

DOI: <https://doi.org/10.11588/ip.2021.2.75717>

Grischa Fraumann, Christian Hauschke*, Svantje Lilienthal, Lambert Heller†

Empfehlungen zum Umgang mit szientometrischen Daten und Visualisierungen

Zusammenfassung

Wir beschreiben 10 Empfehlungen für die Entwicklung szientometrischer Visualisierungen und richten uns dabei vor allem an Betreibende szientometrischer Dienste. Die Grundlage dazu bilden Nutzerstudien mit Forschenden verschiedener Disziplinen in Deutschland, welche in einen iterativen Softwareentwicklungsprozess zur Entwicklung einer szientometrischen Visualisierung im Projekt ROSI eingebunden waren.

Schlüsselwörter

Metriken, Indikatoren, Szientometrie, Visualisierungen, Responsible metrics, Nutzerstudien

Recommendations to handling scientometric data and visualizations

Abstract

We describe 10 recommendations for the development of scientometric visualisations and address primarily the operators of scientometric services. The basis for this are user studies with researchers from different disciplines in Germany, who were involved in an iterative software development process for the development of a scientometric visualisation in the ROSI project.

Keywords

metrics, indicators, scientometrics, visualizations, responsible metrics, user studies

Veröffentlichung

31.08.2021 in Informationspraxis Bd. 7, Nr. 2 (2021)

*Mitglied im Editorial Board und Mitherausgeber der Zeitschrift Informationspraxis

†Mitglied im Editorial Board und Mitherausgeber der Zeitschrift Informationspraxis



Inhaltsverzeichnis

1 Einführung	2
2 Grundlagen der Handreichung	3
3 Empfehlungen an Betreiber szientometrischer Dienste	4
Empfehlung I: Gehen Sie auf fachspezifische Unterschiede ein	5
Empfehlung II: Verwenden Sie offene Datenquellen und machen Sie den Datenfluss transparent	5
Empfehlung III: Ermöglichen Sie die Nachnutzung von Daten	5
Empfehlung IV: Verdeutlichen Sie Grenzen der Aussagekraft von Indikatoren . . .	5
Empfehlung V: Beziehen Sie Nutzende so weit wie möglich in die Entwicklung ein	6
Empfehlung VI: Erlauben Sie den Forschenden Einfluss auf die Darstellung von Profilen	6
Empfehlung VII: Stellen Sie weiterführende Informationen bereit	6
Empfehlung VIII: Stellen Sie Ihre Visualisierungen regelmäßig auf den Prüfstand .	7
Empfehlung IX: Gehen Sie verantwortungsvoll mit szientometrischen Visualisierungen um	7
Empfehlung X: Veröffentlichen Sie eine Dokumentation Ihrer Open-Source-Software und bauen Sie Ihre eigene Community auf	7
4 Fazit	7
5 Quellen	8

1 Einführung

Szientometrie bezeichnet die Lehre vom Vermessen der Wissenschaft. Die vermutlich bekannteste Teildisziplin der Szientometrie ist die Bibliometrie, die sich quantitativ mit Publikationen beschäftigt. Szientometrie geht darüber hinaus und umfasst unter anderem auch die quantitative Auseinandersetzung mit Vortragsaktivitäten oder Hochschulabsolventenzahlen. Zusätzlich zum Forschungsoutput werden also auch Personen und Institutionen untersucht (Mingers und Leydesdorff 2015). Szientometrische Datenquellen, Indikatoren und ihre Visualisierungen spielen eine wichtige Rolle in der digitalen Repräsentation von Forschungsergebnissen. Ob Zitationszählung, *h*-Index oder Erwähnungen von Publikationen in sozialen Medien; diese Zahlen und darauf aufbauende grafische Darstellungen finden sich inzwischen in wissenschaftlichen Zeitschriften, Repositorien, Forschungsinformationssystemen und Online-Profilen von Forschenden. Selbst wenn die vermessene Wissenschaft (BBAW 2005) stets in der Wissenschaftsforschung diskutiert wird, wird zu wenig darüber gesprochen, wie Forschende über diese Indikatoren und Visualisierungen denken, welche Anforderungen und Befürchtungen

sie haben und wie diese konkret in Softwareentwicklungs- und Implementierungsprozessen von neuen Diensten umgesetzt werden können. In dieser Handreichung möchten wir genau dazu Empfehlungen geben, die aus qualitativen Erhebungen mit 38 Forschenden aus den Geistes- und Sozialwissenschaften, Ingenieurwissenschaften, Mathematik und Naturwissenschaften in Deutschland stammen. Diese Handreichung richtet sich an all jene, die Dienste zu szientometrischen Daten und Visualisierungen konzipieren, betreiben oder als Teil eigener Informationsdienstleistungen nutzen.

