigen der Input-Informationen vergleichbar. In einer Vielzahl der in Abschnitt 4.1 erwähnten Publikationen – und darüber hinaus im Bericht zur Leistungsfähigkeit Deutschlands (BMBF 2001b) und den Patentstatistiken des Deutschen Patent- und Markenamts (DPMA 2001) – werden neben Input-Parametern auch Output- und Wirkungs-Größen im Sinne der Tabellen 3 und 4 dokumentiert. Diese Indikatoren beschreiben in der Regel die institutionellen Teilbereiche der deutschen Forschungslandschaft (Hochschulen, außeruniversitäre Forschung, Wirtschaftsforschung) als Ganzes und schlüsseln sie auf nach regionalen, disziplinären und sektoralen Kriterien. So werden im Bundesbericht Forschung (BMBF 2002), dem Grund- und Strukturdaten-Bericht (BMBF 2001) und der Patentstatistik des DPMA (2001) durchaus detaillierte Hinweise zur *Produktivität der Wissensgenese* und

¹⁴ Hornbostel (2000: 11ff) unterscheidet Output- bzw. Wirkungsindikatoren nach ihrer Reputation, Quantität, bibliometrischen Sichtbarkeit, Verwertbarkeit und dem Vermögen, Drittmittel zu attrahieren. Die hier vorgestellte Systematik ist komplementär zu diesen Vorschlägen zu verstehen. Sie deckt vor allem relativ standardisierbare und damit einheitlich erhebbare Indikatoren ab.

der *Produktion von Hochschulabsolventen und Promotionen* gegeben. Weiterführende qualitative oder strukturelle Informationen – so wie sie etwa im US-amerikanischen *Report on Science and Engineering Indicators* zu finden sind (NSB 2000: Kap. 4, 6) – stehen jedoch kaum zur Verfügung. Zu solchen Informationstypen gehören etwa das Verhältnis nationaler/internationaler Publikationen und Patente in unterschiedlichen Technologiesektoren und differenziert nach institutionellen Bereichen (Universitäten, Forschungseinrichtungen etc.). Aus diesen Informationen werden im Science and Engineering Report unterschiedliche Produktivitätsdimensionen (kognitiv, regional) kartiert und übersichtlich dargestellt. Eine solche Kartierung fehlt in Deutschland bisher weitgehend.

Weiterhin fällt auf, dass das Gewicht der Dokumentation eher auf Output-Größen liegt und weniger auf Wirkungsindikatoren. Es gibt unterschiedliche Gründe, warum dies der Fall ist. Zu vermuten ist, dass Outputs mit vergleichsweise weniger Mühe erhebbar und auswertbar sind als Wirkungs-Größen, weil die sachlichen und zeitlichen Zurechenbarkeiten überschaubarer sind. Damit ist zum einen gemeint, dass der materielle Gehalt von "Wirkung" unschärfer als "Output" ist und seine umfassenden Operationalisierung und Messung mit hohem Aufwand verbunden ist, nicht zuletzt wegen ungeklärter kausalsachlicher Zusammenhänge. Zum anderen sind eingetretene "Wirkungen" auch zeitlich nur mittelbar auf geleistete Forschungsinvestitionen zurechenbar (vgl. Airaghi et al. 1999).

Relativ gut dokumentiert ist dagegen die Patentproduktivität der beiden großen Forschungsorganisationen FhG und MPG. Die Patentstelle der Fraunhofer Gesellschaft dokumentiert in ihrem Jahresbericht relativ detailliert Erfindungsmeldungen, Patenterteilungen oder Markenmeldungen nach Technologiegebieten und institutioneller Herkunft (Universitäten, Privatpersonen etc.), und schlüsselt darüber hinaus auch die Entwicklung der Verwertungseinnahmen auf (FhG-Patentstelle 2001). Ebenso dokumentiert die Verwertungsgesellschaft der Max-Planck-Gesellschaft, Garching Innovation GmbH, seit den 1980er Jahren - wenngleich weniger detailliert - die Zahl der Erfindungen, Patente und Lizenzverträge der gesamten MPG.

Für die *Aggregationsebene einzelner Forschungsträgerorganisationen und ihrer Einzelinstitute bzw. Forschungszentren*, welche einzelne Forschungsorganisationen im Sinne von Universitäten oder Einzelinstituten der außeruniversitären Forschungsverbände ausweist, stellt sich die Datenlage sehr heterogen dar. Wiederum - wie bereits bei den Input-Größen festgestellt - sind die außeruniversitären Einzelinstitute umfangreicher beschrieben als Universitäten und zwar einerseits in Form von Jahresberichten, und andererseits im Rahmen der durch die Systemevaluationen angefertigten Fremdbeschreibungen.

Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz: Im Rahmen der Systemevaluation der WGL wurden weder die Gesamtebene, noch die Einzelinstitute hinsichtlich ihrer Forschungsausgaben und -wirkungen beschrieben, zumindest wurden solche evaluativen Informationen keiner breiteren Öffentlichkeit zugänglich gemacht. Diese Situation spiegelt sich auch bei einer konkreten Recherche der Selbstbeschreibungen von fünf Instituten wider (vgl. Fußnote 9). Zwar werden bei einigen Instituten *Publikationen und Projekte* relativ umfangreich ausgewiesen. Allerdings ist die Heterogenität zwischen diesen so hoch, dass von einem einheitlichen Bild nicht die Rede sein kann. Insgesamt ergibt sich derselbe Eindruck, den eine neue Studie zur Systemevaluation der WGL folgendermaßen beschreibt: "Wir haben bei der WZB-Recherche den Eindruck gewonnen, (...) dass in den meisten [WGL]-Instituten dauerhafte Erfahrungen mit Selbstevaluationen, in denen Leistungsmaßstäbe definiert und darauf bezogene Leistungskriterien experimentiert werden,

kaum vorliegen, also auch nicht diskutiert und rasoniert werden. Daraus ergibt sich dann, dass das Wissen dieser Institute über sich selber nicht gut genug ist, um in Fremdevaluierungen für einen selbstbewussten Umgang mit externen Experten und auch zu deren Kontrolle zur Verfügung zu stehen" (Neidhardt 2001: 9).

Helmholtz-Gemeinschaft deutscher Forschungszentren: Im Rahmen der Systemevaluation der HGF wurden die Zentren - im Vergleich zu den Instituten der WGL - relativ ausführlich beschrieben (WR 2001). Der Abschlussbericht hält ein Tableau von "Leistungsdaten" bereit, dass vor allem nicht-wissenschaftliche Leistungsbezüge thematisiert. So werden etwa Patent- und Lizenzdaten, Wirtschaftskooperationen und Ausgründungen nicht nur in quantitativer, sondern auch in struktureller Hinsicht für ein Dreijahreszeitfenster und zentrenbezogen dargestellt. Auffällig ist, dass wissenschaftliche Leistungen nicht zentrenbezogen ausgewiesen werden. Zwar dokumentiert die HGF als Ganzes die Kooperationsanteile ihrer Publikationen nach Internationalität und institutioneller Herkunft, *jedoch nicht zentren- oder einzelinstitutsbezogen*. Dass solche Informationen auch innerhalb der einzelnen Institute bisher offensichtlich nicht gesammelt und aufbereitet werden, wurde bei einer Internet-Recherche bei fünf HGF-Zentren deutlich. Hier konnten auf Einzelinstituts- oder Forschergruppenebene keine outputorientierten Forschungsindikatoren zugänglich gemacht werden.

Fraunhofer-Gesellschaft: Im Rahmen der Systemevaluation der FhG wurden von der Evaluierungskommission keine Informationen hinsichtlich der institutsbezogenen Forschungsleistungen und -wirkungen dokumentiert. Allerdings enthalten die von der FhG-Zentrale publizierten Gesamtjahres- und Finanzberichte nach Instituten aufgeschlüsselte Output- und Wirkungsdaten. Relativ ausführlich sind Informationen zur *Produktivität des Forschungshandelns und der Wissensprodukte* vorhanden, und zwar in *quantitativer und qualitativer* Hinsicht. So werden Projekte auf Institutsebene sehr gut aufgeschlüsselt nach Laufzeit und Finanzierungsgeber, ebenso wie Patentanmeldungen, erteilte Patente und Lizenzeinnahmen nach Technologiesektoren dargestellt werden. Weniger umfangreich sind Daten zu Reputationsaspekten und Aspekte der Sichtbarkeit des generierten Wissens dargestellt. Insgesamt ist auffällig, dass Strukturdaten zu den Forschungswirkungen weniger häufig Gegenstand der Selbstbeschreibungen sind als bei den Forschungsoutputs.

Max-Planck-Gesellschaft: Im Rahmen der Systemevaluation der MPG sind - ebenso wie bei der FhG - keine institutsbezogenen Output- und Wirkungsdaten schriftlich zugänglich gemacht worden. Auch sonst veröffentlicht die MPG keine Gesamtberichte, welche Output- und Wirkungsdaten institutsbezogenen ausweisen würden. Wiederum - wie bereits bei den Input-Größen festgestellt - werden entsprechende Informationen jedoch durch die Jahresberichte der Einzelinstitute vergleichsweise gut abgebildet. So finden sich in den Jahresberichten der fünf analysierten Institute *ausführliche Informationen für alle drei Forschungs-Outputdimensionen* aus Tabelle 3. Zwar sind *Strukturindikatoren nicht so häufig nachweisbar wie quantitative Maße*. Allerdings ergibt sich der Eindruck, dass die Institute der MPG durchaus über ihre outputseitige Produktivität aussagefähige Selbstbeschreibungen anfertigen. Dies gilt eingeschränkt auch für die Wirkungsseite. Insgesamt sind Einzelinstitute somit die geeigneten Adressaten einer zukünftigen Gemeinschaftsstelle zur systematischen Erhebung geeigneter evaluativer Output- und Wirkungsinformationen.

