



Funding Acknowledgements in der Web of Science-Datenbank

Neue Methoden und Möglichkeiten der Analyse von Förderorganisationen

Daniel Sirtes, Mathias Riechert, Paul Donner, Valeria Aman, Torger Möller

Studien zum deutschen Innovationssystem Nr. XX-2014

Institut für Forschungsinformation und Qualitätssicherung

Februar 2015

Diese Studie wurde im Auftrag der Expertenkommission Forschung und Innovation (EFI) erstellt. Die Ergebnisse und Interpretationen liegen in der alleinigen Verantwortung der durchführenden Institute. Die EFI hat auf die Abfassung des Berichts keinen Einfluss genommen.

Studien zum deutschen Innovationssystem

Nr. 9-2014

ISSN 1613-4338

Herausgeber: Expertenkommission Forschung und Innovation (EFI)

Geschäftsstelle, c/o Stifterverband für die Deutsche Wissenschaft, Pariser Platz 6, 10117 Berlin

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie die Übersetzung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der EFI oder der Institute reproduziert oder unter Verwendung elektronischer Systeme gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

Kontakt und weitere Informationen: Daniel Sirtes, iFQ, Schützenstr. 6a, 10117, Berlin.
+4930206417726. sirtes@forschungsinfo.de

Inhaltsverzeichnis

0. Kurzfassung	4
1. Einführung	5
2. Abdeckung	6
2.1 Abdeckung Dokumenttypen	6
2.2 Abdeckung gesamte Datenbank	7
2.3 Abdeckung Disziplinen	8
2.4 Abdeckung Länder.....	11
3. Methoden	12
3.1 Methode der Unifizierung der Aliase von Förderorganisationen	12
3.2 Methode der Impactmessung.....	13
4. Ergebnisse	13
4.1 Aufnahme von Acknowledgements in die Datenbank	13
4.2 Häufigste deutsche Förderorganisationen im WoS	14
4.3 Verteilung der Publikationen auf die wissenschaftlichen Sektoren	15
4.4 Förderprofile nach Disziplin.....	15
4.5 Co-Funding und internationale Kollaborationen nach Förderorganisationen	22
4.6 Impactmessung geförderter Publikationen	24
5. Abschließende Bemerkungen	38
Referenzen	39

0. Kurzfassung

Seit August 2008 befinden sich in der Thomson Reuters (TR)-Zitationsdatenbank „Web of Science“ Funding Acknowledgements. Funding Acknowledgements sind in wissenschaftlichen Artikeln vermerkte Angaben über die finanzielle Förderung der Forschung, die bei der Erstellung der Publikation verwendet wurde. Mit diesen Angaben in der Datenbank lassen sich erstmals systematische, quantitative und globale Aussagen über das Verhältnis von wissenschaftlicher Förderung und wissenschaftlichem Output treffen. Es ist jedoch nicht unproblematisch, valide Schlussfolgerungen aus den Daten zu generieren. Es ist offensichtlich, dass die Daten nicht vollständig sind und fehlerbehaftet sind.

Nur 61,3% aller Publikationen des Jahres 2013 haben in der Datenbank auch Funding Acknowledgements. Eine Analyse der Abdeckung der Publikationen mit Funding Acknowledgements zeigt zudem eine große Heterogenität: Nur Forschungsartikel haben einen relativ hohen Anteil an Funding Acknowledgements (63,5%) während die meisten anderen Dokumenttypen keine solche Angaben besitzen. Die Unterschiede der Abdeckung in den verschiedenen Disziplinen sind besonders groß. Während in den Natur- und Lebenswissenschaften Forschungsartikel mit einem Anteil von bis zu 90% mit Funding Acknowledgements versehen sind, gibt es geisteswissenschaftliche Fächer ohne ein einziges Funding Acknowledgement. Der Anteil von Funding Acknowledgements ist auch in verschiedenen Ländern heterogen. Während z.B. 84,8% der chinesischen Publikationen die Förderorganisation angeben, sind es in der Türkei nur 33,1%. Ob sich diese Unterschiede auf die Förderstrukturen im Land oder auf die Gewohnheiten der AutorInnen beim Erwähnen ihrer Förderer sich zurückführen lassen, kann man auf der Basis der Daten nicht entscheiden.

Darüber hinaus sind die Daten nicht normalisiert und müssen daher sorgfältig gesäubert werden. Es wurde eine vollautomatische Methode zur Normalisierung der vielen Varianten, die für eine Förderorganisation auftauchen, entwickelt und manuell überprüft. Die zehn am häufigsten genannten deutschen und zusätzlich vier europäische Förderorganisationen wurden für die Analyse aufbereitet. Die Deutsche Forschungsgemeinschaft hat dabei den mit Abstand größten Anteil, gefolgt von dem Bundesministerium für Bildung und Forschung. Zusätzlich aufgeführt ist die Alexander von Humboldt-Stiftung, der Deutsche Akademische Austauschdienst, die Deutsche Krebshilfe, der Fonds der Chemischen Industrie, die Helmholtz-Gemeinschaft Deutscher Forschungszentren, die Max-Planck-Gesellschaft, die Studienstiftung des Deutschen Volkes und die Volkswagen-Stiftung. Es ist hierbei zu beachten, dass die HGF und die MPG keine Förderorganisationen im herkömmlichen Sinne sind, trotzdem werden sie oft als die Finanzierer der Forschung an den beteiligten Forschungseinrichtungen genannt. Um die deutschen Ergebnisse in einen internationalen Kontext zu stellen, wurden zusätzlich vier große europäische Förderorganisationen bezüglich ihres Impacts analysiert: die französische Agence Nationale de la Recherche (ANR), der flämische Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO), der österreichische Fonds zur Förderung der wissenschaftlichen Forschung (FWF) und die niederländische Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Um die Möglichkeiten der Analyse anhand von Funding Acknowledgements aufzuzeigen, wurden die deutschen Förderer in verschiedenen Facetten porträtiert: So wurde die Fördertätigkeit an den Universitäten im Vergleich zu der an den außeruniversitären Einrichtungen aufgezeigt, ein disziplinäres Profil der Förderorganisationen anhand verschiedener Indikatoren erstellt, die Häufigkeit des Co-Fundings dieser Organisationen untereinander und mit Förderung aus EU-Geldern wurde festgestellt und der Anteil an internationaler Kollaborationen der Publikationen mit deutscher Förderung gemessen. Während der hohe internationale Anteil der Autoren bei Publikationen gefördert durch die international ausgerichteten Austauschdienst und Alexander von Humboldt-Stiftung wenig überraschend war, ist der hohe Anteil der EU-Finanzierung an Publikationen, die auch von der Helmholtz-Gesellschaft gefördert wurde, weniger vorauszusehen. Bei der Impactmessung der geförderten Publikationen ergaben sich aber die bemerkenswertesten Ergebnisse. Es zeigte sich zum Beispiel, dass die von der Max-Planck-Gesellschaft geförderten Publikati-

onen einen extrem hohen Anteil an hochzitierten Publikationen aufweist. 25,5% der MPG-geförderten Publikationen befinden sich unter den 10% höchstzitierten Publikationen weltweit. Damit gehört die Max-Planck-Gesellschaft zu den erfolgreichsten Institutionen weltweit (vergleichbar mit dem Massachusetts Institute of Technology). Ein Ergebnis, das auch Gegenstand weiterer Forschung sein wird, ist der im Vergleich mit der DFG um 24,5% höhere Anteil an hochzitierten Publikationen des durch das Bundesministeriums geförderten Publikationskorpus‘ (19.3% BMBF, 15.5% DFG).

Diese explorative Studie verdeutlicht anhand beispielhafter Einblicke das große langfristige Potenzial und die breit gefächerten Anwendungsmöglichkeiten der bereinigten Funding Acknowledgement-Daten. Der vorgestellte Bereinigungsprozess ist methodisch-empirisch unterstützbar und für die umfassende Anwendung kompletter Anbieterdatenbanken skalierbar. Zudem wurde im derzeitigen Prozess bereits umfangreiche Expertise hinsichtlich der Konfiguration und methodischen Weiterentwicklung des Prozesses entwickelt. Dadurch sind bereits zum jetzigen Zeitpunkt erkenntnisgenerierende Analysen von und durch Förderorganisationen möglich, deren Durchführung bisher häufig an mangelnder Datenqualität, einem unzureichenden empirischen Kenntnisstand und dem Aufwand der Datenbereinigung scheiterten. Obwohl diese Ergebnisse aufgrund des vorläufigen Entwicklungsstadiums derzeit erst vorsichtige Schlüsse zulassen, wurde bereits in den explorativen Anwendungsszenarien deutlich, dass diese neue Methode und Datenbasis sowohl erwartungsgemäße als auch überraschende Erkenntnisse durch Einblicke in einen bislang kaum untersuchbaren Bereich des Wissenschaftssystems ermöglicht. Schließlich soll noch besondere Erwähnung finden, dass das Ziel der Anwendung der derzeitigen Datenquelle und Methode nicht darin liegt, Evaluationen von Forschungseinrichtungen durchzuführen, da dafür, auch nach der Bereinigung, die Datenlage zu lückenhaft ist. Vielmehr soll das Ziel in der Umsetzung von Fragestellungen aus der Perspektive der Förderorganisationen selbst liegen. So lassen sich Tendenz- und Profilanalysen zur Ableitung von Handlungsszenarien einsetzen. Um die Ergebnisse für diese Förderorganisationen besser zu verstehen, bedürfen die Einflussfaktoren erfolgreicher Förderung sowie ihr organisatorischer und prozessualer Kontext jedoch tieferer Analyse.

1. Einführung

Seit August 2008 befinden sich in der Thomson Reuters (TR)-Zitationsdatenbank „Web of Science“ (WoS) auch Funding Acknowledgements. Funding Acknowledgements (FA) sind in wissenschaftlichen Artikeln vermerkte Angaben über die finanzielle Förderung der Forschung, die bei der Erstellung der Publikation verwendet wurde. Dabei werden entweder in einer allgemeinen Bedankungsfußnote oder in einer speziellen FA-Fußnote, die Förderorganisationen (FO) und oft auch das Aktenzeichen der Förderbewilligung (Grantnumber: GN) angegeben. Mit diesen Angaben in der Datenbank lassen sich erstmals systematische, quantitative und globale Aussagen über das Verhältnis von wissenschaftlicher Förderung und wissenschaftlichem Output treffen. Es ist jedoch nicht unproblematisch, valide Schlussfolgerungen aus den Daten zu generieren. Es gibt eine Vielzahl von Unbekannten: Es ist unklar, seit wann und warum AutorInnen FAs angeben. Weiterhin sind weder die Extraktionsalgorithmen von TR zum Extrahieren der FA-Texte (im Folgenden fundingtext - FT) aus den Artikeln bekannt, noch das Vorgehen bei der weiteren Extraktion der FOs und GNs aus diesem Text. Die verfügbaren Informationen im WoS zeigen dagegen deutlich, dass bei allen diesen Schritten Fehler auftreten. Weiterhin kann auf Basis der derzeitigen Information nicht eindeutig bestimmt werden, ob die AutorInnen diejenigen Fördermaßnahmen angeben, die für die Erstellung der Arbeit bewilligt wurden, oder einfach diejenigen, die sie gerade zur Verfügung hatten (sowohl zeitliche als auch thematische Diskrepanzen können auftreten - siehe (Rigby 2011b; Rigby 2011a; Sirtes 2013) für eine vorläufige Übersicht dieser und weiterer Probleme). Eines der größten Probleme mit dem derzeitigen Stand der Informationen in WoS ist die Vielzahl an Varianten für eine FO. Zum Beispiel wurden allein für die DFG bisher über 12.000 Va-

riantnen gefunden. Um dieses Problem zu adressieren, wurden am iFQ mehrere teil- und vollautomatisierte Methoden entwickelt, manuell überprüft und empirisch validiert (siehe Methoden). Die hier präsentierten Ergebnisse sind das Ergebnis einer vollautomatischen Unifizierungsmethode. Obwohl nicht alle Varianten der Förderorganisationen unifizierbar sind, erlaubt der erreichte Grad an Korrektheit eine Tendenzanalyse auf Basis der Informationen.

2. Abdeckung

Zunächst soll betrachtet werden, für welchen Anteil von Publikationen überhaupt FA verfügbar sind. Das sich ergebende Bild ist hinsichtlich der Dokumenttypen, Disziplinen und der Länder sehr heterogen .

2.1 Abdeckung Dokumenttypen

Die Übersicht der Dokumenttypen mit FA (Tabelle 1) und die Übersicht über die gesamte Datenbank (siehe Tabelle 2) verdeutlicht, dass nur Forschungsartikel (Artikel) und Reviews substanzelle Anteile an FA aufweisen. Da aus Förderersicht vor allem die von ihnen geförderte Forschung von Interesse ist und zudem die FA-Anteile der Artikel auch substanzell höher sind als die der Reviews, werden in der weiteren Analyse nur Journalartikel als Datengrundlage benutzt. Die 4,8% der Review-Publikationen mit FA (hier für das Jahr 2013) werden daher nicht berücksichtigt.