Der Schwerpunkt bei der Entwicklung szientometrischer Visualisierungen im Projekt ROSI liegt auf offen verfügbare Daten. Dabei beziehen wir uns vor allem auf das Konzept von FAIR (findable, accessible, interoperable and reusable) (Wilkinson et al. 2016), welches immer mehr an Bedeutung in der Wissenschaft gewinnt, sobald Forschungsdaten entstehen. Weitere internationale Entwicklungen, auf die im Projekt und in der Folge in diesen Empfehlungen aufgebaut wurde, sind die San Francisco Declaration on Research Assessment (DORA) (Cagan 2013), das Leiden Manifesto for Research Metrics (Hicks et al. 2015) und die Hong Kong Principles for assessing researchers (Moher et al. 2020). Des Weiteren orientieren wir uns an Handreichungen, die im Bereich Bibliotheks- und Informationswissenschaften veröffentlicht wurden, z. B. von der *LIBER Innovative Metrics Working Group* (Slowe & Schwamm 2020).

2 Grundlagen der Handreichung

Das Projekt ROSI (Referenzimplementierung für offene szientometrische Indikatoren) ist aufgeteilt in Softwareentwicklung und Nutzerstudien in einem iterativen Prozess (Hauschke, Cartellieri & Heller 2018). Das Projekt wurde an der TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften von September 2018 bis September 2020 durchgeführt und ist Teil der BMBF-Förderlinie “Quantitative Wissenschaftsforschung”. Auf Seiten der Softwareentwicklung wurde ein Prototyp zur Visualisierung szientometrischer Daten entwickelt: ImpactViz – Open Impact Visualizer. Das Tool steht zu Demonstrationszwecken auf einer Webseite zur Verfügung.¹

Auf der Basis von ImpactViz wurde ein Plugin entwickelt, das auf Artikelebene in OJS (Open Journal Systems, vgl. Willinsky 2005) eingebunden werden kann (Lilienthal 2020). OJS wird als Open Source Software von den Herausgebenden zahlreicher wissenschaftlicher Zeitschriften eingesetzt. Des Weiteren wurde eine Anpassung für das Open-Source-Forschungsinformationssystem VIVO (Conlon *et al.* 2019) erstellt, die den Impact von Artikeln darstellen soll (siehe Abbildung 1).

Ebenfalls im Rahmen des Projektes wurde das *Registry of Scientometric Data Sources*² aufgebaut, das verschiedene szientometrische Datenquellen darstellt, die für szientometrische

