

## Leibniz-Führungskolleg 2023

### Dokumentation

## „Reform der wissenschaftlichen Leistungsbewertung“

6./ 7. Oktober 2023, Geschäftsstelle der Leibniz-Gemeinschaft, Berlin

### Teilnehmende

<i>Mine</i>	<i>Altinli</i>	<i>Bernhard-Nocht-Institut für Tropenmedizin (BNITM)</i>
<i>Matthias</i>	<i>Beller</i>	<i>Leibniz-Institut für Katalyse (LIKAT)</i>
<i>Sophie</i>	<i>Biesenbender</i>	<i>Kommission für Forschungsinformationen in Deutschland (KFID)</i>
<i>Luc</i>	<i>De Meester</i>	<i>Leibniz-Institut für Gewässerökologie und Binnenfischerei (IGB)</i>
<i>Jean-Emmanuel</i>	<i>Faure</i>	<i>European Commission, DG Research and Innovation</i>
<i>Gregor</i>	<i>Kalinkat</i>	<i>Leibniz-Institut für Gewässerökologie und Binnenfischerei (IGB)</i>
<i>Matthias</i>	<i>Kiesselbach</i>	<i>DFG-Geschäftsstelle</i>
<i>Stefan</i>	<i>Kuhlmann</i>	<i>Universität Twente, Niederlande</i>
<i>Anne</i>	<i>Krüger</i>	<i>Weizenbaum-Institut</i>
<i>Ulf</i>	<i>Müller-Ladner</i>	<i>Justus-Liebig Universität Gießen</i>
<i>Isabella</i>	<i>Peters</i>	<i>ZBW - Leibniz-Informationszentrum Wirtschaft</i>
<i>Iris</i>	<i>Pigeot</i>	<i>Leibniz-Institut für Präventionsforschung und Epidemiologie - BIPS</i>

## **Teilnehmende ohne Impulsvortrag**

<i>Artemis</i>	<i>Alexiadou</i>	<i>Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)</i>
<i>Rainer</i>	<i>Danielzyk</i>	<i>ARL - Akademie für Raumentwicklung in der Leibniz-Gemeinschaft</i>
<i>Frank</i>	<i>Ewert</i>	<i>Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF)</i>
<i>Florian</i>	<i>Heider</i>	<i>Leibniz-Institut für Finanzmarktforschung SAFE</i>
<i>Katharina</i>	<i>Helming</i>	<i>Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF)</i>
<i>Michael</i>	<i>Hintermüller</i>	<i>Weierstrass-Institut für Angewandte Analysis und Stochastik (WIAS)</i>
<i>Alice</i>	<i>Hohn</i>	<i>Leibniz-Institut für Agrartechnik und Bioökonomie (ATB)</i>
<i>Oliver</i>	<i>Ibert</i>	<i>Leibniz-Institut für Raumbezogene Sozialforschung (IRS)</i>
<i>Sunhild</i>	<i>Kleingärtner</i>	<i>Deutsches Bergbau-Museum (DBM)</i>
<i>Klaus</i>	<i>Lieb</i>	<i>Leibniz-Institut für Resilienzforschung (LIR)</i>
<i>Jürgen</i>	<i>May</i>	<i>Bernhard-Nocht-Institut für Tropenmedizin (BNITM)</i>
<i>Jörg</i>	<i>Overmann</i>	<i>Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen</i>
<i>Kathrin</i>	<i>Rosing</i>	<i>Leibniz-Institut für Wissensmedien (IWM)</i>
<i>Anna-Maria</i>	<i>Schielicke</i>	<i>Leibniz-Institut für ökologische Raumentwicklung (IÖR)</i>
<i>Raimund</i>	<i>Seidel</i>	<i>Leibniz-Zentrum für Informatik (LZI)</i>
<i>Heike</i>	<i>Solga</i>	<i>Wissenschaftszentrum Berlin für Sozialforschung (WZB)</i>
<i>Veronika</i>	<i>Somoza</i>	<i>Leibniz-Institut für Lebensmittel-Systembiologie (LSB)</i>
<i>Silke</i>	<i>Voigt-Heucke</i>	<i>Museum für Naturkunde (MfN)</i>
<i>Christof</i>	<i>Wolf</i>	<i>GESIS - Leibniz-Institut für Sozialwissenschaften</i>