Universitäten. Forschungsleistungen werden von einzelnen Universitäten in sehr heterogener Weise dokumentiert. Einen ersten Ansatzpunkt bietet das bereits erwähnte DFG-Ranking, welches die Bewilligungsempfänger der DFG nach Hochschulen aufschlüsselt

(DFG 1997, 2000). Neben dem absoluten Gesamtvolumen wird das über die Zahl der an einer Hochschule tätigen Professoren und Hochschullehrer relativierte Volumen ausgewiesen. Interpretiert man eine DFG-Bewilligung nicht als bloßen Mittelzufluss und damit als "Input" im Sinne von Kapitel 4.1, sondern als durch den Begutachtungsprozess zertifizierte Qualität geplanter Forschungsarbeit, dann stellt die DFG – zumindest in Näherung – eine Leistungskennziffer für allgemeine Forschungsleistung zur Verfügung (vgl. DFG 2002b).¹⁵

Generell gilt jedoch, dass eine für Universitäten übergreifende Dokumentation von Output- und Wirkungsindikatoren weitgehend fehlt. Zwar ist im Hochschulrahmengesetz (HRG) seit 1998 die Qualitätskontrolle von Forschung (und Lehre) rechtlich festgeschrieben worden. Allerdings gibt es bei der gesetzlichen und praxisbezogenen Anwendung kein einheitliches Vorgehen in den Bundesländern.

Als derzeitiger "Vorreiter" bei der Implementierung einer systematischen Evaluation von Forschungsleistungen kann Niedersachsen gelten. Hier wurde mit der *Wissenschaftlichen Kommission* eine zentrale, für Forschungsevaluation verantwortliche Institution geschaffen.¹⁶ Diese Kommission evaluiert nach definierten Zielvorstellungen und mithilfe einer methodischen Kombination von zum einen Selbst- und Fremdevaluation, und zum anderen Wissenschaftsindikatorik und "informed peer review" die forschungsbezogene Leistungsfähigkeit niedersächsischer Fachdisziplinen auf *Fakultäts- bzw. Fachbereichsebene und auf der Ebene von Forschergruppen*. Die auf diese Weise generierten evaluativen Informationen sollen dabei künftig auch für Zuwendungsentscheidungen verwendet werden. Die Arbeit der Kommission hat seit 1999 bereits für die Fächer Chemie, Geschichte, Biologie, Bauingenieurwesen/Architektur, Elektrotechnik, Maschinenbau, Wirtschaftswissenschaften, Informatik und Berufswissenschaften Evaluationsberichte zur Lage der Forschung veröffentlicht. Weitere Berichte für Physik, Rechtswissenschaften und Umweltwissenschaften kurz vor ihrer Veröffentlichung stehen.¹⁷

Bei der Durchsicht der Abschlussberichte wird deutlich, dass sich die verwendeten Indikatoren am Forschungsstand der empirischen Wissenschaftsforschung orientieren und gleichzeitig durch die fach- und situationsbezogenen Einschränkungen gutachterlich ausgewogene Urteile gefällt werden. Tabelle 6 fasst zentrale Aspekte zu den Zielen, den Evaluationskriterien- und Verfahren zusammen. Dies niedersächsische Konzept der systematischen Erfassung evaluativer Information zu Forschungsleistungen in einem Berichtssystem zeigt, dass "Bausteine" eines forschungsstrukturbezogenen Informations- und Dienstleistungsangebots des deutschen Forschungssystems bereits existieren – doch nur dann, wenn solche Informationen auch für entsprechende Aggregationszwecke freigegeben werden.

¹⁵ Für eine Diskussion zum In- und Outputcharakter von Drittmitteln vgl. Hornbostel (1997: 211ff).

¹⁶ Baden-Württemberg und Bayern sind derzeit dabei, ein vergleichbares Evaluationssystem einzurichten.

¹⁷ Diese Berichte stehen unter <http://www.wk.niedersachsen.de/Materialien/material.htm> zum Herunterladen bereit.

Tabelle 6: Evaluation von Forschungsleistungen an niedersächsischen Hochschulen

	Niedersachsen
Beginn der Forschungsevaluation	März 1999
Durchführende Einrichtung	Wissenschaftliche Kommission als Teil einer zentralen Evaluationsagentur
Aggregationsebenen der Evaluation	institutionenübergreifende Evaluation von Fachdisziplinen (disaggregiert nach Fachbereich, Arbeitsgruppe, Lehrstuhl)
Ziele	<ul style="list-style-type: none"> - Ermittlung von Forschungsprofilen (landes- und bundesweit sowie international) und deren gezielter Entwicklung - Ermittlung von Maßstäben zur eigenverantwortlichen Planung zur Durchführung von Maßnahmen der Qualitätssicherung (auch Mittelvergabe betreffend) - Schaffung einer validen Informationsgrundlage für unterschiedliche Stakeholdergruppen (Politik, interessierte Öffentlichkeit)
Evaluationskriterien	<ul style="list-style-type: none"> - Qualität und Relevanz (Innovativität, wiss. Ausstrahlung, Interdisziplinarität, wiss. Kooperationsintensität national und international, Effektivität Nachwuchsförderung, Kontaktstrukturen zur Wirtschaft) - Effektivität und Effizienz (Verhältnis eingesetzter Mittel und erreichter, beabsichtigter Ziele) - Strukturpolitische Aspekte
Methoden	informed peer-review; orientiert an anderen nationalen und internationalen Evaluationsverfahren
Verfahren	<ul style="list-style-type: none"> - Selbstbeschreibung der zu evaluierenden Einrichtung anhand eines abgestimmten Datenrasters - Fremdbeschreibung durch Gutachtergremium (eingesetzt von Wiss. Kommission) - Stellungnahme der Hochschule zum schriftlichen Gutachterbericht - Abschließende Ergebnispräsentation und Empfehlungen durch die Wiss. Kommission

Quelle: Wissenschaftliche Kommission Niedersachsen (2002a, 2002b)

4.3 Anbieter evaluativer Informationen

Die Abschnitte 4.1 und 4.2 haben gezeigt, dass evaluative Informationen von einer Vielzahl von Anbietern erzeugt werden, jedoch in unterschiedlichen Formaten und nur begrenzt öffentlich zugänglich. Zum einen kann man die Anbieter danach unterscheiden, ob sie Informationen über sich selbst (Selbstbeschreibung) oder über andere erstellen (Fremdbeschreibung). Die Jahresberichte der Institute der Forschungsorganisationen MPG und FhG sind entsprechend regelmäßig verfasste Selbstbeschreibungen. Hier werden evaluative Informationen z.T. zur Eigeninformation erhoben (strategische Forschungsplanung), z.T. aufgrund bestehender Rechenschaftspflichten als Zuwendungsempfänger von Bund und Ländern, vor allem im Bereich der Input-Daten. Die Systemevaluationen der Forschungssäulen durch Evaluationskommissionen und die genannten Einzelstudien sind dagegen

Fremdbeschreibungen, bei welchen externe Akteure evaluative Informationen in Kooperation mit den und über die zu evaluierenden Einheiten (vgl. Aggregationsebene) erheben.

Eine weitere Unterscheidung betrifft den institutionellen Ort, an dem aus evaluative Informationen erzeugt werden. Wie Daniel (2001) und Hornbostel (2000) zeigen, sind in Westeuropa und den USA entweder eigenständige Körperschaften des öffentlichen Rechts mit Berichtspflichten an parlamentarische Gremien oder Parlamente entstanden, die außerhalb des Wissenschafts- und Forschungssystems stehen. Beispiele dafür sind die Research Assessment Exercises des Higher Education Funding Council of England (HEFCE) im Vereinigten Königreich oder das Centre National de Evaluation (CNE) in Frankreich. Neben dieser unabhängigen Körperschaften gibt es auch solche, die in das Forschungssystem selbst implementiert worden sind, beispielsweise die Untereinheit der National Science Foundation, die den Report on Science and Engineering Indicators herausgibt (Science Resources Statistics, SRS) oder das Observatoire des Sciences et des Techniques (OST) in Frankreich oder die Assessment-Panels der Vereniging van Universiteiten (VSNU) in den Niederlanden.¹⁸

Da Forschungsevaluation in Deutschland bisher keine gemeinsamen Formate verwendet (und erst recht keinen gemeinsamen institutionellen Ort hat), *koexistiert eine Vielzahl heteronomer evaluativer Selbst- und Fremdbeschreibungen*, die mit unterschiedlichen Absichten erstellt werden. Im Folgenden werden in kurzer Form die relevanten Anbieter evaluativer Informationen skizziert. Dabei entstehen zwangsläufig Redundanzen zu den Abschnitten 4.1 und 4.2.

Zentrale Verwaltungen der großen außeruniversitären Forschungsträgereinrichtungen (MPG, HFG, WGL, FhG). In den Zentralverwaltungen der außeruniversitären "Forschungssäulen" Deutschlands laufen traditionell vielfältige Informationen, auch solche zu den Forschungsleistungen der Einzelinstitute zusammen. Dies geschieht jedoch in unterschiedlicher Weise. Die Autoren sind der Auffassung, dass der Umfang der erzeugten Informationen derzeit bei der MPG und der FhG noch höher ist und eine bessere Qualität aufweist als bei den übrigen Einrichtungen. Dies liegt nicht zuletzt daran, dass MPG und FhG relativ klare "Missionen" verfolgen, die in regelmäßiger Weise durch Selbstbeschreibungen und interne Steuerungskommissionen (Beiräte) validiert werden. Ein solcher Zustand wird nach den Systemevaluationen auch bei der HFG und der WGL angestrebt. Derartige Berichtssysteme sind derzeit im Aufbau und noch wenig erprobt. Für die Zukunft steht zu vermuten, dass alle vier Zentralverwaltungen wichtige Informationslieferanten sein werden, insbesondere im Bereich tiefer als bisher aggregierter Daten.

