
Joel Herndon (Hrsg.): *Data Science in the Library: Tools and Strategies for Supporting Data-Driven Research and Instruction*. London: Facet, 2022. XXX, 146 Seiten, ISBN 9781783304592 (Paperback).

Besprochen von **Prof. Dr. Ulrike Wuttke:** Fachhochschule Potsdam, Fachbereich Informationswissenschaften, Kiepenheuerallee 5, D-14469 Potsdam, E-Mail: ulrike.wuttke@fh-potsdam.de. <https://orcid.org/0000-0002-8217-4025>

<https://doi.org/10.1515/bfp-2022-0097>

Das Ziel des vorliegenden von Joel Herndon herausgegebenen Sammelbandes *Data Science in the Library: Tools and Strategies for Supporting Data-Driven Research and Instruction* (2022) ist es darzustellen, wie Bibliotheken durch einen nachhaltigen Ansatz, der die Entwicklung dieses wachsenden Bereichs widerspiegelt, konstruktiv mit den Möglichkeiten der sogenannten *Data Science* (dt. Datenwissenschaft) umgehen können. Der Band richtet sich laut der Einleitung hauptsächlich an Bibliothekarinnen und Bibliothekare bzw. Informationsspezialistinnen und -spezialisten, die daran interessiert sind, Data-Science-Dienstleistungen auf- bzw. auszubauen (S. xxix). Im Klappentext werden zusätzlich Forschende, Studierende und Lehrende mit einem Interesse an der Umsetzung von Modellen für Data-Science-Dienstleistungen, die eine Gemeinschaft schaffen und die Disziplin voranbringen wollen, als Zielgruppe benannt.

Der Herausgeber, Joel Herndon, Ph.D., ist momentan Direktor des Center for Data and Visualization Sciences der University Library der Duke University und leitet ebendort ein Data-Science-Programm.

Bei der Vorstellung des Teams der Autorinnen und Autoren des vorliegenden Sammelbands fällt auf, dass einige der Co-Autorinnen und -Autoren nicht vorgestellt werden. Bei den vorgestellten Co-Autorinnen und -Autoren handelt es sich zumeist um eng am Aufbau relevanter Dienstleistungen beteiligte Bibliothekarinnen und Bibliothekare mit einem Wirkungsschwerpunkt in den Vereinigten Staaten. Der Fokus des Sammelbandes liegt hierdurch bedingt auf den Vereinigten Staaten: Sechs der insgesamt acht Kapitel behandeln forschungsintensive und ressourcenstarke Universitäten, zwei haben einen europäischen Fokus und behandeln kleinere Institutionen in Spanien und Dänemark, weitere Kontinente spielen keine Rolle. Darüber hinaus sind viele der Autorinnen und Autorinnen der Initiative *The Carpentries*¹ verbunden, die hierdurch eine gewisse Prominenz in den Darstellungen einnimmt.

¹ <https://carpentries.org/>.

Der Sammelband ist nach dem einleitenden Kapitel „Introduction: The Rise of Data Science“ zielführend in vier Teile mit jeweils zwei Kapiteln gegliedert: Teil 1 „Data Science and Research Libraries – Perspectives“, Teil 2 „Data Science Instruction“, Teil 3 „Data Science Services“ und Teil 4 „Designing and Staffing Data Science“. Teil 1 beinhaltet „Sustainability and Success Models for Informal Data Science Training within Libraries“ von Elizabeth Wickes (Kapitel 1) und „The Fundación Juan March DataLab: A Data Science Unit within a Research Support Library“ von Luis Martínez-Urbe, Paz Fernández and Fernando Martínez (Kapitel 2). Teil 2 beinhaltet „Toward Reproducibility: Academic Libraries and Open Science“ von Joshua Quan (Kapitel 3) und „Start with Data Science“ von Mine Çetinkaya-Rundel (Kapitel 4). Teil 3 beinhaltet „In Support of Data-Intensive Science at the University of Washington“ Jenny Muilenburg (Kapitel 5) und „From a Data Archive to Data Science: Supporting Current Research“ von Tim Dennis, Zhiyuan Yao, Leigh Phan, Kristian Allen, Jamie Jamison, Doug Daniels and Ibraheem Ali (Kapitel 6). Teil 4 rundet das Bild mit den beiden Kapiteln „In-House Training as the First Step to Becoming a Data Savvy Librarian“ von Jeannette Ekström (Kapitel 7) und „Designing for Data Science: Planning for Library Data Services“ (Kapitel 8) von Joel Herndon ab. Bei einigen Kapiteln wäre vielleicht auch eine andere Zuordnung möglich gewesen, insgesamt bietet sich jedoch ein rundes Gesamtbild. Jedes Kapitel hat seine eigene Bibliografie und es gibt einen funktionalen Index, vermisst wird jedoch eine Gesamtbibliografie.

