

Drs. 2855-13  
Berlin 25 01 2013

---

# Empfehlungen zu einem Kerndatensatz Forschung



## INHALT

---

	<b>Vorbemerkung</b>	<b>5</b>
	<b>Kurzfassung</b>	<b>7</b>
<b>A.</b>	<b>Anlass</b>	<b>8</b>
<b>A.I</b>	<b>Wachsender Informationsbedarf zu Forschungsaktivitäten</b>	<b>8</b>
<b>A.II</b>	<b>Entwicklungen</b>	<b>10</b>
<b>A.III</b>	<b>Zweck eines Kerndatensatzes Forschung</b>	<b>13</b>
	III.1 Systemweite Datenbasis	13
	III.2 Qualitätssicherung	13
	III.3 Interpretierbarkeit	14
	III.4 Begrenzung des Aufwandes	14
	III.5 Begrenzung der Abfragen	14
<b>A.IV</b>	<b>Gegenstand der Empfehlung</b>	<b>14</b>
<b>B.</b>	<b>Grundlegungen</b>	<b>16</b>
<b>B.I</b>	<b>Grundsätze</b>	<b>16</b>
	I.1 Übergreifende Prinzipien	16
	I.2 Prinzip der Datenhoheit und des Datenschutzes	17
	I.3 Prinzipien der Datenvorhaltung	17
	I.4 Prinzipien der Qualitätssicherung	18
	I.5 Prinzipien des Datenaustausches	18
<b>B.II</b>	<b>Auswahl der Inhalte und Festlegung der Definitionen</b>	<b>19</b>
<b>B.III</b>	<b>Pragmatische Abwägungen</b>	<b>19</b>
<b>C.</b>	<b>Empfehlungen</b>	<b>22</b>
<b>C.I</b>	<b>Empfehlungen zum Umgang mit Informationen zu Forschungsaktivitäten</b>	<b>22</b>
<b>C.II</b>	<b>Empfehlung an Forschungsförderer</b>	<b>24</b>
<b>C.III</b>	<b>Empfehlungen zu den Inhalten des Kerndatensatzes</b>	<b>24</b>
	III.1 Beschäftigte	24
	III.2 Nachwuchsförderung	25
	III.3 Drittmittel und Projekte	27
	III.4 Forschungspreise und Auszeichnungen	30
	III.5 Patente	31
	III.6 Publikationen	32

<b>4</b>	<b>C.IV Erweiterungsoptionen ausserhalb des Kerndatensatzes</b>	<b>35</b>
	IV.1 Herausgeberschaften, Mitgliedschaft in Editorial Boards	35
	IV.2 Akademische Transfer-Spinoffs	36
	IV.3 Lizenzeinnahmen	37
	IV.4 Nachwuchsförderung (Erweiterung)	37
	IV.5 Drittmittel (Erweiterung)	39
	<b>C.V Empfehlungen an Bund und Länder</b>	<b>39</b>
	<b>D. Standardisierung des Kerndatensatzes</b>	<b>40</b>
	<b>D.I Akteure</b>	<b>40</b>
	<b>D.II Elemente des Standardisierungsprozesses</b>	<b>42</b>
	II.1 Fächerklassifikation	42
	II.2 Standardisierung von Publikationsformaten	44
	<b>D.III Zeitplan und Meilensteine</b>	<b>44</b>
	III.1 Pilotphase	44
	III.2 Ablauf des Standardisierungsprozesses	44
	<b>D.IV Teilprojekte und Aufgaben (Lastenheft)</b>	<b>50</b>
	IV.1 Entwicklung einer konsensfähigen Fächerklassifikation	50
	IV.2 Standardisierung Publikationsdaten	51
	IV.3 Definitionen der Elemente des Kerndatensatzes	51
	IV.4 Festlegung der Datenformate	51
	IV.5 Erstellung eines konzeptuellen Datenmodells	51
	IV.6 Beurteilung der Machbarkeit	52
	IV.7 Festlegung von Schnittstellen	52
	IV.8 Standards für den Datentransfer	52
	IV.9 Öffentlichkeitsarbeit	53
	<b>D.V Projektabschluss</b>	<b>53</b>
	<b>Anhang</b>	<b>55</b>
	<b>Glossar</b>	<b>57</b>
	<b>Abkürzungsverzeichnis</b>	<b>58</b>
	<b>Abbildungsverzeichnis</b>	<b>59</b>

---

# Vorbemerkung

Zwischen dem legitimen Interesse an möglichst aussagekräftigen Informationen zu Forschungsaktivitäten auf der einen und möglichst geringem Aufwand bei deren Erhebung und Bereitstellung auf der anderen Seite bestehen Zielspannungen. Auf diese hat der Wissenschaftsrat vor kurzem bereits hingewiesen. |<sup>1</sup> Mit den hier vorgelegten Empfehlungen soll ein konkreter Beitrag geleistet werden, um ein vernünftiges und zielführendes Verhältnis zwischen Aufwand und Nutzen der Erhebung von Daten zu Forschungsaktivitäten zu erreichen.

Mit diesen Empfehlungen stößt der Wissenschaftsrat einen Prozess zur Standardisierung von Datenformaten zu Forschungsaktivitäten und Forschungsprodukten an, die den Datenaustausch zwischen verschiedenen Akteuren unterstützen sollen. Im Sinne der gebotenen Datensparsamkeit soll lediglich ein begrenzter Datensatz spezifiziert werden, der möglichst vielseitig Verwendung finden kann. Dieser so genannte Kerndatensatz soll sich in der Regel aus bereits verfügbaren Informationen generieren lassen (vgl. Glossar, S. 57).

Der Wissenschaftsrat wird den empfohlenen Standardisierungsprozess während der Umsetzung begleiten und die resultierende Spezifikation eines Datenaustauschformates veröffentlichen und wissenschaftspolitisch kommentieren.

Der Wissenschaftsrat ist sich bewusst, dass die organisatorische Umsetzung operativer Empfehlungen bei Wissenschaftlerinnen und Wissenschaftlern wie auch bei den Mitarbeiterinnen und Mitarbeitern in den Verwaltungen besondere Anstrengungen erfordert. Die mit diesen Empfehlungen vorgeschlagene Standardisierung wird aber bei allen Beteiligten – datenliefernden forschenden Einrichtungen und Datenabfragern – nach einer Aufbauphase zu einer Aufwandsersparnis führen. Die Voraussetzung dafür ist, dass alle Beteiligten die

|<sup>1</sup> Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, S. 14.

Standardisierungsvorgaben beachten und den Grundsatz der Datensparsamkeit ernst nehmen.

Der Wissenschaftsrat fordert die wissenschaftspolitischen Akteure in den Ländern und im Bund auf, die Implementierungsanstrengungen an Hochschulen und AUF nach Kräften zu unterstützen. Dazu gehört auch, eigene Bedarfsanforderungen zu begrenzen und sich an den Maßgaben der zu entwickelnden Spezifikation des Kerndatensatzes zu orientieren.

Zu operativen Themen wie der Standardisierung des Kerndatensatzes hat sich der Wissenschaftsrat in der Vergangenheit selten geäußert. Die Vielzahl der beteiligten Akteure, die Komplexität und Dringlichkeit sowie der nachhaltig positive Effekt, den ein funktionierender Kerndatensatz Forschung für das Wissenschaftssystem Deutschland haben kann, begründen aber die Befassung des Wissenschaftsrates mit diesem Thema.

In besonderem Maße ist der Wissenschaftsrat der Unterarbeitsgruppe zur „Standardisierung der Datenvorhaltung und Datenabfrage für die Bewertung von Forschungsaktivitäten“ der Steuerungsgruppe „Forschungsrating“ zu Dank verpflichtet. Ihre Vorarbeit ist Grundlage dieser Empfehlungen und war bei ihrer Erarbeitung von großem Wert. Die vorliegenden Empfehlungen übernehmen die von der Unterarbeitsgruppe erarbeiteten Inhalte des Kerndatensatzes mit Modifikationen und Erweiterungen.

Zur Vorbereitung dieser Empfehlungen hat der Wissenschaftsrat im Mai 2012 eine ad hoc Arbeitsgruppe einberufen. Mitgewirkt haben in ihr auch Sachverständige, die nicht Mitglieder des Wissenschaftsrates sind. Ihnen weiß sich der Wissenschaftsrat zu besonderem Dank verpflichtet. Ebenso dankt der Wissenschaftsrat weiteren Sachverständigen, die den Beratungsprozess im Rahmen von Anhörungen und Gesprächen konstruktiv unterstützt haben.

Der Wissenschaftsrat hat die vorliegenden Empfehlungen am 25. Januar 2013 in Berlin verabschiedet.

---

# Kurzfassung

Sammlung, Vorhaltung und Bereitstellung von Daten zu Forschungsaktivitäten gehören inzwischen zu den festen Aufgaben von Hochschulen und außeruniversitären Forschungseinrichtungen (AUF). Diese Daten werden sowohl für interne Steuerungsprozesse, für die Berichterstattung an Mittelgeber und die amtliche Statistik als auch für die Bewertung von Forschungsleistungen durch externe Akteure benötigt. Eine Vielzahl von institutionellen Bewertungsverfahren wie Evaluationen, Akkreditierungen, Rankings oder Ratings stellen damit viele forschende Einrichtungen, wie auch Wissenschaftlerinnen und Wissenschaftler, vor große Herausforderungen.

Mit dieser Empfehlung stößt der Wissenschaftsrat einen Prozess zur Standardisierung von Datenformaten zu Forschungsleistungen an, die den Datenaustausch zwischen verschiedenen Akteuren unterstützen soll. Neben den Inhalten eines Kerndatensatzes Forschung beschreibt sie das Lastenheft eines Projekts, in dem dieser Kerndatensatz zeitnah in allen Details spezifiziert werden soll.

In einem ersten Schritt soll diese Spezifikation den Austausch von Daten zu Beschäftigten, Nachwuchsförderung, Drittmitteln und Projekten, Forschungspreisen und Auszeichnungen, Patenten sowie Publikationen regeln. Für mögliche darüber hinausgehende Erweiterungen werden Vorschläge vorgelegt.

Neben der Klärung von Definitionen soll im Rahmen der Standardisierung auch eine Verständigung über fachspezifische Formate zum Austausch von Publikationsdaten erreicht werden. Ein weiterer Teil des Projektes ist die Erarbeitung einer einheitlichen und konsensfähigen Fächerklassifikation, die eine inhaltlich aussagekräftige Verortung von Forschungsaktivitäten zu einzelnen Fächern ermöglicht aber auch flexibel ist, um interdisziplinäre Entwicklungen darstellen zu können.

Die operative Umsetzung des Projektes begleitet der Wissenschaftsrat durch eine Arbeitsgruppe, die als Beirat agiert.

---

# A. Anlass

## A.1 WACHSENDER INFORMATIONSBEDARF ZU FORSCHUNGSAKTIVITÄTEN

---

Die Berichterstattung über Forschungsaktivitäten, auch für die Bewertung von Leistungen, setzt qualitätsgesicherte und vergleichbare Daten voraus. Die Erfassung aussagekräftiger Daten wurde bislang allerdings durch das Fehlen klarer Standards behindert. Angesichts der weiter wachsenden Bedeutung, Forschungsaktivitäten in deren Qualität und Quantität – auch mit dem Ziel der Steuerung – zu dokumentieren, ist ein nachhaltiger Schritt zur Qualitätssicherung über eine Standardisierung der einschlägigen Datenformate daher unverzichtbar. Mittlerweile liegen vielfältige Erfahrungen aus dem Forschungsrating des Wissenschaftsrates und anderen Bewertungsverfahren vor, auf die eine solche Initiative zur Standardisierung aufbauen kann.

Bereits vor über drei Jahrzehnten hat sich der Wissenschaftsrat in seinen Empfehlungen zur Forschungsberichterstattung der Hochschulen dafür ausgesprochen, „[...] Forschungsberichte ohne unmäßigen Zeit- und Kostenaufwand so abzufassen, dass sie dank einer gewissen Einheitlichkeit einen überregionalen Vergleich ermöglichen [...].“<sup>2</sup> Gerade wegen gestiegener Anforderungen an die Berichtssysteme der Hochschulen hat diese Forderung auch heute noch Gültigkeit. Die vorliegenden Empfehlungen folgen auch den „Empfehlungen zur Bewertung und Steuerung von Forschungsleistungen“. |<sup>3</sup> Mit ihnen hat sich der Wissenschaftsrat für wenige, robuste Indikatoren zur Bewertung von Forschungsaktivitäten ausgesprochen. Grundlage dafür sind standardisierte und qualitätsgesicherte Ausgangsdaten.

Der Status Quo ist vor allem durch die Vielfalt und Heterogenität der Informationsanfragen an Hochschulen und außeruniversitäre Forschungseinrichtungen (AUF) gekennzeichnet. Verlässliche Zahlen zum Umfang der Berichtsbelastun-

<sup>2</sup> Wissenschaftsrat: Empfehlung zur Forschungsberichterstattung der Hochschulen, Köln 1980, S. 2.

<sup>3</sup> Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011.

gen liegen allerdings nicht vor. Typischerweise treten neben Datenanforderungen für die interne Steuerung (z.B. Mittelverteilung und Leistungsbemessung in der W-Besoldung) auch Anfragen der Wissenschaftsministerien der Länder, der amtlichen Statistik, des Wissenschaftsrates, der Akkreditierungsagenturen, des Centrums für Hochschulentwicklung (CHE) oder der Träger nationaler und internationaler Rankings und Benchmarks. |<sup>4</sup> Dabei werden gleiche Informationen oft in unterschiedlicher Form abgefragt (Drittmittel z.B. als Einnahmen oder Ausgaben). Ferner variieren Art und Umfang der gewünschten Daten. So haben z.B. einige Länder die Berichtsanforderungen an ihre Hochschulen wiederholt angepasst.

Neben der zunehmend umfangreichen Berichtsbelastung zeitigt die bestehende Heterogenität der Datenabfragen eine weitere Folge: die auf unterschiedliche Weise erhobenen Daten sind nur eingeschränkt vergleichbar und aggregierbar, wodurch auch die Betrachtung von Entwicklungen über längere Zeiträume hinweg erschwert wird.

Da auch die Pilotstudie Forschungsrating des Wissenschaftsrates als ein Teil der in den letzten Jahren gewachsenen Berichtsanforderungen an Hochschulen und außeruniversitären Forschungseinrichtungen gelten kann, sieht sich der Wissenschaftsrat besonders mit Blick auf eine mögliche Verstetigung seines Forschungsratings in der Pflicht, den Erhebungsaufwand zu reduzieren und ein hohes Maß an Qualität und Wiederverwendbarkeit der Daten zu gewährleisten. Angesichts der genannten Defizite, insbesondere der vielen, nicht abgestimmten Datenabfragen, sieht sich der Wissenschaftsrat in seiner Doppelrolle als Akteur und Beratungsgremium auch aufgefordert, mit diesen Empfehlungen die Initiative zu einer Standardisierung und Implementierung eines Kerndatensatzes zu ergreifen.

Diese Empfehlung stellt die Grundlage für die Entwicklung eines Kerndatensatzes zur Dokumentation von Forschungsaktivitäten dar, aus dessen Daten sich Indikatoren zur Bewertung von Forschungsaktivitäten durch die Datenabfrager ableiten lassen. |<sup>5</sup> Kennzahlen, die sich aus diesen Daten berechnen und zur Bewertung von Forschungsleistungen heranziehen lassen, sind selbst nicht Teil des Kerndatensatzes.

Bei den Elementen des Kerndatensatzes handelt es sich überwiegend um solche Daten, die zum einen an Hochschulen und AUF in der Regel bereits vorliegen.

|<sup>4</sup> Die Erhebungsmethodik internationaler Rankings kann durch eine nationale Standardisierung von Datenformaten allerdings nicht beeinflusst werden.

|<sup>5</sup> Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, S. 39.

Zum anderen sind es Daten, die von vielen datenabfragenden Institutionen angefordert werden. Die Datenabfrager weist der Wissenschaftsrat darauf hin, dass mit den Datenformaten des Kerndatensatzes ein Set an Daten definiert ist, das die wesentlichen Informationsinteressen abdecken sollte. Er fordert sie dazu auf über den Kerndatensatz hinausgehende Anfragen nur in gut begründeten Ausnahmefällen zu stellen.