## **Moderation**

<i>Bettina</i>	<i>Böhm</i>	<i>Generalsekretärin</i>
<i>Claudia</i>	<i>Labisch</i>	<i>Leiterin Europa-Büro</i>
<i>Lisa</i>	<i>Kressin</i>	<i>wiss. Referentin Referat Wissenschaft</i>

## **Geschäftsstelle der Leibniz-Gemeinschaft**

<i>Karin</i>	<i>Effertz</i>	<i>Leiterin Referat Wettbewerb</i>
<i>Anja</i>	<i>Görnitz</i>	<i>wiss. Referentin Referat Wettbewerb</i>
<i>Peter</i>	<i>Heil</i>	<i>Leiter Referat Evaluierung</i>
<i>Britta</i>	<i>Horstmann</i>	<i>Leiterin Referat Wissenschaft</i>
<i>Maria</i>	<i>Migawa</i>	<i>wiss. Referentin Referat Gremien</i>
<i>Gabriel</i>	<i>Sollberger</i>	<i>wiss. Referent Referat Wettbewerb</i>
<i>Veronika</i>	<i>Thanner</i>	<i>Leiterin des Präsidialstabs</i>
<i>Sylvaine</i>	<i>von Franqué</i>	<i>Assistentin der Generalsekretärin</i>
<i>Christine</i>	<i>Wennrich</i>	<i>Leiterin Referat Transfer</i>

## Programm

---

Freitag, 6. Oktober 2023

---

**13:00**    **Gemeinsamer Mittagsimbiss**

**14:00**    **Begrüßung**

**Bettina Böhm**

**14:15**    **Externe Perspektive: Wissenschaftliche Leistungsbewertung in der Forschungsförderung**

**Matthias Kiesselbach (DFG-Geschäftsstelle):** Zur Reform von Forschungsbewertung: Aktuelle Initiativen und Perspektiven in DFG und CoARA

**Jean-Emmanuel Faure (EU-Kommission):** Reforming research assessment: rationale, progress made and role of the European Commission

*Diskussion*

**15:45**    **Interne Perspektive I: Verfahren wissenschaftlicher Leistungsbewertung in der Leibniz-Gemeinschaft**

**Matthias Beller (LIKAT, SAW)** Forschungsbewertung im organisationsinternen Wettbewerb

**Ulf Müller-Ladner (Justus-Liebig Universität Gießen, SAE):** Zum Evaluierungsverfahren des Senats der Leibniz-Gemeinschaft

*Diskussion*

**16:45**    **Pause**

**17:15**    **Interne Perspektive II: Impulse für Veränderung aus der Leibniz-Gemeinschaft**

**Leibniz-StG „Wissenschaftliches Publizieren“: Iris Pigeot (BIPS), Luc De Meester (IGB):** Suggestions on the principles of publication-based research assessment within and outside of Leibniz

**Mine Altinli (BNITM), Christian Nehls (FZB), Gregor Kalinkat (IGB):** Exploring Challenges in Science Evaluation: Perspective from Early Career Researchers

*Diskussion in Gruppen*

**19:30**    **Gemeinsames Abendessen**

**Pascarella**

Chausseestraße 30, 10115 Berlin

---

Sams-  
tag,

7. Oktober 2023

---

9:00 **Reflexion des Vortages**

Alle Teilnehmenden

10:00 **Internationale Wissenschaft oder nationale Wissenschaftssysteme?  
Leistungsbewertung im internationalen Vergleich**

**Stefan Kuhlmann (Prof. Emeritus, Universität Twente):** Recognition and rewards systems for academics: current transformations in the Netherlands

**Isabella Peters (ZBW, G6 Task Force on Research Assessment):** Nationale Perspektiven auf Forschungsevaluation

*Diskussion*

11:00 **Pause mit Fototermin**

11:30 **Wie wir bewerten wollen: (Digitale) Strukturen und Grundlagen der Leistungserfassung**

