

Workshop: „Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science“

Lena Steinmann ¹, Dirk Nowotka ², Lea Oberländer³, Helen Pfuhl⁴, Heiner Stuckenschmidt ⁵ und Rolf Drechsler ⁶

Abstract: In einer zunehmend von Daten geprägten Welt spielt Data Science eine zentrale Rolle, um insbesondere auch hochdimensionale Daten effektiv zu nutzen und daraus wertvolles Wissen zu generieren. Auch das Hochschulsystem hat auf diese Entwicklung reagiert und in den letzten Jahren wurden verschiedene Forschungsstrukturen, wie zum Beispiel Data Science Center, Beratungsstellen, neue Studiengänge sowie Aus- und Weiterbildungsangebote aufgebaut. Der Workshop fördert einen interdisziplinären und standortübergreifenden Austausch zwischen Vertreter:innen verschiedener Data-Science-Initiativen und legt so die Grundlage für eine zukünftige überregionale Zusammenarbeit. Durch den Austausch rund um innovative Ansätze, bewährte Praktiken und Erfahrungen werden Synergien geschaffen und die Weiterentwicklung des Feldes vorangetrieben. Dabei stehen die Herausforderungen und Chancen der datenintensiven Forschung im Fokus der Diskussionen. Der Workshop bietet eine Plattform, um den Dialog zu fördern, Netzwerke zu stärken und Kompetenzen auf nationaler Ebene zu bündeln.

Keywords: Data Science, Forschungsdatenmanagement, Data Literacy, Hochschulen

1 Hintergrund

Data Science wird als Schlüsseldisziplin des 21. Jahrhunderts angesehen und ist ein zentraler Bestandteil der digitalen Transformation. Mit zukunftsweisenden Verfahren wie maschinellem Lernen, können selbst aus komplexen Daten neue Erkenntnisse gewonnen werden, die als Grundlage für datenbasierte Entscheidungen dienen können. Eine

¹ Universität Bremen, Data Science Center, Bibliothekstraße 5, 28359 Bremen, lena.steinmann@uni-bremen.de,  <https://orcid.org/0000-0001-5443-0581>

² Digital Science Center, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, dn@zs.uni-kiel.de,  <https://orcid.org/000-0002-5422-2229>

³ Mannheim Center for Data Science, Universität Mannheim, L5, 68161 Mannheim, lea.oberlaender@uni-mannheim.de

⁴ Digital Science Center, Christian-Albrechts-Universität zu Kiel, Fraunhoferstr. 13, 24118 Kiel, hpfuhl@email.uni-kiel.de

⁵ Mannheim Center for Data Science / Fakultät für Wirtschaftsinformatik und Wirtschaftsmathematik, Universität Mannheim, L5 / B6, 68161 Mannheim, heiner.stuckenschmidt@uni-mannheim.de,  <https://orcid.org/0000-0002-0209-3859>

⁶ Universität Bremen, Data Science Center, Bibliothekstraße 5, 28359 Bremen, drechsler@uni-bremen.de,  <https://orcid.org/0000-0002-9872-1740>

Grundvoraussetzung dafür sind qualitätsgesicherte, maschinenlesbare Daten, die mittels eines nachhaltigen Forschungsdatenmanagements (FDM) gewährleistet werden.

Auch das Wissenschaftssystem hat auf die wachsende Bedeutung von Data Science als interdisziplinäres Wissenschaftsfeld mit vielfältigen Anwendungsfeldern reagiert. Ein ganz wesentlicher Aspekt dabei ist der steigende Bedarf an Datenkompetenzen in Wissenschaft und Wirtschaft, dem durch neue Qualifizierungswege und Studiengänge nachgekommen werden muss, wie unter anderem von der „Expert Group on FAIR“ der Europäischen Kommission [Eu18], dem Stifterverband [Ki18] und dem Rat für Informationsinfrastruktur (RfII) [Ra19] empfohlen. Die Gesellschaft für Informatik e. V. (GI) hat sich ebenfalls seit 2018 mit der Task-Force "Data Science/Data Literacy" und einhergehenden Workshops intensiv mit diesem Thema befasst und entsprechende Empfehlungen erarbeitet [Ge19, Ge23]. Auch der Fakultätentag Informatik e.V. (FTI) widmete im November 2022 einen Workshop speziell dem Thema Data-Science-Studiengänge. Der Wissenschaftsrat formulierte bereits in 2020 acht Leitlinien für die erfolgreiche Umsetzung datenintensiver Forschung und empfiehlt dabei konkret die Errichtung von Data Science Centern als interdisziplinäre Forschungsstrukturen für die datenintensive Wissenschaft, die auch Möglichkeiten zur Weiterbildung und Kompetenzentwicklung bieten [Wi20].

