

Forschungsbewertung neu definiert: FIS, Open Science und KI-gestützte Innovationen

Otmane Azeroual, Ulrich Herb, Joachim Schöpfel

1 Einleitung

Die Art und Weise, wie wissenschaftliche Forschung bewertet wird, unterliegt einem kontinuierlichen Wandel. In Frankreich und Deutschland hat dieser Wandel in den letzten Jahren an Dynamik gewonnen, insbesondere durch die Einführung und Weiterentwicklung von Forschungsinformationssystemen (FIS) oder im Englischen Current Research Information Systems (CRIS). Diese Systeme dienen als zentrale Werkzeuge zur Erfassung, Analyse und Bewertung wissenschaftlicher Aktivitäten. Sie ermöglichen eine umfassende Sicht auf die wissenschaftliche Produktivität und Leistungsfähigkeit von Personen und Forschungseinrichtungen und sind somit unverzichtbar für die interne Steuerung sowie die externe Leistungsbewertung durch Förderorganisationen und staatliche Institutionen. Gleichzeitig wird die Bewegung hin zu Open Science und einer ethisch fundierten Wissenschaft durch Initiativen wie Coalition for Advancing Research Assessment (CoARA)¹ und San Francisco Declaration on Research Assessment (DORA)² unterstützt. Diese fordern eine ganzheitliche und transparente Bewertung wissenschaftlicher Leistungen, die qualitative Aspekte stärker berücksichtigt und den Missbrauch quantitativer Indikatoren vermeidet. Die Wahl der richtigen Indikatoren und die Integration moderner Technologien wie Künstliche Intelligenz (KI) stellen hierbei zentrale Herausforderungen dar. Traditionelle Metriken wie der Journal Impact Factor (JIF) werden u.a. wegen disziplinären Verzerrungen, begrenztem Zitierfenster und Prognoseschwäche kritisiert (Larivière & Sugimoto, 2019). Der h-Index, ebenfalls eine etablierte Metrik, wird wegen seiner Benachteiligung gegenüber jüngeren Forschern, Forschenden in Ländern mit geringen Ressourcen, wegen seines simplen methodischen Charakters und seines begrenzten Nutzens für Forschungsevaluationen kritisiert (Thelwall & Kousha, 2021). Alternative Metriken und KI-gestützte Tools bieten neue Möglichkeiten zur objektiveren und umfassenderen Bewertung wissenschaftlicher Leistungen, dieser Beitrag diskutiert diese Ansätze und ihre Grenzen sowie speziell die Frage, welche Rolle Bibliotheken bei einer FIS-gestützten Reform der Forschungsbewertung einnehmen können.

1 <https://coara.eu/> [28.08.2024]

2 <https://sfedora.org/> [28.08.2024]

Abstract

Die wissenschaftliche Landschaft in Frankreich und Deutschland durchläuft eine tiefgreifende Transformation, die durch die zunehmende Integration von Forschungsinformationssystemen (FIS) geprägt ist. Diese Systeme sind essenziell für die Verwaltung und Bewertung wissenschaftlicher Aktivitäten und spielen eine zentrale Rolle im Übergang zu Open Science und Forschungsethik. Dieser Beitrag untersucht die Entwicklungen und Herausforderungen in der Forschungsbewertung, beleuchtet die Rolle von Initiativen wie Coalition for Advancing Research Assessment (CoARA) und San Francisco Declaration on Research Assessment (DORA) für die Forschungsbewertung und diskutiert die Integration moderner Indikatoren und Künstlicher Intelligenz (KI) in FIS. Zudem wird die Rolle der Bibliotheken in diesem Transformationsprozess hervorgehoben. Die Ergebnisse verdeutlichen, dass die (R)Evolution der Forschungsbewertung durch FIS die wissenschaftliche Leistung fairer und transparenter gestalten kann, trotz bestehender Herausforderungen – wenn es gelingt, Open Science Indikatoren und KI reflektiert einzusetzen.

The scientific landscape in France and Germany is undergoing a profound transformation characterized by the increasing integration of Current Research Information Systems (CRIS). These systems are essential for the management and evaluation of scientific activities and play a central role in the transition to Open Science and research ethics. This article examines the developments and challenges in research assessment, highlights the role of initiatives such as the Coalition for Advancing Research Assessment (CoARA) and the San Francisco Declaration on Research Assessment (DORA) for research assessment and discusses the integration of modern indicators and artificial intelligence (AI) in RIS. It also emphasizes the role of libraries in this transformation process. The results illustrate that the (r)evolution of Current Research Information Systems (CRIS) can make scientific performance fairer and more transparent, despite existing challenges – if Open Science indicators and AI can be used in a reflective way.

2 Leistungsbewertung und FIS

FIS ermöglichen eine detaillierte und differenzierte Leistungsbewertung von Forschenden und Forschungseinrichtungen. Diese Bewertung basiert auf einer Vielzahl von Indikatoren, die sowohl quantitative als auch qualitative Aspekte der Forschung berücksichtigen. Zu den häufig genutzten quantitativen Metriken gehören:

- Personen und Institutionen: Informationen über Forschende, Autoren, Koautoren und andere Beteiligte,

- inklusive der Institutionen, an denen sie tätig sind, und Forschungsprofilen,
- Publikationen: Details zu wissenschaftlichen Veröffentlichungen wie Titel, Abstracts, Autoren, Veröffentlichungsdatum, Zitationsmetriken und DOIs,
 - Projekte und Fördermittel: Informationen über laufende oder abgeschlossene Forschungsprojekte, Fördermittelgeber, Projektbeschreibungen, Budgets und Laufzeiten,
 - Kooperationen und Netzwerke: Angaben zu Forschungs Kooperationen, Partnerschaften und Netzwerken,
 - Erfindungen, geistiges Eigentum und Patente, z.B. Patentnummern, Anmeldedaten und Inhaber,
 - Veranstaltungen und Teilnahme an (internationalen) Konferenzen,
 - Forschungsdaten und -materialien: Informationen über Datensätze, Laborprotokolle, Experimente und andere Forschungsmaterialien sowie deren Metadaten, Zugriffsrechte und Dateiformate,
 - Zitationsmetriken und Sichtbarkeit: z.B. Zitationen, h-Index und Journal Impact Factor Scores,
 - Betreuung von Doktoranden,
 - Zusammenarbeit mit der Industrie.