¹<https://labs.tib.eu/rosi/prototype/>

²<https://labs.tib.eu/rosi/>

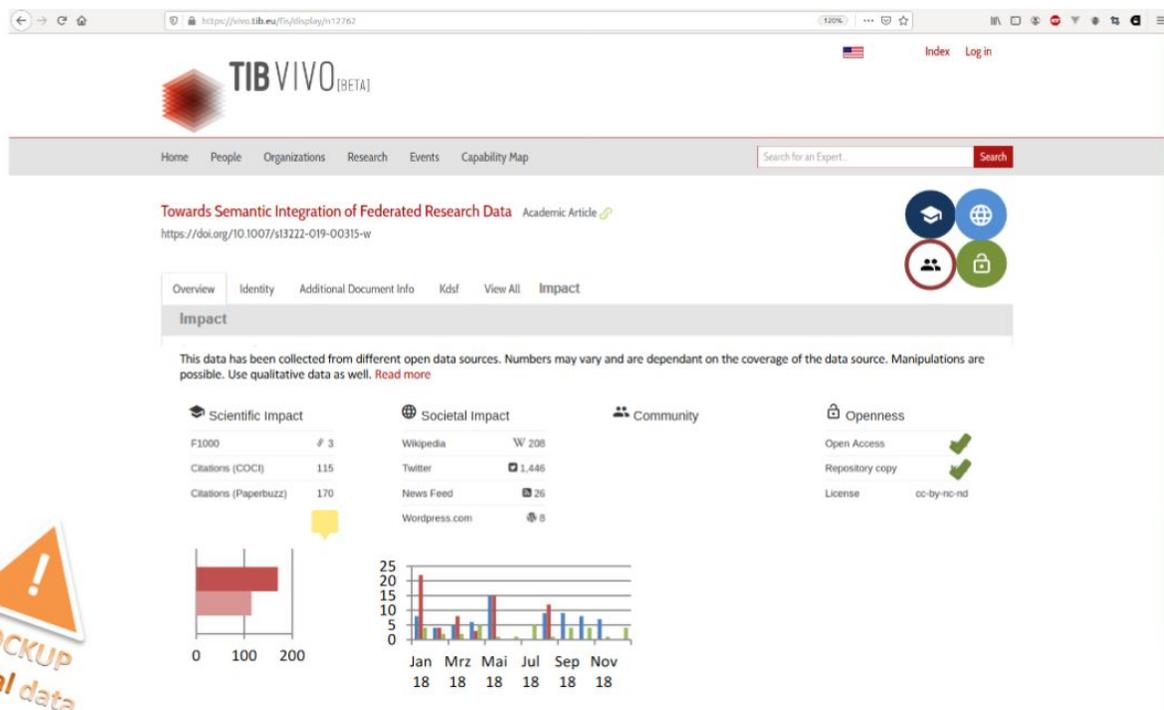


Abb. 1: Mockup einer Implementation im Open-Source-Forschungsinformationssystem VIVO

Anwendungen und Analysen genutzt werden können (Fraumann, Lilienthal & Hauschke, o. J.).

Die Nutzerstudien beinhalteten 16 qualitative Interviews, vier Workshops mit Fokusgruppen bestehend aus 19 Teilnehmenden und die Evaluation der Usability in drei Einzelinterviews.³ Ziel war es, die iterative Softwareentwicklung zu ermöglichen und darüber hinaus Wissenschaftsforschung zu den Bedürfnissen und Anforderungen von Forschenden zu betreiben. Ohne die zahlreichen Teilnehmenden an unser Studie und deren hilfreiches Feedback wäre diese Handreichung nicht möglich gewesen – ihnen gebührt daher unser Dank.

3 Empfehlungen an Betreiber szientometrischer Dienste

Aufbauend auf der oben genannten qualitativen Erhebung (siehe Abschnitt 1 und 2) und den oben erwähnten Prinzipien und Manifesten empfehlen wir die folgenden Punkte für Betreiber szientometrischer Dienste. Selbstverständlich müssen nicht alle Punkte komplett übernommen werden, aber die Empfehlungen können als Grundlage dienen.

³Es ist geplant, eine eingehende Darstellung des empirischen Teils des Projektes zu veröffentlichen.

Empfehlung I: Gehen Sie auf fachspezifische Unterschiede ein

Die Bedürfnisse von Forschenden bezüglich der Darstellung von szientometrischen Indikatoren, Daten und Visualisierungen sind je nach fachlichem Kontext sehr unterschiedlich. Während sich einige gerne mit anderen anhand von Performanz-Indikatoren vergleichen, möchten andere dies umgehen oder halten es in ihrer wissenschaftlichen Disziplin sogar für unmöglich. Die Priorisierung einzelner Indikatoren und die Relevanz verschiedener Datenquellen kann je nach Disziplin sehr unterschiedlich ausfallen. Wenn eine Disziplin z. B. nur in deutschsprachigen Zeitschriften veröffentlicht, sollten entsprechende Datenquellen einbezogen werden. Die Nutzung und Akzeptanz von szientometrischen Visualisierungen kann sich je nach Forschenden und Disziplin auch unterscheiden. Versuchen Sie, die Bedürfnisse Ihrer Nutzerschaft und Ihrer Community zu erheben und nehmen Sie sie ernst.

Empfehlung II: Verwenden Sie offene Datenquellen und machen Sie den Datenfluss transparent

Offene Datenquellen sorgen für Transparenz bei der Visualisierung von szientometrischen Informationen, da die Datenquellen einsehbar, nachnutzbar und für alle Nutzenden frei verfügbar sind. Offene Datenquellen sind z. B. Systeme und Webseiten, die offene Daten zur Verfügung stellen. Somit können Indikatoren anhand der offen verfügbaren Daten durch Nutzende überprüft werden. Die Verwendung offener Datenquellen ermöglicht es zudem die Entstehung von Indikatoren nachvollziehbar zu machen. Dazu können grafische Darstellungen von Datenflüssen hilfreich sein. Idealerweise sollten auch die Algorithmen transparent sein, die zur Berechnung der Daten verwendet werden, was sich leider nicht immer umsetzen lässt.