Tabelle 1: Abdeckung der Dokumenttypen deutscher Publikationen im Jahre 2013¹

	Anzahl Publikationen	Anzahl mit FA	Dokumenttyp mit FA prozentual	Anteil an Publikationen mit FA
Artikel	96.698	61.410	63,5%	95,2%
Meeting Abstract	13.834	0	0,0%	0,0%
Review	5.640	3.084	54,7%	4,8%
Editorial Material	5.244	9	0,2%	0,0%
Letter	1.927	0	0,0%	0,0%
Gesamt	123.343	64.503	52,3%	100,0%

2.2 Abdeckung gesamte Datenbank

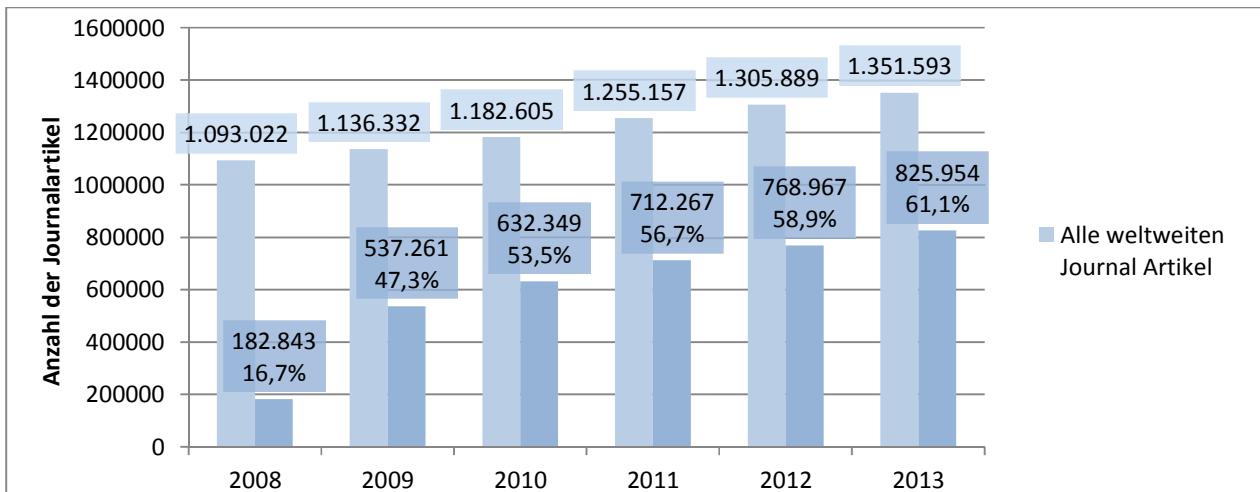
In Tabelle 2 sind alle Publikationen mit und ohne FA aufgelistet. Es wird ersichtlich, dass Publikationen, die nicht in Journals auftauchen (Buchserien, Konferenzbände), so gut wie keine FA aufweisen. Darüber hinaus ist die Abdeckung in den ersten zwei Jahren der Aufnahme von FA in der Datenbank (2008 und 2009) signifikant niedriger im Vergleich zu den letzten drei Jahren. Obwohl das relative Wachstum des Anteils von Artikeln mit FA noch nicht abgeschlossen zu sein scheint, legen die Daten die Beschränkung der Analysen auf die Jahre von 2010 bis 2013 nahe (siehe dazu auch Abbildung 1).

Tabelle 2: Alle deutschen und weltweiten Publikationen im zeitlichen Verlauf jeweils ohne und mit FA

Alle deutschen Publikationen					Alle deutschen Publikationen mit FAs					Alle weltweiten Publikationen	Alle weltweiten Publikationen mit FA
Jahr	Alle Publikationen	Alle Journal-publikationen	Journal Artikel und Reviews	Journal Artikel	Alle Publikationen	Alle Journal-publikationen	Journal Artikel und Reviews	Journal Artikel	Anteil Journal Artikel	Journal Artikel	Anteil Journal Artikel mit FA
2008	123.995	110.495	84.274	79.046	14.350	14.281	14.278	13.492	17,1%	1.093.022	182.843 16,7%
2009	129.117	114.638	87.839	81.829	42.804	42.677	42.647	40.163	49,1%	1.136.332	537.261 47,3%
2010	128.345	116.121	91.145	85.865	51.368	51.346	51.313	48.811	56,8%	1.182.605	632.349 53,5%
2011	131.514	121.398	95.845	90.604	56.988	56.961	56.944	54.368	60,0%	1.255.157	712.267 56,7%
2012	135.266	124.933	99.916	94.158	61.707	61.688	61.678	58.640	62,3%	1.305.889	768.967 58,9%
2013	132.153	125.585	102.338	96.698	64.520	64.503	64.494	61.410	63,5%	1.351.593	825.954 61,1%

¹ Mit „deutschen“ Publikationen sind hier immer Publikationen gemeint, für die mindestens einE AutorIn eine deutsche institutionelle Adresse angegeben hat.

Abbildung 1: Zeitlicher Verlauf der Abdeckung von weltweiten Journalartikeln mit FA



2.3 Abdeckung Disziplinen

Auch bezüglich der Verteilung der Artikel mit FA in verschiedenen Disziplinen ergibt sich ein sehr heterogenes Bild. Aus der aufgelisteten Auswahl der sogenannten Subject Categories (SC) der WoS (Tabelle 3) kann man ersehen, dass in Fächern wie z.B. der Biologie fast alle Journalartikel FA aufweisen. Der höchste Anteil ist jedoch in „Multidisciplinary Sciences“ zu finden. Diese SC beinhaltet die wichtigsten Journals von allgemeinem wissenschaftlichem Interesse (Science, Nature, PNAS). Dagegen haben manche medizinische Fächer sowie Informatik- und Ingenieurwissenschaften nur eine mittlere Abdeckung von etwa 40 bis 60%. Die Sozial- und Geisteswissenschaften, die insgesamt in der Datenbank unterrepräsentiert sind (Nederhof 2006; Chi 2012), haben zusätzlich nur sehr wenige oder gar keine Artikel mit FA. Für diese Unterschiede gibt es drei Erklärungsansätze: (1) Der Anteil drittfinanzierter Forschung unterscheidet sich in den Disziplinen sehr stark, (2) die Standardisierung der Angabe von FA ist in den jeweiligen disziplinären Kulturen verschieden weit fortgeschritten oder (3) die Extraktionsmechanismen von TR in den Journals der verschiedenen Disziplinen sind verschieden. Es ist bisher unklar, welcher dieser Faktoren den maßgeblichen Einfluss auf die Erklärung des Phänomens hat, jedoch sind sie Gegenstand laufender Studien des iFQ.

Tabelle 3: Auswahl der Disziplinen in WoS (SC) und ihr Anteil an weltweiten Journalartikeln mit FA 2013

Disziplin	Journal-Artikel	Journalartikel mit FA	Anteil mit FA
Multidisciplinary Sciences	4187	3797	90,7%
Astronomy & Astrophysics	3384	3047	90,0%
Evolutionary Biology	560	492	87,9%
Biochemistry & Molecular Biology	3502	3072	87,7%
Cell Biology	2032	1781	87,6%
Nanoscience & Nanotechnology	1815	1558	85,8%
Geography, Physical	546	468	85,7%
Ecology	1403	1201	85,6%
Genetics & Heredity	1590	1347	84,7%
Meteorology & Atmospheric Sciences	1204	1016	84,4%
Physics, Condensed Matter	2721	2281	83,8%
Physics, Particles & Fields	1778	1486	83,6%
Geosciences, Multidisciplinary	1938	1610	83,1%
Chemistry, Multidisciplinary	3330	2755	82,7%
Chemistry, Physical	3682	3042	82,6%
Biology	506	411	81,2%
Physics, Multidisciplinary	2545	2063	81,1%
Polymer Science	1167	928	79,5%
Neurosciences	3187	2510	78,8%
Medicine, Research & Experimental	941	734	78,0%
Optics	1777	1379	77,6%
Chemistry, Organic	1415	1096	77,5%
Mathematical & Computational Biology	509	394	77,4%
Materials Science, Multidisciplinary	4641	3571	76,9%
Biochemical Research Methods	1342	1029	76,7%
Physiology	611	467	76,4%
Physics, Mathematical	913	687	75,2%
Electrochemistry	668	501	75,0%
Physics, Fluids & Plasmas	928	691	74,5%
Infectious Diseases	523	379	72,5%
Environmental Sciences	1952	1403	71,9%
Chemistry, Medicinal	681	481	70,6%
Crystallography	500	351	70,2%
Toxicology	501	350	69,9%
Nutrition & Dietetics	445	310	69,7%
Agricultural Engineering	112	78	69,6%
Mathematics, Interdisciplinary Applications	420	291	69,3%
Anatomy & Morphology	158	109	69,0%
Entomology	206	140	68,0%

Disziplin	Journal-Artikel	Journalartikel mit FA	Anteil mit FA
Physics, Nuclear	740	501	67,7%
Energy & Fuels	1108	741	66,9%
Pharmacology & Pharmacy	1640	1083	66,0%
Psychology	590	380	64,4%
Water Resources	749	478	63,8%
Engineering, Multidisciplinary	335	209	62,4%
Mathematics, Applied	1482	920	62,1%
Computer Science, Interdisciplinary Applications	729	441	60,5%
Engineering, Electrical & Electronic	1758	1042	59,3%
Mathematics	1582	932	58,9%
Engineering, Chemical	1293	738	57,1%
Computer Science, Software Engineering	529	284	53,7%
Instruments & Instrumentation	1097	575	52,4%
Computer Science, Information Systems	460	233	50,7%
Clinical Neurology	2316	1167	50,4%
Psychiatry	1441	680	47,2%
Public, Environmental & Occupational Health	990	465	47,0%
Medicine, General & Internal	785	334	42,5%
Veterinary Sciences	826	307	37,2%
Anthropology	155	57	36,8%
Psychology, Experimental	664	234	35,2%
Dentistry, Oral Surgery & Medicine	613	208	33,9%
Nuclear Science & Technology	1064	345	32,4%
Environmental Studies	472	128	27,1%
Surgery	2237	514	23,0%
History & Philosophy of Science	147	26	17,7%
Information Science & Library Science	142	24	16,9%
Psychology, Multidisciplinary	476	57	12,0%
Linguistics	256	30	11,7%
Social Sciences, Interdisciplinary	216	16	7,4%
Economics	1707	116	6,8%
Management	482	25	5,2%
Philosophy	197	8	4,1%
Education & Educational Research	375	12	3,2%
Business	336	7	2,1%
Language & Linguistics Theory	304	4	1,3%
Sociology	256	1	0,4%
Political Science	330	0	0,0%
History	260	0	0,0%
Psychology, Psychoanalysis	132	0	0,0%
Literature	122	0	0,0%

2.4 Abdeckung Länder

Auch zwischen den einzelnen Ländern zeigt sich eine ungleiche Verteilung. Während China mit einer Abdeckung von 84,8% die höchste Abdeckung aufweist (wobei dort die mit Abstand am häufigsten genannte FO die National Science Foundation of China ist, die wohl die gesamte naturwissenschaftliche Forschung unterstützt), sind Jurnalartikel in Ländern wie der Türkei oder dem Iran offenbar nur zu etwa einem Drittel drittmittelfinanziert. Die meisten westeuropäischen Länder und Nordamerika sind jedoch im Weltdurchschnitt von etwa 60–70% zu verorten.

Tabelle 4: Anteil der Jurnalartikel mit FA in den Ländern mit höchsten Artikelzahlen für 2013

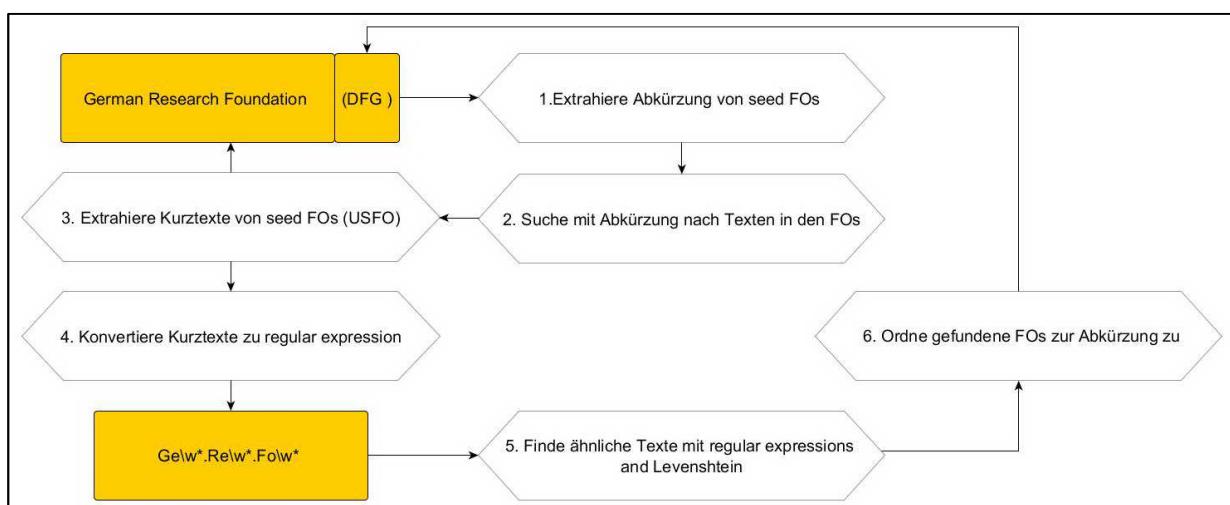
Land	Journal-artikel	Jurnalartikel mit FA	Anteil mit FA
Vereinigte Staaten	354.585	228.214	64,4%
China	213.859	181.253	84,8%
Vereinigtes Königreich	100.243	60.275	60,1%
Deutschland	96.698	61.410	63,5%
Japan	75.573	46.742	61,9%
Frankreich	66.763	41.023	61,4%
Kanada	58.649	39.644	67,6%
Italien	57.480	30.633	53,3%
Spanien	52.492	37.582	71,6%
Südkorea	50.013	37.975	75,9%
Indien	49.873	27.199	54,5%
Australien	49.333	30.771	62,4%
Brasilien	37.514	24.218	64,6%
Niederlande	34.696	21.272	61,3%
Russland	28.532	17.937	62,9%
Taiwan	27.241	19.022	69,8%
Schweiz	25.477	17.081	67,0%
Türkei	25.139	8.331	33,1%
Iran	25.006	10.152	40,6%
Schweden	23.253	16.903	72,7%
Polen	22.398	12.688	56,6%
Belgien	19.238	12.549	65,2%
Dänemark	14.553	10.360	71,2%
Österreich	13.232	8.495	64,2%
Portugal	12.418	9.156	73,7%
Israel	12.169	6.711	55,1%
Norwegen	11.230	7.376	65,7%
Mexiko	11.180	7.516	67,2%
Finnland	11.142	7.993	71,7%
Singapur	11.105	7.488	67,4%