Forschungsförderungseinrichtung DFG. Die Deutsche Forschungsgemeinschaft als zentrale Selbstverwaltungseinrichtung der deutschen Wissenschaft zur Förderung der Forschung an Hochschulen und öffentlich finanzierten Forschungsinstituten stellt mit ihren Informationssystemen GEPRIS und DFG-Ranking zwei Informationstools zur Verfügung, die in den letzten Jahren systematisch ausgebaut wurden und Strukturinformationen zur Forschungslandschaft bereithalten. Durch die im DFG-Ranking in Relation zu Wissenschaftlerstellen (wissenschaftliche Mitarbeiter, Professoren), Fachgebieten und Regionen dargestellten Volumina können die Bewilligungsempfänger mehrdimensional kartiert werden. Dieses

¹⁸ Weitere Information bei den entsprechenden websites:

HEFCE: <http://www.hefce.ac.uk/Research/assessment/>; CNE: <http://www.plan.gouv.fr/organismes/cne/>

SRS: <http://www.nsf.gov/sbe/srs/about.htm>; OST: <http://www.obs-ost.fr/>

VSNU: <http://www.vsnu.nl/show?id=7537&langid=247>.

Informationsangebot leistet damit einen grundlegenden Beitrag zur Diskussion um die Frage nach den "Centers of Excellence" der Forschung in Deutschland (vgl. DFG 2002b).

Der *Stifterverband für die deutsche Wissenschaft* unterstützt die Forschungsberichterstattung der Bundesregierung (Bundesbericht Forschung) und stellt Daten für EUROSTAT, OECD und UNESCO bereit. Die Broschüre "FuE-Info" enthält jeweils Grunddaten und Beiträge über die aktuelle Entwicklung der FuE-Aktivitäten. In der Reihe Materialien zur Wissenschaftsstatistik werden ausgewählte Themen aus dem Bereich der Messung und Beurteilung von FuE, Innovationen, Wissenschaft und Technik aufgegriffen. Der "FuE-Datenreport" enthält die detaillierten Daten der vom Stifterverband zweijährlich durchgeführten Erhebung der FuE-Aktivitäten in der deutschen Industrie.

Das *Hochschulinformationssystem HIS* führt als von Bund und Ländern finanzierte Einrichtung Informationserhebungen durch, die vorwiegend die Entwicklung von Verfahren zur Rationalisierung der Hochschulverwaltung, Mitwirkung bei deren Einführung und Anwendung sowie Gutachten zur Schaffung von Entscheidungsgrundlagen im Hochschulwesen zum Zweck haben. Einen dezidierten forschungspolitischen Auftrag hat das HIS nicht. Zur Zeit laufen einige Projekte zur verbesserten Kosteneffizienz von Hochschulverwaltungen, welche auf das Verhältnis von Input- und Outputgrößen abstellen. Hierzu sind jedoch noch keine hier verwertbaren Publikationen vorhanden.

Der *Wissenschaftsrat (WR)* besitzt als "Fremdbeschreibungsagentur" wissenschaftlicher Einrichtungen eine hohe Reputation. Er hat neben den Querschnittsevaluationen, die zeitlich und sachlich begrenzte evaluative Informationen dokumentieren (vgl. Wissenschaftsrat 1999), kein systematisches Berichtssystem entwickelt, das Forschung in Deutschland beschreibt. Auch die Systemevaluationen der HFG und WGL sind – obgleich umfangreich – nach ihrem Zustandekommen und ihrer Anlage ad-hoc-Untersuchungen, denen kein "langer Atem" zugrunde liegt. Man erhält durch den WR somit durchaus aufschlussreiche, aber ausschnittshafte Beschreibungen der Leistungen ausgewählter Wissenschaftseinrichtungen.

Forschungspolitische Akteure. Das *BMBF* erhebt mit seinen Dokumentationen zur strukturellen Situation des deutschen Forschungssystems (BMBF 2000; 2002) und der technologischen Leistungsfähigkeit Deutschlands (vgl. BMBF 2001b) systematisch wichtige evaluative Informationen. Als Bundeseinrichtung verarbeitet das BMBF diese Fremdbeschreibungen auf hoch aggregierter Ebene.

Seit vielen Jahren führt das Forschungsministerium Programmevaluationen durch (Übersicht bei Kuhlmann/Holland 1995a); bis heute wird aber nur ein Teil dieser von externen Evaluationsexperten durchgeführten Studien (häufig ex post-Wirkungsanalysen) der Öffentlichkeit zugänglich gemacht oder regulär publiziert.

Wenig evaluative Information liegt bisher bezüglich der vom Bund betriebenen Einrichtungen der problemorientierten *Ressortforschung* vor.

Einige *Wissenschaftsministerien der Länder* (insbesondere Baden-Württemberg, Bayern, Nordrhein-Westfalen, Niedersachsen) haben eine Reihe von Initiativen zur Evaluation der Hochschulforschung angestoßen. So stellt etwa das niedersächsische Ministerium für Wissenschaft und Kunst die weiter oben diskutierten Evaluationsberichte auf der Website: <http://www.wk.niedersachsen.de/Materialien/material.htm> zur Verfügung.

Das *Generaldirektorat Forschung der Europäischen Kommission* hat im Rahmen des Auf- und Ausbaus der *europäischen Forschungsrahmenprogramme* nach und nach eine relativ systematische Evaluationspraxis entwickelt (Georghiou 1995). Dazu gehören das Five Year Assessment des Rahmenprogramms durch ein Expertenpanel (z.B. European Commission 2000b), jährliche Monitoring-Berichte, Evaluationen der Spezifischen Programme¹⁹, horizontal Evaluationen übergreifender Fragestellungen und Studien der Wirkungen europäischer Programme im nationalen Forschungssystem (z.B. Reger/Kuhlmann 1995).

Im Rahmen der Kommissionsinitiative zur Schaffung eines "europäischen Forschungsraums" (vgl. Ragnarson 2001; European Commission 2000) gibt es seit etwa einem Jahr Bemühungen, Indikatorenbündel zur Identifikation europäischer wissenschaftlicher "*Centres of Excellence*" zu identifizieren, exemplarisch anzuwenden und in ein öffentlich zugängliches Berichtssystem zu implementieren; hierzu gehören bibliometrische und Patentanalysen sowie Auswertungen vielfältiger Quellen evaluativer Informationen aus nationalen Kontexten. Im Winter 2002/03 sollen entsprechende evaluative Daten für die Gebiete "Economics", "Nanotechnologies" und "Life Sciences" vorliegen. Für die strategische evaluative Analyse ganzer Institutionen-"Landschaften" können solche Datenaggregationen künftig von großer Bedeutung sein.

Außerdem erstellt das Generaldirektorat Forschung – ähnlich wie die OECD (z.B. OECD 2001) – verschiedene *Indikatorenreports* auf der Grundlage evaluativer Information, etwa die "Indicators for benchmarking of national research policies" mit aggregierten Daten zu "human resources in RTD, including the attractiveness of S&T professions; Public and private investment in R&D; Scientific and technological productivity; impact of RTD on economic competitiveness and employment" (European Commission 2001).

Ein wichtigen Beitrag zur Generierung und zum Angebot evaluativer Informationen liefern Forschungseinrichtungen mit einem Schwerpunkt im Bereich *Wissenschafts- und Innovationsforschung*. Hier sind vor allem die folgenden zu nennen (wobei darauf hinzuweisen ist, dass eine ganze Reihe weiterer Forschergruppen gelegentlich Beiträge liefert):

- Das *Institut für Wissenschafts- und Technikforschung (IWT)* an der Universität Bielefeld forscht seit den 1970er Jahren interdisziplinär unter anderem auf den Gebieten Wissenschafts- und Technologiepolitik, Wissenschaftsindikatoren und Evaluation sowie Entwicklung wissenschaftlicher und technischer Institutionen. Das IWT hat seit den 1980er Jahren eine Reihe bibliometrischer Analysen angefertigt, die auch im Kontext evaluativer Verfahren Anwendung fanden, zuletzt eine Strukturanalyse der Publikationsaktivität der medizinischen Fakultät der Universität Münster.

¹⁹ Eine kürzlich abgeschlossene Analyse von 28 (überwiegend publizierten) Evaluationen der "sozio-ökonomischen" Wirkungen verschiedener Stränge des Rahmenprogramms unterscheidet "explorative" and "gründliche" Studien (vgl. European Commission 2002a): (1) "Explorative" Studien behandeln die möglichen künftigen wissenschaftlichen, technologischen, ökonomischen oder sozialen Wirkungen von Fördermaßnahmen. (2) "Gründliche" Studien folgen eher klassischen Evaluationsmethoden, sind ex post-orientiert und betrachten eine große Breite von Effekten der Forschungsförderung; dabei bieten sie eine Fülle quantitativer und qualitativer evaluativer Information auf der Aggregationsebene "Programm". Die Analyse kommt zu dem Schluss, dass die Studien Outputs besser erfassen als längerfristige Wirkungen, dass dabei aber sozio-ökonomischen Wirkungen sogar etwas mehr Aufmerksamkeit geschenkt wird als wissenschaftlichen oder technologischen Effekten. Eine Reihe methodischer Schwächen fielen auf: Die verwendeten Daten waren häufig uni-dimensional (z.B.- Befragungsergebnisse wurden nicht mit unabhängigen Statistiken verglichen); vorhandene Daten wurden nicht erschöpfend analysiert; die Ergebnisse quantitativer Datenanalysen wurden über- sowie die Einstellungen der Forscher unterbewertet.