Empfehlung III: Ermöglichen Sie die Nachnutzung von Daten

Die in szientometrischen Visualisierungen verwendeten Daten sollen transparent und für Analysen – auch durch die Forschenden selbst – nachnutzbar sein. Daher sollten Daten so zur Verfügung gestellt werden, dass sie entweder heruntergeladen werden können, oder dass Nutzende sie anderweitig erhalten können. Durch die Nutzung offener Daten und Lizenzen werden die Möglichkeiten der Nachnutzung eindeutig festgelegt.

Empfehlung IV: Verdeutlichen Sie Grenzen der Aussagekraft von Indikatoren

Szientometrische Visualisierungen werden oftmals für die Evaluierung von Forschungsergebnissen verwendet. Doch quantitative Daten alleine sind für eine angemessene Beurteilung in der Regel nicht hinreichend, eine begleitende qualitative Auswertung ist notwendig. Und auch innerhalb ihres Anwendungsbereiches unterliegen Indikatoren verschiedenen Beschränkungen. Darüber hinaus kann ein Indikator je nach verwendeter Datenquelle unterschiedlich ausfal-

len. Zum Beispiel hängen Zitationszahlen immer davon ab, welche Publikationen von der Datenquelle ausgewertet wurden. Es kann daher von Vorteil sein, diese Zahlen (z. B. Anzahl der Zitationen einer Publikation laut verschiedenen Zitationsdatenbanken) nebeneinander zu stellen. So wird transparent, dass szientometrische Indikatoren keine absolute Wahrheit darstellen, sondern kontextspezifisch zu verstehen sind.

Empfehlung V: Beziehen Sie Nutzende so weit wie möglich in die Entwicklung ein

Feedback kann zum einem über qualitative Interviews, Workshops mit Fokusgruppen oder andere Befragungsmethoden gesammelt werden, zum anderen aber auch informell am Rande von Konferenzen usw. eingeholt werden. Diese Rückmeldungen sollten maßgeblich für die Umsetzung der Visualisierung sein, sofern dies technisch umsetzbar ist. Usability-Tests sind dabei ein entscheidender Schritt, um die Funktion sicher zu stellen und die Akzeptanz der Visualisierung zu erproben (Fraumann & Lilienthal 2020). Der Test bietet auch die Möglichkeit, Nutzende kurz vor der Fertigstellung der Visualisierung nochmals nach allgemeinem Feedback zu fragen.

Empfehlung VI: Erlauben Sie den Forschenden Einfluss auf die Darstellung von Profilen

Auf individueller Ebene unterscheiden sich die Bedürfnisse von Forschenden. Einige Forschende möchten bestimmte Datenquellen aus ihren Forschungsprofilen ausschließen, die andere wiederum unbedingt verwenden möchten. Dazu ist es notwendig, Profile soweit wie möglich anpassbar zu gestalten und Forschenden die Hoheit über ihre eigene szientometrische Repräsentation zu geben (Profilhoheit).

Empfehlung VII: Stellen Sie weiterführende Informationen bereit

Szientometrische Visualisierungen sind oftmals nicht selbsterklärend. Um Missverständnisse zu vermeiden, ist es hilfreich, weiterführende Informationen bereit zu stellen oder auf erläuternde Quellen zu verweisen. Sinnvoll ist zum Beispiel ein Verweis auf die *San Francisco Declaration on Research Assessment* (Cagan 2013), das *Leiden Manifesto for Research Metrics* (Hicks *et al.* 2015) oder die *Hong Kong Principles for assessing researchers* (Moher *et al.* 2020) an der Stelle, wo relevante Informationen für den verantwortungsvollen Umgang mit szientometrischen Indikatoren vermittelt werden.

Empfehlung VIII: Stellen Sie Ihre Visualisierungen regelmäßig auf den Prüfstand

Die Bedürfnisse von Forschenden, Usability-Standards, Datenquellen und szientometrische Indikatoren unterliegen Veränderungen. Stellen Sie sicher, dass Sie regelmäßig überprüfen, ob Sie die hier gegebenen Empfehlungen einhalten, oder ob veränderte Verhältnisse eine Anpassung erfordern.