3. Methoden

3.1 Methode der Unifizierung der Aliase von Förderorganisationen

Um einzelne FOs adäquat analysieren zu können, ist eine Unifizierung des extremen Variantenreichtums der Einträge für die einzelnen Förderorganisationen in WoS notwendig. Zur Erreichung dieses Ziels wurde ein Programm entwickelt, das folgendes Grundkonzept implementiert: Die meisten großen FOs haben standardisierte Abkürzungen, wie z.B. „DFG“ für die Deutsche Forschungsgemeinschaft. Diese Abkürzungen bilden den Ausgangspunkt einer Suche mit „regular expressions“². Auf Basis der Suchergebnisse werden die dazugehörigen Texte extrahiert. Diese Texte bilden wiederum die Grundlage für weitere regular expressions, um ähnliche Texte zu finden. So werden viele Tippfehler und auch Varianten gefunden, die den ausgeschriebenen Namen der FO aufweisen, aber nicht die Abkürzung.

Abbildung 2: Grundkonzept der Unifikation von Varianten der FOs in WOS³



Die tatsächliche Methode ist weit komplizierter, da es eine Vielzahl an Problemen gibt, wie Abkürzungen, die keine FOs sind („USA“), Kurztexte, die nicht aufschlussreich sind („a funding program of“), und Homonyme (wie die amerikanische und die chinesische National Science Foundation (NSF)). Obwohl für die meisten dieser Probleme Lösungen gefunden wurden, ist eine Differenzierung nicht immer möglich, da z.B. in den Rohdaten selbst zwei verschiedene und sogar multinationale FOs zusammengefasst wurden (z.B. „ESF“: European Social Fund und European Science Foundation). Für weitere Details über die Methode sei auf (Sirtes 2013; Sirtes und Riechert 2014; Sirtes und Riechert in preparation verwiesen).

Es konnten 90% aller deutschen Publikationen unifiziert werden (jedoch nur 58% der FOs; 42% der nicht-unifizierten FOs umfassten also nur 10% der Artikel). Durch das angewandte Konzept konnte eine sehr hohe Präzision erreicht werden, wobei die Präzision für die DFG mit 98,5% den niedrigsten Wert aufwies. Bei dem BMBF zum Beispiel lag die Präzision bei 99,8%. In vielen Fällen betrug sie

² Regular expressions oder reguläre Ausdrücke sind Muster, die verwendet werden, um Zeichenkombinationen in Strings zu finden. Für eine theoretische und technische Einführung zu regular expressions siehe <https://www.lrz.de/services/schulung/unterlagen/regul/>

³ Um die Ähnlichkeit von zwei Strings festzustellen (Schritt 5), wurde die relative Levenshtein-Distanz gemessen. (Siehe (Levenshtein 1966))

100%. Da durch die manuelle Überprüfung der Präzision bei den zehn wichtigsten Förderern alle falschpositiven FOs identifiziert und entfernt werden konnten, beträgt die Präzision bei diesen manuell geprüften FOs 100%. Diese vollautomatische Methode wurde anhand der Ergebnisse der semi-automatischen und manuellen Methode (Sirtes 2013) verglichen und zeigte zufriedenstellende Resultate. Bei der DFG wurde ein Recall von 96% der Publikationen erreicht, bei einem untersuchten Sample für die National Institutes of Health im Vergleich zum NIHRePORTER (<http://projectreporter.nih.gov/reporter.cfm>) lag der Recall auch bei 96%.⁴

3.2 Methode der Impactmessung

Als Indikator für den Impact wird der Anteil der Publikationen unter den zehn Prozent der meistzitierten Publikationen des Faches benutzt. Der Weltreferenzwert ist somit stets definiert als ein Anteil von 0,10. Die fachlichen Disziplinen sind über die WoS Subject Categories festgelegt. Für die Publikationen aus 2010 und 2011 werden die Zitationen bis zur 17. Kalenderwoche 2014 gezählt. In den Zitationsverteilungen der Disziplinen wird das 0,9-Quantil als der Schwellenwert für hochzitierte Publikationen bestimmt. Die exakten Anteile an Artikeln über und unter dem Schwellwert werden nach der Methode von (Waltman und Schreiber 2013) berechnet. Publikationen werden außerdem nach Länderanteilen Deutschlands und nach WoS SCs fraktionsiert. In den Abbildungen sind Binomial-Konfidenzintervalle (Konfidenzniveau 0,95%) nach der *Wilson score interval-Methode* gezeichnet. Für weitere Details siehe den EFI-Bericht „Quantilbasierte Indikatoren für Impact und Publikationsstrategie“.

4. Ergebnisse

4.1 Aufnahme von Acknowledgements in die Datenbank

Im Rahmen eines iFQ-Projekts zur Exzellenzinitiative wurde bei 552 deutschen Artikeln in den Ingenieurs- und Lebenswissenschaften der Jahre 2009-2011 überprüft, inwieweit die Fundingtexte in die Datenbank aufgenommen wurden. Dabei zeigte sich, dass es etwa 10% mehr Artikel mit Acknowledgement-Texten mit Informationen über Förderer als in der Datenbank gibt. Inwieweit das Fehlen dieser 10% tatsächlich einen systematischen Fehler erzeugt, ist weiterhin ungeklärt, aber Gegenstand einer weiteren, repräsentativ angelegten Studie am iFQ.

⁴ Für weitere Details siehe wieder oben erwähnte Publikationen.

Tabelle 5: Sample manuell recherchierter deutscher Publikationen nach Vorkommen von Acknowledgements.

	Anzahl der Publikationen	Prozent	Kumulative Prozent
Fundingtext im WoS	453	82,1%	82,1%
Acknowledgement-Text durch Nach-recherche (mit FA)	46	8,3%	90,4%
Acknowledgement-Text durch Nach-recherche (ohne FA)	18	3,3%	93,7%
Kein Fundingtext gefunden	35	6,3%	100,0%
Gesamtergebnis	552	100,00%	

4.2 Häufigste deutsche Förderorganisationen im WoS

Die zehn deutschen Förderorganisationen, die am häufigsten in der Datenbank vorkommen sind mit ihren Publikationszahlen in Tabelle 6 dargestellt. Die Deutsche Forschungsgemeinschaft (DFG) hat dabei den mit Abstand größten Anteil, gefolgt von dem Bundesministerium für Bildung und Forschung (BMBF). Zusätzlich aufgeführt ist die Alexander von Humboldt-Stiftung (AVH), der Deutsche Akademische Austauschdienst (DAAD), die Deutsche Krebshilfe (DKH), der Fonds der Chemischen Industrie (FCI), die Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF), die Max-Planck-Gesellschaft (MPG), die Studienstiftung des Deutschen Volkes (SS) und die Volkswagen-Stiftung (VW). Es ist hierbei zu beachten, dass die HGF und die MPG keine Förderorganisationen im herkömmlichen Sinne sind, trotzdem werden sie oft als die Finanzierer der Forschung an den beteiligten Forschungseinrichtungen genannt.

Tabelle 6: Publikationszahlen deutscher Förderer

FO	2010	2011	2012	2013	Total 2010-13
DFG	19.881	22.061	22.966	23.531	88.439
BMBF	5.066	6.265	7.017	7.408	25.756
DAAD	2.139	2.463	2.418	2.468	9.488
MPG	1.938	2.235	2.589	2.497	9.259
AvH	1.914	2.248	2.454	2.382	8.998
HGF	721	994	1.231	1.351	4.297
FCI	1.239	1.060	922	826	4.047
DKH	560	595	666	658	2.479
VW	618	607	583	634	2.442
SS	334	344	381	411	1.470

4.3 Verteilung der Publikationen auf die wissenschaftlichen Sektoren

Für die Jahre 2010 bis 2012 ist es auch möglich, die Verteilung der Publikationen der einzelnen Förderorganisationen auf die wissenschaftlichen Sektoren in Deutschland darzustellen. Tabelle 7 zeigt diese Verteilung durch Vollzählung der Journalartikel. So werden z.B. Publikationen doppelt gezählt, die sowohl durch AutorInnen am Max-Planck-Institut als auch an einer Universität geschrieben wurden. Als etwas überraschend erweist sich hierbei die Verteilung des FCI. Die einzige der aufgelisteten Förderorganisationen, die von der Privatwirtschaft finanzierten weist den höchsten Anteil an Publikationen mit an einer Hochschule assoziierten AutorInnen auf.

Tabelle 7: Verteilung der Journalartikel nach Sektoren und Förderorganisationen 2010-2012

FO	Hochschulen	Max-Planck-Gesellschaft	Heimholtz-Gemeinschaft	Leibniz-Gemeinschaft	Sonstige	Karlsruher Institut für Technologie	Ressortforschung	Wirtschaft	Fraunhofer-Gesellschaft
DFG	89,5%	11,7%	10,9%	6,5%	4,2%	3,4%	1,7%	1,2%	1,0%
BMBF	85,3%	13,4%	21,8%	9,5%	8,2%	4,5%	3,5%	4,2%	3,2%
DAAD	85,0%	7,8%	9,7%	9,6%	3,1%	2,5%	2,2%	1,0%	1,0%
MPG	45,3%	91,6%	10,9%	3,9%	5,5%	0,6%	1,4%	1,0%	0,4%
AVH	79,5%	22,3%	12,6%	7,2%	2,4%	5,3%	1,2%	0,6%	0,5%
HGF	80,3%	20,2%	77,0%	2,9%	5,6%	13,5%	1,0%	1,2%	0,4%
FCI	92,4%	12,5%	4,1%	2,4%	2,8%	2,8%	1,0%	1,3%	0,8%
DKH	85,9%	4,4%	33,6%	10,4%	10,8%	0,1%	0,9%	1,7%	0,2%
VW	89,6%	15,5%	7,8%	8,3%	7,9%	1,2%	3,7%	0,8%	1,8%
SS	89,9%	11,3%	7,0%	8,6%	3,7%	1,9%	0,9%	1,4%	1,1%

4.4 Förderprofile nach Disziplin

Die verschiedenen FOs unterscheiden sich hinsichtlich der Art der Förderung (z.B. die DAAD als Förderer von Forschungsreisen verglichen mit der DFG, die eine Vielzahl verschiedener Förderinstrumente besitzt, von Projekt- über Personen- bis zu institutioneller Förderung). Sie unterscheiden sich jedoch auch in dem Fächerprofil der Geförderten. Eine Analyse von FOs mit FA-Daten ermöglicht die Erzeugung einer globalen Übersicht über die Förderprofile anhand des Publikationsoutputs. Diese könnte auch dem finanziellen Profil der FOs gegenübergestellt werden⁵. Im Folgenden werden die zehn erwähnten FOs entlang zweier Dimensionen porträtiert. Die jeweils zehn Subject Categories (SCs) mit den meisten Publikationen von 2010 bis 2013 und die zehn mit dem höchsten *Relativen Literaturanteil* werden in den folgenden Tabellen (8-17) dargestellt. Der *Relative Literaturanteil* (RLA) ist ein auf einer Skala von -100 bis 100 renormierter Indikator des Aktivitätsindexes (AI)⁶. Der Aktivitätsindex beschreibt den relativen Anteil der Journalartikel in einer SC einer FO verglichen mit dem Gesamtoutputs

⁵ Für eine ähnliche Gegenüberstellung von DFG-Publikationen mit Fördergeldern jedoch auf Disziplin- und Universitäts-ebene siehe (Sirtes 2013)

⁶ $RLA = 100 \cdot \tanh \ln AI = 100 \cdot \frac{AI^2 - 1}{AI^2 + 1}$

des FOs und dieses Verhältnis wird wiederum ins Verhältnis gesetzt mit dem relativen Anteil der deutschen Journalartikel verglichen mit dem deutschen Gesamtoutput. So lassen sich die wichtigsten geförderten Disziplinen sowohl in absoluten als auch in relativen Zahlen darstellen.

$$AI = \frac{\frac{P_{ij}}{P_i}}{\frac{P_j}{P}} = \frac{\frac{P_{ij}}{P_i}}{\frac{P_j}{P}} \cdot \frac{P}{P_i}$$

P_{ij} Anzahl der Publikationen der FO i im Feld j

P_i Anzahl der Publikationen der FO i in allen Feldern

P_j Anzahl der Publikationen in Deutschland im Feld j

P Anzahl der Publikationen in Deutschland in allen Feldern

Bei den großen Förderern wie DFG (Tabelle 8) und BMBF (Tabelle 9) ist erwartungsgemäß keine ausgeprägte Spezialisierung in den Förderprofilen sichtbar: Die höchste RLA liegt bei 72,4 respektive 85,6 bei Disziplinen aus verschiedenen naturwissenschaftlichen Fächern. Auch die Ausrichtung der DKH (Tabelle 15) auf medizinische Fächer und des FCI (Tabelle 14) auf chemische (RLA bei über 99) sind erwartbar. Überraschender ist jedoch die starke Position der Astronomie und Astrophysik bei der MPG (Tabelle 11), sowohl in absoluten Zahlen als auch in RLA. Diese Zahlen werden jedoch bestätigt durch die MPG-Publikationszahlen, die nur anhand von Affiliationen mit einem Max-Planck-Institut erhoben werden (Daten nicht ausgewiesen). Am schwierigsten zu erklären, ist die starke Ausrichtung der AvH in den physikalischen und chemischen Fächern. Gemäß einer kurzen Recherche nach den StipendiatInnen der AvH (http://www.humboldt-foundation.de/pls/web/pub_hn_query.main) gibt es die höchste Anzahl an StipendiatInnen in der Chemie (4.023), gefolgt von den Biowissenschaften (3.160); beide liegen weit vor allen anderen Fächern. Als nächstes ist die Medizin (1.584) vertreten, deren AutorInnen auch oft in den biologischen Fächern publizieren.