Diese Informationen sind ideal zur Bewertung von Forschungsleistungen und leistungsorientierten Mittelvergaben (LOM) bzw. performance-based-budgeting (PBB) geeignet (Schultz, 2015). Zum Beispiel nutzt die Universitätsmedizin Göttingen (UMG) das FIS FACTScience und beschreibt den Einsatz des LOM auf ihrer Internetseite (Universitätsmedizin Göttingen, o. J.). Die LOM der UMG berücksichtigt vorrangig Publikationen, Hauptautorenschaften und Drittmittelvolumen, gewichtet nach Drittmittelgeber (Universitätsmedizin Göttingen, 2016, 2017)³. Auch die Wirtschaftswissenschaftliche Fakultät der Georg-August-Universität Göttingen sowie die Freie Universität Berlin, die Universitäten Osnabrück und Köln wenden LOM an (Freie Universität Berlin, o. J.; Georg-August-Universität Göttingen, o. J.; Universität Osnabrück, o. J.; Universität zu Köln, 2021). LOM-Vorgaben für Bundesländer sind z.B. für Nordrhein-Westfalen und Schleswig-Holstein zugänglich (Landesrechnungshof Schleswig-Holstein, o. J.; Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen, 2014). Es muss jedoch betont werden, dass nicht die gesamte Mittelvergabe über LOM erfolgt und auch nicht allein über Forschungs-/Publikationsleistung gesteuert wird (Schultz, 2015). An der Freien Universität Berlin werden 20 % der jährlichen Mittel wettbewerblich zugewiesen, wobei 5 % von Gleichstellungserfolgen abhängen (Freie Universität Berlin, o. J.). Ein bedeutender Vorteil von FIS ist ihre Fähigkeit, Daten in Echtzeit zu verarbeiten und zu analysieren, wodurch ein

kontinuierliches Monitoring der Forschungsaktivitäten und eine zeitnahe Bewertung möglich werden. Dies unterstützt nicht nur die interne Steuerung von Forschungseinrichtungen, sondern auch externe Bewertungsverfahren, z.B. durch Förderorganisationen und Akkreditierungsagenturen.

Trotz der zahlreichen Vorteile von FIS gibt es auch Herausforderungen und Kritikpunkte, die mit ihrer Nutzung verbunden sind. Ein zentraler Kritikpunkt betrifft die Überbetonung quantitativer Indikatoren, die zu einem verzerrten Bild der wissenschaftlichen Leistung führen kann. Dies kann Fehlanreize schaffen, wie z.B. die Priorisierung von Quantität über Qualität bei der Publikation von Forschungsergebnissen. Ein weiterer Aspekt ist der Datenschutz, insbesondere die Sicherstellung der Vertraulichkeit und Integrität sensibler Forschungsdaten.

Zudem erfordert die Implementierung und Wartung von FIS erhebliche Investitionen in IT-Infrastruktur und Personal. Die Datenqualität und -konsistenz stellen ebenfalls eine Herausforderung dar, da die Genauigkeit der Leistungsbewertung stark von der Verfügbarkeit und Zuverlässigkeit der eingegebenen Daten abhängt.

Frankreich

In Frankreich wird die Forschung hauptsächlich von einigen großen Forschungsgesellschaften finanziert und organisiert, allen voran dem Centre National de la Recherche Scientifique (CNRS), während die überwiegende Anzahl der Universitäten bislang eine eher untergeordnete Rolle spielt. Zwar hat die Integration von FIS an Bedeutung gewonnen und wird als wesentliches Werkzeug zur Verwaltung und Bewertung wissenschaftlicher Aktivitäten anerkannt. Doch haben vor allem die Forschungsgesellschaften wie CNRS (Grundlagenforschung), INRIA (Informatik), INRAE (Agronomie und Umwelt) und INSERM (Medizin) eigene FIS entwickelt, (bislang) nicht standardisiert und nicht interoperabel, während die Universitäten sich mit mehr oder weniger funktionellen, hauseigenen Lösungen begnügen. Kommerzielle FIS-Angebote wie PURE oder CONVERIS spielen in Frankreich bislang keine Rolle, zumindest nicht im öffentlichen Bereich. Zwei neuere Initiativen tragen dazu bei, die Forschungsbewertung in Frankreich und speziell an den Universitäten zu verbessern. Auf der einen Seite hat die französische AMUE-Agentur, die zentrale IT-Lösungen für den Hochschulbereich anbietet, ein System zur Verwaltung von Forschungsprojekten entwickelt (CapLab), das zumindest teilweise einem FIS entspricht und das zunehmend von Universitäten eingesetzt wird. Auf der anderen Seite wird das nationale Forschungsrepositorium HAL, welches eine zentrale Rolle in der französischen Open-Science-Politik spielt, sowohl

³ Man folgt hier Empfehlungen der Deutschen Forschungsgemeinschaft (DFG) (2004).

von Forschungsgesellschaften als auch von Universitäten und Forschungslabors mehr und mehr zur institutionellen und individuellen Evaluierung benutzt. Diese Initiativen verbessern die Sichtbarkeit und Zugänglichkeit von Forschungsergebnissen. Die Entwicklung und Implementierung von Open-Science-Strategien wird stark gefördert, unterstützt durch Open-Access-Repositorien und Anreize für Forschende. Darüber hinaus fördern Institutionen wie das CNRS und die Agence Nationale de la Recherche (ANR) Forschungsethik und faire Bewertungsrichtlinien. Herausforderungen bestehen in der Harmonisierung der FIS-Plattformen, der IT-Infrastruktur und dem Datenschutz. Frankreich ist auf einem guten, wenngleich eigenen Weg, FIS zu etablieren, Open-Science-Prinzipien zu integrieren und die Forschungslandschaft zu stärken.