Viele deutsche Hochschulen haben auf diese Entwicklungen und Empfehlungen bereits reagiert, indem sie verschiedene Forschungsstrukturen, wie Data Science Center, aber auch innovative Studiengänge, Qualifizierungsangebote und Weiterbildungsprogramme geschaffen haben. Gleichzeitig wurden bedeutende Fortschritte im Bereich der Dateninfrastrukturen und des FDM erzielt, darunter der Aufbau der European Open Science Cloud (EOSC) [Eu23], von Gaia-X [Ga23] und der Konsortien der Nationalen Forschungsdateninfrastruktur (NFDI) [Na23]. Diese Initiativen ermöglichen einen verbesserten Zugang zu Forschungsdaten und fördern ihre Nachnutzbarkeit. Darüber hinaus wurde im Jahr 2022 mit der Verstetigung von fünf durch den Bund geförderten KI-Kompetenzzentren durch den Bund die KI-Forschung im deutschen Wissenschaftssystem gestärkt [Bu23]. Ab Dezember 2023 plant das Bundesministerium für Bildung und Forschung (BMBF) zudem die Förderung des Aufbaus von Datenkompetenzzentren in der Wissenschaft [Bu23b]. Diese Zentren dienen als Orte des Lernens, Forschens und Vernetzens und tragen zur Erweiterung der datenwissenschaftlichen Kompetenzen bei. Das Data-Science-Ökosystem der deutschen Wissenschaftslandschaft ist somit vielfältig aufgestellt und bietet ein hohes Synergiepotenzial.

2 Motivation und Ziele des Workshops

Der Workshop „Aktuelle Entwicklungen und Perspektiven (an Hochschulen) im Bereich Data Science“ knüpft an die zuvor beschriebenen Entwicklungen an und hat das Ziel, verschiedene Akteure im Bereich Data Science, insbesondere der Data Science Center, zusammenzubringen und auf operativer Ebene zu diskutieren, wie zukünftige

Zusammenarbeit gefördert und das Feld der Data Science im wissenschaftlichen Kontext weiter vorangetrieben werden kann. Er bietet ein Forum für den interdisziplinären Austausch zwischen Vertreter:innen unterschiedlicher Data-Science-Initiativen. Auf Grund der engen Verbindung von Data Science und FDM sind auch Organisationen und Projekte aus diesem Bereich herzlich willkommen. So wird eine nachhaltige Vernetzung innerhalb der Community ermöglicht und der Grundstein für zukünftige Kooperationen gelegt. Es soll ein Überblick über erfolgreiche Konzepte und Umsetzungsstrategien gegeben und über zukünftige Herausforderungen und Chancen des digitalen Zeitalters diskutiert werden. Dabei soll auch die Rolle von Data Science im Informationszeitalter und ihr Potenzial zur Bewältigung gesellschaftlicher Herausforderungen beleuchtet werden.

3 Thematische Schwerpunkte

Der Workshop bietet eine Plattform für die Vorstellung von Aktivitäten, Konzepten, Erfahrungsberichten, Lösungsansätzen und Best Practices aus verschiedenen Data-Science-Initiativen. Dies geschieht in Form von Impulsvorträgen und Postern. Die eingereichten Beiträge sollen zur Diskussion anregen und die Schaffung von Synergien fördern.

Es sind Beiträge unter anderem, aber nicht ausschließlich, zu folgenden Themenfeldern eingeladen:

1. Data Science Einrichtungen an Hochschulen, z. B. Data Science Center:
 - a) Tätigkeitsschwerpunkte und Ausrichtung (z. B. in den Bereichen Forschung, Trainings, Services, Beratung, Vernetzung, technische Infrastruktur)
 - b) Erfahrungen aus Aufbau und Etablierung der Einrichtung
 - c) Rolle von Datenmanagement und der Nationalen Forschungsdateninfrastruktur (NFDI)
 - d) Besondere Herausforderungen und deren Bewältigung
 - e) Zukünftige Entwicklungen
2. Data Science in Lehre und Weiterbildung:
 - a) Neue und innovative Data-Science-Studiengänge
 - b) Qualifizierungsangebote für Forschende unterschiedlicher Statusgruppen und Disziplinen (z. B. Zertifikatskurse, Micro-Degrees, Graduiertenschulen, modulare Angebote)
 - c) Einbindung anderer Data-Literacy-Aspekte (Datenmanagement, Critical Thinking, ethische, rechtliche und soziale Aspekte)

- d) Lerninhalte und -formate
- e) Nachnutzbarkeit von Konzepten (z. B. als Open Educational Resources)
- f) Rolle von BMBF-Datenkompetenzzentren
- g) Besondere Herausforderungen und deren Bewältigung
- h) Zukünftige Entwicklungen

4 Ausrichtende Institutionen

Der Workshop wird gemeinsam von dem Data Science Center der Universität Bremen (DSC@UB), dem Digital Science Center der Universität Kiel (DSC@CAU) und dem Mannheim Center for Data Science der Universität Mannheim (MCDS) ausgerichtet. Im Folgenden werden diese drei Einrichtungen als Good-Practice-Beispiele kurz vorgestellt.