**Sophie Biesenbender (Kommission für Forschungsinformationen in Deutschland):** Der KDSF als wissenschaftsgeleiteter Standard für die Erfassung und Nutzung von Forschungsinformationen

**Anne K. Krüger (Weizenbaum Institut):** Kommerzielle und offene Infrastrukturen und ihre Konsequenzen für wissenschaftliche Leistungsmessung

*Diskussion*

12:30 **Schlussbetrachtung und Ausblick**

13:00 **Gemeinsamer Mittagsimbiss**

## Externe Perspektive: Wissenschaftliche Leistungsbewertung in der Forschungsförderung

**Matthias Kiesselbach (DFG-Geschäftsstelle):** *Zur Reform von Forschungsbewertung: Aktuelle Initiativen und Perspektiven in DFG und CoARA*

Die DFG engagiere sich aktiv im internationalen Reformprozess von CoARA, zum einen aufgrund inhaltlicher Übereinstimmung mit den CoARA-Zielen, zum anderen um den Reformprozess mitgestalten zu können. Die DFG unterstütze die Hauptziele der CoARA – die Stärkung der qualitativen Begutachtung, die Abkehr von unsachgemäßen und unreflektierten Verwendungen quantitativer Proxies in der Forschungsbewertung, insbesondere bei Projektanträgen, sowie die Berücksichtigung einer größeren Bandbreite wissenschaftlich wertvoller Beiträge und Praktiken in der Bewertung von Forschenden. In Bezug auf ihr eigenes Förderhandeln sehe sich die DFG bereits auf einem sehr guten Weg. Schon vor ihrem Engagement in der CoARA habe die DFG entsprechende Reformbestrebungen umgesetzt. Seit langem sei etwa in DFG-Anträgen die Anzahl der aufführbaren eigenen Publikationen beschränkt und würden Gutachtende gebeten, die Qualität der Vorarbeiten höher zu gewichten als die Quantität der Publikationen. Zuletzt seien Lebenslaufformulare mit optionalen narrativen Elementen sowie der Möglichkeit der Nennung von Vorarbeiten jenseits klassischer Journalartikel eingeführt worden. Zudem würde von Antragstellenden nun erwartet, dass Vorarbeiten nicht bloß genannt, sondern in inhaltlichen Bezug zum aktuellen Antrag gestellt werden. Insgesamt sei es ein wichtiges Anliegen der DFG, dass die Forschungsbewertung wissenschaftsadäquat sei und keine wissenschaftlich problematischen Anreize setze. Kritiker des Reformprozesses sähen eine Gefahr in der Diversifizierung der Kriterien. So werde befürchtet, dass neue Kriterien die bisherigen Primärkriterien bei der Bewertung überlagern und damit auch das Verständnis von Exzellenz verwässern könnten oder dass zusätzliche Kriterien zu einer Mehrbelastung würden. Die DFG teile diese Sorge nicht. Anstatt die Exzellenzfestlegung zu verwässern sei es der CoARA daran gelegen, durch qualitativ arbeitende Begutachtung exzellente Forschung zielicherer zu identifizieren. Weiterhin befürchteten einige deutsche Wissenschaftsakteure einen Verlust an Autonomie der Wissenschaft durch ein Eingreifen der Politik. Auch diese Gefahr sehe die DFG nicht. Die EU-Kommission etwa sei in der CoARA nur ein Mitglied unter vielen und arbeite mit anderen beteiligten Organisationen auf Augenhöhe zusammen.

**Jean-Emmanuel Faure (EU-Kommission):** *Reforming research assessment: rationale, progress made and role of the European Commission*

Um das Forschungssystem für die jüngere Generation attraktiv zu machen, sei ein grundlegender Kultur- und Systemwandel erforderlich. Dazu müssten alle Akteure – Forscher:innen, Hochschulen, politische Entscheidungsträger etc. – ins Boot geholt werden. Seit Verabschiedung der DORA-Erklärung 2012 habe es nur relativ langsame Fortschritte gegeben. Die Arbeit der ehemaligen Open Science Policy Platform mache jedoch deutlich, wie wichtig eine Reform der Forschungsbewertung sei, um die Praxis der offenen Wissenschaft zu verallgemeinern, die mit Unterstützung der EU-Mitgliedstaaten initiiert wurde. Über 600 Forschungsorganisationen hätten sich auf Grundsätze einer verbesserten Forschungsbewertung und einer gemeinsamen Vision geeinigt. Die Unterzeichner haben die CoARA Koali-



tion gegründet. CoARA solle als Lernplattform fungieren mit dem Ziel, zusammenzuarbeiten, bewährte Verfahren zu identifizieren und dabei die Autonomie der Forschungsorganisationen zu respektieren. Dafür sei eine kritische Masse an Forschungsorganisationen wichtig.