Der Wissenschaftsrat fordert Hochschulen und AUF auf, die im Folgenden genannten Daten für Vergleichszwecke zugänglich zu machen. In den vergangenen zwei Jahrzehnten haben Forschungseinrichtungen ein hohes Maß an zusätzlicher Autonomie gewonnen. Daraus ergibt sich auch die Notwendigkeit einer Auskunftsfähigkeit über die eigenen Forschungsaktivitäten, sowohl nach Außen wie auch nach Innen.<sup>6</sup>

Der Kerndatensatz ist ein Angebot an Hochschulen und AUF, das ihre häufig bereits bestehenden Aktivitäten bei der informationstechnischen Erfassung ihrer Forschungsaktivitäten unterstützen und nicht ersetzen soll. Insbesondere sollen der begrenzte Umfang und die Datenformate des Kerndatensatzes die Grundlage für eine vereinfachte Beantwortung von Datenabfragen. Die Integration des Kerndatensatzes in das einrichtungsinterne Informationsmanagement soll künftig nicht nur einen elektronischen Datenaustausch, sondern vielmehr auch eine weitestgehend rechnerunterstützte Abwicklung von Datenanfragen ermöglichen.

## **A.II ENTWICKLUNGEN**

---

Ein abgestimmter und einheitlicher Umgang mit dem dargestellten wachsenden Informationsbedarf ist noch nicht zu beobachten. Es zeichnen sich insbesondere zwei sich ergänzende Antwortstrategien datenliefernder Forschungseinrichtungen ab: (i) das Bemühen um eine Strukturierung und Standardisierung von Informationen und (ii) der Aufbau von Forschungsinformationssystemen zur Dokumentation von Forschungsaktivitäten und zur weitgehenden und systematischen Unterstützung des Berichtswesens.

Standardisierungsbemühungen zielen darauf ab, einheitliche Definitionen von Daten und Datenformaten festzulegen. Forschungsinformationssysteme stellen einen zunächst technischen Ansatz dar. Mit ihnen verfolgen Hochschulen und AUF das Ziel, Daten z.B. zu Drittmitteln, Personal und Publikationen struktu-

<sup>6</sup> Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, S. 33.

riert vorzuhalten, um einfach und schnell eine Vielzahl unterschiedlicher Berichte zu generieren oder aber auch Serviceleistungen wie die automatisierte Erstellung von Publikationslisten auf den Homepages der Forschungsgruppen bereit zu stellen.

In Deutschland setzt sich neben anderen Akteuren auch eine Arbeitsgruppe der Deutschen Initiative für Netzwerkinformationen (DINI) |<sup>7</sup> für eine Standardisierung von Forschungsdaten ein. Auf europäischer Ebene hat die EU-Kommission seit den frühen 1990er Jahren die Entwicklung des *Common European Research Information Format* (CERIF) |<sup>8</sup>, eines gemeinsamen europäischen Datenaustauschformates für Informationen zu Forschungsaktivitäten, gefördert. Mehrere schon heute an deutschen Hochschulen im Einsatz befindliche Systeme nutzen dieses Datenformat. Auch orientieren sich führende kommerzielle Lieferanten von Publikationsdaten zunehmend an CERIF.

In einem weiteren, für die Abbildung von Forschungsaktivitäten zentralen Bereich, der Klassifikation von Fächern, gibt es ebenfalls Standardisierungsbestrebungen auf internationaler Ebene, wie z.B. das *Frascati Manual* der OECD oder die ISCED der UNESCO.

Die steigende Mitgliederzahl bei EuroCRIS (*European Current Research Information Systems*), der für die Weiterentwicklung von CERIF verantwortlichen Organisation, zeigt das zunehmende internationale Interesse an der Standardisierung von Forschungsinformationen. Der CERIF-Standard wird mittlerweile von Institutionen in über 40 Ländern mitgetragen. Im Jahr 2011 hat EuroCRIS seine Mitgliederzahl gegenüber dem Vorjahr um nahezu ein Drittel ausgebaut. |<sup>9</sup>

Die skizzierten Veränderungen schlagen sich auch in einen Wandel der Einschätzung dazu nieder, welche Bezugsgrößen für aussagekräftige Daten zu Forschungsaktivitäten relevant sind. Während traditionell vor allem Informationen über Fächer oder Einrichtungen – z. B. in der amtlichen Statistik – geliefert wurden, rücken zunehmend die Vorteile flexibel aggregierbarer Datenbestände in den Vordergrund. Daten, die bei den Einrichtungen auf einer möglichst niedrigen Aggregationsebene vorliegen, können mit Hilfe der von Forschungsinformationssystemen angebotenen Funktionalität flexibel und nutzungsspezifisch zusammengefasst werden, um somit eine jeweils angepasste und inhaltlich aussagekräftige Dokumentation zu ermöglichen.

|<sup>7</sup> <http://www.dini.de/ag/fis/> vom 18.12.2012.

|<sup>8</sup> <http://www.eurocris.org/Index.php?page=CERIFreleases&t=1> vom 18.12.2012.

|<sup>9</sup> EuroCRIS hatte 2011 damit 97 institutionelle, 28 individuelle und 18 angegliederte Mitglieder. EuroCRIS. Annual Report 2011, [http://www.eurocris.org/Index.php?page=annual\\_report&t=1](http://www.eurocris.org/Index.php?page=annual_report&t=1), vom 23.11.2012.

Die wachsenden Datenbestände zu Forschungsaktivitäten und deren Bedeutung auch für einzelne Wissenschaftlerinnen und Wissenschaftler führen zu der Frage, wie die individuellen Aktivitäten bei gelegentlichen institutionellen Wechseln dokumentiert werden können. In diesem Zusammenhang wird eine eindeutige Kennung für Wissenschaftlerinnen und Wissenschaftler, die so genannte Forscher- oder Autoren-ID, zum Teil bereits verwendet. Die Niederlande haben beispielsweise mit der *Digital Author Identification* ein solches System bereits eingeführt. Solche Systeme erlauben nicht nur die eindeutige Zuordnung von Forschungsaktivitäten zu Wissenschaftlerinnen und Wissenschaftlern, sie sollen es auch ermöglichen, dass diese die über sie gespeicherten Daten beim Wechsel von einer Einrichtung zur nächsten mitnehmen können.

Der Wissenschaftsrat sieht in einer Kombination der Standardisierung von Datenformaten und der Einführung von Forschungsinformationssystemen den geeignetsten Lösungsansatz für einen professionellen Umgang mit Informationen zu Forschungsaktivitäten. Sie bietet auch die Möglichkeit der Erweiterung, um längerfristig auch Informationen über die Aktivitäten in der Lehre zu integrieren. Mit Hilfe entsprechender Softwaresysteme reduzieren zunehmend mehr Hochschulen und AUF ihren Aufwand für die interne und externe Berichterstellung. Dies gilt sowohl für die gesetzlich vorgeschriebene Weitergaben von Daten an die statistischen Ämter oder an den Bundesbericht Forschung als auch für die Beantwortung weiterer externer Datenabfragen. Berichte aus der Praxis illustrieren, dass solche Systeme nach einer Anlaufphase erhebliche Erleichterungen mit sich bringen. |<sup>10</sup> Nach einer ersten Schätzung, basierend auf Umfragen bei Softwareherstellern und Branchenkennern, wird deutlich, dass Forschungsinformationssysteme zunehmend als die präferierte Lösung angesehen werden. Demnach gibt es über 26 Projekte zur Einführung eines solchen Systems mit durchaus verschiedenen Zuschnitten. Einige Systeme sind auf einzelne Institute einer AUF begrenzt, andere decken ganze Hochschulen ab. Seit 2006 ist in diesem Bereich eine große Dynamik zu verzeichnen. Gut die Hälfte dieser Systeme wurde seit 2011 in Betrieb genommen oder befindet sich derzeit noch im Aufbau. Zwischen dem Beginn der Planungen und dem Beginn der Realisierung eines Forschungsinformationssystems vergehen in der Regel ca. drei Jahre. Somit ist davon auszugehen, dass eine substantielle Anzahl weiterer Systeme derzeit geplant wird. Viele dieser Systeme bieten auch einen Mehrwert für die Wissenschaftlerinnen und Wissenschaftler einer Einrichtung, indem sie Publikationslisten und weitere Einträge für die individuelle Homepa-

|<sup>10</sup> Vgl. Bittner, S.; Hornbostel, S.; Scholze, F.: Forschungsinformationen in Deutschland. Anforderungen, Stand und Nutzen existierender Forschungsinformationssysteme, Berlin, iFQ-Working Paper No. 10, 2012.

ge bereitstellen. Damit wächst nicht zuletzt auch die Akzeptanz der Systeme in den Einrichtungen.

### A.III ZWECK EINES KERNDATENSATZ FORSCHUNG

---

#### III.1 Systemweite Datenbasis

Mit diesen Empfehlungen unterstreicht der Wissenschaftsrat die Notwendigkeit von institutionellen Vergleichen im Wissenschaftssystem. Die Grundlage für Vergleiche bilden belastbare Daten. Ohne qualitätsgesicherte Informationen über die Forschungsaktivitäten kann sich eine Einrichtung nicht selbst, noch können andere sie objektiv und nachvollziehbar einordnen.

Der unmittelbare Zweck des Kerndatensatz Forschung ist es, die Weitergabe vergleichbarer Daten zu ermöglichen. Daher ist eine Standardisierung von Datenformaten auf nationaler Ebene unverzichtbar. Der Kerndatensatz standardisiert Datenformate für den *Vergleich* von Forschungsaktivitäten zwischen (i) verschiedenen Einheiten eines Faches und nach einer geeigneten *Aggregation* auf einer höheren Ebene auch (ii) zwischen institutionellen Einheiten wie z. B. ganzer Hochschulen. |<sup>11</sup> Damit entstehen belastbare Daten, die von den verschiedenen Akteuren des Wissenschaftssystems in Deutschland für Analyse- und Vergleichszwecke (Benchmarking) eingesetzt werden können. Auch können Auswertungen auf Basis dieser Daten wertvolle Informationen für Studierende liefern. Für die Wissenschaftlerinnen und Wissenschaftler wird mit dem Kerndatensatz außerdem ein hohes Maß an Transparenz geschaffen, welche Daten für die Erfassung der eigenen Forschungsaktivitäten herangezogen werden.

Mittelbar erfüllt der Kerndatensatz Forschung dadurch auch die nachfolgend beschriebenen Zwecke.

#### III.2 Qualitätssicherung

Durch die Standardisierung beseitigt der Kerndatensatz definitorische Unsicherheiten |<sup>12</sup>, die heute häufig die Bereitstellung von Daten erschweren. Erreicht wird dadurch eine verbesserte Datenqualität. Zusätzlich entsteht durch die verbreitete Nutzung des Kerndatensatzes auch ein starker Anreiz zur Quali-

| <sup>11</sup> Nicht alle fachspezifischen Angaben lassen sich allerdings sinnvoll über mehrere Fächer aggregieren (z.B. bei unterschiedlichen Drittmittelkulturen).

| <sup>12</sup> So werden bei Drittmitteln teilweise Einnahmen und teilweise Ausgaben abgefragt.

tätskontrolle der Dateninhalte durch die Datenlieferanten auf allen Ebenen – von den Wissenschaftlerinnen und Wissenschaftlern bis zu den Einrichtungen.

### III.3 Interpretierbarkeit

Klar definierte Datenformate erleichtern auch die Interpretation der Daten und erhöhen damit ihre Aussagekraft. Nur mit semantisch eindeutigen Daten lassen sich belastbare Vergleiche über die Zeit durchführen. Nur auf einer solchen soliden Grundlage können verlässliche Indikatoren entwickelt werden, um die Bewertung von Forschungsleistungen durch die Datenabfrager zu unterstützen.

### III.4 Begrenzung des Aufwandes

Durch die Standardisierung reduziert sich der Erhebungsaufwand, da Daten unabhängig von einer speziellen Anfrage vorgehalten werden. Standardisierte Daten erlauben bei einer weitgehend anfrageunabhängigen Definition die Mehrfachnutzung für verschiedenste Anwendungen. Die längerfristige Gültigkeit des Standards vorausgesetzt entfallen auch variierende Anforderungen seitens der Datenabfrager. Spätestens nach einer überschaubaren Anlaufphase können Hochschulen und AUF auf den Kerndatensatz verweisen, um damit – bis auf wenige, gut begründete Ausnahmefälle – nicht mit dem Kerndatensatz kompatible Datenabfragen zurückzuweisen.

### III.5 Begrenzung der Abfragen

Da über den Kerndatensatz die wesentlichen Informationen über die Forschungsaktivitäten einfach bei Hochschulen und AUF abgefragt werden können, werden sich die Daten des Kerndatensatzes als Basis für die Bildung von Indikatoren zur Erfassung und Bewertung von Forschungsleistungen etablieren. Damit werden sich Notwendigkeit und Ausmaß zusätzlicher Datenerhebungen und -abfragen reduzieren.

## **A.IV GEGENSTAND DER EMPFEHLUNG**

---

Nach einigen grundlegenden Hinweisen zum Kerndatensatz Forschung in Abschnitt B werden verschiedene Empfehlungen u. a. zu dessen Inhalten in Abschnitt C formuliert. Der in Abschnitt D dargestellte Prozess ist darauf ausgerichtet, gemeinsam mit allen relevanten Akteuren tragfähige Definitionen der Elemente des Kerndatensatzes zu vereinbaren und Formate für den Austausch der Daten festzulegen. Teil dieses Prozesses ist außerdem die parallele Erarbeitung einer neuen, konsensfähigen Fächerklassifikation. Schließlich sollen im Rahmen dieses Prozesses stufenweise Formate für Publikationsdaten – die fachspezifischen Besonderheiten immer im Blick habend – standardisiert

werden, um eine belastbare Datenbasis zu diesem zentralen Produkt der Forschungstätigkeit zu schaffen.

Beim Kerndatensatz handelt es sich nicht um ein Informationssystem. Damit können bestehende technische Systeme zur Berichterstellung beibehalten werden.

Mit dem Kerndatensatz schafft der Wissenschaftsrat die Grundlage, Forschungsaktivitäten in Deutschland zu dokumentieren, über sie zu informieren, zwischen ihnen zu vergleichen, sie zu analysieren und sie zu reflektieren. Der Kerndatensatz sollte von allen in Deutschland operierenden Hochschulen (nach Maßgabe der jeweiligen Landeshochschulgesetze) und allen außeruniversitären Forschungseinrichtungen bereitgestellt werden können.

---

# B. Grundlegungen

## B.1 GRUNDSÄTZE

---

Der Wissenschaftsrat erklärt folgende Aussagen zu bestimmenden Grundsätzen für die Standardisierung der Spezifikation und für die spätere Arbeit mit dem Kerndatensatz Forschung.

### I.1 Übergreifende Prinzipien

Ziel des Kerndatensatzes ist es, den Austausch von aussagekräftigen Informationen über Forschungsaktivitäten zu ermöglichen.

Der Kerndatensatz standardisiert Datenformate, die für die Dokumentation oder Kontextualisierung von Forschungsaktivitäten bedeutsam sind. Die Daten des Kerndatensatzes können als Grundlage für die Bildung von Indikatoren für die Erfassung und Bewertung von Forschungsleistung dienen. Eine solche Interpretation der Daten über Forschungsaktivitäten einschließlich der von ihnen abgeleiteten Leistungsindikatoren obliegt ihren Nutzern.

Der Kerndatensatz schafft keine neuen Veröffentlichungspflichten.

Die Definition des Kerndatensatzes ist dem Prinzip der Datensparsamkeit verpflichtet. Der Kerndatensatz enthält nur solche Daten, die eindeutig definiert und für eine große Anzahl an Fächern erhoben werden können.