Dem Sammelband ist mit dem Kapitel „Introduction: The Rise of Data Science“ eine Einleitung vorangestellt, in der der Begriff Data Science sowie die Zielsetzung und die Zielgruppe des Werkes sowie die Gliederung näher beschrieben werden. Ansonsten enthält das Kapitel eine Übersicht über alle weiteren Kapitel. In der Einleitung wird unter anderem motiviert, warum sich Bibliotheken in diesem Bereich engagieren sollten und welche Rolle Bibliothekarinnen und Bibliothekare dabei spielen könnten. Es wird deutlich gemacht, dass es im vorliegenden Band – verständlicherweise – vor allem um das Engagement wissenschaftlicher Bibliotheken im Bereich Data Science geht.² Leider wird jedoch Data Science, immerhin ein Kernbegriff des

Sammelbandes, im selbigen nicht eindeutig definiert. Dies ist aus Sicht der Rezensentin eine Schwäche des Werkes, die dazu beitragen kann, dass Lesende, die sich einen Fokus auf professionelle Unterstützung datenintensiver Forschung auf hohem Niveau versprechen, enttäuscht werden. Die beschriebenen Dienstleistungen und Kompetenzen bewegen sich hierdurch relativ undefiniert zwischen Research Data Management, Data Literacy und Data Savviness als Vorstufen zu Data Science bzw. zwischen grundlegenden bzw. basalen bis zu hochkomplexen Data-Science-Methoden und -Anwendungsgebieten, d. h. zwischen (angewandter) Datenwissenschaft im bibliotheks- bzw. informationswissenschaftlichen Kontext und komplexen wissenschaftsgetriebenen Anwendungsgebieten. Während diese Unschärfe den Vorteil hat, dass sich viele Bibliotheken etwas vom sprichwörtlichen „Kuchen abschneiden können“, führt sie jedoch auch dazu, dass nicht immer deutlich ist, inwieweit und warum sich Bibliothekarinnen und Bibliothekare in diesem Bereich engagieren sollten, z. B. weil sie dafür besondere Kompetenzen mitbringen oder weil es gerade angesagt ist? Alles in allem fehlt eine systematische Herausarbeitung der entsprechenden an wissenschaftlichen Bibliotheken (und darüber hinaus) zu verortenden Kompetenzen, an denen diese (zukünftigen) Dienstleistungen anschließen.³

Angesichts der Vielfalt der im Band enthaltenen Kapitel, ist es an dieser Stelle nicht möglich, detailliert auf jedes Einzelkapitel einzugehen, dafür gibt es eine Zusammenfassung im Rahmen der Einleitung sowie Kapitel 8, das die wichtigsten Punkte der vorangegangenen Kapitel wieder aufgreift („reflect on common themes raised by the contributors“, S. xxviii–xxix). Einige Linien sollen hier dennoch zur Sprache kommen.