Bezüglich des Verhältnisses von qualitativer und quantitativer Bewertung sei es wichtig, die richtige Balance zu finden. Das Ziel der Reform dürfe nicht sein, Metriken auszuschließen, sondern sie verantwortungsvoll zu nutzen. So dürfe auch die Gefahr des subjektiven Bias nicht vernachlässigt werden. Bezüglich der Bewertungslast von Wissenschaftler:innen könnten neue Auswahlverfahren, oder eine breitere Verteilung, z. B. unter Einbezug junger Wissenschaftler:innen, Abhilfe schaffen. Im Auge zu behalten sei die Entwicklung der Bedeutung von Künstlicher Intelligenz in wissenschaftlichen Antrags- und Bewertungsverfahren. Eine Position auf EU-Ebene gebe es bislang dazu nicht, die EU-Kommission habe jedoch einige Arbeiten dazu eingeleitet.

## Interne Perspektive I: Verfahren wissenschaftlicher Leistungsbewertung in der Leibniz-Gemeinschaft

**Matthias Beller (LIKAT, SAW):** *Forschungsbewertung im organisationsinternen Wettbewerb*

Drei Thesen werden zur Forschungsbewertung im Allgemeinen und zum Leibniz-Wettbewerb im Spezifischen formuliert: 1) Leistungsbewertung sei eine zentrale Notwendigkeit in der Wissenschaft, 2) sie unterscheide sich je nach Bewertungsgegenstand und 3) eine quantitative Bewertung sei nicht per se schlechter als eine qualitative – im Idealfall würden sie sich ergänzen.

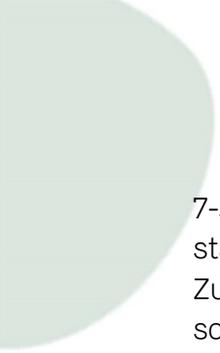
Beim SAW-Verfahren werde bereits eine Vielfalt von möglichen Leistungskriterien berücksichtigt, auch mit Unterschieden zwischen den Programmen. Entscheidungen, die lediglich auf quantitativen oder qualitativen Kriterien beruhen, gebe es nicht. In der Förderlinie Professorinnenprogramm und Nachwuchsgruppen spiele z. B. die Persönlichkeit eine besondere Rolle. Eine hundertprozentig objektive Bewertung könne es nicht geben. Für alle, aber besonders abgelehnte Antragsteller müsse das Verfahren maximal transparent dargestellt werden, um Vertrauen herzustellen. Die starke Berücksichtigung von quantitativen Indikatoren, wie der h-Index, könne man stärker an Hochschulen als an den außeruniversitären Forschungsorganisationen beobachten. Problematisch könne es sein, wenn diese Maßstäbe von extern angelegt würden. Insgesamt sei die Begutachtung im Rahmen des Leibniz-Wettbewerb ein dynamischer Prozess, der sich regelmäßig anpassen müsse.

**Ulf Müller-Ladner (Justus-Liebig Universität Gießen, SAE):** *Zum Evaluierungsverfahren des Senats der Leibniz-Gemeinschaft*

Die Evaluierung beurteile die Leibniz-Einrichtungen in Bezug auf fünf zentrale „Gegenstandsbereiche“. Das Gesamtkonzept eines Instituts und die darauf basierenden Leistungen seien dabei zentral. Mit dem Gesamtkonzept gewichteten die Einrichtungen jeweils die Bedeutung der drei zentralen Aufgabentypen Forschungsinfrastrukturen, Forschung und Wissenstransfer. Zu beurteilen sei dann die innere Logik und Stimmigkeit dieser Gewichtung.