4.4.1 DFG

Tabelle 8: DFG-Förderprofil nach absoluter Häufigkeit und RLA, Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
6607	Biochemistry & Molecular Biology	72,4	559	Developmental Biology
5677	Chemistry, Physical	71,4	3903	Cell Biology
5561	Materials Science, Multidisciplinary	67,5	3526	Physics, Atomic, Molecular & Chemical
5311	Chemistry, Multidisciplinary	67,1	6607	Biochemistry & Molecular Biology
4813	Physics, Condensed Matter	65,1	581	Paleontology
4612	Multidisciplinary Sciences	63,7	5311	Chemistry, Multidisciplinary
4450	Neurosciences	63,4	2345	Chemistry, Inorganic & Nuclear
4123	Physics, Applied	63,1	4078	Physics, Multidisciplinary
4078	Physics, Multidisciplinary	62,6	4813	Physics, Condensed Matter
3903	Cell Biology	62,4	4612	Multidisciplinary Sciences

4.4.2 BMBF

Tabelle 9: BMBF-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
1611	Neurosciences	85,6	403	Mathematical & Computational Biology
1591	Physics, Applied	83,1	136	CELL & TISSUE ENGINEERING
1548	Biochemistry & Molecular Biology	82,8	1334	Physics, Particles & Fields
1418	Materials Science, Multidisciplinary	79,2	1001	Biotechnology & Applied Microbiology
1395	Multidisciplinary Sciences	74,9	981	Genetics & Heredity
1334	Physics, Particles & Fields	71,8	794	Biochemical Research Methods
1192	Physics, Multidisciplinary	70,5	201	Neuroimaging
1070	Chemistry, Physical	68,4	223	Oceanography
1001	Biotechnology & Applied Microbiology	68,3	387	Water Resources
981	Genetics & Heredity	66,0	133	Geriatrics & Gerontology

* (SC mit mindestens 100 Artikeln)

4.4.3 DAAD

Tabelle 10: DAAD-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
639	Chemistry, Physical	93,1	102	Entomology
570	Chemistry, Multidisciplinary	92,3	35	Tropical Medicine
562	Materials Science, Multidisciplinary	90,8	270	Zoology
553	Biochemistry & Molecular Biology	89,2	74	Biodiversity Conservation
439	Plant Sciences	87,2	439	Plant Sciences
384	Physics, Applied	85,6	157	Marine & Freshwater Biology
380	Chemistry, Organic	84,4	94	Soil Science
350	Physics, Condensed Matter	84,0	100	Agronomy
320	Environmental Sciences	83,3	30	Fisheries
311	Physics, Atomic, Molecular & Chemical	83,2	380	Chemistry, Organic

* (SC mit mindestens 30 Artikeln)

4.4.4 MPG

Tabelle 11: MPG-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
2103	Astronomy & Astrophysics	96,8	2103	Astronomy & Astrophysics
878	Biochemistry & Molecular Biology	93,3	67	Anthropology
755	Multidisciplinary Sciences	92,5	220	Evolutionary Biology
577	Chemistry, Multidisciplinary	88,8	97	Developmental Biology
510	Neurosciences	84,1	264	Biology
457	Chemistry, Physical	82,5	755	Multidisciplinary Sciences
457	Cell Biology	82,0	330	Ecology
442	Physics, Particles & Fields	81,3	442	Physics, Particles & Fields
442	Physics, Multidisciplinary	80,3	354	Plant Sciences
378	Materials Science, Multidisciplinary	78,0	878	Biochemistry & Molecular Biology

* (SC mit mindestens 30 Artikeln)

4.4.5 AvH

Tabelle 12: AvH-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
884	Chemistry, Multidisciplinary	90,4	644	Physics, Particles & Fields
826	Materials Science, Multidisciplinary	87,2	799	Physics, Multidisciplinary
799	Physics, Multidisciplinary	83,0	884	Chemistry, Multidisciplinary
797	Chemistry, Physical	82,1	534	Physics, Atomic, Molecular & Chemical
644	Physics, Particles & Fields	80,7	88	Paleontology
613	Astronomy & Astrophysics	80,6	234	Physics, Mathematical
607	Physics, Condensed Matter	80,2	76	Automation & Control Systems
534	Physics, Atomic, Molecular & Chemical	80,0	435	Nanoscience & Nanotechnology
483	Physics, Applied	74,4	797	Chemistry, Physical
435	Nanoscience & Nanotechnology	72,8	306	Mathematics, Applied

* (SC mit mindestens 30 Artikeln)

4.4.6 HGF

Tabelle 13: HGF-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
740	Physics, Particles & Fields	98,4	740	Physics, Particles & Fields
599	Physics, Multidisciplinary	96,3	213	Physics, Nuclear
398	Astronomy & Astrophysics	95,0	599	Physics, Multidisciplinary
227	Multidisciplinary Sciences	88,5	74	Mathematical & Computational Biology
213	Physics, Nuclear	82,8	133	Instruments & Instrumentation
203	Biochemistry & Molecular Biology	82,0	398	Astronomy & Astrophysics
166	Materials Science, Multidisciplinary	77,5	72	Spectroscopy
165	Neurosciences	73,3	40	Oceanography
157	Physics, Atomic, Molecular & Chemical	72,1	90	Biology
143	Physics, Applied	66,7	94	Meteorology & Atmospheric Sciences

* (SC mit mindestens 30 Artikeln)

4.4.7 FCI

Tabelle 14: FCI-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
1133	Chemistry, Multidisciplinary	99,1	773	Chemistry, Organic
773	Chemistry, Organic	98,8	676	Chemistry, Inorganic & Nuclear
676	Chemistry, Inorganic & Nuclear	97,9	1133	Chemistry, Multidisciplinary
582	Chemistry, Physical	92,8	378	Physics, Atomic, Molecular & Chemical
378	Physics, Atomic, Molecular & Chemical	90,1	582	Chemistry, Physical
316	Biochemistry & Molecular Biology	89,2	74	Chemistry, Applied
250	Materials Science, Multidisciplinary	86,0	149	Polymer Science
149	Polymer Science	84,7	88	Chemistry, Medicinal
122	Nanoscience & Nanotechnology	79,6	66	Crystallography
96	Physics, Condensed Matter	68,2	316	Biochemistry & Molecular Biology

* (SC mit mindestens 30 Artikeln)

4.4.8 DKH

Tabelle 15: DKH-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
956	Oncology	99,5	956	Oncology
262	Cell Biology	98,7	257	Hematology
257	Hematology	97,0	90	Pathology
217	Biochemistry & Molecular Biology	96,0	130	Medicine, Research & Experimental
185	Multidisciplinary Sciences	95,1	57	Nutrition & Dietetics
179	Genetics & Heredity	94,9	262	Cell Biology
130	Medicine, Research & Experimental	93,5	179	Genetics & Heredity
98	Immunology	91,3	75	Gastroenterology & Hepatology
90	Pathology	90,1	83	Public, Environmental & Occupational Health
83	Public, Environmental & Occupational Health	86,1	98	Immunology

* (SC mit mindestens 30 Artikeln)

4.4.9 VW

Tabelle 16: VW-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
228	Biochemistry & Molecular Biology	97,3	107	Evolutionary Biology
216	Neurosciences	96,0	122	Zoology
203	Chemistry, Physical	90,0	130	Ecology
197	Chemistry, Multidisciplinary	87,9	88	Biology
194	Multidisciplinary Sciences	87,6	84	Physics, Mathematical
164	Materials Science, Multidisciplinary	86,3	43	Behavioral Sciences
142	Physics, Multidisciplinary	79,1	216	Neurosciences
130	Ecology	78,5	194	Multidisciplinary Sciences
122	Zoology	74,5	107	Nanoscience & Nanotechnology
118	Clinical Neurology	74,4	54	Biophysics

* (SC mit mindestens 30 Artikeln)

4.4.10 SS

Tabelle 17: SS-Förderprofil nach absoluter Häufigkeit und RLA* Journalartikel 2010-13

Anzahl Artikel	SC	RLA	Anzahl Artikel	SC
134	Chemistry, Multidisciplinary	88,5	77	Chemistry, Organic
108	Biochemistry & Molecular Biology	84,2	65	Chemistry, Inorganic & Nuclear
95	Physics, Multidisciplinary	81,4	134	Chemistry, Multidisciplinary
94	Multidisciplinary Sciences	80,6	30	Zoology
89	Chemistry, Physical	79,1	78	Physics, Atomic, Molecular & Chemical
78	Physics, Atomic, Molecular & Chemical	78,2	95	Physics, Multidisciplinary
77	Chemistry, Organic	78,1	48	Ecology
72	Materials Science, Multidisciplinary	77,0	64	Physics, Particles & Fields
70	Neurosciences	72,9	61	Optics
67	Physics, Condensed Matter	71,9	94	Multidisciplinary Sciences

* (SC mit mindestens 30 Artikeln)

4.5 Co-Funding und internationale Kollaborationen nach Förderorganisationen

4.5.1 Internationale Kollaborationen

Eine wichtige Funktion von Förderorganisationen ist auch die Förderung internationaler Kollaborationen. Erstaunlicherweise wurde jedoch in einer kleinen Studie gezeigt, dass durch den Schweizerischen Nationalfonds geförderte Publikationen eher weniger internationale Kollaborationen aufweisen (van den Besselaar, Peter, Inzelt, und Reale, Emanuela 2012). In Tabelle 18 ist der Anteil von Journalartikeln mit internationaler Autorenschaft angegeben. Als Benchmark wurde hier der Anteil an internationalen Artikeln von 2010-2012 angegeben, die deutsche Universitäten als Affiliationen aufweisen. Es muss darauf hingewiesen werden, dass AutorInnen, die neben ihrer deutschen auch eine ausländische Institution als zusätzliche Affiliation haben, ebenfalls als international gewertet werden (für mehrere andere Probleme von Kollaborationszählungen siehe (Katz und Martin 1997; Riechert, Sirtes, und Lietz 2013)).

Tabelle 18: Anteil Journalartikel mit internationaler Beteiligung nach FO. Benchmark: Journalartikel deutscher Universitäten 2010-2012

FO	Artikel	Internationale Artikel	Prozent international
DFG	84774	42188	49.8%
BMBF	25103	12271	48.9%
DAAD	8029	6415	79.9%
MPG	7657	5118	66.8%
AvH	7046	5495	78.0%
HGF	4178	2709	64.8%
FCI	3968	1381	34.8%
DKH	2424	1227	50.6%
VW	2133	1203	56.4%
SS	1327	596	44.9%
Unis 2010-2012	200166	97066	48.5%