Mit der Spezifizierung von Elementen des Kerndatensatzes soll ausdrücklich keine Priorisierung einzelner Aspekte für die Bewertung von Forschungsleistung getroffen werden. Dem Wissenschaftsrat ist dabei allerdings bewusst, dass allein schon die Auswahl der in den Kerndatensatz aufzunehmenden Daten als Wertung verstanden werden kann. So sind beispielsweise Patente ein Bestandteil des Kerndatensatzes. Werden Patente als Hinweis auf Transferleistungen von Forschung aufgefasst, so wären damit automatisch solche Fächer besser dokumentiert, deren Transferleistungen sich über Patente beschreiben lassen, während Interviews in Radio und Fernsehen, die z. B. in den Geistes- und Sozialwissenschaften einen gängigen Weg für Transfer darstellen, nicht Teil des Kerndatensatzes sind. Auch wenn die über den Kerndatensatz standardisierten

Daten für eine breite Zahl von Fächern aussagekräftig sind, lässt sich eine derartige, implizite Wertung nicht vollständig vermeiden. Der Wissenschaftsrat fordert daher alle Nutzer des Kerndatensatzes auf, vor dem Einsatz der Daten, deren Zweckmäßigkeit für die beabsichtigte vergleichende Bewertung einschließlich der erwünschten wie auch unerwünschten Anreizeffekte zu prüfen. |<sup>13</sup>

Über den Kerndatensatz hinaus gibt es eine Vielzahl von Daten zu Forschungsaktivitäten, die entweder für eine kleinere Zahl von Fächern relevant sind oder nicht eindeutig definiert werden können. Es steht jeder Einrichtung frei, solche zusätzlichen (fachspezifischen) „extranuklearen“ Daten zu erfassen. Für einige Erweiterungsoptionen werden in Kapitel C.IV Empfehlungen gegeben. Sie sind aber zunächst nicht Teil der Spezifikation des Kerndatensatzes.

Viele forschungsrelevante Daten, wie insbesondere solche zu Publikationen, liegen für einige Fächer, teilweise auch für ganze Einrichtungen, bereits in kommerziellen oder hauseigenen Datenbanken vor. Bestehende Daten sollen nach Möglichkeit für die Bereitstellung im Kerndatensatz erschlossen werden. Der zusätzliche Erhebungsaufwand wird so für die Einrichtungen minimiert.

Die Definitionen der Datenelemente für den Kerndatensatz müssen so ausgelegt sein, dass jedes Datum nur einmal in den Kerndatensatz eingeht.

Um die internationale Anschlussfähigkeit der Daten des Kerndatensatzes zu gewährleisten sollen alle Angaben, die für eine Verschlagwortung einschlägig sind, sowohl in deutscher als auch in englischer Sprache erfolgen.

## 1.2 Prinzip der Datenhoheit und des Datenschutzes

Jede Einrichtung erhebt und verwaltet ihre eigenen Informationen zu Forschungsaktivitäten.

Der Umgang mit den Daten muss im Einklang mit bestehenden Anforderungen des Datenschutzes stehen.

## 1.3 Prinzipien der Datenvorhaltung

Der Wissenschaftsrat weist darauf hin, dass der Kerndatensatz ein Format für die *Weitergabe* von Daten darstellt. Die den Kerndaten zugrunde liegenden „Grunddaten“ können auf unterschiedliche Art und Weise modelliert und gespeichert werden, um sie für eine spezifische Datenabfrage bedarfsgerecht

|<sup>13</sup> Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, S. 34.

aufzubereiten. So können beispielsweise die Namen aller Wissenschaftlerinnen und Wissenschaftlern und deren fachliche Zuordnung einer Einrichtung gespeichert werden, um daraus die Zahl der Wissenschaftlerinnen und Wissenschaftler in einem Fach als Datum des Kerndatensatzes zu ermitteln. Den Hochschulen und AUF wird empfohlen, ihre Daten in möglichst elementarer Form vorzuhalten, um eine breite Wiederverwendbarkeit zu gewährleisten und flexibel sowie mit überschaubarem Aufwand auf künftige interne oder externe Anfragen reagieren zu können.

Erfahrungen aus der Praxis zeigen, dass Daten in besserer Qualität bei niedrigeren Kosten erhoben werden können, wenn dies im Rahmen der üblichen administrativen Prozesse geschieht. Beispielhaft sei die Erfassung von Projektdaten bei der Einrichtung eines Kontos der Kostenstelle durch die Finanzbuchhaltung genannt. Entsprechend sollten die Hochschulen und AUF die Datenerhebung in die bestehenden administrativen Prozesse integrieren.

Die Verantwortung für die Bereitstellung der Daten sollte bei einer zentralen Stelle innerhalb einer Einrichtung liegen.

Sofern Daten an Einrichtungen bereits nach dem CERIF-Standard vorgehalten werden, sind sie mit dem Kerndatensatz kompatibel.

Der Datenbestand ist regelmäßig, aber *mindestens* einmal jährlich zum Stichtag 31.12. zu aktualisieren.

#### 1.4 Prinzipien der Qualitätssicherung

Jede Einrichtung implementiert und verantwortet ein System zur regelmäßigen Qualitätssicherung der im Kerndatensatz enthaltenen Daten sowie der dazugehörigen Erfassungsprozesse.

Die Verwendung des Kerndatensatzes an sich sowie die Kommunikation der Nutzer über die Daten des Kerndatensatzes stellen wichtige Teile der Qualitätssicherung dar.

Alle Nutzer der Daten des Kerndatensatzes sind aufgefordert, zur Qualitätssicherung beizutragen. Fallen bei einer Nutzung der Daten Inkonsistenzen oder Fehler auf, so sollten diese den Einrichtungen zurückgemeldet werden.

Ebenfalls sollten mögliche Fehler der Spezifikation an den zuständigen Beirat (s. D.I) zurückgemeldet werden.

#### 1.5 Prinzipien des Datenaustausches

Bei der Beantwortung einer Datenabfrage übermittelt die angefragte Einrichtung den Kerndatensatz für ein Fach oder für mehrere Fächer im Regelfall über

ein definiertes Austauschformat auf elektronischem Wege. Somit wird eine maschinengestützte Weiterverarbeitung der Daten erleichtert.

Der Kerndatensatz soll den Aufwand bei der Erhebung und Bereitstellung von Daten begrenzen. Daher ist der Austausch über Softwareschnittstellen ein wesentliches Element. Ihre Entwicklung und Implementierung soll als Teil des Standardisierungsprozesses geprüft werden.

## **B.II AUSWAHL DER INHALTE UND FESTLEGUNG DER DEFINITIONEN**

---

Die vorliegende Empfehlung steckt einen Rahmen für die Erarbeitung einer Spezifikation des Kerndatensatzes Forschung ab. Sie spricht sich auch für bestimmte Inhalte des Kerndatensatzes aus.

Die Auswahl dieser Inhalte erfolgte mit dem Ziel, durch die Daten aussagekräftige Informationen über Forschungsaktivitäten und deren Produkte zu erhalten. Die Auswahl basiert auf den Erfahrungen aus der Pilotstudie Forschungsrating sowie auf Konsultationen mit Experten. Sie orientiert sich daher am Stand der Technik und zielt auf eine hohe Aussagekraft der Daten des Kerndatensatzes zu zentralen Aspekten der Charakterisierung von Forschungsaktivitäten ab. |<sup>14</sup> Die Elemente des Kerndatensatzes (vgl. Glossar, S. 57) betreffen die Bereiche Beschäftigte, Nachwuchsförderung, Drittmittel, Forschungspreise und andere Auszeichnungen, Patente sowie Publikationen.

Für die operative Umsetzung müssen die Elemente des Kerndatensatzes klar definiert und ggf. abgegrenzt werden. Diese semantische Klärung geschieht im Rahmen der Standardisierung des Kerndatensatzes (vgl. Abschnitt D). Beispielsweise ist es Aufgabe einer Projektgruppe zu klären, wie umfassend die Informationen zu Drittmittelprojekten zu erfassen sind und ob beispielsweise aus Landesmitteln außerhalb des Grundhaushaltes finanzierte Projekte zu dieser Kategorie zu zählen sind.

## **B.III PRAGMATISCHE ABWÄGUNGEN**

---

Mit dem Kerndatensatz unternimmt der Wissenschaftsrat einen Vorstoß in einem Bereich, auf dem bislang sowohl die Vorgehensweisen der verschiedenen Hochschulen und AUF als auch deren technische Ausstattung äußerst hetero-

|<sup>14</sup> Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, S. 39-41.

gen sind. Des Weiteren hat sich das Feld der Forschungsberichterstattung in den letzten Jahren sehr dynamisch entwickelt und demonstriert schon heute die Vorteile flexibel aggregierbarer Informationen (vgl. A.II). Der Wissenschaftsrat nimmt die Zielspannungen zwischen inhaltlich Wünschenswertem und technisch Machbaren, die sich aus dieser Situation ergeben, ernst. Im Bestreben, einer für alle Beteiligten gangbaren Lösung, schlagen diese Empfehlungen einen pragmatischen Kompromiss vor. Die wichtigsten Abwägungen sollen im Folgenden dargestellt werden.

Nicht nur die in jüngerer Zeit sich stark verändernden Anforderungen an die Erfassung der Daten zu Forschungsaktivitäten haben deutlich gemacht, dass sich ein zukunftssicheres Dokumentations- und Berichtssystem vor allem durch Flexibilität und Anpassbarkeit auszeichnen muss. Auch neue Formen der Forschungsk Kooperationen, z. B. im Rahmen von Forschungsclustern oder von anderen interdisziplinär koordinierten Forschungsprojekten zeigen die hohe Dynamik heutiger Forschung. Sie belegen somit, dass die Dokumentation von Forschungsaktivitäten dementsprechend flexibel anpassbar sein muss. Flexibilität bedeutet in diesem Kontext in erster Linie, Daten je nach beabsichtigter spezifischer Nutzung unterschiedlich aggregieren zu können. Dazu ist es notwendig, dass Informationen bei den Einrichtungen in möglichst elementarer Form vorliegen. Informationen über Publikationen oder Forschungsprojekte von Wissenschaftlerinnen und Wissenschaftlern lassen sich nur so flexibel beispielsweise für ein Institut, ein Fach oder ein interdisziplinäres Forschungsfeld zusammenfassen. Eine derartige Datenvorhaltung führt zu einer nachhaltigen Aufwandsersparnis, da alle Daten unabhängig von einer späteren Nutzung nur einmal erfasst werden müssen.

Konsequent wäre es ein Datenmodell festzulegen, das elementare Daten und ihre Beziehungen untereinander vorhält. Allerdings ist sich der Wissenschaftsrat bewusst, dass eine solche Forderung für viele forschende Einrichtungen kurzfristig eine personelle, technische und finanzielle Überforderung darstellen würde. Daher ist der Kerndatensatz als Datenaustauschformat angelegt, das unabhängig von der Art der Datenvorhaltung in bestehenden oder geplanten, einrichtungsspezifischen Informationssystemen verwendet werden kann. Für die Weitergabe der Daten muss eine Abbildung von den Datenformaten der Einrichtung auf die Datenformate des Kerndatensatzes vorgenommen werden. Diese Abbildung ist umso einfacher, je ähnlicher sich die dem Kerndatensatz und den Informationssystemen der Einrichtungen zugrunde liegenden Datenformate sind. Der Kerndatensatz bezieht sich außerdem immer auf die Aggregationsebene eines Faches (s. Glossar, S. 57), um damit eine größtmögliche Vergleichbarkeit zu gewährleisten. Grundlage für die Zuordnung der personenbezogenen Forschungsaktivitäten ist die im Zuge der Standardisierung zu erarbeitende Fächerklassifikation (vgl. Kapitel D.II.1).

Angesichts der Anlage des Kerndatensatzes als Austauschformat weist der Wissenschaftsrat ausdrücklich darauf hin, dass für die Bereitstellung der Daten des Kerndatensatzes keine spezielle technische Ausstattung nötig ist, wenn die Abbildung aus den in der Einrichtung gehaltenen Daten händisch erfolgt. Somit sollten alle Hochschulen und AUF in der Lage sein, Informationen zu ihren Forschungsaktivitäten nach der Spezifikation des Kerndatensatzes ohne weitere Aufwendungen mit einem geringeren Aufwand als bisher bereitzustellen. Eine weitere Reduktion des Aufwandes kann durch die rechnergestützte Abbildung zwischen eigener Datenhaltung und Kerndatensatz erreicht werden.

Obwohl der Wissenschaftsrat die praktischen Restriktionen anerkennt, erachtet er es für unverzichtbar, dass neue informationstechnische Lösungen der Datenhaltung an forschenden Einrichtungen sich an einem mit dem CERIF-Standard kompatiblen Datenmodell ausrichten. Damit sind die Einrichtungen auf künftige Entwicklungen und gegebenenfalls bedarfsgerechte Erweiterungen gut vorbereitet.

---

# C. Empfehlungen

## **C.1 EMPFEHLUNGEN ZUM UMGANG MIT INFORMATIONEN ZU FORSCHUNGSAKTIVITÄTEN**

---

Die Vorhaltung von Informationen zu Forschungsaktivitäten und ihrem Kontext ist in den verschiedenen Hochschulen und AUF in unterschiedlicher Art und Weise organisiert. Der Wissenschaftsrat ist sich bewusst, dass derzeit eine große Zahl von Einrichtungen daran arbeitet, bestehende Datenquellen in so genannte Forschungsinformationssysteme zu integrieren. |<sup>15</sup> Forschungsinformationssysteme können vor allem Berichts- und Arbeitsprozesse erheblich vereinfachen und tragen zu einer verbesserten Datenqualität bei. Auch für Wissenschaftlerinnen und Wissenschaftler ist die Pflege der damit verbundenen Daten nützlich, da sie sich z. B. für aktuelle Literaturverzeichnisse nutzen lassen. Ebenso lässt sich die Erstellung von Forschungsberichten weitgehend unterstützen. Um angesichts der beachtlichen Komplexität solche Systeme zukunftssicher zu gestalten, gibt der Wissenschaftsrat folgende Empfehlungen:

Jedes Forschungsinformationssystem sollte auf einem expliziten Datenmodell aufbauen. |<sup>16</sup> Nur so lässt sich eine flexible Anpassung an mögliche zukünftige Datenerfordernisse sicherstellen. Ein Vorschlag für ein solches Datenmodell wird im Zuge der Spezifikation des Kerndatensatzes zu entwickeln sein. Auch wenn der Kerndatensatz selbst ein Datenaustauschformat spezifiziert, kann ein Datenmodell auch leitend für die Realisierung eines in der Planung befindlichen Forschungsinformationssystems einer Einrichtung sein.

| <sup>15</sup> Bittner, S.; Hornbostel, S; Scholze, F.: Forschungsinformationen in Deutschland. Anforderungen, Stand und Nutzen existierender Forschungsinformationssysteme, Berlin, iFQ-Working Paper No. 10, 2012.

| <sup>16</sup> Vergleiche beispielsweise: Herwig, S.; Becker, J.: Einführung eines Forschungsinformationssystems an der Westfälischen Wilhelms-Universität Münster. Von der Konzeption bis zur Implementierung in: Bittner, S.; Hornbostel, S; Scholze, F. (Hrsg.): Forschungsinformationen in Deutschland. Anforderungen, Stand und Nutzen existierender Forschungsinformationssysteme, Berlin, iFQ-Working Paper No. 10, 2012, S. 46.

Es ist absehbar, dass die Notwendigkeit des Austausches von Informationen zu Forschungsaktivitäten zwischen forschenden Einrichtungen, aber auch zwischen Forschenden und Forschungsförderern, national aber auch auf europäischer und internationaler Ebene zunehmen wird. Der Wissenschaftsrat rät daher dazu, Informationen zu Forschungsaktivitäten nach dem auf europäischer Ebene entwickelten und innerhalb der EU bereits verbreiteten CERIF-Standard abzulegen. Auch die Definition der Daten im Kerndatensatz wird sich an diesen Standard anlehnen.

Der Wissenschaftsrat empfiehlt des Weiteren, beim Aufbau von Forschungsinformationssystemen darauf zu achten, dass diese Systeme neben Finanz- und Personaldaten auch weitere im Rahmen der etablierten administrativen Prozesse anfallenden Daten integrieren. Isolierte Systeme, die allein auf eine Domäne, wie z. B. auf die Verwaltung von Projektdaten abzielen, haben ungleich geringeren Nutzen.