Die Institute seien gebeten, ihre Leistungen in allen drei Aufgabenbereichen darzulegen. Während die qualitativen und quantitativen Hinweise zum Bereich Forschung in der Regel gut strukturiert und damit bewertbar seien, ließen sich die Darlegungen zu forschungsinfrastrukturellen und Transfer-Leistungen vielfach noch verbessern. Um die qualitative Bewertung zu erleichtern, bäten SAE und Senat die Institute seit fünf Jahren darum, „Highlights“ ihrer Leistungen zu benennen. Dieses Vorgehen bewähre sich. Es bleibe aber eine Herausforderung für die Leibniz-Gemeinschaft, institutsübergreifend für FI- und Transfer-Leistungen Standards bzw. neue Leistungsparameter zu definieren.

In der anschließenden Diskussion werden Diversität und Interdisziplinarität der Bewertungsgruppen angesprochen. Es wird darauf hingewiesen, dass die SAE-Mitglieder, die die Zusammensetzung der Bewertungsgruppen verantworteten und sich dabei u. a. auf Vorschläge der Institute stützten, eine möglichst breite personelle Diversität anstrebten. Im



7-Jahres-Zyklus von Evaluierungen aller Institute (2016–2023) seien 40% der Sachverständigen im Ausland tätig gewesen, ebenfalls 40% seien Frauen gewesen. Die disziplinäre Zusammensetzung entspreche jeweils dem Fachspektrum eines Instituts. Das akademische Alter der Sachverständige werde nicht erfasst, leitungserfahrene Sachverständige würden aber tendenziell in höherem Maße beteiligt.

## Interne Perspektive II: Impulse für Veränderung aus der Leibniz-Gemeinschaft

**Iris Pigeot (BIPS), Luc De Meester (IGB) (Leibniz-StG „Wissenschaftliches Publizieren“):** *Suggestions on the principles of publication-based research assessment within and outside of Leibniz*

Zwei Prinzipien der wissenschaftlichen Leistungsbewertung scheinen zentral und Konsens zu sein: 1) ein ausgewogener Einsatz quantitativer und qualitativer Indikatoren und 2) die Anerkennung einer Vielfalt von wissenschaftlichen Aktivitäten als Leistung.

In der Diskussion werde deutlich, dass sowohl ein blindes Vertrauen auf Messgrößen als auch ein blindes Vertrauen auf Erzählungen vermieden werden sollte. Die richtige Balance zu finden, bleibe eine große Herausforderung, z. B. im Hinblick auf den Anspruch möglichst objektiver Verfahren, knappe Zeitressourcen und vielfältige Karrierewege. Es bestehe Einigkeit darüber, dass der Qualität der Veröffentlichungen Vorrang vor ihrer Quantität eingeräumt werden sollte. Eine vordefinierte Checkliste mit zu stellenden Fragen könnte eine Orientierung für die inhaltliche Gestaltung der qualitativen Bewertung bieten. Es sei nicht einfach, die Qualität von Zeitschriften angemessen zu berücksichtigen, wenn man sich nicht blind an den Impact-Faktoren kommerzieller Zeitschriften orientieren wolle. Fachspezifische oder institutionsinterne Positivlisten von Zeitschriften mit hoher Qualität innerhalb der Disziplin könnten hier Unterstützung bieten, sollten aber eine Bewertung der Qualität jeder einzelnen Publikation nicht ersetzen. Neben der Erfahrung als Leser:innen und Publizierende könnten auch die Erfahrungen als Editor:innen und Reviewer einfließen, wobei z. B. Zugriffszahlen ein Indikator für die Qualität von Peer-Review-Zeitschriften sein könnten. Eine Alternative zu Positivlisten könnte eine abgestufte Kategorisierung von Zeitschriften sein (d. h. verschiedene Listen mit einer gewissen Qualitätsdifferenzierung). Solche Positivlisten von Zeitschriften könnten mit Unterstützung des wissenschaftlichen Beirats des jeweiligen Instituts erstellt bzw. genehmigt werden und müssten aktuell gehalten werden. Für hoch interdisziplinäre Institute seien derartige Positivlisten jedoch schwer zu erstellen. Bei einer Reform sei zu berücksichtigen, dass die Leibniz-Gemeinschaft im Einklang mit der internationalen Praxis bleiben müsse, da sie nicht als Insel funktionieren könne.