4.5.2 Co-Funding deutscher Förderorganisationen und Förderung aus der EU.

Eine seriöse Analyse der Publikationen, die gemeinsam durch verschiedene Organisationen gefördert wurde (Co-Funding), würde eine ähnliche Datenbereinigung aller internationalen Förderorganisationen voraussetzen, wie wir sie für die zehn deutschen Förderorganisationen durchgeführt haben. Da dies im Rahmen dieser Studie nicht umsetzbar ist, haben wir uns auf zwei Aspekte des Co-Fundings konzentriert: inwieweit werden Publikationen von den verschiedenen deutschen Förderern gemeinsam unterstützt? Wie oft werden Publikationen der deutschen Förderer auch aus Geldern der EU gefördert? Es ist in diesem Zusammenhang wichtig, anzumerken, dass die Funding Acknowledgements nicht autorengebunden in der Datenbank vorkommen. Daher ist nicht feststellbar, ob eine gemeinsame Förderung entweder auf eine Autorin zurückzuführen ist, die verschiedene Geldquellen besitzt, oder darauf,

dass verschiedene Autoren jeweils ihre Förderung angeben, oder sogar die Förderung selbst sich aus zwei verschiedenen Quellen zusammensetzt (sogenanntes Joint-funding). In Tabelle 21 sind die prozentualen Anteile der Publikationen, die sich aus den jeweiligen zwei Quellen speisen, angegeben. Dabei wurden die Prozentzahlen zeilenweise berechnet. So weist beispielsweise die DFG bei allen anderen FOs jeweils den höchsten Anteil auf, da sie die größte Förderorganisation ist. Die interessantere Lesart ist jedoch die spaltenweise Dimension (so wurden auch die Farben von rot nach grün verteilt), die die Affinität der Förderorganisationen zueinander darstellt. So sieht man z.B., dass die DFG zur FCI die höchste Affinität hat, d.h. FCI und DFG-Förderungen kommen öfter als alle anderen Kombinationen gemeinsam vor. Es ist auch ablesbar, dass die DKH zu sechs der neun anderen FO die niedrigste Co-Funding-Affinität besitzt, jedoch häufig mit dem BMBF gemeinsam gefördert wird, da Publikationen mit DKH-Förderung fast ausschließlich auch durch BMBF- und DFG-Mittel gefördert sind. In den zwei nachfolgenden Spalten sind auch der Anteil der Publikationen angegeben, die nicht von den anderen neun Förderorganisationen unterstützt werden, und der Anteil an Publikationen, für die mehr als zwei der zehn FO als Geldgeber angegeben sind. Für die HGF lassen sich folgende interessante Ergebnisse finden: Zu drei FOs besitzt sie die höchste Affinität (AVH, BMBF und MPG) und auch mit der DFG ist sie stark gekoppelt. Das lässt sich auch durch die zwei weiteren Spalten erklären: Die HGF hat eine niedrige Anteil ohne (deutsches) Co-Funding und den höchsten Anteil an Publikationen mit mehr als zwei Förderern aus diesen zehn. In der letzten Spalte ist eine von diesen neun anderen FOs unabhängige Zahl angegeben, nämlich der Anteil an Publikation mit genau einer angegebenen FO, d.h. diese Journalartikel haben keine auch hier nicht aufgeführten Förderer (aber möglicherweise mehrere Grants von der gleichen FO). Aus dem Vergleich der Spalte ohne Co-Funding unter diesen zehn und der Spalte des Anteils mit nur einer FO lässt sich ableiten, dass im Durchschnitt etwa zwei Drittel der Publikationen mit nur einer FO dieser zehn mindestens eine nicht aufgeführte FO als Förderer aufweisen kann (dieser Anteil schwankt relativ wenig mit einer Standardabweichung von $\pm 7\%$).

Tabelle 21: Co-Funding deutscher Förderorganisationen

FO	AVH	BMBF	DAAD	DFG	DKH	FCI	HGF	MPG	SS	VW	kein Co-funding	mehr als 2 FO	Nur eine FO
AVH		10,8%	3,7%	30,3%	0,2%	2,3%	5,4%	6,8%	0,5%	0,9%	60,2%	21,0%	14,4%
BMBF	3,8%		2,5%	35,2%	2,2%	0,9%	7,1%	5,2%	0,5%	1,1%	55,6%	14,2%	23,6%
DAAD	3,5%	6,8%		28,0%	0,3%	1,7%	1,0%	1,6%	0,8%	1,0%	63,0%	7,9%	15,8%
DFG	3,1%	10,3%	3,0%		1,0%	3,1%	2,4%	3,2%	0,8%	1,1%	77,1%	5,0%	33,6%
DKH	0,7%	22,6%	1,3%	34,7%		0,4%	2,9%	1,1%	0,2%	0,1%	49,5%	13,5%	18,0%
FCI	5,1%	5,9%	4,1%	67,9%	0,3%		0,6%	5,2%	2,9%	1,5%	23,8%	17,3%	8,1%
HGF	11,3%	42,5%	2,1%	49,2%	1,7%	0,6%		12,9%	0,7%	0,5%	32,5%	54,0%	10,4%
MPG	6,6%	14,4%	1,7%	30,7%	0,3%	2,3%	6,0%		0,5%	1,0%	58,3%	21,6%	12,2%
SS	2,9%	9,4%	5,3%	45,9%	0,3%	8,0%	1,9%	2,9%		1,8%	40,1%	18,5%	11,3%
VW	3,3%	12,0%	4,1%	40,7%	0,1%	2,5%	0,9%	3,8%	1,1%		48,7%	17,2%	17,0%

Eine besonders begehrte Form des Co-Fundings ist diejenige aus EU-Fördergeldern. Grants aus dem europäischen Rahmenprogramm oder aus dem European Research Council (ERC) gelten als besonders prestigeträchtig. Bei der Durchsicht der FA-Angaben ist jedoch festgestellt worden, dass eine Unterscheidung nach verschiedenen EU-Töpfen kaum möglich ist. Die Angaben bestehen oft nur aus Abkürzungen wie EU oder EC oder aus Abkürzungen mit ganz unterschiedlichen sich überlappenden Kurztexten. Eine saubere Trennung ist nur anhand von Grantnumbers und eines Abgleichs mit EU Daten-

banken wie CORDIS (<http://cordis.europa.eu/>) oder der ERC-Seite <http://erc.europa.eu/projects-and-results/> möglich. Aus diesem Grund wurden alle Förderorganisationen mit EU-Bezug zusammengefasst. In Tabelle 22 sind die Anteile der Publikationen erfasst, die neben einem der deutschen Förderer auch aus EU-Mitteln gefördert wurden. Neben dem sehr hohen Anteil der HGF ist besonders auffällig, dass, verglichen mit dem Anteil der internationalen Kollaborationen (Tabelle 18), die AvH keinen besonders hohen Anteil an EU-Publikationen aufweist. Bemerkenswert ist auch der relativ niedrige Anteil bei der DFG.

Tabelle 22: Anteil der Publikation mit deutscher FO und EU-Förderung

Deutsche FO	Anteil EU-Publikationen
AVH	19,1%
BMBF	22,8%
DAAD	11,3%
DFG	14,5%
DKH	19,6%
FCI	11,9%
HGF	40,0%
MPG	22,5%
SS	13,6%
VW	15,6%

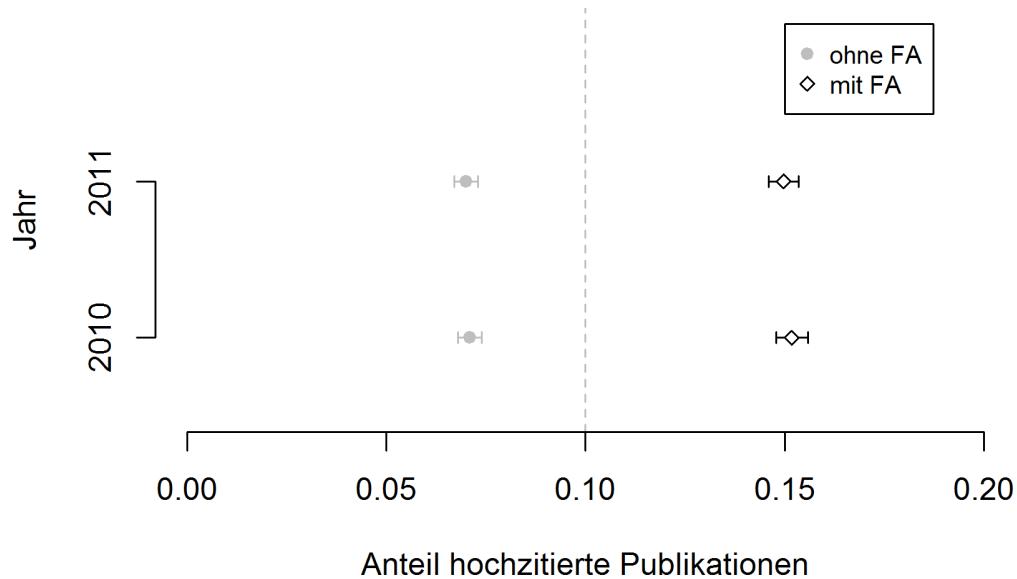
4.6 Impactmessung geförderter Publikationen

Für die Jahre 2010 und 2011 können zurzeit bereits Berechnungen bezüglich der Zitationszahlen der geförderten Publikationen angestellt werden, da ein Zitationsfenster von drei Jahren als ein Minimum angesehen werden kann, um valide Aussagen zu Langzeitperspektiven des Impacts der Artikel zu treffen (siehe Wang 2013).

4.6.1 Impactmessung Journalartikel mit und ohne FA

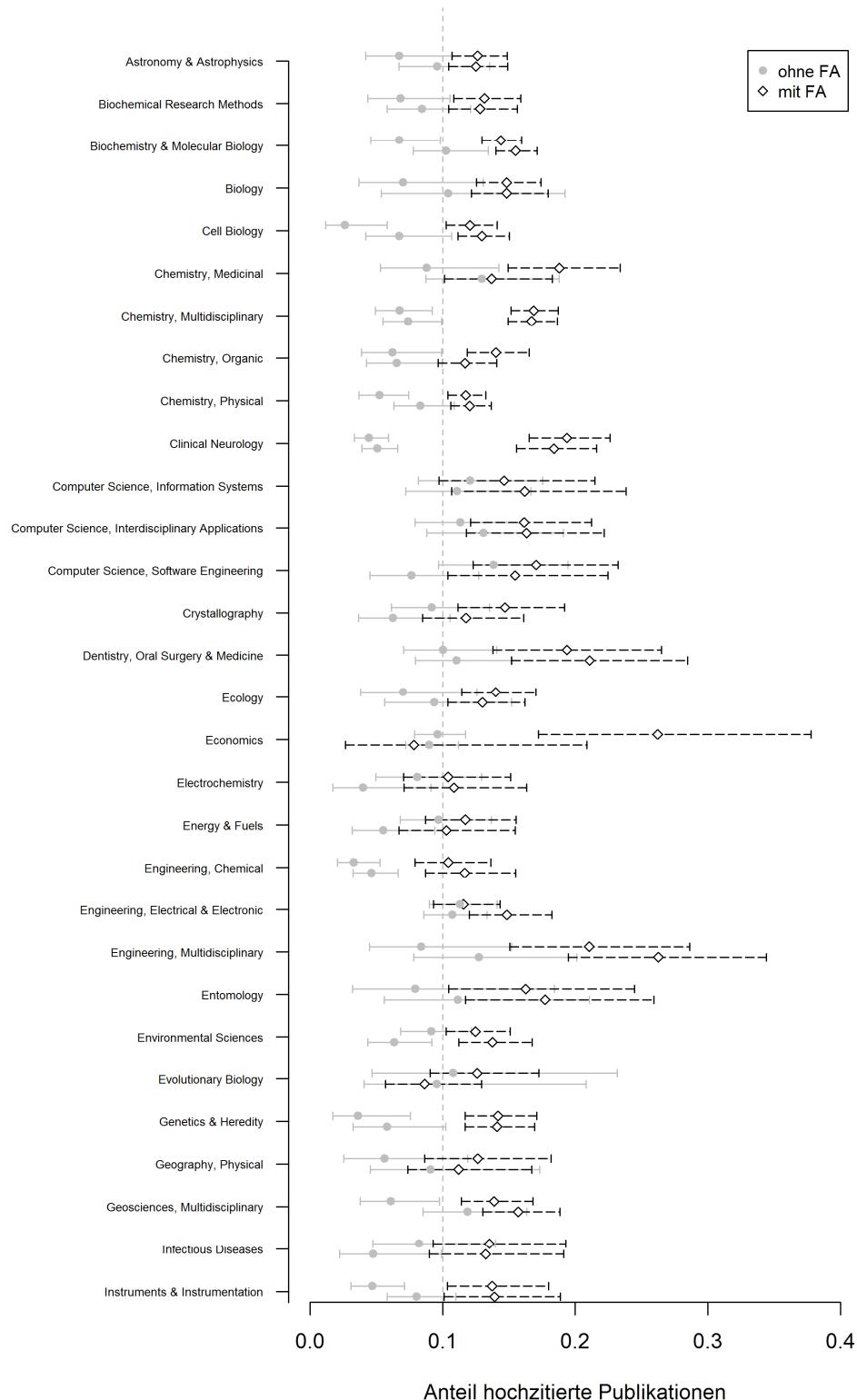
Für diese zwei Jahre wurde zuerst ein Vergleich zwischen deutschen Publikationen mit FA und deutschen Publikationen ohne FA angestellt (Abbildung 3). Es zeigt sich ein deutlicher Unterschied: Mehr als doppelt so viele Journalartikel mit FA finden ihren Weg in die zehn Prozent der meistzitierten ihres Faches.