Informationen sollten einrichtungsintern in möglichst elementarer Form vorgehalten werden. Dies erlaubt eine flexible, nutzungsgerechte Aggregation bei geringem Aufwand. Vorzugsweise sollten die Daten für jede forschende Person, zumindest aber auf die Ebene einer Professur aggregiert vorgehalten werden.

Sofern Hochschulen oder AUF individuelle Daten über ihre Wissenschaftlerinnen und Wissenschaftler (wie z.B. Publikationen oder Projekte) vorhalten, sollten sie diese im Fall eines Wechsels an eine andere Einrichtung in Form eines mit dem Kerndatensatz kompatiblen, maschinenlesbaren Formates für die Mitnahme zur Verfügung stellen.

Beim Aufbau neuer Forschungsinformationssysteme sollten Hochschulen und AUF sich am Stand der Technik orientieren. Der Wissenschaftsrat weist dabei darauf hin, dass am Markt etablierte Systeme gegenüber Eigenentwicklungen erhebliche Vorteile im Hinblick auf Betrieb, Wartung und Erweiterbarkeit haben.

An die Länder und den Bund richtet der Wissenschaftsrat die Empfehlung, die Einführung integrierter Forschungsinformationssysteme an Hochschulen und AUF nach Kräften zu unterstützen. Nicht zuletzt ermöglichen diese Systeme durch die regelmäßige Nutzung der gleichen Daten, sowohl in internen Dokumentationsprozessen der Einrichtungen als auch in der Berichterstattung nach Außen, eine nachhaltige Qualitätssicherung bei reduzierten Aufwendungen.

Die im Kerndatensatz Forschung spezifizierten Daten sollten unabhängig von Art und Reifegrad ihrer Forschungsinformationssysteme von allen Hochschulen und außeruniversitären Forschungseinrichtungen zeitnah nach einer Anfrage ohne großen Aufwand bereitgestellt werden können.

Forschungsförderer wie beispielsweise die DFG oder die Alexander von Humboldt-Stiftung halten qualitätsgesicherte Daten über Forschungsprojekte oder Stipendien vor. Diese Daten beinhalten wichtige Informationen über die Forschungsaktivitäten einzelner Hochschulen oder AUF. Sofern diese Daten nicht bereits öffentlich einsehbar sind, wären mit den Förderorganisationen Gespräche zu führen, wie die Daten den Einrichtungen rechtskonform in elektronischer Form zur Verfügung gestellt werden können. Der Wissenschaftsrat empfiehlt den Forschungsförderern, Informationen zu gewährten Förderungen durch gut dokumentierte Schnittstellen für die Datenabfrage durch forschende Einrichtungen automatisiert nutzbar zu machen. Sie sichern damit eine verlässliche Darstellung ihrer Förderleistungen an den Hochschulen und AUF. Darüber hinaus wäre die Bereitstellung der Daten in den Systemen der Förderorganisationen in einer zur Spezifikation des Kerndatensatzes kompatiblen Form wünschenswert.

### C.III EMPFEHLUNGEN ZU DEN INHALTEN DES KERNDATENSATZES

Eindeutige und operationale Definitionen der nachfolgend beschriebenen Daten sowie die Standardisierung von Datenformaten werden im Rahmen der Spezifikation des Kerndatensatzes festgelegt.

Diese Daten definieren einen Kern. Es steht selbstverständlich allen forschenden Einrichtungen frei, darüber hinaus weitere Daten vorzuhalten (siehe u. a. Kapitel C.IV). Alle Daten beziehen sich auf ein Kalenderjahr.

#### III.1 Beschäftigte

##### III.1.a Kerndaten

Die folgenden Kontextinformationen zur Einordnung von Forschungsaktivitäten sind Teil des Kerndatensatzes:

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Anzahl Personen (Kopfzahlen)	Geschlecht, Personalkategorie, Finanzierungsform (grund-/drittmittelfinanziert), Qualifikation (promoviert/habilitiert).	Fach
Vollzeitäquivalente	Geschlecht, Personalkategorie, Finanzierungsform (grund-/drittmittelfinanziert), Qualifikation (promoviert/habilitiert).	Fach
Anzahl und Anteil befristeter Verträge	Geschlecht, Personalkategorie, Finanzierungsform (grund-/drittmittelfinanziert), Laufzeit (bis 1 Jahr, 1 bis 2 Jahre, 2 bis 3 Jahre, mehr als 3 Jahre).	Fach
Doppelberufungen	Name der weiteren Einrichtung, Art des Beschäftigungsverhältnisses.	Fach

### III.1.b Begründung und Konkretisierung

Schon heute halten Hochschulen gemäß dem Gesetz über die Statistik für das Hochschulwesen (Hochschulstatistikgesetz - HStatG) Personaldaten entlang der Personalkategorien |<sup>17</sup> des Statistischen Bundesamtes vor. Daten über Beschäftigte geben zwar keine direkten Hinweise auf Forschungsaktivitäten, sind aber z. B. für eine auf die vorhandenen Ressourcen relativierte Bewertung der Forschungsleistung bedeutsam.

Der Wissenschaftsrat empfiehlt, zusätzlich zu den nach den Personalkategorien des Statistischen Bundesamtes differenzierten Personalzahlen insbesondere auch die Informationen über Doppelberufungen bzw. gemeinsame Berufungen von Professorinnen und Professoren sowie differenzierte Informationen über die Arbeitsverhältnisse von wissenschaftlichen und künstlerischen Mitarbeiterinnen und Mitarbeitern (befristete/unbefristete Verträge, grund- und drittmitelfinanziertes Personal) für den Kerndatensatz bereit zu halten. Informationen darüber, ob Beschäftigte promoviert bzw. habilitiert sind, können zusätzliche Hinweise auf den Umfang möglicher Forschungsaktivitäten geben. Qualität und Quantität von Forschungsaktivitäten hängen häufig unmittelbar von den vorhandenen Ressourcen ab. Aus diesem Grund sollten für alle Kategorien neben Informationen über die Anzahl der Beschäftigten (Kopfzahlen) auch Informationen über die Vollzeitäquivalente (VZÄ) vorgehalten werden.

### III.2 Nachwuchsförderung

#### III.2.a Kerndaten

Zur Dokumentation der Förderung des wissenschaftlichen Nachwuchses gehören folgende Informationen zum Kerndatensatz:

|<sup>17</sup> Vgl. bspw. Statistisches Bundesamt (Hrsg.): *Bildung und Kultur. Personal an Hochschulen*, Fachserie 11 Reihe 4.4, Wiesbaden 2009, S. 12-13.

Kerndatum	Aufschlüsselung	Aggregationsebene
Absolventenzahlen	Geschlecht, Herkunft (Inland, Ausland), Abschluss (Bachelor, Master, Staatsexamen, Diplom, Magister)   <sup>18</sup>	Fach
Grund- und drittmittelfinanzierte Promotionsprogramme und Graduiertenschulen   <sup>19</sup>	Name, Typ des Programms (Grund-/Drittmittel), Mittelgeber (DFG, EU, Bund, Länder, Unternehmen, Stiftungen u. a.), Antragstellerin oder Antragssteller, Sprecherin oder Sprecher, Laufzeit, Zahl der weiblichen Studierenden, Zahl der männlichen Studierenden, Herkunft (Inland, Ausland).	Programm
Zahl der abgeschlossenen Promotionen	Jahr, Geschlecht, Herkunft (Inland, Ausland)	Fach

### III.2.b Begründung und Konkretisierung

Die Förderung des wissenschaftlichen Nachwuchses gehört zu den genuinen Aufgaben insbesondere der Universitäten. Sie ist fester Bestandteil der wissenschaftlichen Tätigkeiten. Auch andere Hochschulen und die außeruniversitären Forschungseinrichtungen engagieren sich im Bereich der Nachwuchsförderung. Aspekte der Diversität haben hierbei in den vergangenen Jahren zunehmend an Bedeutung gewonnen und werden von Hochschulen und AUF schon heute regelmäßig in der Berichterstattung dokumentiert. Die Förderinstrumente in diesem Bereich wurden in den letzten Jahren stark ausgebaut. Dazu gehören sowohl die Ausweitung des Stipendienwesens und der (drittmittelfinanzierten) Promotionsprogramme und Graduiertenkollegs sowie die Einführung der Juniorprofessur. Erfolge bei der Einrichtung von Graduiertenschulen, bei der Einwerbung von Stipendien oder bei der Promotion von Nachwuchsforscherinnen und Nachwuchsforschern geben Auskunft über das Engagement der Wissenschaftlerinnen und Wissenschaftler in der Förderung des wissenschaftlichen Nachwuchses. Daher werden Informationen über die Erfolge der Nachwuchsförderung sowohl bei internen Steuerungsentscheidungen von Hochschulen und AUF als auch für externe Bewertungen von Forschungsaktivitäten herangezogen.

Abhängig vom jeweiligen Interesse der Datenabfrager werden dabei unterschiedliche Aspekte der Nachwuchsförderung und unterschiedliche Karriere-

|<sup>18</sup> Für interdisziplinäre und Doppelfach-Studiengänge sind klare Zurechnungskonventionen zu definieren, welche eine eindeutige und doppelungsfreie Zuordnung garantieren.

|<sup>19</sup> Unter drittmittelfinanzierten Promotionsprogrammen und Graduiertenschulen werden dabei solche drittmittelfinanzierte Programme der Doktorandenförderung verstanden, die mindestens zwei der folgenden Bedingungen erfüllen: (1) Strukturiertes Veranstaltungsprogramm für alle Teilnehmerinnen und Teilnehmer, (2) gemeinsame Verantwortung für die Betreuung der Doktorandinnen und Doktoranden durch die beteiligten Hochschullehrerinnen und -lehrer, (3) offenes, wettbewerbliches Aufnahmeverfahren mit Ausschreibung, (4) Stipendien oder Stellen für zumindest einen Teil der teilnehmenden Doktorandinnen und Doktoranden.

phasen betrachtet. Die hierzu vorhandene Datenbasis ist jedoch in der Regel lückenhaft. Während Angaben über die Absolventenzahlen, über eingerichtete Promotionsprogramme und Graduiertenschulen und über abgeschlossene Promotionen in der Regel bei unterschiedlichen Organisationseinheiten innerhalb der Hochschule vorliegen, werden Informationen über die Anzahl der laufenden Promotionen, über die Dauer der Promotionen, über in Anspruch genommene Promotions- und Postdoc-Stipendien oder über angenommene Erstrufe bisher noch selten systematisch gesammelt. |<sup>20</sup> Diese Informationen sind daher Teil der Erweiterungsoptionen (vgl. Kapitel C.IV).

### III.3 Drittmittel und Projekte

#### III.3.a Kerndaten

Drittmitteldaten sollen als *verausgabte Drittmittel* für bereits *laufende Projekte*, bezogen auf die einzelnen Antragstellerinnen und Antragsteller ausgewiesen werden. Eine einschlägige Projektdefinition, die auch Verbundprojekte mit einschließt, wird im Zuge der Standardisierung erarbeitet. Folgende Informationen sind Teil des Kerndatensatzes:

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Antragstellerin/Antragsteller, Sprecherin/Sprecher, Projektleitung	Geschlecht	Projekt
Titel des Projekts	Bei Verbundprojekten: einrichtungsintern / -extern	Projekt
Ggf. Titel des Teilprojekts		Projekt
Englischer Titel		Projekt
Förderer/Mittelgeber	Unterscheidbar sein müssen: DFG, EU, ERC, Bund   <sup>21</sup> , Länder, AIF, Unternehmen, Stiftungen, andere	Projekt
Förderkennzeichen		Projekt
Beginn		Projekt
Ende		Projekt

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Summe verausgabter Drittmittel	Jahr, Mittelgeber	Antragstellerin / Antragsteller

|<sup>20</sup> Der Wissenschaftsrat hat bereits mehrfach auf die Wichtigkeit dieser Informationen hingewiesen. Vgl. Wissenschaftsrat: Anforderungen an die Qualitätssicherung der Promotion, Köln 2011.

|<sup>21</sup> Hier sollen, unabhängig vom Ressort, alle durch den Bund geförderten Projekte aufgeführt werden.

Zu einem erheblichen Teil werden Drittmittel im Rahmen von *peer-review*-Verfahren vergeben. Dieser Vergabemodus impliziert in der Regel eine Begutachtung, die auch Vorarbeiten und damit eine Beurteilung der Forschungsqualität bisheriger Forschungsaktivitäten der Antragstellerinnen und Antragsteller berücksichtigt. Drittmittelsummen und drittmittelfinanzierte Projekte liefern so wichtige Hinweise auf Forschungsaktivitäten der einzelnen Wissenschaftlerinnen und Wissenschaftler. Daher werden Informationen zu Drittmitteln häufig als ein wichtiger Indikator für die Bewertung von Forschungsaktivitäten in Evaluationen, Ratings, Rankings, aber auch im Rahmen von internen Steuerungsprozessen an Hochschulen und AUF herangezogen.

Viele Forschungsvorhaben und größere Themenzusammenhänge erfordern zudem eine kooperative und arbeitsteilige Herangehensweise. Kooperative Forschung intensiviert den wissenschaftlichen Austausch und die interdisziplinäre Zusammenarbeit, die sich positiv auf die Forschungsarbeit auswirken. Darüber hinaus tragen Forschungsverbände oft zur besseren (internationalen) Sichtbarkeit der Forschung bei. Auch für die Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler kann sich die Einbindung in größere Projekte als vorteilhaft erweisen.

Forschungsverbände werden in der Regel in einem kompetitiven *peer-review*-Verfahren vergeben, so dass Informationen über geförderte Forschungsverbände häufig als Indikator im Rahmen von Forschungsbewertungen herangezogen werden. Die Sprecherrollen in Forschungsverbänden können auch Aufschluss über die Rolle einer Einrichtung in der Vernetzung der Forschung geben.

Unter Verwendung der Definition des Statistischen Bundesamtes werden unter „Drittmitteln“ solche Mittel verstanden, „[...] die zur Förderung von Forschung und Entwicklung sowie des wissenschaftlichen Nachwuchses und der Lehre zusätzlich zum regulären Hochschulhaushalt (Grundausstattung) von öffentlichen oder privaten Stellen eingeworben werden. Drittmittel können der Hochschule selbst, einer ihrer Einrichtungen (z.B. Fakultäten, Fachbereichen, Instituten) oder einzelnen Wissenschaftlern im Hauptamt zur Verfügung gestellt werden. Nicht zu den Drittmitteln zählen z. B. Mittel aus Zentral- und Fremdkapiteln des Trägerlandes, Mittel nach dem Überlastprogramm, Mittel der rechtlich selbstständigen Institute an Hochschulen. Zu beachten ist auch, dass die Regelungen zur Abwicklung von Drittmittelprojekten (z. B. Einstellung in den Haushalt, Verwahr- oder Sonderkonten) in den einzelnen Bundesländern

unterschiedlich sind.“ |<sup>22</sup> Auch auf die besonderen Bedingungen an nichtstaatlichen Hochschulen hat der Wissenschaftsrat bereits hingewiesen. |<sup>23</sup> Drittmittel für AUF sind entsprechend zu verstehen. Über Erfolge in der Einwerbung von Drittmitteln können verschiedene Informationen für interne Steuerungszwecke und zur Bewertung von Forschungsaktivitäten vorgehalten und herangezogen werden, die mit unterschiedlichen Vor- und Nachteilen verbunden sind: Drittmittelbewilligungen, Drittmittelleinnahmen und Drittmittelausgaben.

Informationen über Drittmittelbewilligungen werden z.B. von der DFG erfasst und könnten auch anderen Nutzern (Hochschulen, AUF oder datenabfragenden Institutionen) zur Verfügung gestellt werden. Die von der DFG erfassten Informationen zu bewilligten Drittmitteln sind jedoch unvollständig, da die von der Industrie/Privatwirtschaft und von privaten Stiftungen bewilligten Drittmittel nicht berücksichtigt werden.