In Bezug auf die Anerkennung einer Vielfalt von wissenschaftlichen Aktivitäten und nicht nur die wissenschaftlich hochrangige Publikation als Leistung sei wichtig anzumerken, dass von Wissenschaftler:innen nicht verlangt werden könne, die gesamte Bandbreite wissenschaftlicher Leistungen abdecken zu können. Zudem sollen auch zukünftig wissenschaftliche Publikationen ein besonderes Gewicht bekommen.

Leibniz-intern könnte sich eine Gruppe zusammenfinden, die eine Konsensliste von „möglichen wissenschaftlichen Leistungen“ mit illustrativen Beispielen formuliere. Eine solche Liste könne als „living document“ stetig erweitert werden.

Allerdings sei es schlussendlich Aufgabe der Institutsleitungen, den mit den oben genannten Punkten zusammenhängenden Kulturwandel voranzubringen.

**Mine Altinli (BNITM), Christian Nehls (FZB), Gregor Kalinkat (IGB):** *Exploring Challenges in Science Evaluation: Perspective from Early Career Researchers*

Für viele Early Career Researcher (ECR) basierten die Leistungserwartungen, denen sie gerecht werden müssten, teilweise auf anachronistischen Vorstellungen von Wissenschaft und der Person, die sie betreibt. Durch diesen Widerspruch zur gelebten Praxis der ECR würden zum einen fragwürdige Anreize gesetzt, z. B. der Fokus auf die individuelle statt kollaborative Leistung, zum anderen würden die eigentlichen Leistungen der ECR, z. B. im Rahmen von Promotionsbetreuungen, nicht hinreichend gewürdigt. Zudem fehle es an Transparenz und Konsistenz bzgl. der Leistungen, die für einen erfolgreichen Verbleib in der Wissenschaft zu erbringen seien. Eine Reform der Leistungsbewertung müsse diese Aspekte auch vor dem Hintergrund der aktuellen Diskussionen um das Wissenschaftszeitvertragsgesetz (WissZeitVG) mitdenken. Neben veränderten Anforderungen an die ECR sei auch die höchste Leitungsebene stärker im Hinblick auf Leitungs- und Mentoringkompetenzen zu bewerten.

In der anschließenden Diskussion zeigen sich die unterschiedlichen Erwartungen an Wissenschaftler:innen, von denen sowohl ein außergewöhnliches Maß an Kreativität als auch die Bereitschaft zur hohen Arbeitsbelastung erwartet werde. ECR müsse die Möglichkeit gegeben werden, ihre Selbstständigkeit zu entwickeln und unter Beweis zu stellen. Dies setze jedoch die entsprechende Entlastung von zusätzlichen Aufgaben und die Möglichkeit voraus, jenseits von Projektförderung auf Haushaltsmittel zugreifen zu können. Eine stärkere Berücksichtigung der Vielfalt an Personen, die Wissenschaft betreiben, und eine Vielfältigkeit der berücksichtigten Leistungen könne zudem nicht das grundlegende Problem der wenigen unbefristeten Stellen im System lösen. Leistungen der Leitungen im Bereich des Mentoring könnten daher auch eine realistische Beratung im Hinblick auf alternative Karrierewege umfassen. In der Evaluierung der Leibniz-Institute gebe es bisher im „Kopfloren Gespräch“ die Möglichkeit für ECR, die Situation an ihren Einrichtungen zu kommentieren. Dies könnte um die Möglichkeit anonymer Rückmeldungen erweitert werden.