Deutschland

In Deutschland sind FIS tief in die IT-Infrastruktur von Universitäten und Forschungseinrichtungen integriert. Sie fungieren als zentrale Datenbanken, die Informationen aus verschiedenen Quellen konsolidieren und den Zugriff auf umfassende Forschungsinformationen erleichtern. Diese Systeme sind meist modular aufgebaut und können an die spezifischen Bedürfnisse und Anforderungen der jeweiligen Institution angepasst werden. Zu den häufig genutzten FIS-Plattformen in Deutschland gehören PURE, HISinOne-RES und CONVERIS, die eine Vielzahl von Funktionen bieten, wie z.B. die Verwaltung von Forschungsprojekten, die Erstellung von Berichten und die Unterstützung von Open Access.

Die Weiterentwicklung von FIS in Deutschland zielt darauf ab, diese Herausforderungen zu adressieren und die Systeme noch benutzerfreundlicher und leistungsfähiger zu gestalten. Die Integration von Open-Science-Prinzipien und die Nutzung von Künstlicher Intelligenz zur Verbesserung der Datenanalyse und -interpretation sind vielversprechende Ansätze, um eine fairere und transparentere Forschungsbewertung zu gewährleisten. In diesem Kontext spielen auch Kooperationen zwischen Forschungseinrichtungen, IT-Dienstleistern und politischen Entscheidungsträgern eine entscheidende Rolle, um die Rahmenbedingungen für die Nutzung von FIS weiter zu verbessern und die deutsche Forschungslandschaft im internationalen Wettbewerb zu stärken.

3 Aktuelle Entwicklungen und Herausforderungen

3.1 CoARA, DORA und Open-Science-Metriken

CoARA und DORA adressieren die erwähnten Mängel der klassischen und in FIS genutzten Indikatorik vor allem, indem sie

- die Vermeidung der übermäßigen Abhängigkeit von Journal Impact Factors,

- die Einbeziehung von alternativen Bewertungsmetriken wie Altmetrics und Open Science / Open-Science-Indikatorik,
- die Anpassung der Leistungsbewertung an die individuellen Kontexte und Beiträge der Forschenden,
- die Förderung der Transparenz und Offenlegung der Bewertungsprozesse

fordern. Kurz gesagt: DORA und CoARA rufen zur Abkehr von der Verwendung kritizierter quantitativer Metriken für Forschungsqualität und -produktivität (z.B. JIF, h-Index, Anzahl der Zitate, Anzahl der (Ko-)Autorenschaften, Anzahl der Veröffentlichungen) auf, da diese nur lose mit wissenschaftlicher Qualität verbunden seien, und fordern eine Änderung der sich primär an Publikationsinformationen orientierenden Anreizstruktur in der Wissenschaft.

Diese Kritik an Metriken wie JIF oder h-Index wird ergänzt durch die Kritik an der (im Falle des JIF ausschließlich und des h-Index häufig) genutzten Datenbasis des Web Of Science bzw. anderer Zitationsdatenbanken wie Scopus, die z.B. nicht-englisch Publizierende hinsichtlich der Indexierung in den Datenbanken und wegen der Unterrepräsentanz nicht-englischsprachiger Literatur in den Datenbanken bei der Ermittlung der Zitationszahlen benachteiligen (Vera-Baceta et al., 2019). Zudem werden diese Zitationsdatenbanken mit Skepsis betrachtet (Herb, 2016) aufgrund

- ihrer Eindimensionalität, denn sie verengen die Idee des wissenschaftlichen Impacts auf Zitationen und lassen z.B. Verbreitung in Social Media (via Altmetrics), Downloads von Inhalten, Leistungen in der Lehre/ Betreuung außer Acht, und
- der Proprietät ihrer Daten, die eine Überprüfung der ermittelten Indikatoren nicht zulässt.

Die Barcelona Declaration on Open Research Information (Bianca Kramer et al., 2024) erweitert folgerichtig CoARA-/DORA-Forderungen, indem sie die Verwendung von offenen (transparenten, überprüfbar, nachnutzbar) Informationen bei der Forschungsbewertung postuliert. Ganz im Einklang hiermit befindet sich Open Science Netherlands (12.5 million euros for Open Science Infrastructure | Open Science NL, o. J.), das im Juni 2024 ein Fördervolumen von 12,5 Millionen Euro für die Förderung von Open-Science-Infrastrukturen ausrief. Sowohl die von DORA und CoARA geforderte Abkehr vom alleinigen Vertrauen auf Zitationsdaten zugunsten einer Mehrdimensionalität der Bewertung als auch die Offenheit und Überprüfbarkeit der Daten kann die Manipulation von Zitationsmetriken erschweren bzw. deren Aufdeckung erleichtern. Das gilt z.B. für das von Żuradzki & Wroński (2024) geschilderte Manipulieren von Scopus-basierten Rankings durch Kreuzzitationen außerhalb des Zitationsnetzwerks wenig zitierter Journale.