Drittmittelleinnahmen werden vollständig, das heißt einschließlich der Zuwendungen von der Industrie/Privatwirtschaft und von privaten Stiftungen, durch die amtliche Statistik erfasst. Entsprechende Informationen werden von den Hochschulen selbst aufbereitet und an das Statistische Bundesamt übermittelt. Damit können diese Informationen grundsätzlich einer mehrfachen Nutzung zugänglich gemacht werden. Eine zusätzliche Abfrage an Hochschulen wäre somit obsolet. Allerdings sind Informationen über Drittmittelleinnahmen insbesondere bei Forschungsverbänden unzuverlässig. Beträge werden häufig nur bei der jeweiligen Sprecherin bzw. beim jeweiligen Sprecher eines Verbundprojektes verbucht und von diesem an andere Projektpartner weitergeleitet. Werden Informationen über die Drittmittelleinnahmen für die Bewertung von Forschungsaktivitäten herangezogen, kann dies zudem immer dann zu verzerrten Ergebnissen führen, wenn die Drittmittel bereits im ersten Jahr für alle geförderten Folgejahre ausgezahlt werden.

Informationen über Drittmittelausgaben schließlich liegen häufig nur mit zeitlicher Verzögerung und nur bei den jeweiligen Hochschulen und AUF vollständig vor. Ihr Vorteil ist jedoch, dass sie auch einzelnen Personen eindeutig zugeordnet werden können. Darüber hinaus können Drittmittelausgaben einzelnen Jahren genau zugeordnet und somit präziser erfasst werden als Drittmittelleinnahmen und Bewilligungssummen. In der Regel sind auch die datenabfragenden Institutionen an Informationen über Drittmittelausgaben

|<sup>22</sup> Statistisches Bundesamt (Hrsg.): Bildung und Kultur. Monetäre hochschulstatistische Kennzahlen, Fachserie 11 Reihe 4.3.2, Wiesbaden 2010, S. 8.

|<sup>23</sup> Wissenschaftsrat: Private und kirchliche Hochschulen aus Sicht der Institutionellen Akkreditierung, Köln 2012, S. 114-115.

interessiert, da diese Hinweise auf tatsächliche Forschungsaktivitäten geben und somit für die Erfassung von Forschungsaktivitäten besser geeignet sind.

In Abwägung der jeweiligen Vor- und Nachteile, spricht sich der Wissenschaftsrat daher dafür aus, Informationen über die Drittmittelausgaben in den Kerndatensatz aufzunehmen. Diese Informationen sollten die, nach Gebern differenzierten, jährlich ausgegebenen Drittmittel auf der Ebene der Antragstellerinnen und Antragsteller abbilden. Der Wissenschaftsrat erachtet es als elementar, weitergereichte Drittmittel nicht doppelt zu erfassen.

### III.4 Forschungspreise und Auszeichnungen

#### III.4.a Kerndaten

Der Kerndatensatz enthält Informationen über *bedeutende Forschungspreise und Auszeichnungen*, die den an der Einrichtung beschäftigten Personen verliehen wurden. Folgende Daten sind Teil des Kerndatensatzes:

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Name der Preisträgerin bzw. des Preisträgers	Geschlecht	Preis
Bezeichnung des Preises/der Auszeichnung	Insbesondere Akademiemitgliedschaften, Nobelpreise, Gottfried Wilhelm Leibniz-Preise, Max-Planck-Forschungspreise, Heinz-Maier-Leibnitz-Preise, Otto-Hahn-Medaillen, Humboldt-Professuren, ERC-Grants, <i>invited</i> , <i>keynote</i> , sowie <i>named lectures</i> und weitere herausgehobene Auszeichnungen.	Preis
Jahr der Verleihung		Preis
Name der verleihenden Einrichtung		Preis
Land der verleihenden Einrichtung		Preis

#### III.4.b Begründung und Konkretisierung

Der Verleihung von bedeutenden wissenschaftlichen Forschungspreisen und Auszeichnungen geht in der Regel ein Bewertungsprozess voraus. Dabei werden Forschungsaktivitäten im Hinblick auf ihre Qualität, Bedeutsamkeit, Innovation, Ausstrahlung und andere Aspekte beurteilt. Forschungspreise und Auszeichnungen geben somit Auskunft über die Sichtbarkeit, Würdigung und Anerkennung von erbrachten Forschungsaktivitäten durch die *scientific community* und somit letztlich über die Reputation einer Wissenschaftlerin oder eines Wissenschaftlers. Aus diesen Gründen werden Daten zu Forschungspreisen und Auszeichnungen häufig im Rahmen der Bewertung von Forschungsleistungen herangezogen. Obwohl diese Informationen wertvoll sind und auch häufig nachgefragt werden, werden sie bislang an vielen Einrichtungen nicht systematisch gesammelt und sind daher vielfach nicht verfügbar. In der Regel werden diese Informationen ausschließlich von den Wissenschaftlerinnen und Wissenschaftlern selbst vorgehalten.

Der Wissenschaftsrat empfiehlt den Hochschulen und AUF, Informationen über die Verleihung wissenschaftlicher Preise und Auszeichnungen der an ihrer Einrichtung beschäftigten Wissenschaftlerinnen und Wissenschaftler für den Kerndatensatz vorzuhalten. Damit der Erhebungsaufwand begrenzt bleibt, sollen nur herausgehobene Forschungspreise und Auszeichnungen im Kerndatensatz berücksichtigt werden.

Eine erschöpfende und vollständige Liste aller bedeutenden Spitzenpreise ist nicht leistbar. Die vorliegende Liste stellt damit nur einen ersten, pragmatischen, Ansatz dar. Im Zuge der Standardisierung ist es Aufgabe einer Projektgruppe eine begrenzte Liste herausgehobener Auszeichnungen vorzulegen.

### III.5 Patente

#### III.5.a Kerndaten

Informationen über die durch das zuständige Patentamt veröffentlichten Anmeldungen oder Erteilungen von Patenten sind Teil des Kerndatensatzes:

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Anzahl veröffentlichter Patente		Fach
Anzahl erteilter Patente		Fach
Erfinderin/Erfinder	Geschlecht, Herkunft (Inland, Ausland)	Patent
Anmelderin/Anmelder	Bezug zur Einrichtung	Patent
Titel des Patents		Patent
Patentamt		Patent
Status	<i>Veröffentlichte</i> Anmeldung oder Erteilung	Patent
Datum der Anmeldung oder Erteilung		Patent
Family-ID		Patent

Sofern die veröffentlichten oder erteilten Patente nicht von der Einrichtung selbst, sondern von Partnern aus dem Erfinderfeld angemeldet wurden, muss der Bezug zur Einrichtung ersichtlich sein. Ggf. sollte ein Abgleich mit Kooperationspartnern an anderen Einrichtungen erfolgen.

#### III.5.b Begründung und Konkretisierung

Nach §1 des Patentgesetzes werden „Patente [...] für Erfindungen auf allen Gebieten der Technik erteilt, sofern sie neu sind, auf einer erfinderischen Tätigkeit beruhen und gewerblich anwendbar sind.“ Patentanmeldungen und -erteilungen geben somit Auskunft über das Innovations- und Leistungspotenzial der Forschenden und werden daher insbesondere in natur- und technikkwissenschaftlichen Fächern häufig als Indikator für Relevanz und Anwendbarkeit der Forschung herangezogen.

Dabei können im Rahmen der Dokumentation von Forschungsaktivitäten drei verschiedene Arten von Informationen über Patente berücksichtigt werden, die mit unterschiedlichen Vor- und Nachteilen verbunden sind: Patentanmeldun-

gen, durch das Patentamt veröffentlichte Patentanmeldungen und Patentbewilligungen/-erteilungen.

Patentbewilligungen/-erteilungen geben Auskunft darüber, ob die im angemeldeten Patent beschriebene Erfindung tatsächlich als innovativ anzusehen ist. Patentbewilligungen/-erteilungen sind somit ein zuverlässigeres Datum zu Forschungsaktivitäten. Sie können das Bild jedoch verzerren, da sie teils erst mit einer großen Zeitverzögerung von mehreren Jahren nach der Patentanmeldung zur Verfügung stehen. Informationen über Patentanmeldungen liegen hingegen sehr früh vor, haben jedoch zunächst eine geringe Aussagekraft. Darüber hinaus werden viele angemeldete Patente noch vor der Veröffentlichung zurückgezogen.

Aus diesen Gründen empfiehlt der Wissenschaftsrat den Hochschulen und AUF die Vorhaltung der Informationen über die durch das zuständige Patentamt *veröffentlichten* Anmeldungen von Patenten sowie der Informationen über Patenterteilungen.

### III.6 Publikationen

Informationen über das Publikationsverhalten von Wissenschaftlerinnen und Wissenschaftlern haben sowohl im Rahmen von internen Steuerungsentscheidungen als auch bei externen Bewertungen von Forschungsaktivitäten eine herausgehobene Bedeutung. Darüber hinaus sind vollständige und qualitätsgesicherte Informationen zu Publikationen die Grundlage für die internationale Sichtbarkeit der deutschen Wissenschaft sowie für die Berechnung von bibliometrischen Kennzahlen, die in einigen Disziplinen immer häufiger nachgefragt werden.

Zugleich ist die Abfrage dieser Daten häufig mit großem Aufwand für die Einrichtungen verbunden. Trotz dieses Aufwands erscheint der Nutzen einer systematischen Erfassung von Publikationen erheblich. Der Wissenschaftsrat empfiehlt daher ihre Einbeziehung in den Kerndatensatz. Er legt mit diesen Empfehlungen darüber hinaus die Grundlage dafür, Standards für die Erfassung bibliometrischer Informationen, die fachspezifischen Besonderheiten gerecht werden, zu schaffen.

#### III.6.a Kerndaten

Hochschulen und AUF sollten Informationen zu Publikationen der Wissenschaftlerinnen und Wissenschaftler der jeweiligen Einrichtung systematisch sammeln und vorhalten. Als zentrales Kriterium gelten Zitierfähigkeit und Wiederauffindbarkeit.

Die Einträge von Publikationen sollten von den einzelnen Wissenschaftlerinnen und Wissenschaftlern selbst vorgenommen werden (können), da in der Regel nur diese über vollständige Publikationslisten verfügen.

Zur Sicherung der Datenqualität sollte die Leitung einer Hochschule oder AUF eine Stelle (z. B. innerhalb der Universitätsbibliothek) beauftragen, die Regeln für die Publikationserfassung festzulegen und die Publikationsmeldungen bzw. -eingaben zu prüfen.

Folgende Publikationstypen sind in den Kerndaten zu erfassen:

- \_ Zeitschriftenartikel,
- \_ Herausgeberwerke, wie z. B. Sammelbände, Konferenz-Proceedings, Sonderhefte von Zeitschriften,
- \_ Monographien |<sup>24</sup>,
- \_ Sammelbandbeiträge, Buchkapitel,
- \_ begutachtete Konferenzbeiträge,
- \_ weitere Kategorien (z.B. elektronische Veröffentlichungen) entsprechend der Profilvorstellung des jeweiligen Faches.

In Abhängigkeit vom jeweiligen Publikationstyp sind mindestens folgende Informationen zu einzelnen Publikationen Teil des Kerndatensatzes:

<b>Kerndatum</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Namen der Autorinnen und Autoren		Publikation
Erscheinungsjahr		Publikation
Erscheinungsort	Verlag, Zeitschriftentitel, elektronische Publikation	Publikation
Titel der Publikation		Publikation
Publikationstyp	Zeitschriftenartikel, Herausgeberwerke, Monographien, Sammelbandbeiträge, Konferenzbeiträge, Buchkapitel, weitere Kategorien	Publikation
ISSN / ISBN / DOI		Publikation

### III.6.b Begründung und Konkretisierung

Publikationsdaten zählen bei Bewertungen von Forschungsaktivitäten zu den wichtigsten Output-Indikatoren. Die Bedeutung einzelner Publikationstypen bzw. Publikationskategorien ist dabei stark fachabhängig. Welche Publikationsmerkmale (Anzahl der Publikationen, Publikationslisten, Inhalt, Autorenzahl, Autorenrollen, Zitationen, u. a.) der Bewertung zugänglich gemacht werden, ist dabei ebenfalls von Fach zu Fach unterschiedlich. Insbesondere für die

|<sup>24</sup> Monographien sind im Kontext des Kerndatensatzes Forschung ausschließlich als Forschungsmonographien zu verstehen.

interne Hochschulsteuerung (LOM, W-Besoldung) sind jedoch fächerübergreifende (quantitative) Informationen über die Publikationsaktivität von Wissenschaftlerinnen und Wissenschaftlern von großem Interesse. Sie bieten häufig einen Einstieg in die fachspezifische Betrachtung der Publikationsleistungen.

Es kann auf unterschiedliche Publikations- und Zitationsdatenbanken zurückgegriffen werden, in denen qualitätsgesicherte Datenbestände gespeichert sind und die sowohl Informationen zur Publikationsaktivität einzelner Personen als auch weiterführende Analysen anbieten. Die gängigen kommerziellen Publikations- und Zitationsdatenbanken (z. B. Scopus, Web of Science) sind als einzige externe Datenquelle jedoch unzureichend. Ihre Qualität und ihr Abdeckungsgrad variieren stark zwischen Fächern und sind besonders in vielen geistes- und sozialwissenschaftlichen Fächern mangelhaft. Auch sind Monographien, Buch- und Konferenzbeiträge sowie nicht-englischsprachige Veröffentlichungen und damit wesentliche Teile der in Deutschland erbrachten Forschungsaktivitäten in solchen Datenbanken nur unzureichend erfasst.

Der Wissenschaftsrat empfiehlt den Hochschulen und AUF eine systematische Sammlung und Vorhaltung von Informationen zu Publikationen ihrer Wissenschaftlerinnen und Wissenschaftler. Zur Vereinfachung der Publikationserfassung wird Hochschulen und AUF die Implementierung einer entsprechenden Softwarelösung empfohlen, die die im Folgenden genannten Mindestanforderungen erfüllt. Das zu implementierende System sollte:

- \_ die Erfassung auch fachspezifischer Publikationsformen ermöglichen,
- \_ eine einfache Dateneingabe ermöglichen,
- \_ eine automatisierte Hilfestellung bei der Eingabe der Daten bieten, z. B. durch die Übernahme von Daten aus anderen, bereits vorhandenen, Datenquellen,
- \_ einen Datenimport und -export von und in unterschiedliche Datenformate (z.B. Zitationssoftware) erlauben,
- \_ für die einzelnen Wissenschaftlerinnen und Wissenschaftler einen Mehrwert, z. B. in Form einer automatischen Erstellung von Literaturlisten für den Lebenslauf, die Homepage, Forschungsberichte, institutionelle Repositorien u.ä., schaffen,
- \_ eine Autorenidentifikation (Zuordnung) anbieten, um die Verknüpfung mit anderen Informationen zu erleichtern,
- \_ zur Vermeidung von Verwechslungen innerhalb einer Einrichtung zumindest auf Einrichtungsebene mit einer eindeutigen Autoren-ID arbeiten und dabei die Vereinbarkeit mit dem Datenschutz sicherstellen.

Eine große Zahl von Daten zu Forschungsaktivitäten lässt sich nicht so eindeutig definieren wie die oben aufgeführten Inhalte des Kerndatensatzes. Auch sind etliche Daten nur für eine begrenzte Anzahl von Fächern einschlägig. Einige dieser Daten schlägt der Wissenschaftsrat im Folgenden als anzustrebende Erweiterungen für die standardisierte Datenvorhaltung vor.

Auch diese Daten können für einige Fächer von großer Wichtigkeit sein. Sie vorzuhalten, steht jeder Einrichtung frei. Im Zuge des Standardisierungsprozesses sollten nach Möglichkeit auch Empfehlungen für die Datenvorhaltung dieser „extranuklearen“ Informationen zu Forschungsaktivitäten entwickelt werden.