- Das *Centrum für Hochschulentwicklung (CHE)* erforscht und entwickelt Modelle zur Modernisierung der deutschen Hochschulen und zielt dabei auf die konkrete Umsetzung und Erprobung neuer Organisations- und Steuerungsmodelle in der Hochschulwirklichkeit. Besonders bekannt wurde das CHE durch sein Hochschulranking, das derzeit detaillierteste Ranking der Studienangebote deutscher Hochschulen (Uni/FH). Es enthält Fakten und Urteile von Studierenden und Professoren zu verschiedenen Aspekten von Lehre und Forschung. Die Daten werden nicht zu einer Gesamtpunktzahl verrechnet, sondern zeichnen ein differenziertes Bild Fachhochschulen. Außerdem führt das CHE Evaluationen der Forschungsleistungen von Hochschuleinrichtungen durch (z.B. "Förderung von Exzellenz in Forschung, Lehre und Organisation an der Fakultät für Mathematik der TU München"; siehe http://www.che.de/html/das_che.htm)
- Das *Wissenschaftszentrum Berlin (WZB)* führt unter anderem Forschungen zum Verhältnis von Aufgabenstellung und Organisationsstruktur wissenschaftlicher Einrichtungen und zu Zielen, Verfahren und Kriterien von Evaluationen in Wissenschaft und Forschung durch. Am Beispiel der Evaluation der Einrichtungen der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL) wurden die Möglichkeiten und Grenzen "reflexiver" Evaluationen basierend auf einer selbstkritischen Stärken-Schwächen-Analyse untersucht (Röbbecke/Simon 2001).
- Das *Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI)* forscht seit nahezu dreißig Jahren interdisziplinär zu Strukturveränderungen im Forschungs- und Innovationssystem. Die Evaluation forschungspolitischer Maßnahmen sowie die konzeptionelle Weiterentwicklung forschungspolitischer Ansätze bilden einen zentralen Bestandteil der Tätigkeiten des ISI. Vor allem untersucht es mit Hilfe methodisch-basierter Evaluationen die erreichten und absehbaren Wirkungen von Fördermaßnahmen. Dabei haben die Arbeiten am Institut traditionell sowohl theoretisch als auch praktisch zur Weiterentwicklung von Evaluationsmethoden und zur Umsetzung von Evaluation in innovative Politikkonzepte beigetragen (z.B. Kuhlmann/Holland 1995). In jüngerer Zeit ist das ISI zusehends in europäischen Projekten eingebunden, die die Methodik zur Evaluation und Konzeption von Forschungspolitik international weiter führen (Airaghi et al. 1999; Kuhlmann et al. 1999; PREST et al. 2002). Ein Schwerpunkt der Evaluationsstudien in den letzten Jahren bildete die netzwerkorientierte Analyse von staatlichen Forschungsprogrammen, in denen der Staat Initiator und Moderator von Kooperationsstrukturen ist, wie etwa bei der Evaluation interdisziplinärer, klinischer Forschungszentren (Braun et al. 1997) oder der Evaluation der Kompetenzzentren der Nanotechnologie (Bührer et al. 2002). Des Weiteren wurde eine größere Studie zum Wissens- und Technologietransfer durchgeführt, in der die Fragen der auf industrielle Bedürfnisse orientierten Forschung, der Modi der Kooperation sowie der Verwertungsstrategien von öffentlichen Forschungsinstituten analysiert wurden (Schmoch et al. 2000).
- Das *Wissenschaftliche Zentrum für Berufs- und Hochschulforschung* an der Universität Gesamthochschule Kassel (gegründet 1978) beschäftigt sich vorrangig mit Fragen des Zusammenhangs von Hochschulausbildung und Beruf, im weiteren Rahmen auch von Hochschule, Staat und Gesellschaft. Evaluative Information zu Leistungen des Forschungssystem werden dabei nur am Rande behandelt.

Zunehmend tragen auch ausländische Forschungseinrichtungen evaluative Studien zum deutschen Forschungssystem bei. Zu nennen sind hier Einrichtungen wie

- das *Centre for Science and Technology Studies (CWTS)* an der Universität Leiden, Niederlande hat sich auf die quantitative Analyse wissenschaftlicher und technologischer Leistungsfähigkeit sowie die kognitive and organisatorische Struktur von Wissenschaft

und Technologie spezialisiert. CWTS führt Studien für Regierungen, Forschungsträgerorganisationen, Universitäten und Unternehmen durch (z.B. kürzlich eine bibliometrische Studie zur medizinischen Forschung in Deutschland, siehe Tijssen et al. 2001).

- das *Department of Philosophy of Science and Technology der Universität Twente*, Niederlande, erforscht die Dynamik moderner Forschungssysteme, u.a. mit Bezug auf Charakteristika des deutschen Systems (z.B. Rip/van der Meulen 1996).
- *Technopolis* ist ein internationales Forschungs- und Beratungsunternehmen mit dem Schwerpunkt Wissenschafts-, Forschungs-, Technologie- und Innovationspolitik (Hauptsitz Brighton, UK). Das Unternehmen führt gelegentlich Evaluationen im deutschen Forschungssystem durch (z.B. gemeinsam mit einem Konsortium deutscher Partner eine Evaluation des Mikrosystemtechnik 2000+ Programms des BMBF; siehe <http://www.technopolis.co.uk>)

5 Informationsnutzer und Desiderata der Informationsbereitstellung

Vor dem Hintergrund der explorativen Skizze der Erzeugung und des "Angebots" evaluativer Information im deutschen Forschungssystem werden nun Desiderata der Bereitstellung solcher Information mit Blick auf mögliche nachfragende Nutzer behandelt. Die Angaben zur potentiellen Nachfrage beruhen teils auf Hinweisen, die in den Hintergrundgesprächen gegeben wurden, teils aber auch auf Plausibilitätserwägungen der Autoren.

5.1 Nutzerinteressen

Die *Deutsche Forschungsgemeinschaft* hat im Jahr 2000 einen Senatsausschuss "Perspektiven der Forschung" ins Leben gerufen. Er soll als eine Art "Frühwarnsystem" im Rahmen wissenschaftlicher Bedarfsanalysen innovative Forschungsfelder identifizieren und die forschungsstrategischen Initiativen des Senats vorbereiten. Damit ergänzt die DFG ihre bisherige Praxis, dass zu jedem Thema zu jeder Zeit ein Antrag gestellt werden kann. Der neue Ansatz soll zur Identifikation wissenschaftlicher "Leuchttürme" beitragen. Die erforderlichen Bedarfsanalysen können nur auf der Basis evaluativer Information erstellt werden. Hierzu gehören Input-Informationen (Wer fördert bereits was, national und international?), Output-Informationen auf mittlerem und höherem Niveau (Welche Einrichtungen leisten bereits exzellente Forschung in bestimmten Gebieten? Welche Einrichtungen sollten gestärkt werden?) sowie Information über erzielte oder potentielle Wirkungen von Forschung und Förderung (z.B. in der akademischen Ausbildung; im Bereich Technologie und Innovation; zur Lösung sozialer oder ökologischer Probleme, etc.).

Strategische Planung großer Forschungseinrichtungen: Die Forschungsträgerorganisationen wurden im Rahmen der Systemevaluationsen aufgefordert, deutlicher als bisher explizite strategische Entwicklungsplanungen zu betreiben. Sie müssen daher künftig – auf der Grundlage einer regelmäßig aktualisierten kritischen Analyse ihres eigenen Leistungsspektrums – ihr Forschungsportfolio im Kontext der nationalen und internationalen Forschungslandschaft positionieren. Ähnlich wie die DFG benötigen sie dabei Input-Informationen (Wer investiert bereits in welchem Bereich, national und international?), Output-Informationen auf mittlerem und höherem Niveau (Welche anderen Einrichtungen leisten bereits exzellente Forschung in bestimmten Gebieten?) sowie Information über erzielte oder potentielle Wirkungen von Forschung und Förderung (z.B. im Bereich Technologie und Innovation; zur Lösung sozialer oder ökologischer Probleme, etc.).

Bundes- und Landeseinrichtungen der Forschungspolitik (BMBF, BMWi; Bundestag; Wissenschaftsministerien der Länder) sind verantwortlich für die Pflege der regionalen, nationalen und institutionellen Forschungslandschaft im internationalen Kontext. Strategieentwicklung und Planungsarbeiten können in dieser Hinsicht bisher nur auf fragmentarische oder ad hoc erstellte Aggregate evaluativer Information zurückgreifen. Weder in der Interaktion mit Forschungsträgerinstitutionen und Universitäten noch gegenüber den Integrationsversuchen der Europäischen Union (Europäischer Forschungsraum; "Mapping of

Centres of Excellence") können Bund und Länder bisher ihre Interessen, Entscheidungen und Vorgehensweisen auf systematische Analysen stützen, die mittels repräsentativer evaluativer Information erstellt wurden (abgesehen von – zum Teil guten, aber aufwändigen – ad hoc-Studien durch Einrichtungen der Wissenschafts- und Innovationsforschung). Benötigt werden Input-Informationen (Welche Forschungskapazitäten gibt es bereits, und wer fördert was, national und international?), Output-Informationen auf mittlerem und höherem Aggregationsniveau (Welche Einrichtungen leisten bereits exzellente Forschung in bestimmten Gebieten? Welche Einrichtungen sollten gestärkt werden?) sowie Information über erzielte oder potentielle Wirkungen von Forschung und Förderung (z.B. in der akademischen und sonstigen Bildung; im Bereich Technologie, Innovation und Wirtschaftsentwicklung; zur Lösung sozialer oder ökologischer Probleme, etc.).