Open-Science-Indikatoren in der Forschungsbewertung

Forderungen, wie durch CoARA oder DORA dargestellt, führen nun aber nicht zu einer Abkehr der Forschungsbewertung mittels Metriken; im Gegenteil findet sich geradezu eine Inflation neuer Metriken (oder Überlegungen zu diesen), die vor allem auf der Einführung von Open-Science-Indikatoren fußen. Diese Indikatoren dienen der Ermittlung der Open Science Compliance/ Adaption einer Person, eines Journals, einer Research Performing Organisation (RPO), einer Research Funding Organisation (RFO) oder einer anderen Entität. Mögliche Itemtypen/ Messobjekte zur Ermittlung der Compliance sind etwa: Quantität der

- Datenmanagement Pläne (DMPs),
- Open-Access-Text-Publikationen,
- Open-Access-/Daten-Publikationen,
- Open Review Reports,
- FAIR-lizenzierten Items,
- Preprints,
- Prä-Registrierungen,
- Replikationsstudien,
- Downloads,
- Zitationen,
- Social Media Mentions,
- Diversität,
- Kollaborationen,
- Citizen-Science-Aktivitäten,
- Medien-Aufmerksamkeit.

Die genannten Itemtypen entstammen dem Open Science Indicator Handbook (*Open Science Indicator Handbook*, o. J.), Publikationen und Informationen über die Integration von Open-Science-Indikatoren in ein CRIS der Region Flandern (Dengis & De Bal, 2022), das von der Europäischen Union geförderte Projekt *GraspOS: next Generation Research Assessment to Promote Open Science* (Hyrkkänen et al., 2023), einer Publikation des Projekts *Tools to Advance Research Assessment (TARA)*⁴, das die Entwicklung neuer Strategien und Verfahren für akademisches Career Assessment zum Ziel hat (Schmidt, 2022) sowie den Berufungsmodalitäten z.B. des Fachbereichs Psychologie an der TU Dresden (Gärtner, 2023; Schönbrodt et al., 2022). Hier und auch an anderen Standorten, speziell in der Psychologie, werden Bewerbungen auf Professuren in einer ersten Auswahl u.a. nach dem Ausmaß der Open-Science-Aktivitäten und inhaltlicher Qualität (sofern sie sich durch Kriterien messen lässt) gerankt⁵, die auf der Shortlist ver-

bleibenden Personen werden anschließend narrativ (inhaltlich) anhand der Qualität ihrer Arbeit beurteilt. Wissenschaftliche Qualität soll demnach in beiden Phasen bewertet werden; in Phase eins wird sie eher als Methodological Rigor / Research Integrity interpretiert, in Phase zwei eher als Innovativität und Originalität der Forschung.

Bei vielen, wenn auch nicht allen dieser Vorschläge (bleibt wie bei den klassischen Metriken) auch der Verdacht bestehen, dass sich diese Vorhaben mit dem „Messen des leicht Meßbaren“ (Fröhlich, 1999) begnügen, zumal – wie bei den etablierten Metriken – fast ausnahmslos quantitative Informationen verarbeitet werden. Unklar bleibt zuvorderst (genau wie bei JIF und h-Index), welches Merkmal erfasst werden soll: Offenheit der Wissenschaft? Nachnutzbarkeit? Diversität? Impact vulgo: Wirkung oder Einfluss? Qualität? Zweckerreichung? Nachdem man zu Recht die Benachteiligung z.B. sprachlicher Diversität bei der Metrik JIF kritisiert hat, ist es vielleicht überlegenswert, ob es sinnvoll ist, Diversität selbst zu einem Indikator einer Metrik zu machen. Die Frage nach dem Zweck von Forschung lässt sich unterschiedlich beantworten, ebenfalls die danach, was Qualität von Forschung ausmacht. Je nach Verständnis sind Offenheit, Nachnutzbarkeit und Diversität vielleicht synonym mit Qualität, allerdings dürfte diese Wahrnehmung in den Communitys und auch der Zivilgesellschaft nicht unbedingt dominieren, sondern eher vielleicht die Zielerreichung relevant sein. Offenheit und Nachnutzbarkeit können Forschung fraglos beschleunigen, auch helfen, Fehler zu entdecken und damit Qualität sichern, allerdings sind sie keine *conditio sine qua non* für Qualität; z.B. korrelieren auch die Verfügbarkeit von Forschungsdaten und Retractions nicht (Berberi & Roche, 2022). Das Verfahren der (z.B. Dresdener) Psychologie bemüht sich um Ausgewogenheit zwischen Open-Science-Inzentivierung, quantitativen und qualitativen Informationen, räumt der Offenheit der Forschung und der wissenschaftlichen Strenge allerdings das Primat über die inhaltliche Qualität ein und setzt damit starke disziplinäre Impulse für Open Science. Mit Erfolg: In den letzten Jahren nahm in Deutschland bei Ausschreibungen für Professuren der Open-Science-affinen-Psychologie die Erwähnung von Open-Science-Aktivitäten deutlich zu (Schönbrodt et al., 2023). Auch Open Science Netherlands (*Call to include open science in hiring and promotion policies | Open Science NL*, o. J.) zielt auf die Berücksichtigung von Open Science Compliance bei Berufsverfahren und lobte im Juni 2014 Förderung von bis zu 50.000 € für Organisationen, die Pläne vorlegen, wie sie diese Kriterien in ihren

4 Mittels des Service Reformscape (als einem TARA-Baustein) können bereits implementierte Verfahren des Responsible Research Assessments recherchiert werden, <https://sfdora.org/reformscape/> [28.08.2024].

5 Man nutzt ein Framework namens RESQUE, <https://nicebread.github.io/RESQUE/web/> [28.08.2024]. Andere Parameter sind z.B. Beiträge zur Theorieentwicklung und -testung sowie wissenschaftlicher Impact. Letzterer wird über age- und field-corrected BIP! scores (Vergoulis et al., 2022) ermittelt. Ein passendes Webinterface zum BIP! findet sich unter: <https://bip.imsi.athenarc.gr/search> [28.08.2024]

Berufungsverfahren berücksichtigen wollen, aus. Der Ansatz von Open Science NL (und der Psychologie) scheint: Inzentivierung von Open Science, weil offen verfügbare Wissenschaft bessere Wissenschaft ermöglicht: "For open science to become the norm, it is essential that researchers are recognised and rewarded for putting it into practice." Allerdings sollte der Impuls der Inzentivierung von Open Science nicht konfundiert werden mit der Qualität der Forschung und es stellt sich die Frage, welches Moment bei Berufungsverfahren stärker gewichtet werden sollte. Generell gilt für Open-Science-Indikatoren überdies, dass sie, wie alle „Indikatoren (...) die Gefahr in sich [bergen], dass das Handeln der beteiligten Akteure sich nicht mehr an der Optimierung der Leistung, sondern an einer Optimierung der Indikatoren orientiert." (Schultz, 2015).