#### IV.1 Herausgeberschaften, Mitgliedschaft in Editorial Boards

##### IV.1.a Vorschlag zur Erweiterung

Hochschulen und AUF sollten Informationen über die Herausgeberschaften und die Mitgliedschaften ihrer Wissenschaftlerinnen und Wissenschaftler in Editorial Boards erfassen. Dazu zählen insbesondere:

Datum für Erweiterung	Aufschlüsselung	Aggregationsebene
Herausgeberin oder Herausgeber		Publikation
Titel der Zeitschrift oder Reihe		Publikation
Typ der Herausgeberschaft	Hauptausgeber, Mitglied des Editorial Boards	Publikation
Ort	ggf. Online-Publikation	Publikation
Verlag		Publikation
Amtsdauer		Publikation
ISSN / ISBN / DOI		Publikation

##### IV.1.b Begründung und Konkretisierung

Herausgeberschaften – sowohl von Buchreihen als auch von Zeitschriften – wie auch die Mitgliedschaft in Editorial Boards stellen anerkannte Dienstleistungen für die wissenschaftliche Gemeinschaft dar. Sie unterliegen häufig Auswahlprozessen und sind in der Regel Ausweis der fachlichen Anerkennung einer einzelnen Wissenschaftlerin oder eines einzelnen Wissenschaftlers.

## IV.2.a Vorschlag zur Erweiterung

Hochschulen und AUF sollten Informationen über akademische Spinoffs aus den Einrichtungen systematisch erfassen. Folgende Informationen sollten zu Ausgründungen aus den Hochschulen oder AUF erfasst werden:

<b>Datum für Erweiterung</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Namen der Gründerinnen oder Gründer	Geschlecht	Spinoff
Name der Ausgründung		Spinoff
Branche/Zweig		Spinoff
Jahr der Ausgründung		Spinoff
Bezug der Gründung zur Einrichtung		Spinoff

## IV.2.b Begründung und Konkretisierung

Transfer von Forschungsaktivitäten, ob als Technologietransfer oder Wissensvermittlung, gehört zu den wichtigsten Aufgaben von Hochschulen und AUF. Disziplinabhängig erfolgt der Forschungstransfer in unterschiedliche gesellschaftliche Bereiche und in unterschiedlicher Form.

Ausgründungen bzw. akademische Spinoffs aus Hochschulen geben Auskunft über Praxisrelevanz sowie über die Anwendbarkeit der Forschung und können somit als Indikator für erfolgreichen Forschungstransfer herangezogen werden. Trotz des großen Interesses an Informationen zu Ausgründungen von akademischen Spinoffs aus den Hochschulen und AUF ist die Datenlage sehr lückenhaft.

Grundsätzlich können Kompetenz-Spinoffs (z.B. Anwaltskanzlei) von Transfer-Spinoffs unterschieden werden. Die EU definiert in Anlehnung an das ZEW<sup>25</sup> und internationale Standards bedeutender Transfervereinigungen (AUTM, ASTP) Transfer-Spinoffs als „neue Firmen, die zur Verwertung von geistigem Eigentum oder Know-how der Wissenschaftseinrichtung unter Abschluss einer formalen Vereinbarung mit der Wissenschaftseinrichtung gegründet wurden (Lizenz- und/oder Beteiligungsvertrag)“. <sup>26</sup> Nur solche Transfer-Spinoffs geben Auskunft über den Transfer von Forschungsaktivitäten, da für ihre Gründung die an einer

<sup>25</sup> Das Zentrum für europäische Wirtschaftsforschung definiert akademische Spinoffs als Neugründungen von Unternehmen, „[...] die mit dem Zweck, neues Wissen, das in öffentlichen Forschungseinrichtungen (Universitäten, Fachhochschulen, außeruniversitären staatlichen Forschungseinrichtungen) erarbeitet wurde, in Marktangebote und damit in Wertschöpfung und Beschäftigung umzusetzen, gegründet wurden. Vgl. hierzu Egel, J.; Fryges, H.; Gottschalk, S.; Rammer, C.: Dynamik von akademischen Spinoff-Gründungen in Österreich, Discussion Paper No. 07-021, Mannheim, ZEW, 2007, S. 1.

<sup>26</sup> European Commission's Expert Group on Knowledge Transfer Metrics: Report Metrics for Knowledge Transfer from Public Research Organisations in Europe, European Commission, 2009.

Einrichtung erarbeiteten Forschungsergebnisse oder Methoden unverzichtbar waren. Gründungen durch Beschäftigte der Hochschulen bzw. der AUF ohne vertragliche Regelung mit der Forschungseinrichtung oder Gründungen aus dem Studium heraus (wie sie z. B. in der IT-Branche häufig vorkommen) sollten nicht herangezogen werden.

Der Wissenschaftsrat empfiehlt Hochschulen und AUF Informationen über akademische Transfer-Spinoffs systematisch zu erfassen. Dabei erscheinen beratende Transfer-Servicezentren als ein geeignetes Instrument sowohl zur Förderung als auch zur Nachverfolgung akademischer Spinoffs.

#### IV.3 Lizenzeinnahmen

Lizenzeinnahmen stellen neben Patenten eine weitere mögliche Konkretisierung des Leistungspotenzials von Hochschulen und AUF dar.

##### IV.3.a Vorschlag zur Erweiterung

Hochschulen und AUF sollten folgende Angaben zu ihren im Laufe eines Kalenderjahres erzielten Lizenzeinnahmen vorhalten:

<b>Datum für Erweiterung</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Lizenzeinnahmen für Nutzung	Patente, Gebrauchsmuster, Know-how, Software	Fach

##### IV.3.b Begründung und Konkretisierung

Ähnlich wie Patente können Lizenzeinnahmen Hinweise auf den Transfer von Forschungsaktivitäten geben, wenngleich häufig erst nach einer längeren Verzögerung. Die Verbreitung der Nachfrage nach einer Lizenz lässt sich über die Anzahl der Lizenznehmerinnen und Lizenznehmer abschätzen.

#### IV.4 Nachwuchsförderung (Erweiterung)

##### IV.4.a Vorschlag zur Erweiterung

Hochschulen und AUF sollten über die Kerndaten zur Nachwuchsförderung hinaus Möglichkeiten schaffen, um die folgenden Informationen zu erfassen:

Datum für Erweiterung	Aufschlüsselung	Aggregationsebene
Anzahl laufender Promotionen	Geschlecht, Herkunft (Inland, Ausland)	Fach
Anzahl Nachwuchsgruppenleistungen	Geschlecht, Herkunft (Inland, Ausland)	Fach
Anzahl angenommener Erstrufe	Art der Professur (Juniorprofessur, Dauerstelle), berufende Einrichtung, Land der berufenden Einrichtung	Fach
Auszeichnungen für den wissenschaftlichen Nachwuchs	z.B. Lichtenberg-Professuren der Volkswagenstiftung, Heisenberg-Professuren, Humboldt-Professuren, Emmy Noether-Nachwuchsgruppen, ERC Starting Grant	Fach
Anzahl extern finanzierte Promotionsstipendien	Geschlecht	Fach
Anzahl extern finanzierte Postdoc-Stipendien	Geschlecht	Fach

Die befasste Projektgruppe ist aufgefordert zu prüfen, inwieweit sich die Anzahl der extern finanzierten Promotionsstipendien sowie die Anzahl der extern finanzierten Postdoc-Stipendien für eine sofortige Aufnahme in den Kerndatensatz eignen.

#### IV.4.b Begründung und Konkretisierung

Sowohl Hochschulen und AUF als auch datenabfragende Einrichtungen sind an Informationen über laufende Promotionen sowie über Promotions- und Postdoc-Stipendien interessiert. Die Kenntnis dieser Informationen könnte nicht zuletzt zur besseren Betreuung von Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftlern beitragen. Der Wissenschaftsrat ist sich zugleich der Schwierigkeiten bewusst, die mit Abfrage und Bewertung dieser Daten verbunden sind. An anderer Stelle hat er in diesem Zusammenhang bereits die Schaffung eines eigenen Doktorandenstatus empfohlen. |<sup>27</sup> Die Aussagekraft solcher Kennzahlen für die Bewertung von Forschungsleistung ist bislang insofern als eingeschränkt einzuschätzen, als die Informationen über laufende Promotionen sowie über Promotionsstipendien keinen Aufschluss über die Qualität der Promotionen liefern. Darüber hinaus ist die Vollständigkeit dieser Informationen häufig nicht gesichert, da viele Promovierende sich erst kurz vor Abgabe ihrer Dissertation an den Hochschulen anmelden. Zugleich lassen einige Promovierende, die an Hochschulen als solche angemeldet sind, das Verfahren aus unterschiedlichen Gründen ruhen. Die Datenlage zu laufenden Promotionen und zu Promotions- und Postdoc-Stipendien erweist sich somit als lückenhaft und schwer interpretierbar; sie führt zu Verzerrungen in der Präsentation der Einrichtungen im Hinblick auf die Nachwuchsförderung. Es wird daher dringend empfohlen, Möglichkeiten zur besseren Erfassung dieser Informationen zu

|<sup>27</sup> Vgl. Wissenschaftsrat: Anforderungen an die Qualitätssicherung der Promotion, Köln 2011, S. 17.

schaffen. Als geeignetes Instrument werden Promotionsvereinbarungen, wie sie an einigen Einrichtungen bereits praktiziert werden, empfohlen. |<sup>28</sup>

Informationen zu angenommenen Erstrufen können wichtige Hinweise darauf geben, welche Einrichtungen den wissenschaftlichen Nachwuchs ausbilden, der andernorts auf Professuren berufen wird. Auch bei diesem Datum ist eine systematische Erfassung wünschenswert. Herausgehobene Auszeichnungen für den wissenschaftlichen Nachwuchs wie beispielsweise die Leitung einer Emmy Noether-Nachwuchsgruppe sollten ebenfalls systematisch erfasst werden.

#### IV.5 Drittmittel (Erweiterung)

##### IV.5.a Vorschlag zur Erweiterung

<b>Datum für Erweiterung</b>	<b>Aufschlüsselung</b>	<b>Aggregationsebene</b>
Drittmittelausgaben	Personal- und Sachausgaben	Projekt

##### IV.5.b Begründung und Konkretisierung

Die Kostenarten, für die Drittmittel verausgabt werden, unterscheiden sich von Disziplin zu Disziplin zumindest der Proportion nach. So fließt ein großer Teil der Drittmittel in den Naturwissenschaften häufig für Gerätschaften, während in den Geisteswissenschaften ein hoher Anteil der Gelder für das Personal ausgegeben wird. Um eine fachspezifische Einordnung der verausgabten Drittmittel zu ermöglichen (vgl. auch Kapitel B.I.2) sollten die Einrichtungen auch differenzierte Informationen über die aus den Drittmitteln getätigten Personal- und Sachausgaben vorhalten.

## **C.V EMPFEHLUNGEN AN BUND UND LÄNDER**

---

Der Wissenschaftsrat fordert Bund und Länder auf, gegenüber Hochschulen und AUF ihre Unterstützung für den Kerndatensatz deutlich zu machen und auf eine zeitnahe Umsetzung dieses Standards hinzuwirken.

Der Wissenschaftsrat empfiehlt Bund und Ländern, ihre forschenden Einrichtungen bei der Einführung von Forschungsinformationssystemen nach Möglichkeit zu unterstützen.

Er erwartet auch von Bund und Ländern, die eigenen Datenabfragen an den Daten des Kerndatensatzes auszurichten und darüber hinausgehende Anfragen selten und nur in gut begründeten Fällen zu stellen.

|<sup>28</sup> Vgl. Wissenschaftsrat: Anforderungen an die Qualitätssicherung der Promotion, Köln 2011.

---

# D. Standardisierung des Kerndatensatzes

Dem Wissenschaftsrat ist es ein Anliegen, über seine Empfehlungen zu den Inhalten des Kerndatensatzes hinaus, eine solide Basis für eine erfolgreiche Implementierung an Hochschulen und AUF zu schaffen. Im Folgenden legt er daher einen Prozess zu Entwicklung einer Spezifikation des Kerndatensatzes vor. Diese Festlegungen binden die mit der Umsetzung zu beauftragenden Akteure.

Die Erarbeitung einer Spezifikation auf Basis dieser Empfehlung erfolgt in Form eines Projektes und wird von einem Projektmanagement-Team verantwortet. Der Wissenschaftsrat bittet das Bundesministerium für Bildung und Forschung, die Entwicklung der Spezifikation eines Kerndatensatzes durch eine entsprechende Förderung des Projektes und der Projektgruppen zu unterstützen.

## D.1 AKTEURE

---

Der Prozess der Entwicklung einer Spezifikation des Kerndatensatzes liegt maßgeblich in den Händen des *Projektmanagements* (vgl. Abbildung 1).

Eine Arbeitsgruppe des Wissenschaftsrates begleitet das Projekt in der Funktion eines *Beirats*. Sie übt eine Kontrollfunktion aus. Der eingangs genannten Spannung zwischen Datensparsamkeit und einer (zu) umfassenden Datenerfassung wird dabei besondere Aufmerksamkeit zu schenken sein. Änderungen, die sich auf die Inhalte der Kapitel C.III und C.IV beziehen, bedürfen der Zustimmung des Beirates. Der Beirat wählt die *Leitung* des Projektmanagements in Abstimmung mit dem BMBF aus. Ebenso muss der Beirat der Besetzung der Projektgruppen und deren Leitung zustimmen. Nach Ablauf der Projektphase legt der Beirat dem Wissenschaftsrat den Spezifikationsentwurf zusammen mit einer wissenschaftspolitischen Stellungnahme zur Beratung und Verabschiedung vor. Nach Ablauf des in dieser Empfehlung festgelegten Standardisierungsprozesses setzt der Beirat Ziele für die Weiterführung des Projektes zur Einführung eines

Kerndatensatzes Forschung in Deutschland und erwägt einen sinnvollen Zeitpunkt für eine erste Evaluation.

Dem Beirat gehören neben Vertretern der Wissenschaft auch Vertreter der Politik aus Bund und Ländern, Experten für Datenverarbeitung, sowie Personen mit Erfahrungen in der Einführung von Forschungsinformationssystemen an. Auch ein Experte für Datenschutz ist Mitglied des Beirates.

Das *Projektmanagement* beruft und koordiniert drei Projektgruppen (PG). Dies sind die PG „*Definition und Dateiformate*“, die PG „*Technik, Schnittstellen und Dateiformate*“ sowie die PG „*Bibliometrie*“. Eine vierte Projektgruppe, die PG „*Fächerklassifikation*“ wird direkt durch den Beirat eingesetzt. Ihre Vorsitzende oder ihr Vorsitzender ist aber genau wie seine Kolleginnen und Kollegen der anderen PG Mitglied des Projektmanagements. Die Leiterinnen und Leiter der Projektgruppen wählt die Leitung des Projektmanagements in Abstimmung mit dem Beirat aus. Das Projektmanagement bindet insbesondere die DFG in den Standardisierungsprozess ein.

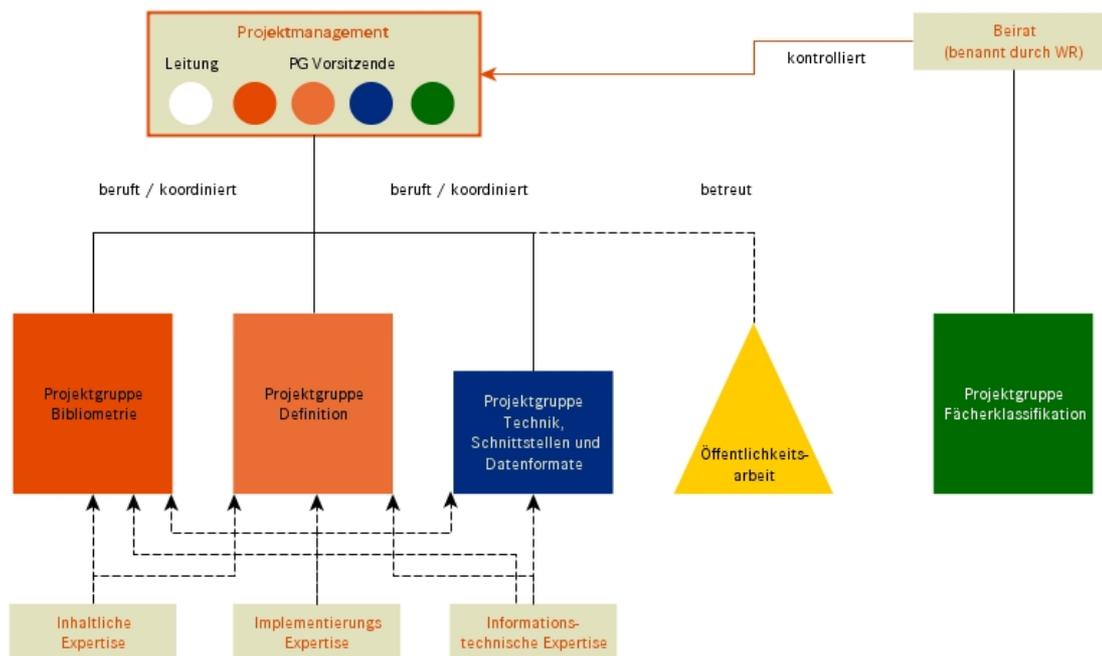
Neben der Projektleitung besteht das Projektmanagement aus den Vorsitzenden der vier Projektgruppen. Das Projektmanagement hat eine komplexe und koordinationsintensive Aufgabe zu erfüllen und setzt sich aus diesem Grund aus einer Reihe ausgewiesener Expertinnen und Experten in den jeweiligen Gebieten zusammen. Die Mitglieder des Projektmanagements treffen sich regelmäßig, um die Fortschritte des Projektes zu dokumentieren und die für den Projektverlauf nötigen Absprachen zu treffen.