Das *Generaldirektorat Forschung der Europäischen Kommission* wird im Rahmen der Initiative "Europäischer Forschungsraum" die Arbeiten an der Identifikation wissenschaftlicher "Centres of Excellence" fortsetzen und entsprechende Ergebnisse einem öffentlich zugänglichen Berichtssystem zuführen. Hier entsteht künftig eine dauerhafte Nachfrage nach verlässlichen, regelmäßig aktualisierten mittel bis hoch aggregierten evaluativen Informationen aus nationalen Kontexten. Je hochwertiger diese Nachfrage gedeckt wird, desto adäquater wird die Forschung in Deutschland in der europäischen Landschaft abgebildet.

Auch der *Wissenschaftsrat* wird bei künftigen Evaluationen von Forschungsinstitutionen oder Themenfeldern ebenso wie bei möglichen weiteren Systemevaluationen Aggregatinformationen, wie sie die anderen, oben behandelten Einrichtungen benötigen, heranziehen müssen. Ein besonders dringender Bedarf besteht hinsichtlich Strukturinformationen zur Forschung in Universitäten.

Einrichtungen der *Wissenschafts- und Innovationsforschung* konnten ihre – teils aus wissenschaftlichem Interesse, teils im Auftrag forschungspolitischer Akteure durchgeführten – Studien zu Stärken, Schwächen und Dynamik des deutschen Forschungssystems bisher nicht auf ein verlässliches, allgemein akzeptiertes System evaluativer Information stützen; vielfach müssen Grundlagen-Daten aufwändig für die spezifischen Zwecke einer Studie erhoben werden.

5.2 Desiderata

Zusammenfassend lassen sich die möglichen Verwendungszwecke (potentieller) evaluativer Information, die nach kompatiblen Formaten generiert, systematisch und regelmäßig gesammelt, nach festgelegten Fragestellungen strukturiert und unterschiedlichen Nutzungen aggregiert wurde, wie in Tabelle 7 darstellen. Die Tabelle ordnet die Relevanz von evaluativen Input-Informationen, Output-Informationen sowie Wirkungsinformationen nach den in Kap. 3 vorgestellten Evaluationszwecken und Aggregationsebenen.

Tabelle 7: Zwecke und Aggregationsebenen des Bedarfs an evaluativer Information; Schätzung auf Basis explorativ gewonnener Angaben sowie Plausibilitätsüberlegungen

Aggregationsebenen evaluativer Information	Evaluationszwecke				
	Qualitäts- sicherung	Nutzen- optimierung	Selbst- steuerung	Effizienz- kontrolle	Struktur- verbesserung
<i>Vertikal</i>					
Einzelwissen- schaftler	☉				
Forschungs- gruppe; Lehrstuhl	☉ → ⇨	⇨	☉ → ⇨		
Fakultät; Institut; Abteilung	☉ → ⇨	⇨	☉ → ⇨		
Universität; Forschungs- träger- Organisation	☉ → ⇨	⇨	☉ → ⇨	☉ → ⇨	⇨
<i>Horizontal</i>					
Disziplin; Themenfeld	☉ → ⇨	⇨	⇨	⇨	☉ → ⇨
Programm		→ ⇨		☉ → ⇨	
Institutionen- "Landschaft"	☉ → ⇨	⇨	⇨	☉ → ⇨	☉ → ⇨

Erläuterung:

☉ = Input-Informationen; → = Output-Informationen; ⇨ = Wirkungsinformationen

Der Bedarf an aufbereiteter evaluativer, forschungsstrukturbezogener Information ist um so größer, je höher aggregiert die Betrachtungs- und Entscheidungsebene der Informationsnutzer ist. Ein großer Teil des erkennbaren Bedarfs wird vom heute verfügbaren Angebot im deutschen Forschungssystem *nicht gedeckt*. Desiderata der Informationsbereitstellung bestehen vor allem in folgenden Bereichen:

- (1) Es fehlen bisher vor allem *kompatible Formate und Rhythmen der Erfassung evaluativer Information* quer zu den Forschungseinrichtungen, mit Blick auch auf die internationale Anschlussfähigkeit wichtiger Daten- und Informationsbestände, sowohl bei Input-Informationen, Output-Informationen wie Wirkungsinformationen. Entsprechende Formatierung und Verfahrensabstimmung erfordert ein koordiniertes Vorgehen aller Verantwortlichen im deutschen Forschungssystem.
- (2) Kompatible Formate bilden die Voraussetzung für *Vergleiche* (institutionelle, internationale, ...) auf *aggregierten Niveaus* – besonders hier ist der Bedarf ausgeprägt. Dabei müssen die unterschiedlichen Aufgabenstellungen und fachlichen Orientierungen heterogener Forschungseinrichtungen unbedingt berücksichtigt werden – dies erforder-

dert *intelligente Aggregationsstrategien*. Aggregation durch pure Addition von Daten wäre demgegenüber "dumm". Intelligente, Institutionen übergreifende, strategisch orientierte Vergleiche gibt es bisher nur vereinzelt.

- (3) Falls die vorgenannten Desiderata befriedigt werden können, werden künftig auch informationelle *Verknüpfungen von Forschungsinput, Forschungsoutput und Forschungsstrukturdaten* mit verwandten Zielgrößen wie bildungs-, regional- oder europa-politischen Fragestellungen und Informationen erleichtert.

6 Mögliche organisatorische Antworten

Die folgenden Vorschläge für eine organisatorische *Verbesserung der Bereitstellung evaluativer Information im deutschen Forschungssystem* bleiben aufgrund des explorativen Charakters des Gutachtens *bewusst vorläufig* und sollen vor allem zur Diskussion anregen. Dabei wollen wir betonen, dass die Aufmerksamkeit vor allem der Bereitstellung aggregierter Information gilt, hier also nicht Vorschläge zum institutionellen Ort künftiger Durchführung von Evaluationen oder Forschungsstrategieplanung unterbreitet werden: diese Aufgaben sind und bleiben in der Verantwortung der Forschungsförderungsinstitutionen, der Forschungsträgereinrichtungen, der Forschungspolitik sowie – als weithin anerkannter neutraler Instanz – des Wissenschaftsrates.

Anregend kann dabei ein Blick auf mögliche Modelle im Ausland wirken. Exemplarisch seien drei Einrichtungen genannt, die verwandte Aufgaben erfüllen, jedoch mit unterschiedlicher Ausprägung:

- *Überinstitutionelle Datensammlung.* Das Observatoire des Sciences et des Techniques (OST) wurde 1990 als öffentliche Einrichtung etabliert. Etwa 15 Personen arbeiten an Projekten für die Trägereinrichtungen des OST, vor allem nationale Ministerien und öffentliche und private Forschungsträgereinrichtungen. OST publiziert jährlich einen FuT-Indikatorenbericht (Science & Technologie – Indicateurs, éditions Economica, Paris) und erstellt Indikatoren, die Entscheidungsträgern der Forschungspolitik und des Forschungsmanagements bei der Strategieentwicklung unterstützen sollen. Dabei werden evaluative Daten bisher überwiegend nur sehr hoch aggregiert präsentiert, kaum nach Institutionen differenziert und wenig analytisch aufbereitet.
- *Überinstitutionelle Datensammlung, -erhebung und Analyse.* Die Abteilung "Science Resources Statistics (SRS)" der National Science Foundation (NSF) der USA erfüllt die gesetzliche Aufgabe "to provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources, and to provide a source of information for policy formulation by other agencies of the Federal Government" (SRS 2002). Zu diesem Zweck erstellt SRS etwa 14 periodische Erhebungen sowie eine Reihe weiterer Studien und Projekte. Alle US-Bundesbehörden, die FuE betreiben, liefern SRS relevante Daten. Außerdem arbeitet SRS eng mit Universitäten, Industrieunternehmen, Berufsorganisationen und internationalen Organisationen zusammen und hält evaluative Information anderer Erzeuger bereit. Der Schwerpunkt der Aktivität liegt bei SRS allerdings auf der Aggregation und Analyse deskriptiver Daten, weniger hingegen bei der strategischen Auswertung evaluativer Information.
- *Überinstitutionelle Datensammlung, -erhebung und Analyse kombiniert mit Wissenschafts- und Innovationsforschung, Evaluation, Vorausschau, Analyse und Beratung.* Das schweizerische Zentrum für Wissenschafts- und Technologiestudien (CEST) steht unter Oberaufsicht des Eidgenössischen Departements des Innern sowie des Eidgenössischen Volkswirtschaftsdepartements. CEST umfasst ein Kompetenzzentrum Scientometrie, Einheiten für Evaluation, Technologie und Innovation, für Prospektion sowie eine Dokumentationsstelle für Wissenschaftspolitik. CEST beschafft und überprüft Grundlagen für die schweizerische Forschungs-, Hochschul- und Technologiepolitik. Es untersucht den aktuellen Stand und zeigt mögliche Entwicklungen auf. Zu diesem Zweck erstellt es sowohl Fallstudien als auch vergleichende Analysen. Es führt Evaluati-

onen durch von Disziplinen, von Institutionen sowie von durch den Bund finanzierten oder mitfinanzierten Wissenschafts- und Technologieprogrammen.