Die Liste der möglichen Itemtypen verdeutlicht weitere Herausforderungen: Wie gelingt die Standardisierung und Vergleichbarkeit von Open-Science-Indikatoren? Müssen diese Indikatoren für unterschiedliche Entitäten (Journal, Person, Organisation) unterschiedliches Design haben? Wie gestalten sich Datenerhebung und -validierung? Zumindest die Nachnutzbarkeit der Daten wird durch die erwähnte Barcelona Declaration adressiert. Besonders, jedoch nicht nur für Organisationen stellt sich die Frage, des Erhebungsaufwands: Das mit Unterstützung des Stifterverbandes entwickelte und von diesem beworbene⁶ Tool Road2Openness (R2O),⁷ das RPOs durch Benchmarking und Toolkits Anreize zur Umsetzung von Open Science geben will, erfordert durch das Abfragen von bis zu 172, teils aufwändig zu erhebenden Informationen und Daten zu Themen wie Demografie der Einrichtung, Open Research Outputs, Citizen Science, Open Innovation, Open Educational Resources, Research Quality Management, Open Governance, Open Research Assessment und Open-Science-Strategie nicht zu unterschätzenden Aufwand. Der Nutzen eines solchen Werkzeugs, in diesem Fall die Formulierung von Hilfen zur Steigerung der Open Science Compliance, muss angesichts dessen klar erkennbar sein.

3.2 Datenschutz, KI und Ethik

Die Integration von KI in FIS bietet erhebliche Vorteile, wie die Automatisierung komplexer Datenanalysen, die Identifizierung von Mustern in großen Datensätzen und die Unterstützung bei der Vorhersage wissenschaftlicher Trends.

Speziell hinsichtlich der Forschungsbewertung ermöglicht KI eine Verbesserung der Datenanalyse und -bewertung. Maschinelles Lernen erlaubt die Identifikation von Forschungstrends und -mustern, Konkurrenz- und Kollaborationsanalysen und die Entwicklung von Algorithmen

zur Unterstützung der Entscheidungsfindung in der Forschungsbewertung.

Diese technologischen Fortschritte bringen jedoch auch eine Reihe von ethischen Herausforderungen mit sich, die sorgfältig adressiert werden müssen, um verantwortungsbewusste und faire Praktiken in der Forschungsbewertung zu gewährleisten.

Ein zentrales ethisches Anliegen bei der Nutzung von KI in FIS betrifft die Art und Weise, wie Daten verarbeitet und genutzt werden. Die Algorithmen der KI benötigen umfangreiche Datensätze, um präzise und zuverlässige Analysen durchführen zu können. Dabei besteht das Risiko, dass personenbezogene Daten und sensible Forschungsinformationen ohne ausreichenden Schutz vor unbefugtem Zugriff verwendet werden. Dies erfordert strenge Datenschutzrichtlinien und Maßnahmen zur Sicherstellung der Vertraulichkeit und Integrität der Daten. Forschungseinrichtungen müssen sicherstellen, dass alle rechtlichen und ethischen Vorgaben zum Datenschutz eingehalten werden, um die Privatsphäre von Forschenden und Probanden zu schützen.

Ein weiteres ethisches Problem ist die potenzielle Verzerrung (Bias) in den KI-Algorithmen. Algorithmen lernen aus den Daten, die ihnen zur Verfügung gestellt werden, und wenn diese Daten verzerrt oder unvollständig sind, können die Ergebnisse der KI ebenfalls verzerrt sein. Dies kann zu ungerechten und diskriminierenden Bewertungen führen, die bestimmte Gruppen von Forschenden benachteiligen. Es ist daher essenziell, dass die Daten, die zur Schulung von KI-Systemen verwendet werden, repräsentativ und divers sind. Zusätzlich sollten Mechanismen zur Überprüfung und Korrektur von Bias implementiert werden, um die Fairness und Gerechtigkeit der Bewertungen sicherzustellen.

Ein wichtiger Aspekt der ethischen Nutzung von KI in FIS ist die Transparenz und Erklärbarkeit der Algorithmen und ihrer Entscheidungen. Forschende und Administratoren müssen verstehen können, wie die KI zu ihren Schlussfolgerungen gelangt, um die Ergebnisse nachvollziehen und überprüfen zu können. Dies ist besonders wichtig in der Forschungsbewertung, wo Entscheidungen erhebliche Auswirkungen auf die Karriere und Finanzierung von Wissenschaftlern haben können. Die Entwicklung von erklärbaren KI-Modellen, die ihre Entscheidungsprozesse transparent machen, ist daher eine wesentliche Herausforderung.

Die Einführung von KI in die Forschungsbewertung wirft auch Fragen der Verantwortung und Haftung auf. Wer trägt die Verantwortung für Entscheidungen, die von KI-Systemen getroffen werden? Und wer ist haftbar, wenn

6 <https://www.stifterverband.org/road2openness-strategiewerkstatt> [28.08.2024]

7 Disclaimer: Einer der Autoren, Ulrich Herb, war an der Entwicklung von R2O als Fellow des Stifterverbandes beteiligt.

diese Entscheidungen fehlerhaft oder ungerecht sind? Forschungseinrichtungen müssen klare Richtlinien und Verantwortlichkeiten festlegen, um sicherzustellen, dass es bei Problemen mit KI-gestützten Bewertungen eine klare Rechenschaftspflicht gibt.