Insgesamt eröffnet der Sammelband ein interessantes Kaleidoskop an Aspekten des bibliothekarischen Engagements im Bereich Data Science. Die Berichte sind mehrheitlich aus der Praxis gegriffen und zeigen unterschiedliche Entwicklungspfade auf, die für eine breitere Leserschaft interessant sein könnten. Wie bereits erwähnt, hantieren die Autorinnen und Autoren dabei eine unscharfe Definition von Data Science, was zu sehr unterschiedlichen Dienstleistungsspektren, von technologischen Lösungen bis zu didaktischen Interventionen führt. Technologisch avancierter und wegen der europäischen Perspektive lesens-

2 Im englischen Text werden unterschiedliche Begriffe verwendet, u. a. *academic libraries* und *research libraries*. Auch wenn diese Begriffe nicht eins-zu-eins mit dem Begriff wissenschaftliche Bibliothek übersetzbar sind, wird dieser Begriff der Intention als Sammelbegriff für eng an wissenschaftlicher Forschung affilierte Bibliotheken zu dienen momentan am ehesten gerecht. Vgl. die Überlegungen von Ceynowa, Klaus (2018): *Research Library Reloaded? Überlegungen zur Zukunft der geisteswissenschaftlichen Forschungsbibliothek*. In: *ZfBB*, 65 (1), 3–7. <http://dx.doi.org/10.3196/186429501865112>.

3 Vgl. z. B. Directorate-General for Research and Innovation (European Commission), EOSC Executive Board, Natalia Manola, Emma Lazzeri, Michelle Barker, Iryna Kuchma, Vinciane Gaillard, Lennart Stoy (2021): *Digital skills for FAIR and Open Science: Report from the EOSC Executive Board Skills and Training Working Group*. Luxembourg: Publications Office of the European Union: <https://data.europa.eu/doi/10.2777/59065>, insb. 16–22.

wert ist Kapitel 2 (Luis Martínez-Urbe, Paz Fernández und Fernando Martínez), das jedoch drängende ethische Fragen bzgl. der Auswertung von Nutzendendaten offen lässt.

Viele der Autorinnen und Autoren reflektieren über die Auswirkungen dieser Entwicklungen auf das Selbstverständnis bzw. das Berufsbild von Bibliothekarinnen und Bibliothekaren. Diesbezüglich wiederkehrende Punkte sind u. a., inwieweit Bibliothekarinnen und Bibliothekare dazu prädestiniert sind, im Bereich Data Science tätig zu werden, und welche formellen und informellen Aus- und Weiterbildungsbedarfe bestehen. Insbesondere die angesichts des raschen Technologiewandels bestehende Notwendigkeit des sogenannten Lebenslangen Lernens, *lifelong learning*, *upskilling*, *reskilling* (siehe Elizabeth Wickes, Kapitel 1, S. 25), erfordert sowohl auf der Bibliotheksleitungsebene aber auch bei den einzelnen Mitarbeitenden Agilität, Unterstützung, Freiräume, aber auch Motivation, eine gewisse *data savviness* als Teil des professionellen Selbstverständnisses zu betrachten (siehe Kapitel 7, Jeannette Ekstrøm). Diese neuen Anforderungen müssen letztlich auch in Hochschul-Curricula übersetzt werden.⁴

Eine weitere wichtige Rolle spielen erfreulicherweise Fragen der Nachhaltigkeit, die unter verschiedenen Gesichtspunkten beleuchtet werden. Positiv herauszustreichen ist, dass immer wieder herausgestellt wird, dass Data Science kein Spielfeld für sogenannte „unicorns“, sondern „team based“ ist (beide Zitate, S. 5) und das Synergien und Netzwerke innerhalb und außerhalb der Bibliothek, im Fall von Hochschulen campusweit, eine wichtige Rolle spielen.

Eine erstaunlich geringe Rolle kommt im Sammelband den im europäischen Kontext omnipräsenten FAIR-Prinzipien zu, sie tauchen noch nicht einmal im Gesamtindex auf.⁵ Zumindest heben aber einige der Autorinnen und Autoren auf Open-Science-Prinzipien, insbesondere Reproduzierbarkeit (*reproducibility*), ab. Besonders anregend ist diesbezüglich Kapitel 3 von Joshua Quan, das tiefe Einblicke in ein auf datenkritisches Denken ausgerichtetes Universitäts-Curriculum gibt (auch interessant in diesem Kontext ist das andere Kapitel in dieser Sektion von Mine Çetinkaya-Rundel).