## Internationale Wissenschaft oder nationale Wissenschaftssysteme? Leistungsbewertung im internationalen Vergleich

**Stefan Kuhlmann (Prof. Emeritus, Universität Twente):** *Recognition and rewards systems for academics: current transformations in the Netherlands*

Das niederländische Wissenschaftssystem und die Governance seiner Hochschulen sei zehn Jahre vom New Public Management (NPM) geprägt gewesen. Dadurch seien Hochschulen wie Unternehmen behandelt und die Leistungsbewertung, etwa durch den Niederländischen Forschungsrat (NWO), einseitig auf quantitative Indikatoren ausgerichtet worden. Wenige Förderprogramme wurden zum Flaschenhals der akademischen Karriere. So sei das Erreichen einer Professur in den Niederlanden ohne Förderung im Programm „Veni, Vidi, Vici“ kaum möglich. Die Kritik an diesem System sei 2015 in einem Protest von Studierenden der „Jungen Akademie“ (der Koninklijke Nederlandse Akademie van Wetenschappen) kulminiert. Es fokussiere einseitig auf Konkurrenz statt Kooperation und sei daher schädlich für die Qualität von Forschung. Zudem fände die Diversität von Leistungstypen und Karrierewegen keine Anerkennung. Verbunden mit der Durchsetzung von Open Science sei es in den vergangenen Jahren zu einer Neuausrichtung der Leistungsbewertung an niederländischen Hochschulen gekommen. Die großen Wissenschaftsorganisationen hätten mit dem Konzept „Room for everyone’s Talent“ eine neue Leitorientierung für Leistungsbewertung in der Wissenschaft entwickelt, die verschiedene Dimensionen gleichwertig in den Blick nehme (z. B. Impact von Forschung, Lehre, Kooperationsfähigkeit). Klassische Leistungen in der Forschung seien dabei wichtig geblieben (inklusive Publikationen und Zitationen), die Fokussierung auf quantitative Metriken habe jedoch nachgelassen. Dagegen sei das Impact Narrativ wichtig geworden. Jedoch habe auch diese Neuorientierung Probleme und Kritik mit sich gebracht. So bestehe die Gefahr, Exzellenz werde nun in allen Leistungsdimensionen erwartet. Zudem müsse man anschlussfähig an internationale Standards bleiben. Insbesondere in den Naturwissenschaften verlief der Wandel schleppend.

**Isabella Peters (ZBW, G6 Task Force on Research Assessment):** *Nationale Perspektiven auf Forschungsevaluation*

Zur besseren Einordnung der wissenschaftlichen Leistungsbewertung in Deutschland und zur informierten Beteiligung am europäischen CoARA-Prozess lohne sich der vergleichende Blick auf die Leistungsbewertung in anderen europäischen Staaten. Große Unterschiede bestünden beispielsweise zwischen Spanien, Frankreich, Italien und Deutschland im Hinblick auf die Kriterien, Verfahren und zuständige Akteure der Leistungsbewertung sowie bezüglich der Umsetzung von Karriereförderung. Damit unterschieden sich zum einen die Startbedingungen für die Beteiligung an CoARA, zum anderen mache dies deutlich, dass das Ziel dieses Prozesses zwar die gesteigerte Abstimmung hierzu jedoch kein „One size fits all“ sein könne. Mit dem Grad der Freiheit der Wissenschaft in einem Land unterscheide sich auch die Motivation zur Beteiligung an CoARA. Je stärker der Staat gegenwärtig „reinregiere“, desto größer sei die Hoffnung, durch CoARA stärker wissenschaftsgeleitet agieren zu können. In Deutschland sei mit dem Primat der Selbstverwaltung sowie der Freiheit von Forschung und Lehre hingegen die Sorge vor politischem Einfluss größer.

## Wie wir bewerten wollen: (Digitale) Strukturen und Grundlagen der Leistungserfassung

**Sophie Biesenbender (Kommission für Forschungsinformationen in Deutschland):** *Der KDSF als wissenschaftsgeleiteter Standard für die Erfassung und Nutzung von Forschungsinformationen*

In der Diskussion um das Verständnis wissenschaftlicher Leistung und Indikatorik seien die Prozesse, Systeme und digitalen Infrastrukturen zur Erfassung und Verarbeitung von Forschungsinformationen mitzudenken. Forschungsinformationen (wie z. B. Angaben zu Publikationen, Forschungsdaten oder Patenten als Outputs wissenschaftlicher Leistungsprozesse) bilden in vielen Kontexten die Grundlage für Evaluations- und Bewertungsverfahren in der Forschung. Die Erfassung und Verarbeitung von Forschungsinformationen auf der Basis von Definitionen seien Voraussetzung für eine transparente, faire und adäquate Nutzung bzw. Nicht-Nutzung dieser Angaben im Rahmen der Leistungsbewertung.