Die Nutzung von KI in FIS bietet großes Potenzial zur Verbesserung der Effizienz und Genauigkeit der Forschungsbewertung. Gleichzeitig bringt sie jedoch erhebliche ethische Herausforderungen mit sich, die nicht ignoriert werden dürfen. Datenschutz, Bias, Transparenz und Verantwortung sind zentrale Aspekte, die bei der Implementierung von KI in FIS sorgfältig berücksichtigt werden müssen. Nur durch einen verantwortungsbewussten und ethisch fundierten Ansatz kann sichergestellt werden, dass KI-basierte Systeme fair und gerecht sind und das Vertrauen der wissenschaftlichen Gemeinschaft gewinnen.

4 Résumé

CoARA, DORA, Open-Science-Indikatoren und KI schicken sich an, die Forschungsbewertung zu (r)evolutionieren. CoARA und DORA finden Berücksichtigung in Policies und Statements, kaum jedoch in konkreten Anwendungsszenarien für die Bewertungen wissenschaftlicher Leistung⁸: 27 deutsche Mitgliedsorganisationen der CoARA haben ein nationales Chapter gegründet, um die Ausgestaltung der Reform des Research Assessments mitzubestimmen und diese in Einklang mit nationalen Gegebenheiten zu bringen. Diese Zusammenarbeit soll sicherstellen, dass die spezifischen Anforderungen und Besonderheiten des deutschen Wissenschaftssystems in den Reformprozess einfließen. Neben dieser Initiative gibt es auch Projekte wie „Entwicklung eines Referenzmodells zum Reporting in wissenschaftlichen Einrichtungen anhand von DORA“ (ERRED)⁹. Dieses Projekt zielt darauf ab, ein standardisiertes Modell für das Reporting in wissenschaftlichen Institutionen zu entwickeln, welches sich an den Prinzipien der DORA orientiert. Diese Initiativen und Projekte tragen maßgeblich dazu bei, die Bewertung von Forschung auf eine transparentere und gerechtere Basis zu stellen und gleichzeitig die internationale Wettbewerbsfähigkeit deutscher Forschungseinrichtungen zu stärken. In Frankreich wurde ebenfalls ein nationales CoARA Chapter gegründet, dazu finden sich wie in Deutschland mit ERRED Forschungsprojekte zu Open Access und dessen Verwendung in der Evaluierung, und bereits 2022 wurde der Paris Call on Research Assessment veröffentlicht, als Erklärung und Appell für eine qualitative, vielfältige und inklusive Forschungsbewertung.

Abseits der Policy-Ebene erfolgen Entwicklung und An-

wendung beider Treiber derzeit ungesteuert, dies gilt auch für die Integration der Open-Science-Indikatoren und von KI in FIS und somit in die praktizierte Forschungsbewertung an RPOs.¹⁰ Dass es dabei bleibt, ist nicht zu vermuten, allein schon, weil findige Anbieter kommerzieller FIS (z.B. Elsevier als Anbieter des FIS Pure) zügig die Integration beider Werkzeuge angehen, nicht zuletzt aus Wettbewerbsgründen – zumal der Markt ja nach einer vielfältigen, diverseren und Open Science fördernden Indikatorik verlangt und die Verheißungen der KI zu verlockend sind, um sie nicht zur Vermarktung zu nutzen. Damit die Integration auf Seiten der Open-Science-Indikatorik den Ansprüchen eines Responsible Research Assessment genügt, ist es dringend nötig,

- Messdimensionen, Operationalisierungen und Indikatoren wissenschaftlich fundiert zu bestimmen (um nicht in der besten Absicht, Open Science zu fördern, nicht-valide Metriken wie JIF/ h-Index durch neue, ebenso wenig valide Metriken abzulösen),
- die Qualität der zur Ermittlung der Metriken benötigten Daten sicherzustellen,
- den verantwortungsvollen, sich ihrer Grenzen bewussten Umgang mit diesen Metriken zu praktizieren.

Bei der Integration der KI stellen sich ähnliche Herausforderungen wie

- Datenschutz,
- Vertraulichkeit,
- Integrität der Daten,
- Qualität der Trainingsdaten,
- Transparenz der Algorithmen,
- Nachvollziehbarkeit der Entscheidungsprozesse,
- Fairness und Vermeidung von Bias,
- Skalierbarkeit und Performance,
- Kompatibilität mit bestehenden Systemen,
- Sicherheit und Schutz vor Manipulationen,
- Einhaltung regulatorischer Anforderungen,
- Benutzerfreundlichkeit und Akzeptanz der Nutzer,
- Kontinuierliche Überwachung und Anpassung der Modelle.

Open-Science-Metriken und KI können vorteilhafte Effekte auf die Forschungsbewertung haben, sofern ihre Implementierung den oben genannten Anforderungen entspricht. Trotz dieser Herausforderungen bietet ihre Integration in FIS und die Forschungsbewertung die Chance, die wissenschaftliche Leistung fairer und transparenter zu bewerten und die Qualität der Forschung insgesamt zu verbessern. Bibliotheken sind hierbei unverzichtbare Part-

8 wie erwähnt zu recherchieren via Reformscape, <https://sfdora.org/reformscape/>

9 <https://www.bibliothek.kit.edu/ERRED.php> [28.08.2024]

10 Eine Ausnahme bildet die erwähnte Abbildung von Open-Science-Indikatoren im CRIS der Region Flandern (Dengis & De Bal, 2022).

ner, die durch ihre Expertise und Infrastruktur maßgeblich zum Erfolg beitragen. Dies betrifft ihre Rolle als

- Open-Science-Agenten, die Open-Science-Plattformen bereitstellen, Open-Science-Praktiken unterstützen und Open Science politisch vorantreiben,
- Spezialisten in Sachen Datenqualität durch Anwendung von Standards und Nutzung sowie Implementierung von Persistent Identifiers PID,
- Berater bei der Implementierung von Open-Science-Indikatorik und KI in FIS, speziell in Abgrenzung zu den wissenschaftssteuernden Toolkits kommerzieller Anbieter (Herb, 2018), um einen verantwortungsvollen Umgang mit Metriken sicherzustellen,
- Experten bei der Auswahl in RPOs anzuwendender KI-Tools durch Auditing/ Bewertung von KI-Tools, die zur Auswertung in FIS herangezogen werden,
- Schnittstelle zwischen Forschenden, IT-Abteilungen und Forschungsförderern, die zur erfolgreichen Im-

plementierung neuer Bewertungsansätze durch Open Science und KI beiträgt.