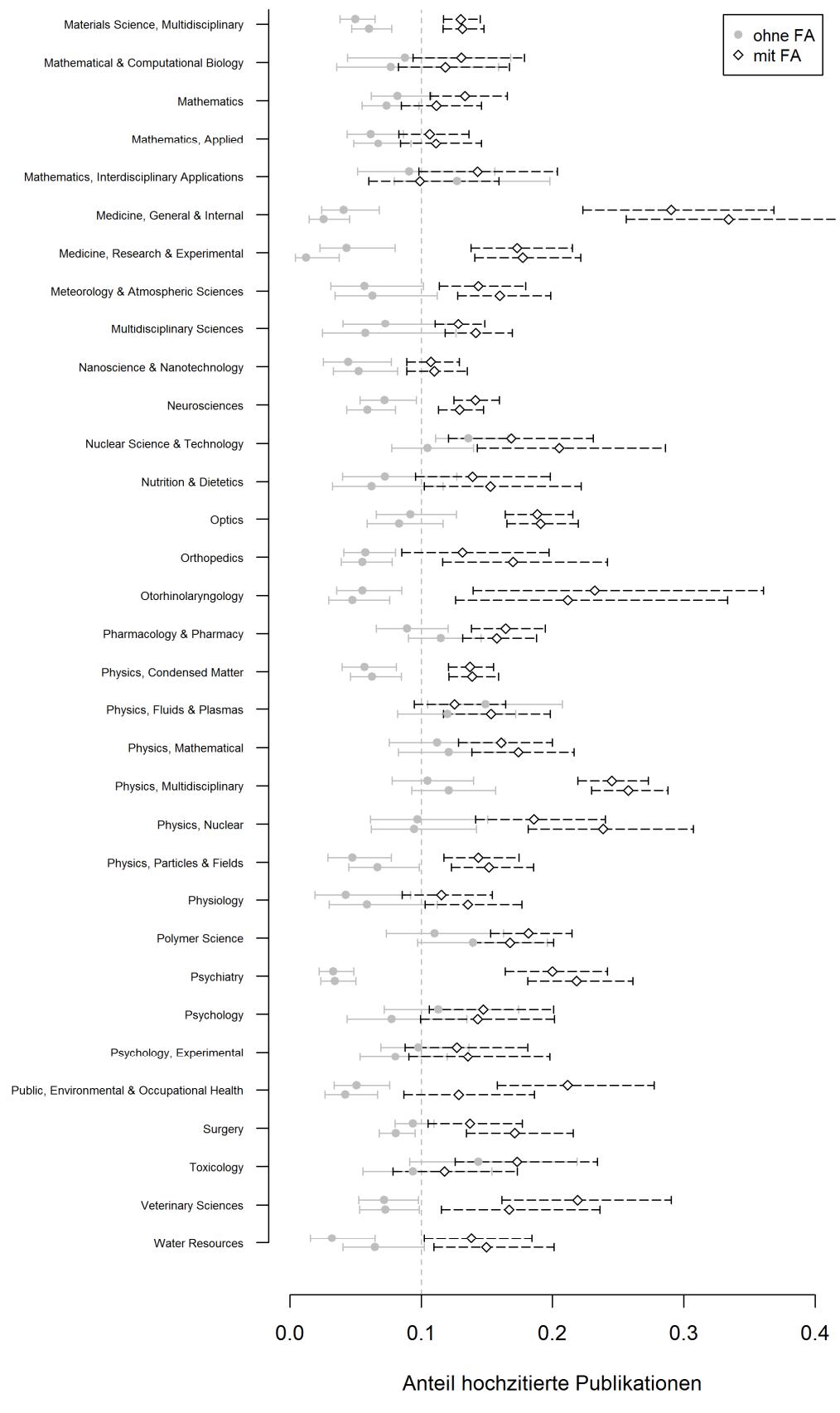
Abbildung 3: Anteil hochzitierter deutscher Publikationen mit und ohne FA für 2010 und 2011



Wenn man dieses Verhältnis genauer nach einzelnen Disziplinen differenziert (Abbildung 4), sieht man, dass die Diskrepanz zwischen Publikationen mit FA und ohne FA meist bestehen bleibt und die einen über, die anderen unter 10% hochzitierte Publikationen aufweisen. Es stellt sich aber heraus, dass der Unterschied zwischen den beiden Gruppen sehr stark divergieren kann. Während sich in Fächern wie Allgemein- und interne Medizin etwa 30% der explizit geförderten Publikationen unter den hochzitierten befinden und diejenigen ohne FA nur etwa 4% hochzitierte aufweisen, ist kein Unterschied in z.B. Flüssigkeiten- und Plasmaphysik ersichtlich.

Abbildung 4: Hochzitierte Journalartikel einer Auswahl publikationsstarker Disziplinen (SCs) mit und ohne FA für 2010 (untere) und 2011 (obere).





4.6.2 Impactmessung nach Förderorganisationen

Mit den Publikationsdaten der einzelnen FOs lassen sich nun für die Jahre 2010 und 2011 Impactmessungen durchführen. In Abbildung 5 werden die Ergebnisse für die zehn Organisationen aggregiert dargestellt. Dabei sind vor allem zwei Punkte herauszustellen: Die MPG ist mit 25,5% hochzitierter Journalartikel tendenziell die Organisation mit dem höchsten Impact und ist damit vergleichbar mit Universitäten wie MIT und Harvard, die mit 25,2% respektive 23% hochzitierter Publikationen im Leiden-Ranking angegeben werden (<http://www.leidenranking.com/ranking/2014>). Bemerkenswert ist jedoch das bessere Abschneiden des BMBF (19,3%) im Vergleich zur DFG (15,5). Da erst eine tiefgehende Analyse der Einflussfaktoren dieser Diskrepanz eine abschließende Bewertung erlaubt, sind diese Resultate vorläufiger Natur. Eine mögliche Erklärung für den Unterschied könnte eine besondere relative Stärke des BMBF in besonders publikationsstarken SCs sein. In Abbildung 6 sind alle SCs aufgeführt, in denen das BMBF für das Jahr 2011 mindestens 70 Publikationen aufweist. Das publikationsstärkste SC der DFG, „Biochemistry and Molecular Biology“ zeigt dabei einen markanten Unterschied zwischen BMBF und DFG. Obwohl in vielen SCs die beiden größten Förderorganisationen kaum voneinander zu unterscheiden sind, könnte eine Handvoll solcher Fälle den Gesamtunterschied erklären.

Abbildung 5: Anteil hochzitierter Journalartikel nach Förderorganisationen für 2010 und 2011

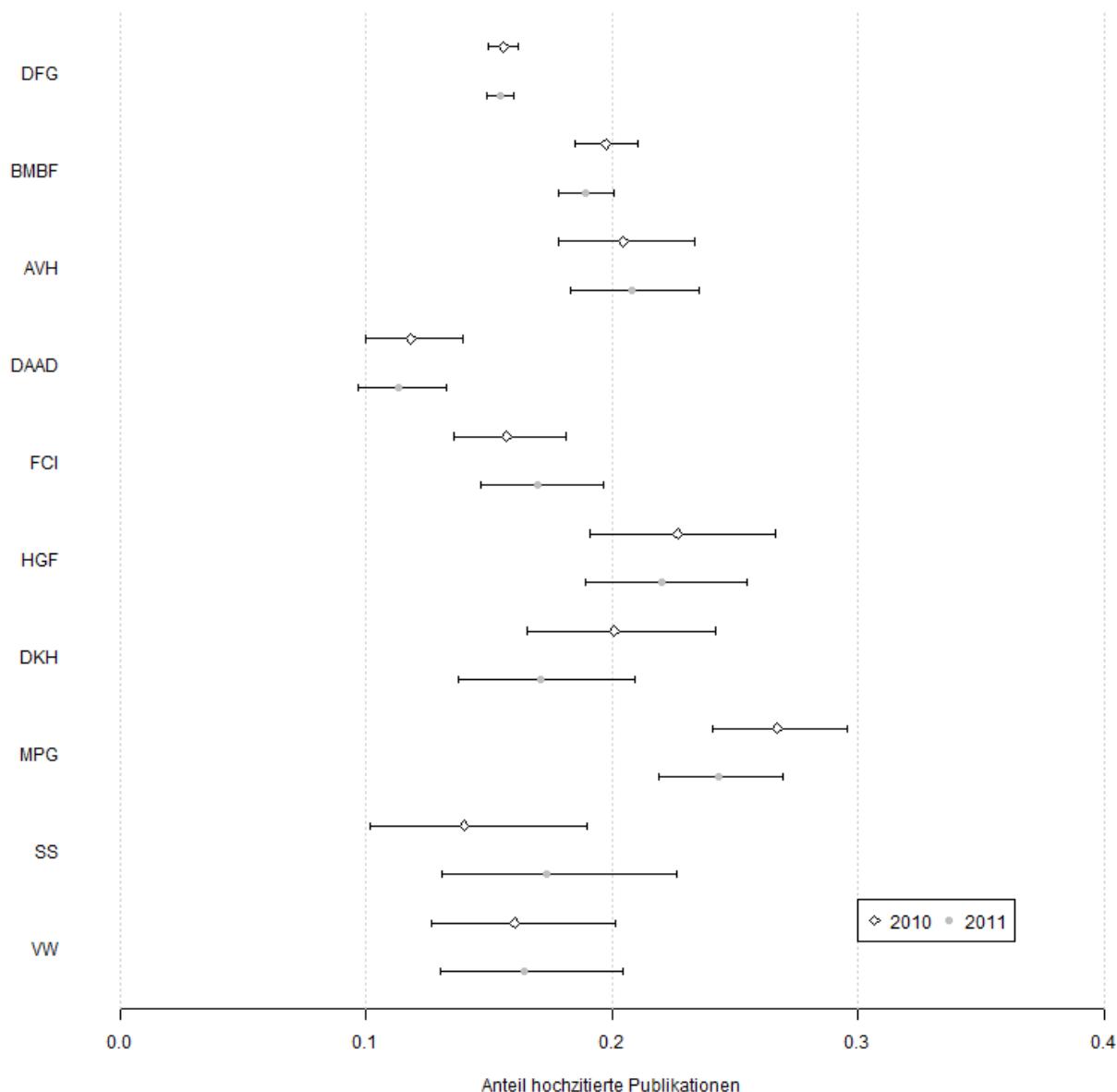
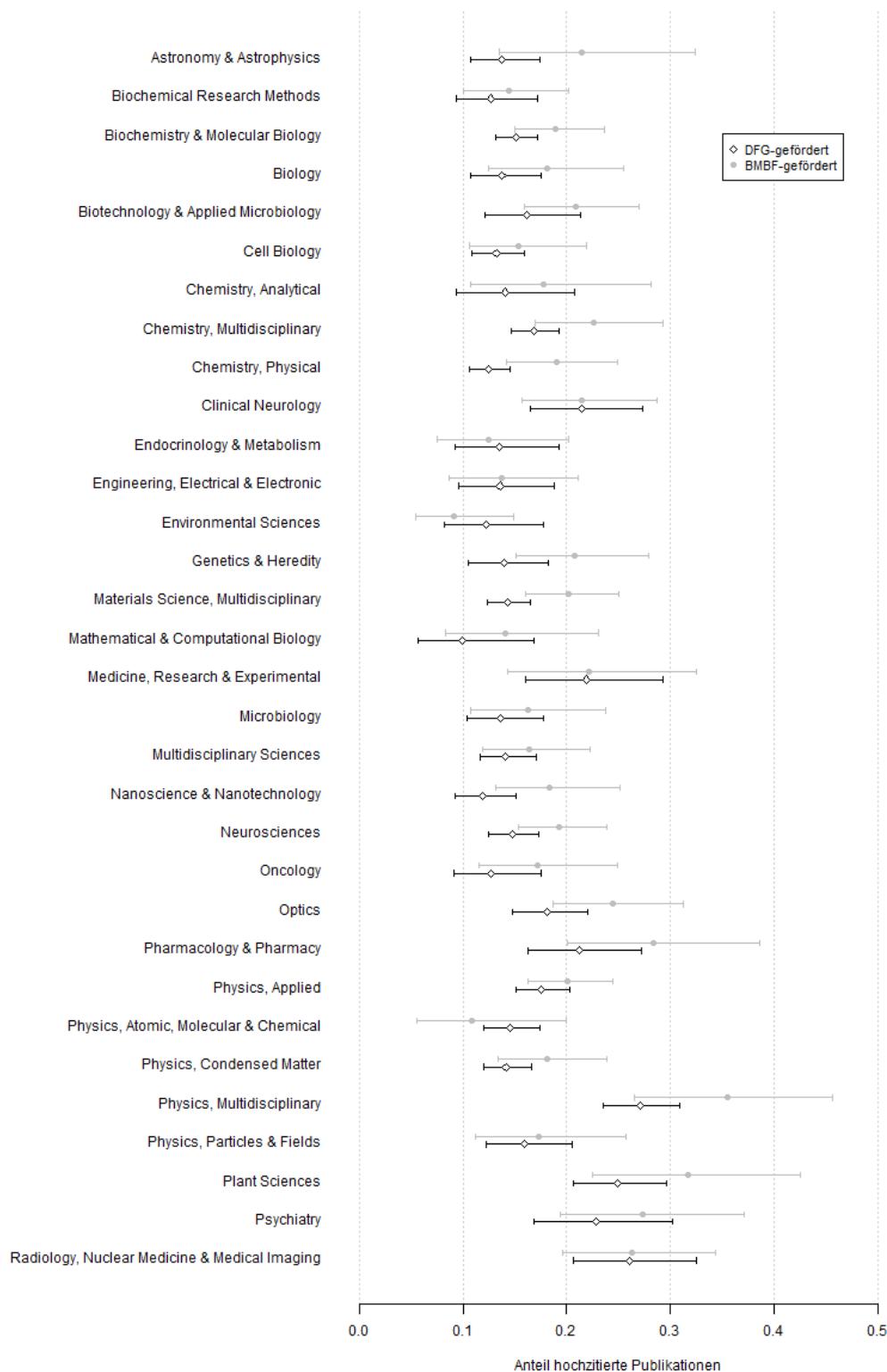


Abbildung 6: Vergleich des Anteils hochzitierter Journalartikel für DFG und BMBF des Jahres 2011 in publikationsstarken SCs



Ein weiteres potenzielles Anwendungsgebiet des im Projekt bereits erreichten Unifizierungsstandes bildet die Analyse hochzitierter Fächer. So können Förderorganisationen ihr eigenes Förderprofil gezielt analysieren und entwickeln. In den folgenden Abbildungen (7-11) wird dargestellt, wie Förderorganisationen die FO-Daten nutzen können, um herauszufinden, in welchen Fächern sie besonders hoch zitierte Publikationen fördern. Exemplarisch seien hier die fünf größten FOs mit den jeweils bezüglich der absoluten und RLA Werte wichtigsten SCs (siehe Abschnitt 4.4) dargestellt.