Angesichts der Natur des Projektes achtet das Projektmanagement darauf, die PG vor allem mit Mitgliedern mit ausgewiesener Expertise in folgenden Gebieten zu besetzen.

- 1 – Inhaltliche Expertise zur Erfassung und Nutzung von Informationen zu Forschungsaktivitäten,
- 2 – Informationstechnische Expertise in den Bereichen Datenmodellierung und (Forschungs-)Informationssysteme, sowie
- 3 – Implementierungsexpertise, die die rechtlichen und praktischen Rahmenbedingungen der Implementierung in den Hochschulen und AUF einzuschätzen vermag.

*Inhaltliche Expertise* sichert den Projektgruppen einen breiten Wissens- und Erfahrungsschatz im Umgang mit den vorgeschlagenen Kerndaten. Hier richtet sich der Blick zusätzlich zu Hochschulen und AUF zunächst auf die politische Seite – Bund und Länder, ebenso aber auch das Statistische Bundesamt. Sodann hat neben der DFG auch der Wissenschaftsrat selbst größere Datenbedarfe. Für die Standardisierung der Publikationserfassung ist auch die Expertise der Nationalbibliothek und anderer Bibliotheken einschlägig.

Abbildung 1: Organigramm des Projekts zur Spezifikation des Kerndatensatzes



*Informationstechnische Expertise* sichert im Implementierungsprozess die Integration bestehender Datenbestände und die Konformität zu einem einheitlichen Datenaustauschformat. Auch sollten erfahrene Anwender verschiedener kommerziell verfügbarer Forschungsinformationssysteme oder anderer in Hochschulen eingesetzter Softwarelösungen eingebunden werden.

*Implementierungsexpertise* der Hochschulverwaltungen und der AUF ist schließlich von zentraler Bedeutung, um die Umsetzbarkeit der einzelnen Aspekte dieser Empfehlung abzuschätzen. Mithilfe dieser Expertise verschafft sich das Projektmanagement einen entsprechenden Überblick. Der Wissenschaftsrat empfiehlt, die Setzungen durch den Kerndatensatz soweit als möglich im Einklang mit den bestehenden Möglichkeiten der Datenlieferanten zu gestalten. In der Pilotphase ist der Machbarkeitsaspekt primär für ausgewählte Piloteinrichtungen zu betrachten.

## D.II ELEMENTE DES STANDARDISIERUNGSPROZESSES

Neben der Ausarbeitung einer Spezifikation der Inhalte des Kerndatensatzes sieht der Wissenschaftsrat auch Handlungsbedarf, eine konsensfähige Fächerklassifikation und eine fächerspezifische Standardisierung von Publikationsformaten zu erarbeiten.

### II.1 Fächerklassifikation

Zum jetzigen Zeitpunkt werden in Deutschland vor allem die Fächerklassifikationen des Statistischen Bundesamtes und der DFG verwendet. Beide sind nicht

mit dem Fokus entwickelt worden, Fächer als inhaltliche – und damit über Einrichtungen hinweg vergleichbare – Einheiten zu definieren. Klassifikationen, die auf internationaler Ebene vorliegen haben eher einen fachzentrierten Ansatz. |<sup>29</sup> Der Wissenschaftsrat sieht es als unverzichtbar an, über eine konsensfähige Fächerklassifikation zu verfügen, die die bestehende Heterogenität auflöst. Sie muss dabei sowohl klare und überschneidungsfreie Zuordnungen (u. a. von Finanzen) zulassen, die Verortung von Forschungsaktivitäten inhaltlich aussagekräftig einzelnen Fächern zuordenbar machen aber auch flexibel sein, um interdisziplinäre Entwicklungen darstellen zu können. Dabei ist es zusätzlich wünschenswert, dass eine Vergleichbarkeit der Informationen gegeben ist, die bislang auf Basis bestehender Klassifikationen zusammengestellt wurden.

Der Wissenschaftsrat begrüßt daher das Vorhaben des Statistischen Bundesamtes, seine eigene Klassifikation im Zuge einer Angleichung an internationale Standards zu überarbeiten. Nach der Vorlage entsprechender Vorschläge wird zu prüfen sein, inwieweit diese mit den Zielen des Kerndatensatzes kompatibel sind. Soweit möglich erscheint dem Wissenschaftsrat ein gemeinsames Vorgehen bei der Erarbeitung einer neuen Fächerklassifikation erstrebenswert.

Die bestehenden nationalen und internationalen Klassifikationen können nach Ansicht des Wissenschaftsrates dabei helfen, zunächst eine Grobklassifikation nach inhaltlichen Kriterien aufzubauen. Es ist dabei essentiell, auch eine größere internationale Vergleichbarkeit sicher zu stellen. Jenseits der Grobkategorien muss eine Klassifikation aber auch in der Lage sein, Fächer mit ihren distinkten Spezialisierungen abzubilden. Zu diesem Zweck hält es der Wissenschaftsrat für unverzichtbar, mit den Fachgesellschaften gemeinsam an der Integration der jeweiligen Klassifikationen in ein gröberes, übergeordnetes Schema zu arbeiten. Um eine überschneidungsfreie Zuordnung zu Elementen einer solchen Klassifikation zu ermöglichen, muss diese hierarchisch gestaltet sein.

Die Erfassung von Forschungsaktivitäten quer zu den Grenzen etablierter Fächer wird immer wichtiger. Gerade Interdisziplinarität lässt sich im Rahmen notwendig hierarchischer Klassifikationen aber kaum abbilden. Die Dynamik in der Entwicklung von neuen Fächern wie auch wechselnde fächerübergreifende Kooperationen lassen sich als Folge nicht zuverlässig dokumentieren und vergleichen.

Der Wissenschaftsrat regt daher an, die Bemühungen um eine konsensfähige Fächerklassifikation dazu zu nutzen, die hierarchische Klassifikation durch

|<sup>29</sup> Vgl. die *International Standard Classification of Education* (ISCED) und OECD: *Frascati Manual. Proposed Standard Practice for Surveys on Research and Experimental Development*, 6. Auflage. Paris 2002.

einen fachspezifischen und vordefinierten Thesaurus zu ergänzen. Auch hier sollten die Thesauri der Fachgesellschaften und ggf. solche von Zeitschriften als Vorlage dienen. Forscherinnen und Forscher können dann sich und ihre Projekte mithilfe von Schlagworten verorten. Dadurch werden interdisziplinäre Vergleiche über unterschiedliche Kombinationen der Schlagwörter möglich.

## II.2 Standardisierung von Publikationsformaten

Jedes Fach pflegt einen eigenen Umgang mit Publikationen. Während in einem Fach Zeitschriftenartikel Norm und Maßstab darstellen, erfreuen sich in anderen Fächern Monographien oder Beiträge in Tageszeitungen großen Ansehens. Die Auswahl der in den Kerndatensatz aufzunehmenden Publikationen muss somit die verschiedenen Fachkulturen berücksichtigen. Der Wissenschaftsrat erkennt an, dass heute nicht in allen Fächern ein Konsens über relevante Publikationen existiert. Er empfiehlt daher in diesen Fächern einen Konsens unter Beteiligung der relevanten Akteure herzustellen. Die an einigen Hochschulen verwendeten Grundsätze zur Auswahl von Publikationen können hier Hilfestellung sein. Auch erscheint es geboten, die reichhaltige Erfahrung von Bibliotheken auf diesem Feld zu nutzen.

## **D.III ZEITPLAN UND MEILENSTEINE**

---

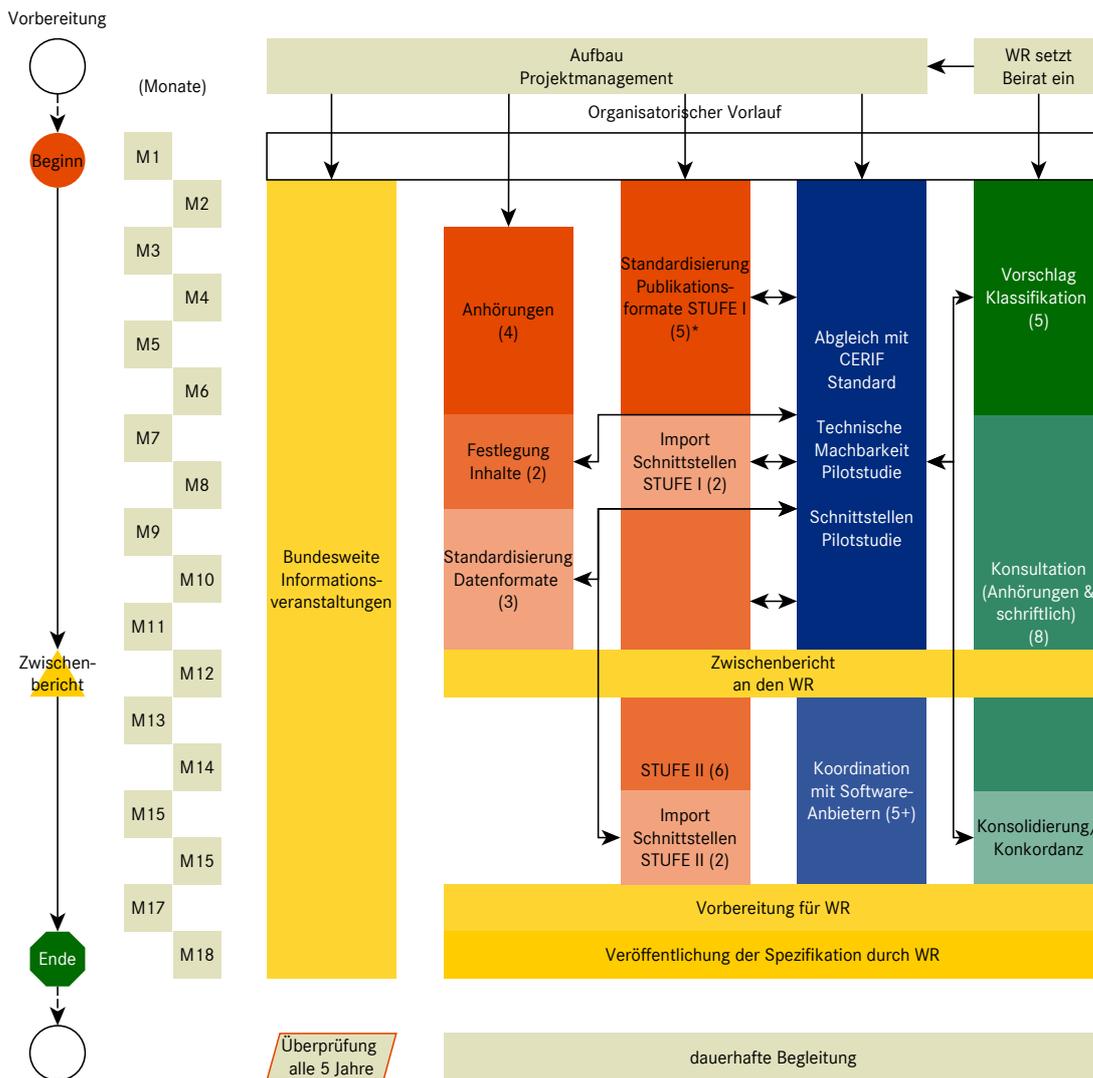
Für die Durchführung des Projektes spricht sich der Wissenschaftsrat für die nachfolgend dargestellte Zeitplanung aus.

### III.1 Pilotphase

Dem Wissenschaftsrat ist daran gelegen, sicher zu stellen, dass sich die Spezifikation in der Praxis von Hochschulen und AUF möglichst unproblematisch umsetzen lässt. Die Umsetzbarkeit der Spezifikation soll daher im Zuge einer Pilotphase schon während des Projektes getestet werden. Zu diesem Zweck wählt das Projektmanagement in Absprache mit dem Beirat zwei bis drei Pilot-einrichtungen aus, in denen die Implementierung getestet wird. Diese Piloteinrichtungen sollten möglichst repräsentativ sein, so dass die gewonnenen Erfahrungen sich leicht auf andere Einrichtungen übertragen lassen.

### III.2 Ablauf des Standardisierungsprozesses

Im Folgenden wird der Projektablauf, wie in Abbildung 2 dargestellt, erläutert.



III.2.a Vorbereitungen

WR setzt Beirat ein

Vor dem Start des Projektes setzt der Wissenschaftsrat den Beirat als Arbeitsgruppe ein. Zur Zusammensetzung des Beirats wird auf Kapitel D.I verwiesen.

Aufbau  
Projekt-  
management

Durch den Beirat wird

1 – eine Leiterin oder ein Leiter des Projektmanagements benannt.

Diese oder dieser gewinnt

2 – Expertinnen und Experten für die Besetzung der Projektgruppen und stellt

3 – Mitarbeiterinnen und Mitarbeiter für die Koordination und Betreuung der Projektgruppen ein.

Die Benennung der Expertinnen und Experten für die Projektgruppen und deren Leitungen bedarf der Zustimmung des Beirates. Der Zeitbedarf für diese Vorbereitungen lässt sich nicht verlässlich abschätzen.

III.2.b Projektgruppe „Definitionen und Datenformate“ (Zeitangaben in Monaten)

Anhörungen  
(4)

Basierend auf den Empfehlungen zu den Inhalten des Kerndatensatzes führt die Projektgruppe zunächst Anhörungen durch. Diese haben *Konsultationscharakter*. Die Beteiligten nehmen schriftlich zu den Vorschlägen der Projektgruppe Stellung. Die Anhörungen binden relevante Akteure ein.

Festlegung  
Inhalte (2)

Nach den Anhörungen legt die PG verbindliche und eindeutige Definitionen zu allen Kerndaten fest. Die Definitionen sollen weitestgehend an bestehende Standards anknüpfen. Sie sollten auch auf den CERIF-Standard abbildbar sein.

Die technische Realisierbarkeit ist durch personelle Verbindungen zur Projektgruppe „Technik, Schnittstellen und Datenformate“ im Blick zu behalten.

Standardisierung  
Datenformate  
(3)

Unter Rückbindung an die implementierende Hochschulen und AUF erfolgt eine Festlegung der Datenformate. Diese werden durch Beispiele ergänzt.

Die technische Realisierbarkeit ist durch personelle Verbindungen zur Projektgruppe „Technik, Schnittstellen und Datenformate“ im Blick zu behalten.

### III.2.c Projektgruppe „Bibliometrie“ (Zeitangaben in Monaten)

#### Standardisierung Publikations- formate STUFE I (5)\*

Die Projektgruppe „Bibliometrie“ erarbeitet einheitliche Standards für die Erfassung von Publikationsdaten. Sie beteiligt dabei insbesondere auch Bibliotheken. In einer ersten Stufe geschieht dies nur für solche Fächer, in denen bereits ein substantieller Konsens über relevante Arten von Publikationen besteht. Die PG bereitet eine Liste mit Fächern und eine Analyse der fachspezifischen Einordnung von Publikationen vor. Dabei wird sichergestellt, dass bestehende Literaturverzeichnisse und Datenbanken möglichst weitgehend genutzt werden können.