Für das deutsche Forschungssystem lassen sich grundsätzlich vier alternative Richtungen für eine organisatorische Verbesserung der Bereitstellung evaluativer Information im skizzieren:

- (1) *Fortsetzung des Status quo*: Die aktuelle Situation der Nachfrage nach, der Erzeugung von und des Angebots an evaluativer, forschungsstrukturbezogener Information wird als akzeptabel bewertet; nach dieser Sichtweise besteht also kein Bedarf nach einer neuen Dienstleistung.
- (2) *Einigung* aller relevanten Akteure im deutschen Forschungssystem (einschließlich Repräsentanten der Wissenschafts- und Innovationsforschung) auf ein gewisses *Minimum kompatibler Formate* und Rhythmen der Erfassung evaluativer Information quer zu den Forschungseinrichtungen, bei gleichzeitiger Sicherstellung internationaler Kompatibilität wichtiger Daten- und Informationsbestände, sowohl bei Input-Informationen, Output-Informationen wie Wirkungsinformationen. Die kompatibel formatierte Information kann die Grundlage für Aggregation, Vergleiche etc. bilden, durchführbar durch alle interessierten Stellen.
- (3) Einrichtung einer *Gemeinschaftsstelle* der deutschen Forschung für die Aufbereitung und Vermittlung kompatibel formatierter Forschungsstrukturinformationen. Die Stelle sollte als intelligenter "Vermittlungsknoten" mit der Funktion eines "Clearinghouse" in einem Netz von Anbietern und Abnehmern konzipiert werden, also nicht als monopolistischer Informationsverwalter; sie sollte auch die Moderation zwischen den interessierten Akteuren des Forschungssystem übernehmen.
- (4) Einrichtung einer zentralen, *übergreifenden "Evaluationsagentur"* ggf. in Verbindung mit einer zentralen Strategieentwicklungseinheit des deutschen Forschungssystems.

Die Alternative 1 (keine Aktion) muss angesichts des diskutierten wachsenden Problemdrucks und Analysebedarfs als nachteilig bewertet werden. Die Alternative 4 ("Evaluationsagentur") würde kaum Akzeptanz bei der Mehrheit der Akteure finden, nicht zuletzt weil stets die Gefahr einseitiger politischer Vereinnahmung besteht; außerdem drohen stark institutionalisierte Maßnahmen schneller zu "vergreisen" als flexible, reversible.

Die Autoren des Gutachtens empfehlen eine *Kombination der Alternativen 2* (mehr Koordination von Formaten und Verfahren der Erfassung evaluativer Information) *und 3* ("intelligenter Knoten als Clearinghouse"). Die Aufgabe einer solchen Gemeinschaftsstelle "Clearinghouse" besteht in der *Unterstützung* ihrer Trägerorganisationen bei der Erfüllung ihrer – je nach Institution verschieden ausgeprägten – Zwecke der Verwendung evaluativer Information: Qualitätssicherung, Nutzenoptimierung, Effizienzkontrolle, Selbststeuerung, Strukturverbesserung (siehe Tabelle 8).

Die Gemeinschaftsstelle "Clearinghouse" wird gesteuert und überwacht von einem *Konsortium* (d.h. einer freiwilligen Kooperation) bestehend aus allen relevanten deutschen Einrichtungen der Forschungspolitik, der Forschungsförderung, der Forschungsträgerinstitutionen sowie ausgewählten Einrichtungen der Wissenschafts- und Innovationsforschung. Dieses Konsortium einigt sich auf den Umfang und die Formate quantitativer und qualitativer evaluativer Information, welche Forschungsorganisationen, Forschungsförderer und andere Konsortialmitglieder beitragen, sowie die Zielsetzungen, Formate und Verwen-

dungszwecke aggregierter strategischer Information der Gemeinschaftsstelle; diese modelliert die erforderlichen Abstimmungsprozesse. Informative Produkte der Gemeinschaftsstelle stehen grundsätzlich allen Konsortialmitgliedern zur Verfügung (siehe Abbildung 1); Ausnahmen hiervon sind möglich, jedoch einvernehmlich zu regeln. Möglichst viele Informationsprodukte sollen der allgemeinen, interessierten Öffentlichkeit zugänglich gemacht werden (z.B. Internet-Website).

Tabelle 8: Denkbare Trägerkonsortium und wichtige Unterstützungszwecke einer Gemeinschaftsstelle "Clearinghouse"

Träger-Einrichtung	Unterstützungszwecke				
	Qualitäts-sicherung	Nutzen-optimierung	Effizienz-kontrolle	Selbst-steuerung	Struktur-verbesserung
DFG	<input type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
MPG	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
FhG	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HGF	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
WGL	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
BMBF		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
BLK		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
WR	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Stifter-verband				<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Wiss.- u. Innovations-forschung					<input checked="" type="checkbox"/>

Erläuterung: = Evaluationsaktivitäten in eigener Regie; = Zusatznutzen durch Gemeinschaftsstelle

Dabei sind mehrere Gestaltungsvarianten denkbar:

- Arbeit auf der Grundlage *freiwilliger Selbstverpflichtung* aller relevanten Institutionen des deutschen Forschungssystems oder im Rahmen *regulierter Berichtspflichten* (z.B. auf gesetzlicher Grundlage).
- Ansiedlung der Stelle an einer *wissenschaftliche Einrichtung* (Universitätslehrstuhl; WGL-Einrichtung nach dem Vorbild der GESIS; Bindung an eine andere Forschungsträgerorganisation, z.B. eine – noch zu gründende – nationale Akademie der Wissenschaften) oder bei einer forschungspolitischen Institution (etwa beim *Wissenschaftsrat* als intermediärer Beratungseinrichtung).

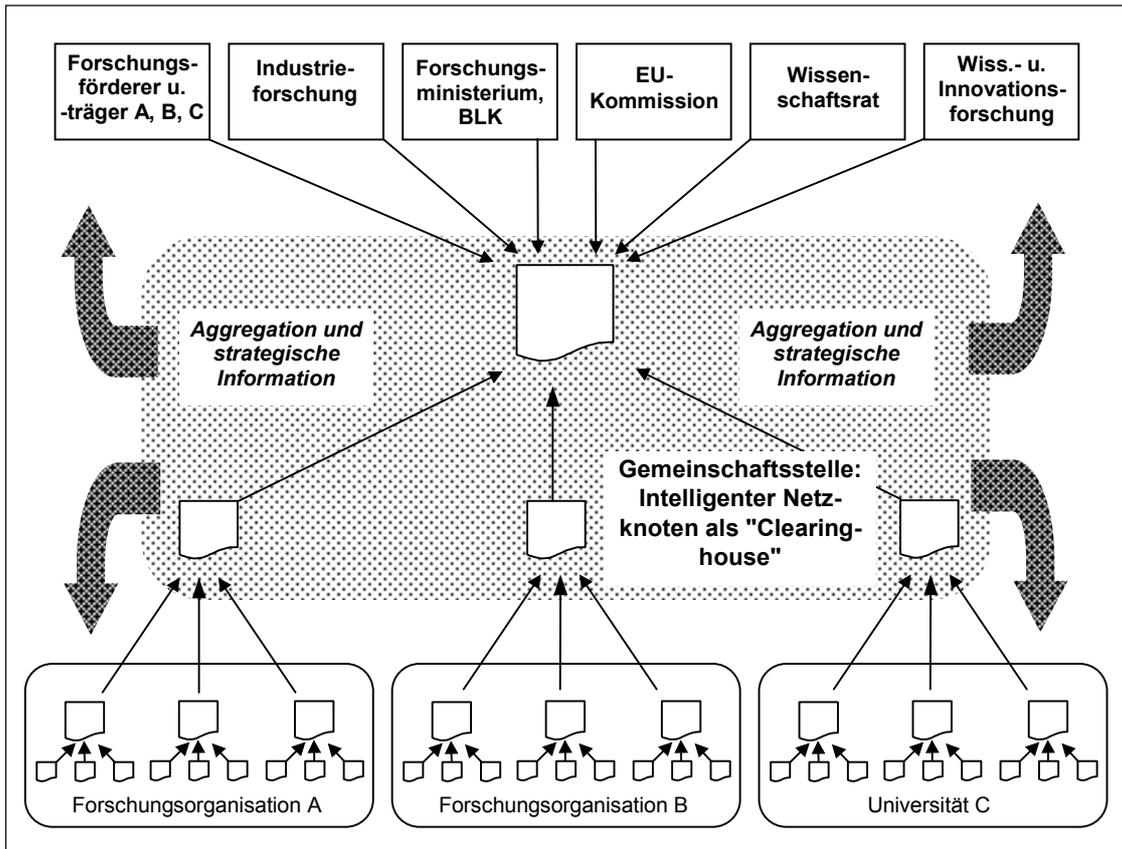
Zur *Arbeitsweise* einer Gemeinschaftsstelle "Clearinghouse": Im Rahmen der o.a. Zwecke werden aggregierte strategische Informationen auf der Basis evaluativer Information erarbeitet. Die realisierten Aggregationsebenen müssen in Relation zu den Verwendungszwecken der generierten Information stehen. Typische Aufgaben und Verfahrensweisen sind

- Erarbeitung und Pflege *kompatibler Formate und Rhythmen der Erfassung evaluativer Information* quer zu den Forschungseinrichtungen, mit Blick auch auf die internationale

Kompatibilität wichtiger Daten- und Informationsbestände, sowohl bei Input-Informationen, Output-Informationen wie Wirkungsinformationen.

- Entsprechende Formatierung und Verfahrensabstimmung erfordert die *Moderation* der Perspektiven und Interessen der verschiedenen Institutionen im deutschen Forschungssystem.