Abbildung 7: Anteil hochzitierter Journalartikel der DFG in den publikationsstärksten Disziplinen (SCs) für 2010 und 2011

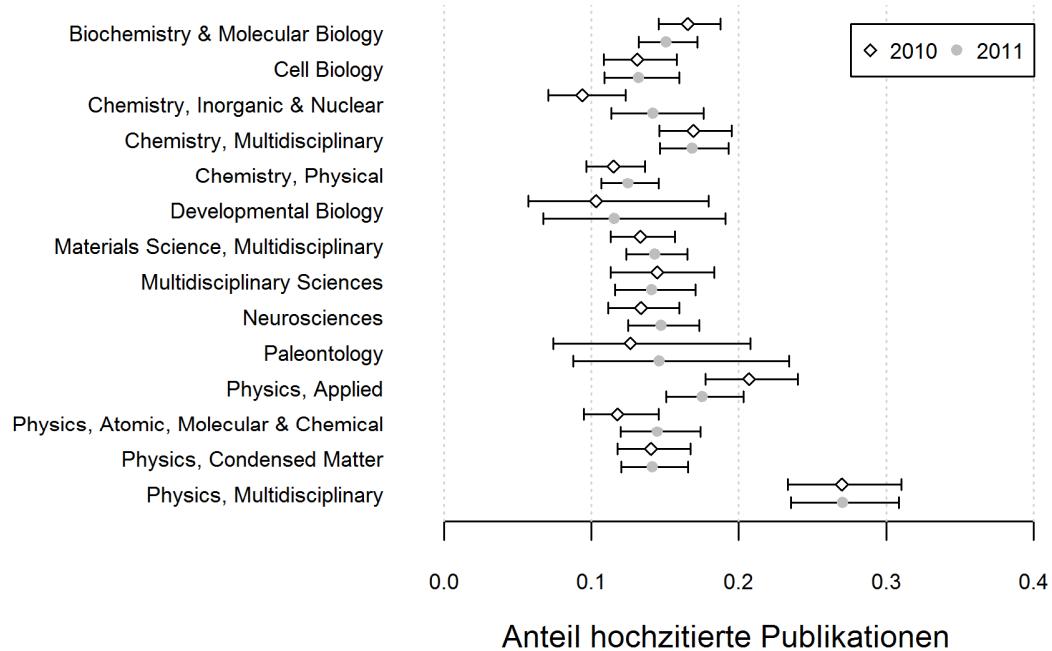


Abbildung 8: Anteil hochzitierter Journalartikel der BMBF in den publikationsstärksten Disziplinen (SCs) für 2010 und 2011

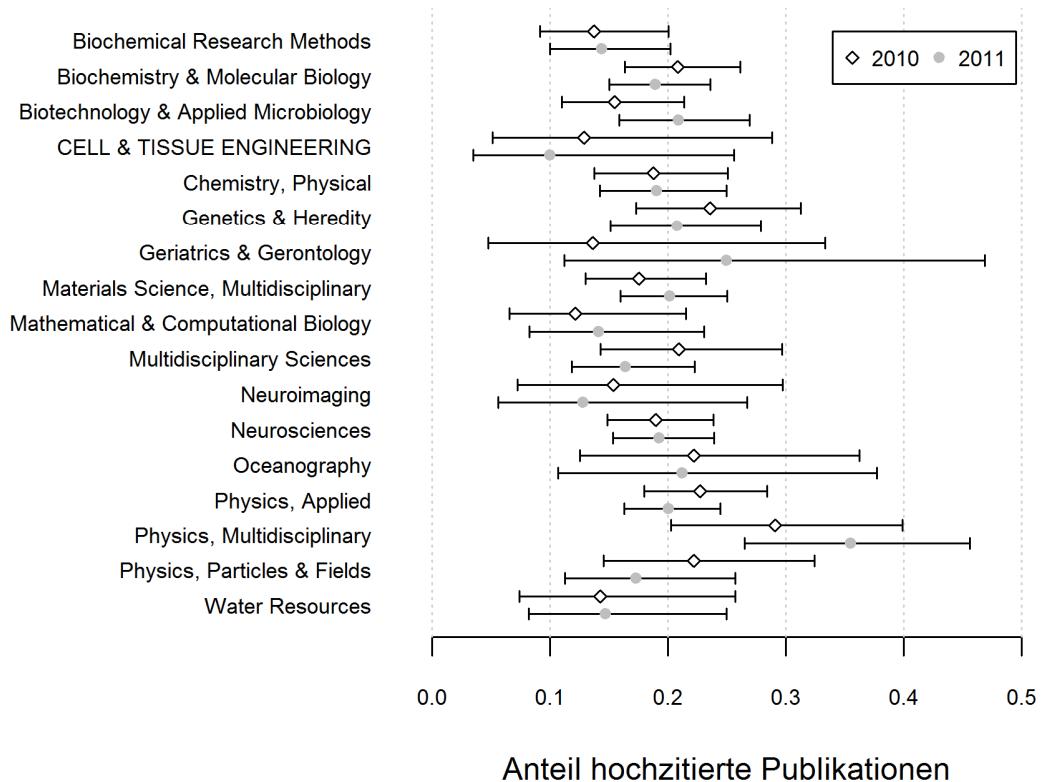
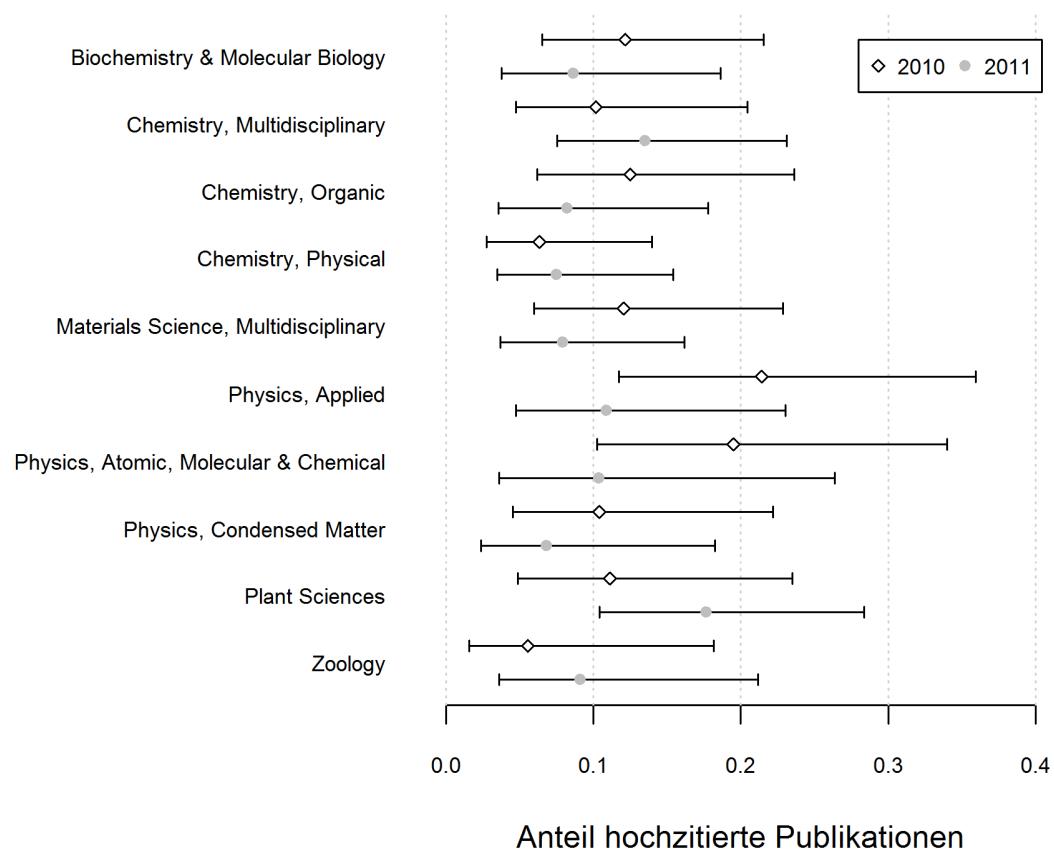
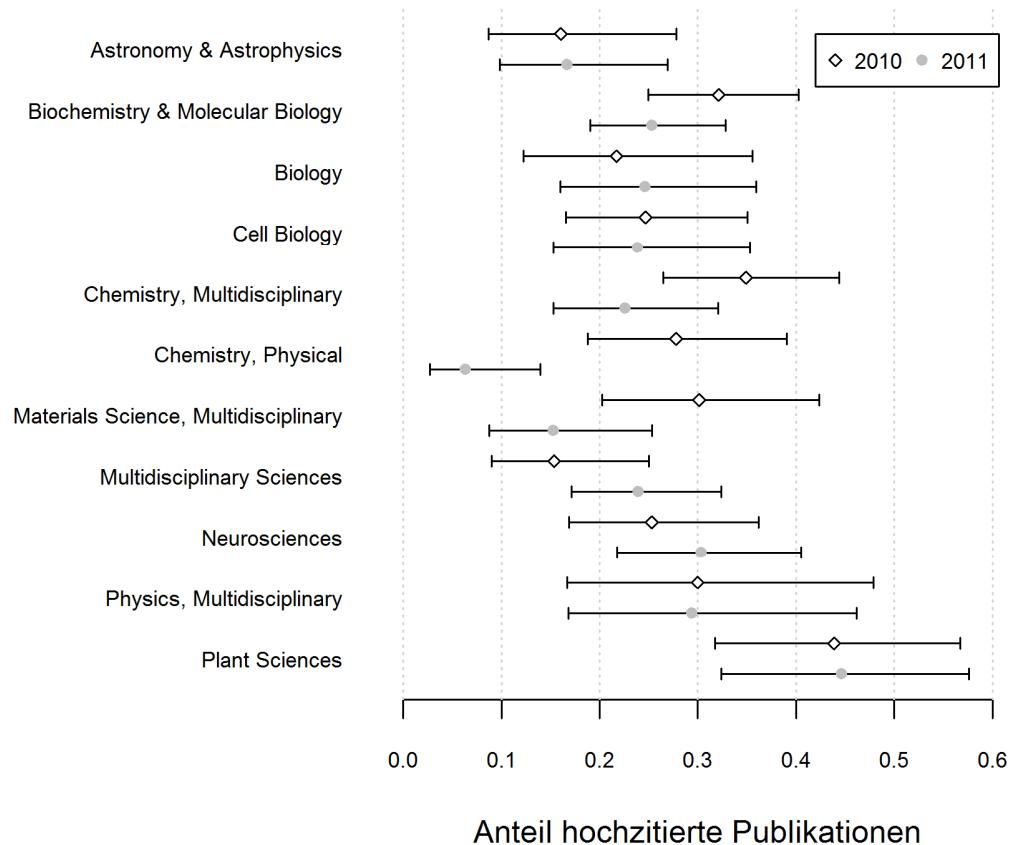


Abbildung 9: Anteil hochzitierter Journalartikel der DAAD in den publikationsstärksten Disziplinen (SCs) für 2010 und 2011*



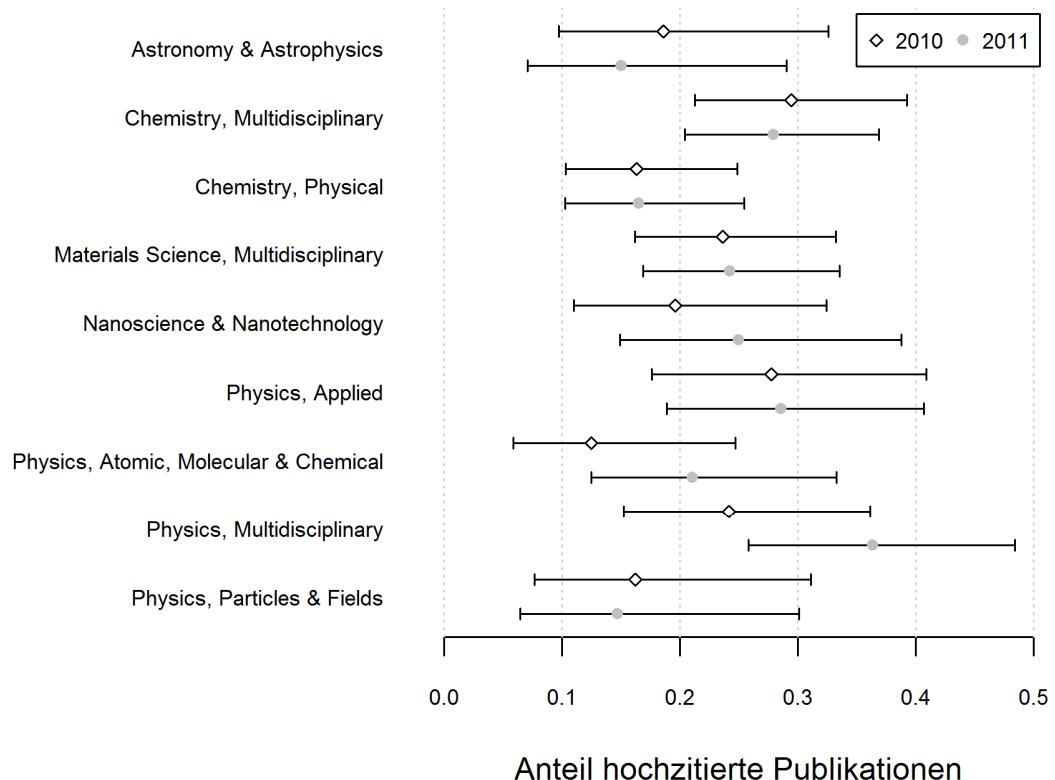
*mit SCs, in denen jährlich mindestens 30 Publikationen veröffentlicht werden

Abbildung 10: Anteil hochzitierter Journalartikel der MPG in den publikationsstärksten Disziplinen (SCs) für 2010 und 2011*



*mit SCs, in denen jährlich mindestens 30 Publikationen veröffentlicht werden

Abbildung 11: Anteil hochzitierter Journalartikel der AvH in den publikationsstärksten Disziplinen (SCs) für 2010 und 2011*



*mit SCs, in denen jährlich mindestens 30 Publikationen veröffentlicht werden

4.6.3 Impactmessung im internationalen Vergleich

Um die Impactmessung in einen internationalen Kontext zu stellen, wurden zusätzlich vier große europäische Förderorganisationen analysiert: die französische Agence Nationale de la Recherche (ANR), der flämische Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO), der österreichische Fonds zur Förderung der wissenschaftlichen Forschung (FWF) und die niederländische Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). In Tabelle 19 ist die Anzahl der gefundenen Journalartikel für die Jahre 2010 bis 2013 aufgeführt.