Um diese Rolle ausfüllen zu können, benötigen Bibliotheken zweierlei: Erstens Ressourcen und die Einsicht, diese in benötigte Technik, Personal und interne Schulungen zu investieren. Gerade, was KI angeht, haben vermutlich viele Mitarbeiter und Forscher nur ein begrenztes Verständnis der grundlegenden Konzepte und Prinzipien der KI. Weiterhin fehlt es an Wissen darüber, wie KI in ihren spezifischen Forschungsfeldern und für ihre individuellen Bedürfnisse eingesetzt werden kann. Besonders, was KI, und speziell im Zusammenspiel mit Forschungsbewertung angeht, besteht Schulungsbedarf. Zweitens muss es Bibliotheken gelingen, Entscheidungsträgern auf Organisationsebene und darüber hinaus zu vermitteln, wie wichtig und unerlässlich ihr Beitrag zu einer Forschungsbewertung, die KI und Open-Science-Indikatorik verantwortlich und ethisch einsetzt, ist. **I**

5 Literatur

- Berberi, Ilias / Roche, Dominique G. (2022). No evidence that mandatory open data policies increase error correction. *Nature Ecology & Evolution*, 6(11), 1630-1633. <https://doi.org/10.1038/s41559-022-01879-9>
- Kramer, Bianca / Neylon, Cameron / Waltman, Ludo (2024). *Barcelona Declaration on Open Research Information*. <https://barcelona-declaration.org/> [28.08.2024]
- Call to include open science in hiring and promotion policies | Open Science NL*. (o. J.). <https://www.openscience.nl/en/news/call-to-include-open-science-in-hiring-and-promotion-policies> [28.08.2024]
- Dengis, Pascale B., & De Bal, Ils (2022). Monitoring Open Science policy using a regional CRIS – the Flanders case with FRIS. *Procedia Computer Science*, 211, 101-110. <https://doi.org/10.1016/j.procs.2022.10.181>
- Deutsche Forschungsgemeinschaft. (2004). *Empfehlungen zu einer »Leistungsorientierten Mittelvergabe« (LOM) an den Medizinischen Fakultäten*. <https://www.dfg.de/resource/blob/169106/eb4c72d6c6514800b4e2c83cf6e7641b/stellungnahme-klinische-forschung-04-data.pdf> [28.08.2024]
- Freie Universität Berlin. (o. J.). *Instrumente • Team geschlechter:gerecht der Freien Universität Berlin • Freie Universität Berlin*. Abgerufen 26. Juni 2024, von <https://www.fu-berlin.de/sites/frauenbeauftragte/gleichstellung/instrumente/index.html> [28.08.2024]
- Fröhlich, Gerhard (1999). Das Messen des leicht Meßbaren: Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? In J. Becker & W. Göhring (Hrsg.), *Kommunikation statt Markt: Zu einer alternativen Theorie der Informationsgesellschaft* (S. 27-38). GMD- Forschungszentrum Informationstechnik GmbH. <http://eprints.rclis.org/9115/> [28.08.2024]
- Gärtner, Anne (2023). *Open Science as a good practice in doctoral training and shifting incentives for Open Science in Academia*. https://www.bib.uni-mannheim.de/media/Einrichtungen/Universitaetsbibliothek/Dokumente/Lehren_u_Forschen/Open_Science_Day_2023/Open-Science-Day_Gaertner_OSF_4_.pdf [28.08.2024]
- Georg-August-Universität Göttingen. (o. J.). *Leistungsorientierte Mittelvergabe*. Abgerufen 26. Juni 2024, von <https://www.uni-goettingen.de/de/415681.html> [28.08.2024]
- Herb, Ulrich (2016). Impactmessung, Transparenz & Open Science. *Young Information Scientist*, 59-72. <https://doi.org/10.25365/YIS-2016-1-5>
- Herb, Ulrich (2018). Zwangsehen und Bastarde. *Information – Wissenschaft & Praxis*, 69(2–3), 81-88. <https://doi.org/10.1515/iwp-2018-0021>
- Hyrkkänen, Anna-Kaisa/ Ivanović, Dragen / Pölonen, Janne/ Kari, Marita / Pylvänäinen, Elina (2023). *GraspOS Deliverable D2.1 „OS-aware RRA approaches landscape report“*. <https://zenodo.org/doi/10.5281/zenodo.11098095>
- Landesrechnungshof Schleswig-Holstein. (o. J.). *Leistungsorientierte Mittelverteilung an die Hochschulen*. <https://landesrechnungshof-sh.de/file/24leistung.pdf> [28.08.2024]
- Larivière, Vincent / Sugimoto, Cassidy R. (2019). The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects. In Glänzel, Wolfgang/ Moed, Henk F. / Schmoch, Ulrich / Thelwall, Mike (Hrsg.), *Springer Handbook of Science and Technology Indicators* (S. 3-24). Springer International Publishing. https://doi.org/10.1007/978-3-030-02511-3_1
- Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen. (2014). *Leistungsorientierte Mittelverteilung*. https://www.gleichstellung.uni-wuppertal.de/fileadmin/gleichstellung/pdf/Arbeitspapiere_zur_Gleichstellung/15_Leistungsorientierte_Mittelverteilung.pdf [28.08.2024]
- Open Science Indicator Handbook*. (o. J.). <https://handbook.pathos-project.eu/>
- Schmidt, Ruth (2022). *Building Blocks for Impact*. Zenodo. <https://doi.org/10.5281/ZENODO.7249187>