Ein bewusstes und zielorientiertes Forschungsinformationsmanagement sei somit die Grundlage für Forschungsevaluation und ein strategisches Thema einer jeden Forschungseinrichtung.

Ein aus der Wissenschaft heraus entwickelter Standard für Forschungsinformationen und ihre Verarbeitung sei der KDSF-Standard („Kerndatensatz Forschung“), der verschiedene Informationsbedarfe und Anwendungsfälle berücksichtige. Im Kontext des Reformprozesses unterstütze der KDSF eine verantwortungsvolle Forschungsevaluation, indem er die Betrachtung vielfältiger Forschungsaktivitäten und -outputs sowie fachlicher und struktureller Spezifika ermögliche und eine transparente Grundlage für die Heranziehung quantitativer Forschungsinformationen böte.

Der KDSF setze somit einen Impuls für einen effizienten, professionellen und verantwortungsvollen Umgang mit Forschungsinformationen, die Minimierung von diesbezüglichen Abhängigkeiten von Dritten sowie die Stärkung von Datenkompetenz und Datensouveränität im Kontext digitaler Transformationsprozesse im Wissenschaftssystem.

**Anne K. Krüger (Weizenbaum Institut):** *Kommerzielle und offene Infrastrukturen und ihre Konsequenzen für wissenschaftliche Leistungsmessung*

Durch digitale Infrastrukturen der Forschungsinformation und -evaluation können eine Vielzahl neuer Daten über wissenschaftliche Praxis und Leistung erhoben werden. Hiermit seien jedoch auch eine Vielzahl an Problemen verbunden: Erstens würden diese Infrastrukturen von kommerziellen Anbietern vertrieben, wodurch in der Nutzung dieser Datenbanken und Softwareprodukte neue finanzielle Abhängigkeiten entstehen könnten. Zweitens stelle sich die Frage der Qualität der zur Information und Evaluation genutzten Daten sowie nach ihren adäquaten Verwendungsmöglichkeiten. Drittens seien die Auswertungsmöglichkeiten vor allem auf die Praxis der Naturwissenschaften zugeschnitten, was anderen Fächern in der Erhebung und Darstellung von Leistung nicht gerecht werde. Viertens nähmen die Bewertungsmöglichkeiten hinsichtlich des *societal impact* von Forschung kontinuierlich zu, was dessen Relevanz im Vergleich z. B. zur Grundlagenforschung erhöhe. Und fünftens werde deutlich, dass sich dadurch nicht nur die Bewertungskriterien verändern,

sondern auch, wer – jenseits der wissenschaftlichen peers – eigenständig Informationen generieren und damit Forschung bewerten könne. Zudem werde dadurch die quantitative Leistungsmessung gestärkt bis hin zu Überlegungen der Prognose von Forschungserfolg („predictive analytics“). Dadurch werde deutlich, dass die digitalen Infrastrukturen der Forschungsinformation und -evaluation einen zentralen Anteil daran hätten, wer Wissenschaft wie bewerten kann. Dies gelte es kritisch zu reflektieren.

## Schlussbetrachtung und Ausblick

Dieses Führungskolleg sei als Auftakt des Leibniz-internen Austausches zur Reform der wissenschaftlichen Leistungsbewertung zu verstehen. Es sollte auch einen ersten Überblick über mögliche Themen geben, die in den kommenden Jahren in verschiedenen Formaten des Austauschs weiter vertieft würden. Dabei seien keine neuen Leitlinien oder Manifeste das Ziel, sondern strukturierte Diskussionen, die die Leibniz-internen Evaluierungsverfahren voranbringen und zu einer gemeinschaftlichen Position – in Rückkopplung an den CoARA-Prozess – beitragen.

Ein wichtiges Forum hierfür werde zukünftig das Leibniz-Strategieforum Research Assessment sein.

## Gruppenfoto