Was können die Leserinnen und Leser von den verschiedenen Beiträgen lernen, in deren Fokus mehrheitlich Entwicklungen an gut ausgestatteten amerikanischen Campus-Universitäten stehen? Die generelle Tendenz zur größeren Rolle von Data Science im Forschungskontext (nicht nur als sogenanntes Buzzword) ist nicht zu verleugnen. Die Gretchenfrage ist, wie diese Forschungsaktivitäten am besten unterstützt werden können, rund um den Daten- und Forschungskreislauf und für verschiedene Karrierestufen. Hier bietet der vorliegende Band einige interessante Anknüpfungspunkte. Die Ansiedlung dieser Dienstleistungen im Kontext der Bibliothek könnte laut den Autorinnen und Autoren nämlich den Vorteil bieten, fast automatisch einen interdisziplinären Ausgangspunkt zu bieten (so z. B. „play an essential, discipline-agnostic-role in campus support“, Kapitel 5, Jenny Muilenburg). In den meisten Fällen wird die Bibliothek auch nicht der einzige Akteur sein, denn – wie bereits oben erwähnt – „Data science is a team endeavor“ (Kapitel 8, Joel Herndon), was übrigens auch noch einmal eindrücklich die Rolle von sozialen Kompetenzen für Bibliothekarinnen und Bibliothekare neben fachlichen, technologischen und didaktischen Kompetenzen für diejenigen, die sich in Schulungen und Trainings engagieren wollen, unterstreicht.

Obwohl jedes Kapitel für sich interessante Einsichten bietet, sind der Ton und die Qualität der Beiträge wechselhaft. Einige Beiträge sind aus der Praxis gegriffen und kommen stark erzählend ohne tiefergehende theoretische oder didaktische Reflexionen bzw. Nachweise rezipierter Quellen oder vergleichbarer Entwicklungen daher. Das schmälert die Lektüre und den wissenschaftlichen Wert neben der bereits genannten schwammigen Definition des Begriffs Data Science.

Zu wünschen wäre für die Zukunft, dass sich die Spannung, ob Bibliotheken Data-Science-Aktivitäten als eine Art letzte Zuflucht betrachten sollten, um nicht ihre Relevanz zu verlieren, oder ob es sich hier um eine natürliche Entwicklung handelt, zum letzteren auflöst, denn es ist viel zu tun. Natürlich braucht eine solche Verschiebung der Aktivitäten und Dienstleistungen Zeit und Willen, um eine passende Rollenverteilung zwischen den verschiedenen Akteuren auszubalancieren und auch fortwährend auszutarieren. Hier liegt eine Stärke des Bandes, die Etablierung von Data-Science-Dienstleistungen wird aus unterschiedlichen Perspektiven betrachtet, wie z. B. notwendige organisatorische Wandel und didaktische Ansätze inkl. inklusiver, community-basierter, offener Ansätze wie *The Carpentries*.

Das vorliegende Buch *Data Science in the Library: Tools and Strategies for Supporting Data-Driven Research and Instruction* ist trotz der kritischen Anmerkungen relevant für Bibliothekarinnen und Bibliothekare, auch auf

⁴ Wie diese sich bereits langfristig abzeichnenden Trends in die bibliothekarische Hochschulausbildung integriert werden können, ist u. a. exemplarisch dargelegt in Wuttke, Ulrike, Alexandra Claasen, Antje Michel, Heike Neuroth, Ellen Euler (2022): Bibliothekswissenschaftliche Zukunftsfelder im Kontext der Neuausrichtung des Bachelorstudiengangs Bibliothekswissenschaft der Fachhochschule Potsdam. In: *BIBLIOTHEK – Forschung und Praxis*, 46 (3), 414–21. <https://doi.org/10.1515/bfp-2022-0039>.

⁵ <https://www.go-fair.org/fair-principles/>.

der Leitungsebene, die strategisch mit der Entwicklung von Data-Science-Dienstleistungen befasst sind bzw. sich für entsprechende didaktische Ansätze interessieren. Für neu in die Materie einsteigende Lesende ist es keine besonders geeignete Lektüre, weil Grundlagen vorausgesetzt werden.