Empfehlung IX: Gehen Sie verantwortungsvoll mit szientometrischen Visualisierungen um

Im Kontext der Szientometrie ist ethisches und verantwortungsvolles Handeln notwendig. Stellen Sie sicher, dass verschiedene und angemessene Indikatoren verwendet werden. Berücksichtigen Sie, dass Personen aufgrund von Visualisierungen bewertet werden können, auch wenn es nicht Ihr Ziel ist. Machen Sie darauf aufmerksam, dass Datenquellen und Indikatoren manipuliert werden können (Wilsdon *et al.* 2015), und welche Folgen dies für Forschende haben kann.

Empfehlung X: Veröffentlichen Sie eine Dokumentation Ihrer Open-Source-Software und bauen Sie Ihre eigene Community auf

Open-Source-Software bietet Dritten die Möglichkeit zur Anpassung. Dokumentation der Software ist ein entscheidender Punkt, um die Wahrscheinlichkeit der Nachnutzung zu erhöhen. Da diese Software meistens nur von einer aktiven Community getragen wird, sollten neue Entwickler*innen die Möglichkeit haben, alles Wichtige von Anfang an lesen zu können. Außerdem lebt Open-Source-Software von der Community, die sie betreibt. Deshalb ist es sinnvoll, andere Institutionen und Interessenten frühzeitig in die Nachnutzung und Weiterentwicklung mit einzubeziehen.

4 Fazit

Mit dieser Handreichung möchten wir in vereinfachter Form Hinweise für die Entwicklung und den Umgang mit szientometrischen Daten und Visualisierungen in Profilen von Publikationen und Forschenden anbieten. Diese Empfehlungen fassen aus unserer Sicht auf einen Blick die wesentlichen Punkte zusammen, die wir aus Interviews und Workshops mit Forschenden sowie zahlreichen Diskussionen mit Expert*innen zu diesem Thema mitnehmen konnten. Ihre Einhaltung ist aus unserer Sicht notwendig, aber nicht unbedingt hinreichend, und können nur ein Baustein von vielen für einen verantwortungsvollen Umgang mit szientometrischen Indikatoren sein. Die Beschäftigung mit den Wünschen und Befürchtungen der Forschenden

hat sich als äußerst wertvoll erwiesen und es wäre wünschenswert diese Perspektiven in weitere Studien einzubeziehen.

5 Quellen

Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) (2005). *19. Zeit Forum der Wissenschaft "Die vermessene Wissenschaft"*. Verfügbar unter: <https://www.bbaw.de/veranstaltungen/veranstaltung-19-zeit-forum-der-wissenschaftdie-vermessenene-wissenschaft>

Cagan, Ross (2013). The San Francisco Declaration on Research Assessment. *Disease Models & Mechanisms* 6(4), 869–870. <https://doi.org/10.1242/dmm.012955>

Conlon, Michael, Woods, Andrew, Triggs, Graham, O’Flinn, Ralph, Javed, Muhammad, Blake, Jim, Gross, Benjamin, Ijaz Ahmad, Qazi Azim, Ali, Sabih, Barber, Martin, Elsborg, Don, Fofack, Kitio, Hauschke, Christian, Ilik, Violeta, Khan, Huda, Lawless, Ted, Levernier, Jacob, Lowe, Brian, Luis Martin, José, McKay, Steve, Porter, Simon, Walther, Tatiana, White, Marijane, Wolff, Stefan & Younes, Rebecca (2019). VIVO: A System for Research Discovery. *Journal of Open Source Software* 4(39), 1182. <https://doi.org/10.21105/joss.01182>

Fraumann, Grischa & Lilienthal, Svantje (2020). Usability-Tests in der Praxis – ein Werkstattbericht des Projekts ROSI. *TIB-Blog*. Verfügbar unter: <https://blogs.tib.eu/wp/tib/2020/09/15/usability-tests-in-practice-a-progress-report-of-the-rosi-project/>

Fraumann, Grischa, Lilienthal, Svantje & Hauschke, Christian (eingereicht). Towards the Registry of Scientometric Data Sources.