Die technische Realisierbarkeit wird durch personelle Verbindungen zur Projektgruppe „Technik, Schnittstellen und Datenformate“ im Blick behalten.

#### Import Schnittstellen STUFE I (2)

In Absprache mit den Piloteinrichtungen identifiziert das Projektmanagement wichtige Zielsysteme für Schnittstellen zum Import bestehender Publikationsdaten gemäß den Vorgaben. Die konkrete Umsetzung bleibt zunächst auf die Piloteinrichtungen beschränkt.

#### STUFE II (6)

Nach der ersten Stufe wendet sich die Projektgruppe „Bibliometrie“ den Fächern zu, in denen bislang eine standardisierte Erfassung von Publikationsdaten schwieriger erscheint.

Die technische Realisierbarkeit wird durch personelle Verbindungen zur Projektgruppe „Technik, Schnittstellen und Datenformate“ im Blick behalten.

#### Import Schnittstellen STUFE II (2)

Auch für Stufe II stellt die Projektgruppe sicher, dass die Grundlagen für Schnittstellen zu bestehenden Literaturdatenbanken bestehen. Die konkrete Umsetzung bleibt zunächst auf die Piloteinrichtungen beschränkt.

## III.2.d Projektgruppe „Fächerklassifikation“ (Zeitangaben in Monaten)



Vorschlag  
Klassifikation  
(5)

In enger Abstimmung mit DFG und dem Statistischen Bundesamt legt der Beirat einen Vorschlag für eine Fächerklassifikation vor. Dieser orientiert sich maßgeblich an bestehenden Klassifikationen, z. B. von Fachgesellschaften sowie anderen internationalen oder nationalen Organisationen.

Die Klassifikation wird ergänzt durch Schlüsselwortlisten, einen Thesaurus also, der auf gegebenenfalls bestehenden Fachthesauri aufbaut.



Konsultation  
(Anhörungen &  
schriftlich)  
(8)

Ein breit angelegtes Konsultationsverfahren (z. B. über die Homepage des Wissenschaftsrates) schließt sich an. Dieser ist aufgrund ggf. erforderlicher interner Abstimmungsprozesse in Fachgesellschaften und anderen Organisationen zeitaufwändig, ist aber zur Erreichung der notwendigen Akzeptanz unerlässlich. Ergänzt wird das Konsultationsverfahren um die Anhörung politisch wichtiger Akteure.



Konsolidierung/  
Konkordanz

Die Vorschläge werden abschließend zusammengefasst und die Konkordanz zu bestehenden bedeutsamen Klassifikationen dokumentiert. Daraus resultiert eine verbindliche Übersetzungstabelle zu bestehenden Klassifikationen.

Die technische Realisierbarkeit wird durch personelle Verbindungen zur Projektgruppe „Technik, Schnittstellen und Datenformate“ im Blick behalten.

### III.2.e Projektgruppe „Technik, Schnittstellen und Datenformate“ (Zeitangaben in Monaten)

Abgleich mit  
CERIF  
Standard

Technische  
Machbarkeit  
Pilotstudie

Schnittstellen  
Pilotstudie

Bei allen Standardisierungsschritten wird die technische Realisierbarkeit mit bedacht. Parallel baut das Projektmanagement Kontakte zu relevanten Softwareherstellern und potentiellen Datenlieferanten (wie z.B. DFG, CORDIS, AvH) auf, damit alle wichtigen Schnittstellen zum Import von bereits bestehenden Daten und zum Export der Daten des Kerndatensatzes zeitnah nach der Verabschiedung der Spezifikation umgesetzt werden können.

Die konkrete Umsetzung bleibt zunächst auf die Piloteinrichtungen beschränkt.

Koordination  
mit Software-  
Anbietern (5+)

Nach der Veröffentlichung der Spezifikation des Kerndatensatzes arbeitet das Projektmanagement sowie die Projektgruppe „Technik, Schnittstellen und Datenformate“ eng mit Anbietern von Informationssystem-Software zusammen. Hauptziel ist ein *Import-File* nach den Vorgaben des Kerndatensatzes für Datenabfrager aus weit verbreiteten Softwaresystemen heraus zur Verfügung zu stellen.

### III.2.f Veröffentlichung und begleitende Öffentlichkeitsarbeit

Zwischenbericht  
an den WR

Nach dem Ablauf des ersten Projektjahres legen die PG dem Wissenschaftsrat über den Beirat einen Zwischenbericht zum erreichten Stand des Projektes vor. Die Definitionen der Kerndatenformate sollten hierfür bereits vorliegen.

Veröffentlichung  
der  
Spezifikation  
durch WR

Zum Abschluss des Projektes veröffentlicht der Wissenschaftsrat eine verbindliche und operationale Spezifikation zu

- 1 – den Inhalten des Kerndatensatzes mitsamt eindeutigen Definitionen aller seiner Elemente,
- 2 – Vorgaben für die Datenformate,
- 3 – Vorgaben für Publikationsdaten, sowie

---

4 – einen Vorschlag für eine Fächerklassifikation.

---

Bundesweite  
Informations-  
veranstaltungen

Von Beginn des Projektes an verantwortet das Projektmanagement eine proaktive Öffentlichkeitsarbeit. So können bspw. Informationsveranstaltungen

- 1 – die Transparenz des Standardisierungsprozesses und
- 2 – die Bekanntheit des Kerndatensatzes erhöhen.

Nach der Veröffentlichung der Spezifikationen des Kerndatensatzes befördert die Öffentlichkeitsarbeit eine zeitnahe Implementierung des Kerndatensatzes in der Fläche.

### III.2.g Projektabschluss

---

Überprüfung  
alle 5 Jahre

Einzelne Teile des Kerndatensatzes bedürfen in unterschiedlichem Maße der Ergänzung und Aktualisierung. Definitionen sollten so spätestens nach fünf Jahren überprüft und gegebenenfalls angepasst werden. Entsprechende Überlegungen werden zum Abschluss des Projektes angestellt.

## D.IV TEILPROJEKTE UND AUFGABEN (LASTENHEFT)

---

### IV.1 Entwicklung einer konsensfähigen Fächerklassifikation

Unter Einbeziehung der DFG und des statistischen Bundesamtes soll ein Konsultationsverfahren mit dem Ziel durchgeführt werden, eine konsensfähige Fächerklassifikation zu entwickeln, die

- \_ sich an internationalen und nationalen Vorlagen orientiert,
- \_ auf drei Ebenen (Wissenschaftsbereich, engere Fachbezeichnung, detaillierte Fachbezeichnung) standardisiert ist,
- \_ über Fächer hinweg vergleichbar ist,
- \_ unterhalb der detaillierten Fachbezeichnung jeweils fachspezifische Klassifikationen für weitere Spezialisierungen nutzt,
- \_ hierarchisch aufgebaut ist und eindeutige Zuordnungen ermöglicht.
- \_ die Abbildung interdisziplinärer Forschung ermöglicht.
- \_ durch einen Thesaurus ergänzt wird, der auch fächerübergreifende Zuordnungen einzelner Wissenschaftlerinnen und Wissenschaftler ermöglicht.

Für alle Fächer erfolgt eine Festlegung der für den Kerndatensatz relevanten Publikationsformen sowie der vorzuhaltenden Informationen (gemäß den Vorgaben dieser Empfehlung). Diese basiert auf einer Übersicht über die gegenwärtige Situation in den Fächern. Anschließend werden zunächst die Fächer bearbeitet, in denen weitgehender Konsens über Umgang mit Publikationen besteht. Alle weiteren Fächer werden in einer anschließenden zweiten Stufe bearbeitet. Die Umsetzung orientiert sich dabei an CERIF. Da Publikationsverzeichnisse nur dann belastbar sind, wenn sie eine zweifelsfreie Zuordnung von Autorinnen und Autoren zu Publikationen ermöglichen, wäre zu prüfen, inwieweit sich eine eindeutige ID zur Lösung von Zuordnungsproblemen von Publikationen zu Autorinnen und Autoren gleichen Namens eignet.

#### IV.3 Definitionen der Elemente des Kerndatensatzes

Es bedarf einer Verständigung auf eindeutige und klar operationalisierbare Definitionen für alle Elemente des Kerndatensatzes, die zusammen mit den dazugehörigen Erläuterungen in Abschnitt C zusammengestellt sind. Insbesondere müssen die dort erwähnten relevanten Bezugsgrößen, wie beispielsweise eine Professur (z.B. Umgang mit Honorarprofessuren, außerplanmäßige Professuren) oder ein Projekt (Umgang mit Verbundprojekten, ...) semantisch eindeutig definiert werden. Dabei ist nach Möglichkeit die Kompatibilität zu gängigen Definitionen, z. B. des Statistischen Bundesamtes oder des CERIF-Standards, sicher zu stellen. Auch ist eine Verständigung auf herausgehobene Forschungspreise und Auszeichnungen in den verschiedenen Fachgebieten Teil der Definitionsleistung. Ebenso erfolgt eine eindeutige Festlegung der Zurechnung von Forschungsaktivitäten (*current-potential* oder *work-done-at*). In einem Glossar sind die Definitionen der Elemente des Kerndatensatzes mit Erläuterungen zu dokumentieren.

#### IV.4 Festlegung der Datenformate

Die Datenformate für alle Elemente des Kerndatensatzes sind festzulegen. Die Umsetzung orientiert sich am CERIF-Standard. Künftige mögliche Erweiterungen sind dabei im Blick zu behalten.

#### IV.5 Erstellung eines konzeptuellen Datenmodells

Als Grundlage für eine weitergehende Nutzung des Kerndatensatzes ist ein Datenmodell auf Basis einer Ontologie zu erarbeiten. Es schafft u. a. die Grundlagen für die Abbildung zwischen den Datenmodellen bestehender Forschungsinformationssysteme und dem Kerndatensatz. Damit lassen sich Schnittstellen zum automatisierten Export des Kerndatensatzes und zur Verknüpfung der Forschungsinformationssysteme verschiedener Einrichtungen zur Unterstüt-

zung institutioneller Kooperationen leichter entwickeln. Die Umsetzung orientiert sich dabei an CERIF. Das Datenmodell wird mit der Spezifikation veröffentlicht.

#### IV.6 Beurteilung der Machbarkeit

Es ist eine Übersicht über unterschiedliche Gegebenheiten bei den Piloteinrichtungen zu erstellen. Es ist zu prüfen, ob und in welchem Rahmen eine Implementierung des Kerndatensatzes erfolgen kann bzw. welche Hilfestellung ggf. vonnöten ist. Ziel ist es eine Einschätzung abzugeben, bis zu welchem Zeitpunkt eine flächendeckende Einführung erwartet werden kann.

#### IV.7 Festelegung von Schnittstellen

Es sind Schnittstellen zum Import bestehender Informationen für die Verwendung im Rahmen des Kerndatensatzes zu definieren. Dazu sind mit den Piloteinrichtungen die Zielsysteme für die Entwicklung der Schnittstellen festzulegen. Beispielhaft kommen folgende Informationen für den Import infrage:

- \_ Publikationsdaten aus
  - \_ Bibliographischer Software (z.B. EndNote oder Zotero),
  - \_ Bibliothekskatalogen,
  - \_ kommerziellen Datenbanken (z.B. Web of Science oder PubMed),
  - \_ bestehenden Forschungsinformationssystemen jeweils für die Piloteinrichtungen.
- \_ Projektdaten aus
  - \_ GEPRIS der DFG
  - \_ CORDIS der EU
  - \_ Datenbanken der Volkswagenstiftung.
- \_ Stipendien z. B. der AvH.

Sofern Förderinformationen nicht bereits öffentlich zugänglich sind, sind Gespräche zu führen, wie sie den Einrichtungen rechtskonform zur Verfügung gestellt werden können.

#### IV.8 Standards für den Datentransfer

Zu leisten ist die Festlegung eines Verfahrens und eines Datenformates, für die Weitergabe von Kerndaten an Datenabfrager oder andere Einrichtungen sowie an Wissenschaftlerinnen und Wissenschaftler, die eine Einrichtung verlassen. Alternativ zu einem elektronischen Austauschformat sollte auch die Festlegung für eine papiergebundene Weitergabe der Daten in Form von Tabellen und Listen erfolgen.

Parallel zur inhaltlichen Arbeit obliegt dem Projektmanagement die Begleitung des Projektes durch eine proaktive Öffentlichkeitsarbeit. Deren Ziele sind:

- \_ Offenlegung des Prozesses,
- \_ Kommunikation der Inhalte,
- \_ Verweis auf den Umgang mit bestehenden Systemen,
- \_ Gewinnung der Einrichtungen für schnelle Umsetzung,
- \_ Vermittlung von Kontakten zum Stand der Technik der Forschungsinformationssysteme.

#### **D.V    PROJEKTABSCHLUSS**

---

Nach Abschluss des in dieser Empfehlung festgelegten Standardisierungsprozesses setzt der Beirat Ziele für eine mögliche Fortführung dieses Projektes und erwägt einen sinnvollen Zeitpunkt für eine erste Evaluation. Der Wissenschaftsrat geht davon aus, dass einzelne Teile der Spezifikation des Kerndatensatzes spätestens nach fünf Jahren überprüft und gegebenenfalls angepasst werden müssen. Zu Art, Umfang und einem Verfahren der nötigen Anpassungen wird sich der Wissenschaftsrat zu gegebener Zeit äußern.



---

# Anhang

<b>Glossar</b>	<b>57</b>
<b>Abkürzungsverzeichnis</b>	<b>58</b>
<b>Abbildungsverzeichnis</b>	<b>59</b>



Dieses Glossar definiert zentrale Konzepte dieser Empfehlung.

Ein **Kerndatum** repräsentiert eine einzelne Klasse von Informationen über Forschungsaktivitäten. Kerndaten können entweder individualisierte Informationen (Empfängerin oder Empfänger eines Forschungspreises, Autorin oder Autor einer Publikation, ...) oder aggregierte Informationen enthalten (Anzahl wissenschaftlicher Mitarbeiterinnen und Mitarbeiter).

Der **Kerndatensatz** eines *Faches* ist die Gesamtheit aller Kerndaten des Faches für den Zeitraum eines *Kalenderjahres*, die für den Export an Datenabfrager bereitgehalten werden.

Die **Spezifikation** der *Kerndaten* bzw. des *Kerndatensatzes* ist die Menge der Definitionen, die die Datenformate für den Austausch von Kerndaten beschreiben.

Die **Standardisierung** des *Kerndatensatzes* ist der Prozess, der, unter Einbeziehung relevanter Akteure, zur Erstellung und regelmäßigen Aktualisierung der Spezifikation führt.

Ein **Forschungsinformationssystem** verbindet eine Vielzahl von Datenquellen zu Forschungsaktivitäten und Kontextinformationen und stellt eine Schnittstelle für standardisierte Abfragen zur Verfügung.

ASTP	Association of European Science and Technology Transfer Professionals
AUF	Außeruniversitäre Forschungseinrichtung(en)
AUTM	Association of University Technology Managers
AvH	Alexander von Humboldt Stiftung
BMBF	Bundesministerium für Bildung und Forschung
CERIF	Common European Research Information Format
CHE	Centrum für Hochschulentwicklung
CORDIS	Community Research and Development Information Service (Forschungsaktivitäten im Bereich der EG)
CRIS	Current Research Information System
DFG	Deutsche Forschungsgemeinschaft
DINI	Deutsche Initiative für Netzwerkinformationen
EuroCRIS	European Current Research Information Systems
GEPRIS	Geförderte Projekte Informationssystem der DFG
HStatG	Hochschulstatistikgesetz
ISCED	International Standard Classification of Education
KDS	Kerndatensatz Forschung
PG	Projektgruppe
VZÄ	Vollzeitäquivalente
WR	Wissenschaftsrat

Abbildung 1:	Organigramm des Projekts zur Spezifikation des Kerndatensatzes	42
Abbildung 2:	Projektablauf zur Spezifikation des Kerndatensatzes	45