Abbildung 1: Funktionen einer deutschen Gemeinschaftsstelle "Clearinghouse" für Forschungsstrukturinformationen: "intelligenter Netzknoten"



- Erarbeitung von thematischen, international kompatiblen *Forschungslandkarten*, die eine Grundlage für Vergleiche bilden, aber auch helfen Lücken zu identifizieren sowie die Vernetzung von Forschungsaktivitäten erleichtern.
- Auf der Grundlage kompatibler Formate Durchführung von intelligenten, Institutionen übergreifenden, *strategisch orientierten Vergleichen* (institutionelle, internationale, ...) auf aggregierten Niveaus. Dabei müssen die unterschiedlichen Aufgabenstellungen und disziplinären Umwelten heterogener Forschungseinrichtungen unbedingt berücksichtigt werden – dies erfordert intelligente Aggregationsstrategien. Aggregation durch pure Addition von Daten wäre demgegenüber "dumm".
- *Stärken-Schwächen-Analysen* als Basis für Entscheidungen über Mittelallokationen durch Forschungsförderungseinrichtungen (z.B. DFG, BMBF, Landesministerien)

- Daten werden *nicht* "blindlings" gesammelt, aggregiert, analysiert sondern nur im Rahmen von Aufgabenstellungen, die mit den Trägern (allen oder einzeln, in diesem Falle aber konsensual) abgestimmt wurden. Dies schließt ein, dass die Träger an sich "vertrauliche" Informationen gezielt für vereinbarte strategische Analysen bereit stellen.
- Hierbei kann die Gemeinschaftsstelle nur in begrenztem Umfang eigene Forschungsaufgaben wahrnehmen: anspruchsvolle Analysen sollen in der Regel von Einrichtungen der *Wissenschafts- und Innovationsforschung* ausgeführt werden (aus Eigeninteresse oder im Rahmen von Auftragsforschung).
- Ein rollierendes System von *Personalabstellungen* aus Nutzereinrichtungen (wie bei KoWi) erleichtert zusätzlich die Mobilisierung externer Intelligenz.
- Bei ihren Aggregations- und Analysearbeiten soll die Gemeinschaftsstelle die intellektuellen *Grenzen von Evaluationsverfahren* (siehe Kap. 2) bewusst halten. "Blinde Flecken" des Nichtwissens sollen nicht allein als Risiko sondern auch als Chance für Kreativität akzeptiert werden.

7 Verwendete Literatur

- Airaghi, A. / Busch, N. E. / Georghiou, L. / Kuhlmann, S. / Ledoux, M. J. / van Raan, A. F.J. / Viana Baptista, J., 1999: Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes. Report to the European Commission, DG XII, Evaluation Unit, Brussels/Luxembourg (Office for Official Publications of the European Communities). – ISBN 92-828-3721-1
- Arnold, E. und K. Guy, 1997: 'Technology Diffusion Programme and the Challenge for Evaluation,' In: OECD (ed.): Policy Evaluation in Innovation and Technology, Towards Best Practices. Paris, S. 65-87
- Barré, R., 2001: Sense and Nonsense of S&T Productivity Indicators, *Science and Public Policy*, vol. 28 (4), S. 1-8
- Becher, G. und S. Kuhlmann, Hrsg., 1995: Evaluation of Technology Policy Programmes in Germany, Boston/ Dordrecht/ London
- Block, H.-J. / Krull, W., 1990: What are the consequences? Reflections on the impact of evaluations conducted by a science policy advisory body, in: *Scientometrics*, Vol. 19, Nos. 5-6, 427-437.
- Bozeman, B. / Melkers, J. (Hg.), 1993: Evaluating R&D Impacts: Methods and Practice, Boston / Dordrecht / London (Kluwer Academic Publishers)
- Braun, D. / Hinze, S. / Hüsing, B. / Kuhlmann, S. / Menrad, K. / Peter, V., 1997: Interdisziplinäre Zentren für Klinische Forschung in der Frühphase der Bundesförderung - Vergleichende Analyse. In Zusammenarbeit mit Université de Lausanne. Stuttgart (Fraunhofer IRB Verlag)
- Brockhoff, K., 1999: Forschung und Entwicklung. Planung und Kontrolle, München / Wien (R. Oldenbourg), 5. erg. u. erw. Aufl.
- Buchholz, K., 1995: Criteria for the Analysis of Scientific Quality, in: *Scientometrics* 32 (2), S. 195-218
- Bührer, S. / Bierhals, R. / Heinze, Th. / Hullmann, A. / Studer, Th. / Erlinghagen, R. / Lang, C., 2002: Die Kompetenzzentren der Nanotechnologie in der Frühphase der Bundesförderung: Ein Bericht der begleitenden Evaluation. Karlsruhe (ISI)
- Bundesministerium Bildung und Forschung, BMBF (Hrsg.), 2000: Bundesbericht Forschung 2000, Bonn
- Bundesministerium Bildung und Forschung, BMBF (Hrsg.), 2001a: Grund- und Strukturdaten 2000/2001, Bonn
- Bundesministerium Bildung und Forschung, BMBF (Hrsg.), 2001b: Zur technologischen Leistungsfähigkeit Deutschland. Zusammenfassender Endbericht 2000, Bonn
- Bundesministerium Bildung und Forschung, BMBF (Hrsg.), 2002: Faktenbericht Forschung, Bonn
- Bundesministerium für Bildung und Forschung und Medizinischer Fakultätentag (Hrsg.), 2002: Forschungslandkarte Hochschulmedizin. (Bearbeitung: Susanne Bührer, Fraunhofer ISI). Fraunhofer IRB-Verlag. Stuttgart
- Bürgel, H. D. /Haller, C. / Binder, M.; 1996: F & E-Management, München (Vahlen)
- Callon, M., 1992: The Dynamics of Techno-Economic Networks. In: Coombs, R. / Saviotti, P. / Walsh, V. (eds.): Technological Change and Company Strategies: Economic and sociological perspectives, London et al. (Academic Press Limited), 72-102
- Callon, M. / Larédo, Ph. / Mustar, Ph., 1995: La gestion stratégique de la recherche et de la technologie. L'évaluation des programmes, Paris (Economica)
- Campbell, D. F.J. / Felderer, B., 1997: Evaluating Academic Research in Germany. Patterns and Policies, Vienna (Institute for Advanced Studies) (Political Science Series, No. 48) Daniel (2001)
- Daniel, H.-D., 1993: Guardians of Science – Fairness and Reliability of Peer Review, Weinheim

- Daniel, H.-D., und R. Fisch, Hrsg., 1988: Evaluation von Forschung, Konstanz
- Daniel, H.-D., 2001: Wissenschaftsevaluation. Neuere Entwicklungen und heutiger Stand der Forschungs- und Hochschulevaluation in ausgewählten Ländern, CEST, Bern
- Derlien, H.-U., 1976: Die Erfolgskontrolle staatlicher Planung. Eine empirische Untersuchung über Organisation, Methode und Politik der Programmevaluation, Baden-Baden (Nomos)
- Deutsches Patent- und Markenamt, DPMA, 2001: Jahresbericht 2000, München
- DFG, Deutsche Forschungsgemeinschaft, 1997: Bewilligungen an Hochschulen 1991 bis 1995, Bonn
- DFG, Deutsche Forschungsgemeinschaft, 2000: Bewilligungen an Hochschulen und außeruniversitäre Forschungseinrichtungen 1996 bis 1998, Bonn
- DFG, Deutsche Forschungsgemeinschaft, 2002a: Reform des Begutachtungssystems, siehe: <http://www.dfg.de/organisation/fachgutachter/reform.html>
- DFG, Deutsche Forschungsgemeinschaft, 2002b: Förder- und Forschungsmonitoring durch die DFG: Aktueller Stand und Perspektiven. In: Perspektiven der Forschung (im Druck)
- Etzkowitz, H. / Leydesdorff, L., 2000: The dynamics of innovation: from National Systems and "Mode2" to a Triple Helix of university-industry-government relations. In: Research Policy 29, 109-123
- European Commission (Hrsg.), 2002a: Assessing of the Socio-economic Impacts of the Framework Programme. Final Report, Manchester (mimeo; im Erscheinen)
- European Commission (Hrsg.), 2002b: RTD Evaluation Toolbox, a report prepared by the Thematic Network 'Socio - Economic Evaluation of Public RTD Policies (EPUB)' (mimeo, im Erscheinen)
- European Commission, 2000: 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, Brussels, 18 January 2000, COM (2000) 6
- European Commission, 2001: Indicators for benchmarking of national research policies. Luxembourg: Office for Official Publications of the European Communities, ISBN 92-894-1183-X
- Evaluierungskommission (Hrsg.), 1998: Systemevaluierung der Fraunhofer-Gesellschaft. Bericht der Evaluierungskommission, München (<http://www.fraunhofer.de/german/company/fhmodel/evaluierungsbericht.pdf>)
- Fisch, R. und H.-D. Daniel (Hrsg.), 1986: Messung und Förderung von Forschungsleistung, Konstanz
- Fraunhofer Patentstelle, 2001: Jahresbericht der Fraunhofer Patentstelle für die Deutsche Forschung, 2000/2001, München
- Fraunhofer-Gesellschaft, FhG (Hrsg.), 2001: Jahresbericht 2000, München
- Funtowicz, S. O. / J. R. Ravetz, 1993: Science for the post-normal age. In: Futures, 25:7, 739-755.
- Funtowicz, S. O., 2001: Post-Normal Science. Science and Governance under Conditions of Complexity. In: Politeia 2001, 77-85
- Georghiou, L., 1995: Assessing the Framework Programmes - a meta-evaluation, Evaluation Vol. 1, No.2, October 1995 pp.171-188
- Gerpott, T. J., 1999: Strategisches Technologie- und Innovationsmanagement. Eine konzentrierte Einführung, Stuttgart (Schäffer-Poeschel)
- Gibbons, M. / Limoges, C. / Nowotny, H. / Schwartzman, S. / Scott, P. / Trow, M., 1994: The new production of knowledge, London
- Grupp, H., 1998: Foundations of the Economics of Innovation: Theory, Measurement and Practice. Cheltenham (Edward Elgar Publishing)
- Grupp, H. / Kuntze, U. / Schmoch, U., 1995: New Technology Indicators for the Evaluation of Research and Development Programmes. In: Becher, G. / Kuhlmann, S. (Hg.): Evaluation of Technology Policy Programmes in Germany, Boton/Dordrecht/London (Kluwer Academic Publishers), 243-284