Tabelle 19: Publikationszahlen internationaler Förderer

FO	2010	2011	2012	2013	Total 2010-2013
ANR	5.544	6.813	7.820	8.386	28.563
FWO	1.862	2.197	2.214	2.248	8.521
FWF	2.198	2.551	2.683	3.080	10.512
NWO	3.662	4.123	4.572	4.783	17.140

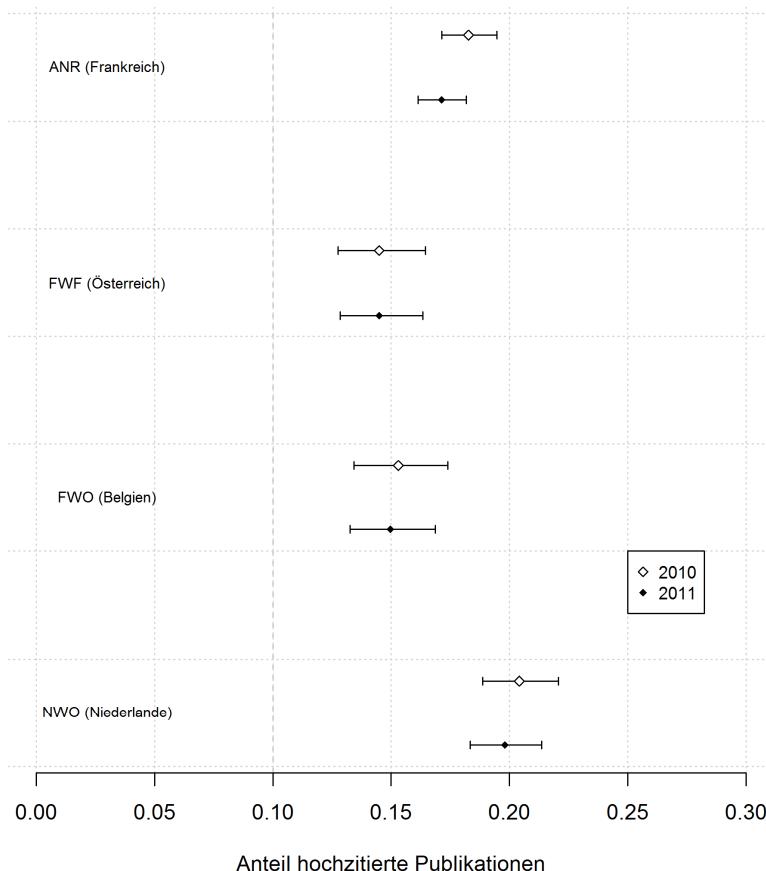
Um die Dominanz dieser Förderorganisationen in ihrem nationalen Wissenschaftssystem abzuschätzen, wurde in Tabelle 20 die Anzahl der geförderten Journalartikel durch die jeweils größten Organisationen im Land mit allen Journalartikeln und allen Artikeln mit Funding Acknowledgements in den Jahren 2010 bis 2013 verglichen. Für Deutschland wurde hier die DFG ausgewählt. Hierbei zeigt sich, dass die großen nationalen Förderorganisationen in Deutschland und Österreich eine wichtigere Rolle spielen als in den anderen Ländern. Anzumerken ist, dass es in Belgien zwei große Organisationen gibt – eine für Flandern und eine für die Wallonische Region (Fonds de la Recherche Scientifique – FNRS). Die Suche nach Artikeln des FNRS ergab 3.238 Journalartikel für diese Periode. Es sei hier aber nochmals erwähnt, dass obwohl die Präzision unserer Methode nahezu hundert Prozent beträgt, der Recall, also der Anteil an Falschnegativen, nicht im Einzelnen überprüft wurde (siehe Abschnitt 3.2).

Tabelle 20: Anteil der größten nationalen Förderer am nationalen Output 2010-2013

Land	Alle Journal- artikel	Größte FO	Anteil FO am Gesamtoutput	Journalartikel mit FA	Anteil FO an FA
Frankreich	257.733	28.563	11,1%	148.235	19,3%
Belgien	71.904	8.521	11,9%	44.720	19,1%
Österreich	49.150	10.512	21,4%	30.276	34,7%
Niederlande	129.202	17.140	13,3%	74.861	22,9%
Deutschland	367.325	88.439	24,1%	223.229	39,6%

Die Impactmessung dieser vier europäischen Förderer wird in Abbildung 12 gezeigt. Es sind auch hier signifikante Unterschiede zu vermerken. Während der belgische FWO und der österreichische FWF mit der DFG in etwa gleichauf liegen, sind die französische ANR und der niederländische NWO erfolgreicher bei der Dissemination ihres Wissens.

Abbildung 12: Anteil hochzitierter Journalartikel nach Förderorganisationen für 2010 und 2011



In Tabelle 23 wurden schließlich die Ergebnisse für die beiden Jahre zusammengefasst und mit den zehn deutschen Förderern nach Anteil hochzitierter Journalartikel geordnet aufgelistet. Daneben wurde jeweils die untere und obere Grenze der 95%-Konfidenzintervalle angegeben. Wenn man den DAAD ausklammert, fördern alle diese Organisationen mindestens anderthalbmal so viele hochzitierte Publikationen verglichen zum Weltdurchschnitt (10%). Wenn man diese Ergebnisse mit dem Leidenranking vergleicht, wären diese Organisationen alle mit den Top 50-Universitäten der Welt vergleichbar. Dabei ist aber anzumerken, dass das Leidenranking auch die im Durchschnitt höher zitierten Reviews beinhaltet, während sie hier ausgeschlossen wurden. Ein Vergleich auf gleicher Dokumenttypbasis würde noch bessere Ergebnisse für die FOs hervorbringen.

Tabelle 23: Anteil hochzitierter Journalartikel deutscher und internationaler Förderorganisationen für 2010 und 2011 zusammengefasst

FO	unterer KI-Wert	Anteil hochzitierter Journalartikel	oberer KI-Wert
MPG	23,6%	25,5%	27,4%
HGF	19,9%	22,3%	24,9%
AVH	18,8%	20,6%	22,6%
NWO	19,0%	20,1%	21,2%
BMBF	18,5%	19,3%	20,2%
DKH	16,1%	18,6%	21,3%
ANR	16,9%	17,6%	18,4%
FCI	14,7%	16,3%	18,1%
VW	13,8%	16,2%	19,0%
SS	12,7%	15,7%	19,2%
DFG	15,1%	15,5%	15,9%
FWO	13,8%	15,1%	16,5%
FWF	13,3%	14,5%	15,8%
DAAD	10,3%	11,6%	13,0%

5. Abschließende Bemerkungen

Die hier präsentierten Daten sind die ersten Resultate einer neu entwickelten Methode zur Bereinigung der Funding Acknowledgement-Einträge der Web of Science Datenbank. Diese explorative Studie verdeutlicht anhand beispielhafter Einblicke das große langfristige Potenzial und die breit gefächerten Anwendungsmöglichkeiten der bereinigten Funding Acknowledgement-Daten. Der vorgestellte Bereinigungsprozess ist methodisch-empirisch unterstützbar und für die umfassende Anwendung kompletter Anbieterdatenbanken skalierbar. Zudem wurde im derzeitigen Prozess bereits umfangreiche Expertise hinsichtlich der Konfiguration und methodischen Weiterentwicklung des Prozesses entwickelt. Dadurch sind bereits zum jetzigen Zeitpunkt erkenntnisgenerierende Analysen von und durch Förderorganisationen möglich, deren Durchführung bisher häufig an mangelnder Datenqualität, einem unzureichenden empirischen Kenntnisstand und dem Aufwand der Datenbereinigung scheiterten. Obwohl diese Ergebnisse aufgrund des vorläufigen Entwicklungsstadiums derzeit erst vorsichtige Schlüsse zu lassen, wurde bereits in den explorativen Anwendungsszenarien deutlich, dass diese neue Methode und Datenbasis sowohl erwartbare als auch überraschende Erkenntnisse durch Einblicke in einen bislang kaum untersuchbaren Bereich des Wissenschaftssystems ermöglicht.

Schließlich soll noch besondere Erwähnung finden, dass das Ziel der Anwendung der derzeitigen Datenquelle und Methode nicht darin liegt, Evaluationen von Forschungseinrichtungen durchzuführen, da dafür, auch nach der Bereinigung, die Datenlage zu lückenhaft ist. Vielmehr soll das Ziel in der Umsetzung von Fragestellungen aus der Perspektive der Förderorganisationen selbst liegen. So lassen sich Tendenz- und Profilanalysen zur Ableitung von Handlungsszenarien einsetzen. Um die Ergebnisse für diese Förderorganisationen besser zu verstehen, bedürfen die Einflussfaktoren erfolgreicher Förderung sowie ihr organisatorischer und prozessualer Kontext jedoch tieferer Analyse.

Referenzen

- Chi, Pei-Shan. 2012. „Bibliometric Characteristics of Political Science Research in Germany“. *Proceedings of the American Society for Information Science and Technology* 49 (1): 1–6. doi:10.1002/meet.14504901115.
- Katz, J. Sylvan, und Ben R. Martin. 1997. „What is research collaboration?“. *Research Policy* 26 (1): 1–18. doi:10.1016/S0048-7333(96)00917-1.
- Levenshtein, V. 1966. „Binary Codes Capable of Correcting Deletions, Insertions, and Reversals“. *Soviet Physics-Doklady* 10 (8): 707–10.
- Nederhof, Anton J. 2006. „Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review“. *Scientometrics* 66 (1): 81–100. doi:10.1007/s11192-006-0007-2.
- Riechert, Mathias, Daniel Sirtes, und Haiko Lietz. 2013. „Collaboration Benefits and International Visibility of Two-Country-Collaborations“. In *Translational twists and turns: Science as a socio-economic endeavor. Proceedings of STI 2013 Berlin – 18th International Conference on Science and Technology Indicators.*, 302–12. Berlin.
- Rigby, John. 2011a. „Systematic Grant and Funding Body Acknowledgment Data for Publications: An Examination of New Dimensions and New Controversies for Bibliometrics“. *SSRN eLibrary*, Januar. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1742845.
- . 2011b. „Systematic Grant and Funding Body Acknowledgement Data for Publications: New Dimensions and New Controversies for Research Policy and Evaluation“. *Research Evaluation* 20 (5): 365–75. doi:10.3152/095820211X13164389670392.
- Sirtes, Daniel. 2013. „Funding Acknowledgements for the German Research Foundation (DFG). The Dirty Data of the Web of Science Database and How to Clean It Up“. In *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference*, herausgegeben von Juan Gorraiz, Edgar Schiebel, Margit Hörlesberger, und Henk Moed, 1:784–95. Vienna: AIT GmbH.
- Sirtes, Daniel, und Mathias Riechert. in preparation. „Problems and (some) Solutions for Funding Acknowledgements“
- . 2014. „A Fully Automated Method for the Unification of Funding Organizations in the Web of Knowledge“. In *Context counts: pathways to master big and little data. Proceedings of the 19th International Conference on Science and Technology Indicators, 2014 Leiden*, herausgegeben von Ed Noyons, 594–97.
- van den Besselaar, Peter, Annamaria Inzelt, und Reale, Emanuela. 2012. „Measuring Internationalization of Funding Agencies“. In *Proceedings of 17th International Conference on Science and Technology Indicators*, herausgegeben von Éric Archambault, Yves Gingras, und Vincent Larivière, 121–30. Montreal.
- Waltman, Ludo, und Michael Schreiber. 2013. „On the Calculation of Percentile-Based Bibliometric Indicators“. *Journal of the American Society for Information Science and Technology* 64 (2): 372–79. doi:10.1002/asi.22775.
- Wang, Jian. 2013. „Citation Time Window Choice for Research Impact Evaluation“. *Scientometrics* 94 (3): 851–72. doi:10.1007/s11192-012-0775-9.