Hauschke, Christian, Cartellieri, Simone & Heller, Lambert (2018). Reference implementation for open scientometric indicators (ROSI). *Research Ideas and Outcomes (RIO)* 4: e31656. <https://doi.org/10.3897/rio.4.e31656>

Hicks, Diana, Wouters, Paul, Waltman, Ludo, de Rijcke, Sarah & Ràfols, Ismael (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520(7548), 429–431. <https://doi.org/10.1038/520429a>

Lilienthal, Svantje (2020). ImpactViz – Ein OJS-Plugin für offene szientometrische Informationen. *TIB Blog*. <https://blogs.tib.eu/wp/tib/2020/09/25/impactviz-ein-ojs-plugin-fuer-offene-szientometrische-informationen/>

Mingers, John & Leydesdorff, Loet (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research* 246(1), 1–19. <https://doi.org/10.1016/j.ejor.2015.04.002>

Moher, David, Bouter, Lex, Kleinert, Sabine, Glasziou, Paul, Sham, Mair Har, Barbour, Virginia, Coriat, Anne-Marie, Foeger, Nicole & Dirnagl, Ulrich (2020). The Hong Kong

Principles for assessing researchers: Fostering research integrity. *PLOS Biology* 18(7): e3000737. <https://doi.org/10.1371/journal.pbio.3000737>

Slowe, Sarah & Schwamm, Hardy (2020). *Why Do Measures Fluctuate? Metrics Report – Guidelines for Talking to Management. Report by LIBER’s Innovative Metrics Working Group.* <https://doi.org/10.5281/zenodo.3902348>

Wilkinson, Mark D., Dumontier, Michel, Aalbersberg, IJsbrand, J., Appleton, Gabrielle, Axton, Myles, Baak, Arie, Blomberg, Niklas, Boiten, Jan-Willem, Bonino da Silva Santos, Luiz, Bourne, Philip E., Bouwman, Jildau, Brookes, Anthony J., Clark, Tim, Crosas, Mercè, Dillo, Ingrid, Dumon, Olivier, Edmunds, Scott, Evelo, Chris T., Finkers, Richard, Gonzalez-Beltran, Alejandra, Gray, Alasdair J.G., Groth, Paul, Goble, Carole, Grethe, Jeffrey S., Heringa, Jaap, 't Hoen, Peter A.C, Hooft, Rob, Kuhn, Tobias, Kok, Ruben, Kok, Joost, Lusher, Scott J., Martone, Maryann E., Mons, Albert, Packer, Abel L., Persson, Bengt, Rocca-Serra, Philippe, Roos, Marco, van Schaik, Rene, Sansone, Susanna-Assunta, Schultes, Erik, Sengstag, Thierry, Slater, Ted, Strawn, George, Swertz, Morris A., Thompson, Mark, van der Lei, Johan, van Mulligen, Erik, Velterop, Jan, Waagmeester, Andra, Wittenburg, Peter, Wolstencroft, Katherine, Zhao, Jun and Mons, Barend (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

Willinsky, John (2005). Open Journal Systems. *Library Hi Tech* 23(4), 504–519. <https://doi.org/10.1108/07378830510636300>

Wilsdon, James, Allen, Liz, Belfiore, Eleonora, Campbell, Philip, Curry, Stephen, Hill, Steven Jones, Richard, Kain, Roger, Kerridge, Simon, Thelwall, Mike, Tinkler, Jane, Viney, Ian, Wouters, Paul, Hill, Jude & Johnson, Ben (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management.* London, UK: HEFCE. <https://doi.org/10.13140/RG.2.1.4929.1363>

AutorInnen

- Grischa Fraumann, gfr@hum.ku.dk

TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften, Welfengarten 1
B, D-30167 Hannover

<https://orcid.org/0000-0003-0099-6509>

- Christian Hauschke⁴, christian.hauschke@tib.eu

TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften, Welfengarten 1
B, D-30167 Hannover

⁴Mitglied im Editorial Board und Mitherausgeber der Zeitschrift Informationspraxis

<https://orcid.org/0000-0003-2499-7741>

- Svantje Lilienthal, svantje.lilienthal@tib.eu

TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften, Welfengarten 1
B, D-30167 Hannover

<https://orcid.org/0000-0003-1537-2862>

- Lambert Heller⁵, lambert.heller@tib.eu

TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften, Welfengarten 1
B, D-30167 Hannover

<https://orcid.org/0000-0003-0232-7085>

⁵Mitglied im Editorial Board und Mitherausgeber der Zeitschrift Informationspraxis