- Hellstern, G.-M. / Wollmann, H., 1984: Handbuch zur Evaluierungsforschung, Bd. 1, Opladen (Westdeutscher Verlag)
- Hochschulrektorenkonferenz, HRK (Hrsg.), 2001: Hochschulgesetzliche Regelungen zur Qualitätssicherung. Hochschulrahmengesetz und Hochschulgesetze der Länder, Bonn
- Hornbostel, S., 1997: Wissenschaftsindikatoren. Bewertungen in der Wissenschaft, Opladen
- Hornbostel, S., 2000: Kurzgutachten für die DFG zum Stand der Forschung im Bereich Forschungsevaluation, Jena
- Internationale Kommission, 1999: Forschungsförderung in Deutschland. Bericht der internationalen Kommission zur Systemevaluation der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft. Studie im Auftrag der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK), herausgegeben von der Volkswagen-Stiftung, Hannover
- Jasanoff, Sh., 1997: Introduction. In: Jasanoff, Sh. (ed.): Comparative Science and Technology. The International Library of Comparative Public Policy, vol. 5, Cheltenham (Edward Elgar), xiii-xxiii
- Kommission Systemevaluation der Wirtschaftsintegrierenden Forschungsförderung, 2001: Endbericht der Kommission, Berlin (mimeo, download: <http://www.bmwi.de/Homepage/download/technologie/Systemevaluation.pdf>)
- Küchler, T., et al., 2000: Hochschulreform Baden-Württemberg 2000. Stellungnahme und Empfehlungen, Centrum für Hochschulentwicklung, Gütersloh
- Kuhlmann, S., 1998: Politikmoderation. Evaluationsverfahren in der Forschungs- und Technologiepolitik, Baden-Baden: Nomos
- Kuhlmann, S., 2000 : Evaluation in der Forschungs- und Innovationspolitik. In: Stockmann, Reinhard (Hg.): Evaluationsforschung, Leverkusen 2000, S. 287-307.
- Kuhlmann, S. / Holland, D., 1995a: Evaluation von Technologiepolitik in Deutschland - Konzepte, Anwendung, Perspektiven, Heidelberg: Physica-Verlag.
- Kuhlmann, S. / Holland, D., 1995b: Erfolgsfaktoren der wirtschaftsnahen Forschung, Heidelberg: Physica-Verlag.
- Kuhlmann, S. / Boekholt, P. / Georghiou, L. / Guy, K. / Héraud, J.-A. / Laredo. Ph. / Lemola, T. / Loveridge, D. / Luukkonen, T. / Polt, W. / Rip, A. / Sanz-Menendez, L. / Smits, R., 1999: Improving Distributed Intelligence in Complex Innovation Systems. Brussels/Luxembourg (Office for Official Publications of the European Communities) (http://www.isi.fhg.de/abtIlg/ti/pb_html/final.pdf)
- Liefner, Ingo, 2001: Leistungsorientierte Ressourcensteuerung in Hochschulsystemen. Ein internationaler Vergleich, Berlin
- Max-Planck-Gesellschaft, MPG (Hrsg.), 2001: Jahresbericht 2000, München
- Mayntz, R., 1980: Implementation politischer Programme. Empirische Forschungsberichte, Königstein (Verlagsgruppe Athenäum/Hain/Scriptor; Neue Wissenschaftliche Bibliothek)
- Mayntz, R., 1983: Implementation politischer Programme II. Ansätze zur Theoriebildung, Opladen (Westdeutscher Verlag)
- Meyer-Krahmer, F., 1989: Der Einfluß staatlicher Technologiepolitik auf industrielle Innovationen. Baden-Baden (Nomos)
- National Science Board, NSB; Hrsg., 2000: Science and Engineering Indicators, Washington
- Neidhardt, F., 1988: Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG, Opladen: Westdeutscher Verlag
- Neidhardt, F., 2001: Vorwort, in: Röbbelcke/Simon, S. 7-12
- OECD, 2001: OECD Science, Technology and Industry Scoreboard 2001, Towards a knowledge-based economy. <http://www1.oecd.org/publications/e-book/92-2001-04-1-2987/>
- Patton, M. Q., 1997: Utilization-Focused Evaluation. The New Century Text, Thousand Oaks et al. (Sage)PREST et al. 2002

- Quoniam, L. et al., 1995: Treating Bibliometric Indicators with Caution: Their Dependence on the Source Database, in: *Research Evaluation* 5 (3), S. 177-181
- Ragnarson, R., 2001: Zukünftiger europäischer Forschungsraum, in: *FuE-Info* 2/2001, S. 8-11
- Reger G. / Kuhlmann S., 1995: *European Technology Policy in Germany. The Impact of European Community Policies upon Science and Technology in Germany*, Heidelberg: Springer
- Rip, A., 1990: Implementation and Evaluation of Science & Technology Priorities and Programs, in: Cozzens, S. (ed.): *The Research System in Transition*, Boston/Dordrecht/London: Kluwer Academic Publishers, 263-280
- Rip, A. / van der Meulen, B.J.R. (1996): 'The Post-Modern Research System'. In: *Science & Public Policy* 23 (5), S. 343-352
- Rip, A., 2002: *Science for the 21st Century*. Essay prepared for the 'Sciences and Arts Debate' of the *Hollandsche Maatschappij der Wetenschappen*, 16 February 2002
- Röbbেকে, M. und D. Simon, 2001: *Reflexive Evaluation. Ziele, Verfahren und Instrumente der Bewertung von Forschungsinstituten*, Berlin
- Schmoch, U. / Licht, G. / Reinhard, M., 2000: *Wissens- und Technologietransfer in Deutschland. Studie im Auftrag des Bundesministerium für Bildung und Forschung*. In Zusammenarbeit mit IZentrum für Europäische Wirtschaftsforschung (ZEW) und Ifo-Institut für Wirtschaftsforschung, Stuttgart (Fraunhofer IRB Verlag)
- Schmoch, U., 2001: *Akademische Forschung in der Interaktion mit industrieller Forschung. Zur sozialen Vermittlung von Theorie und Praxis in der Technikgenese*, Habilitationsschrift, Karlsruhe
- Schön, D. / Rein, M., 1994: *Frame Reflection. Toward the Resolution of Intractable Policy Controversies*, New York (BasicBooks)
- Shapira, P. und Kuhlmann, S. (Hrsg.), 2002: *Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe*, Cheltenham, forthcoming
- Statistisches Bundesamt, StBA (Hrsg.), 2001a: *Wissenschafts- und Technologieindikatoren*, Wiesbaden
- Statistisches Bundesamt, StBA (Hrsg.), 2001b: *Zur finanziellen Lage der Hochschulen*, in: *Wirtschaft und Statistik* 12/2001, Wiesbaden, S. 1021-1026
- Statistisches Landesamt Baden-Württemberg, StLABW (Hrsg.), 2001: *FuE-Monitor Baden-Württemberg*, Stuttgart
- Stock, W. G., 1994: *Wissenschaftsevaluation. Die Bewertung wissenschaftlicher Forschung und Lehre*, ifo Diskussionsbeiträge, München
- Tijssen, R. J.W. et al. (2001): *Mapping the Scientific Performance of German Medical Research. An International Comparative Bibliometric Study* (initiated and funded by the German Federal Ministry of Education and Research, BMBF), Stuttgart (Schattauer für Medizin und Naturwissenschaften),
- van Raan, A. F.J. (Hg.), 1988: *Handbook of Quantitative Studies of Science and Technology*, Amsterdam (Elsevier)
- Weingart, P., 1997: From "Finalization" to "Mode 2": old wine in new bottles, in: *Social Science Information* 36, 4, 591-613
- Weingart, P. / Sehringer, R. / Winterhager, M. (Hrsg.), 1991: *Indikatoren der Wissenschaft und Technik*
- Widmer, Th., 1996: *Meta-Evaluation. Kriterien zur Bewertung von Evaluationen*, Bern et al. (Haupt)Wissenschaftliche Kommission Niedersachsen, Hrsg., 1999: *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen. Chemie. Bericht und Empfehlungen*, Hannover
- Wissenschaftliche Kommission Niedersachsen, Hrsg., 2001: *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen. Wirtschaftswissenschaften. Bericht und Empfehlungen*, Hannover

- Wissenschaftliche Kommission Niedersachsen, Hrsg., 2002a: Forschungsevaluation an niedersächsischen Hochschulen. Grunddaten, <http://www.wk.niedersachsen.de>
- Wissenschaftliche Kommission Niedersachsen, Hrsg., 2002b: Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen. Grundzüge des Verfahrens, <http://www.wk.niedersachsen.de>
- Wissenschaftsrat, WR (Hrsg.), 1994: Stellungnahme zur Umweltforschung in Deutschland, Bd.1 und 2, Köln
- Wissenschaftsrat, WR (Hrsg.), 1996: Stellungnahme zur außeruniversitären Materialwissenschaft, Köln
- Wissenschaftsrat, WR (Hrsg.), 1999: Stellungnahme zur Energieforschung, Köln
- Wissenschaftsrat, WR (Hrsg.), 2000a: Systemevaluation der Blauen Liste - Stellungnahme des Wissenschaftsrates zum Abschluss der Bewertung der Einrichtungen der Blauen Liste, Köln
- Wissenschaftsrat, WR (Hrsg.), 2000b: Drittmittel und Grundmittel der Hochschulen 1993-1998, Köln
- Wissenschaftsrat, WR (Hrsg.), 2001: Systemevaluation der HGF - Stellungnahme des Wissenschaftsrates zur Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren, Köln
- Wollmann, H., 1998: Evaluation research and politics: Between a science-driven and a pluralist controversy-responsive policy-making model. Potential and limitations. Paper given to the Conference of the European Evaluation Society (EES), Roma, October 1998
- Ziman, J. M., 1995: *Of One Mind: The Collectivization of Science*, Woodbury NY (AIP Press)
- Ziman, J. M., 2001: *Real science: what it is, and what it means*, Cambridge: Cambridge University Press