- Schönbrodt, Felix / Gärtner, Anne / Frank, Maximilian / Gollwitzer, Mario / Ihle, Malika / Mischkowski, Dorothee / Phan, Le Vy / Schmitt, Manfred / Scheel, Anne M. / Schubert, Anna-Lena / Steinberg, Ulf / Leising, Daniel (2022). *Responsible Research Assessment I: Implementing DORA for hiring and promotion in psychology*. <https://doi.org/10.23668/PSYCHARCHIVES.8162>
- Schönbrodt, Felix / Gärtner, Anne / Leising, Daniel. (2023). *Responsible Research Assessment*. Open Science Framework. <https://osf.io/https://osf.io/763f4> [28.08.2024]
- Schultz, Volker (2015). *Mittelvergabe nach dem Leistungsprinzip*. *duz – Deutsche Universitätszeitung*. 21.08.2015. <https://www.duz.de/beitrag/!/id/339/mittelvergabe-nach-dem-leistungsprinzip> [28.08.2024]
- TheWall, Mike / Kousha, Kayvan (2021). Researchers' attitudes towards the h-index on Twitter 2007–2020: Criticism and acceptance. *Scientometrics*, 126(6), 5361–5368. <https://doi.org/10.1007/s11192-021-03961-8>
- Universität Osnabrück. (o. J.). *Leistungsbezogene Mittelzuweisung – Universität Osnabrück*. <https://www.uni-osnabrueck.de/universitaet/zahldatenfakten/leistungsbezogene-mittelzuweisung/> [28.08.2024]
- Universität zu Köln. (2021). *Leistungsorientierte Mittelverteilung (LOM)*. https://strategy.uni-koeln.de/strategisches_controlling__informationsmanagement/lom/index_ger.html [28.08.2024]
- Universitätsmedizin Göttingen. (o. J.). *Forschungscontrolling: FACTScience*. <https://www.umg.eu/ueber-uns/vorstand/ressort-forschung-lehre/forschungscontrolling/facts-science/> [28.08.2024]
- Universitätsmedizin Göttingen. (2016). *Leistungsindikator Publikationen*. https://www.umg.eu/fileadmin/Redaktion/Dachportal/006_Ueber_uns/id143_Vorstand/id148_Zentrale_Bereiche/id1044_Forschungscontrolling/id1045_Forschungsevaluation/02_Leistungsindikator_Publikationen.pdf [28.08.2024]
- Universitätsmedizin Göttingen. (2017). *Leistungsindikator Drittmittel*. https://www.umg.eu/fileadmin/Redaktion/Dachportal/006_Ueber_uns/id143_Vorstand/id148_Zentrale_Bereiche/id1044_Forschungscontrolling/id1045_Forschungsevaluation/03_EVA_Leistungsindikator_Drittmittel.pdf [28.08.2024]
- Vergoulis, Thanasis / Kanellos, Ilias / Atzori, Claudio / Mannocci, Andrea / Chatzopoulos, Serafeim / La Bruzzo, Sandro / Manola, Natalia / Manghi, Paolo. (2021). BIP! DB: A Dataset of Impact Measures for Scientific Publications. *Companion Proceedings of the Web Conference 2021*, 456–460. <https://doi.org/10.1145/3442442.3451369>
- Żuradzki, Tomasz / Wroński, Leszek (2024). *How a widely used ranking system ended up with three fake journals in its top 10 philosophy list – Retraction Watch*. <https://retractionwatch.com/2024/06/12/how-a-widely-used-ranking-system-ended-up-with-three-fake-journals-in-its-top-10-philosophy-list/> [28.08.2024]



Dr. Otmane Azeroual

ist Senior Wissenschaftlicher Mitarbeiter am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) in Berlin. Er promovierte in Ingenieurinformatik mit Schwerpunkten auf Datenbanken, Informationssystemen, Künstliche Intelligenz und Data Science. Neben seiner Rolle am DZHW lehrt er an der Fernuniversität Hagen und der University of Europe for Applied Sciences (UE) in Berlin. Parallel verfolgt er seine Habilitation und ist als Chief Technology Officer (CTO) bei der SBI Group AG tätig. Dr. Azeroual leitet internationale Forschungs- und Beratungsprojekte und verbindet akademische Forschung mit praktischer Anwendung.

azeroual@dzhw.eu

ORCID: <https://orcid.org/0000-0002-5225-389X>



Dr. Ulrich Herb

ist Diplom-Soziologe, promovierter Informationswissenschaftler, Leiter der Abteilung Publikations- und Forschungsunterstützung an der Saarländischen Universitäts- und Landesbibliothek, Mitglied in verschiedenen Arbeitsgruppen zu wissenschaftlichem Publizieren, Vorstandsmitglied im Hochschulverband Informationswissenschaft, Gesellschafter von scidecode science consulting (Berlin) und freiberuflicher Berater u.a. für OUROUK Consulting (Paris).

u.herb@sulb.uni-saarland.de

ORCID: <https://orcid.org/0000-0002-3500-3119>



Dr. Joachim Schöpfel

ist Professor der Informationswissenschaften an der Universität Lille (Frankreich) und Wissenschaftler am Gériico-Forschungslabor, Vorstandsmitglied von euroCRIS und freiberuflicher Berater für OUROUK Consulting (Paris) und scidecode science consulting (Berlin).

joachim.schopfel@univ-lille.fr

ORCID: <https://orcid.org/0000-0002-4000-807X>