4.1 Data Science Center der Universität Bremen (DSC@UB)

Das Data Science Center (DSC@UB) wurde 2019 mit Unterstützung vom Land Bremen als interdisziplinäres Institut an der Universität Bremen gegründet und dient als Knotenpunkt für die datenintensive Forschung. Die übergeordneten Ziele des DSC sind die wertschöpfende Nutzung von Daten in allen Forschungsbereichen nachhaltig zu ermöglichen und den kulturellen Wandel im Sinne einer „FAIR-Data-Kultur“⁷ voranzutreiben. Dafür werden die Aspekte FDM und Data Science eng verzahnt und abgestimmte Angebote entwickelt. So ist es möglich, Forschende im gesamten Datenlebenszyklus zu unterstützen und eine bestmögliche Wertschöpfung aus Daten zu erzielen [SD21].

Das DSC@UB bietet umfassende Services für datenintensiv Forschende, darunter Rechenkapazitäten wie GPU Server, finanzielle Unterstützung durch „DSC Seed Grants“, Data-Literacy-Trainings, Unterstützung bei der Implementierung von Data-Science-Methoden wie maschinelles Lernen und Datenmanagement-Beratung durch Data Stewards (siehe auch [St23] für eine detaillierte Beschreibung der Aktivitäten und Services). In seinem interdisziplinären Expert:innen-Netzwerk werden Kompetenzen gebündelt, wodurch die Vernetzung und der Transfer von Wissen über fachliche Grenzen hinweg gefördert und disziplinäre Silos aufgebrochen werden. Dies schafft neue Wege für die kooperative Forschung.

Somit repräsentiert das DSC@UB einen wichtigen strukturellen Pfeiler für die datenintensive Forschung an der Universität Bremen und fungiert als Bindeglied zwischen Forschenden und Infrastruktureinrichtungen wie Bibliotheken oder Rechenzentren.

⁷ Gemäß der FAIR-Prinzipien sollen Forschungsdaten auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und nachnutzbar (reusable) sein [Wi16].

Außerdem ist es an der Strategieentwicklung an der Universität Bremen und in der U Bremen Research Alliance - dem 2016 gegründeten Kooperationsnetzwerk der Universität Bremen und zwölf außeruniversitärer Forschungsinstitute - beteiligt. Als Beispiel sei hier ein Whitepaper erwähnt, das den aktuellen Stand sowie zukünftige Handlungsfelder für die disziplinübergreifende Etablierung eines kooperativen FDM skizziert [Pi21].

4.2 Digital Science Center der Universität Kiel (DSC@CAU)

Das Digital Science Center an der Universität Kiel wurde 2022 gegründet mit dem Ziel die digitalen Kompetenzen an der Universität zu vernetzen und Methoden gemeinsam mit den Anwendern weiter zu entwickeln. Projekte und Programme rund um datenintensive Forschung finden hier ein gemeinsames Dach zum Netzwerken und Erfahrungsaustausch, sodass Methoden und Erkenntnisse aus der Informatik und der Mathematik zu den Fachbereichen und Anwendungen gebracht werden können. Die intensive und gemeinschaftliche Weiterentwicklung der Datenwissenschaft wird in die Lage versetzt, einen Mehrwert in der Forschung, in der Lehre und für den Wissens- und Technologietransfer zu schaffen.

Auch die Ausbildung der Fachkräfte von morgen ist ein wichtiges Ziel des Digital Science Centers. Hier werden künftig Expertinnen und Experten ausgebildet, die sowohl Kompetenzen in der Methodik der Datenwissenschaft wie auch in einem gewählten Anwendungsfeld haben. Durch die Verschränkung von Methodik und Anwendung in Fachdisziplinen aller Fakultäten werden die Potenziale der Datenwissenschaften in die gesamte Universität und darüber hinaus getragen. Studiengänge nach dem Prinzip der 2-Fächer-Kombination von Datenwissenschaften und Anwendungsfach mit einer Profilierung im Bereich Künstlicher Intelligenz, werden ergänzt durch Angebote für Ausgründungen und Start-ups mit einem KI-Fokus, genauso wie für Unternehmen mit einem Interesse an KI-Anwendungen.

4.3 Mannheim Center for Data Science der Universität Mannheim (MCDS)

Das Mannheim Center for Data Science (MCDS) bündelt die Expertise der Universität Mannheim in den Wirtschafts-, Sozial- und Verhaltenswissenschaften. Es wurde mit Beschluss des Senats vom 4. März 2020 als Institut der Universität Mannheim eingerichtet. Mitglieder des Gründungsdirektoriums sind Prof. Dr. Frauke Kreuter, Prof. Dr. Florian Stahl, Prof. Dr. Heiner Stuckenschmidt und Prof. Dr. Jutta Mata. Die Profilprofessur „Data Science in den Wirtschafts- und Sozialwissenschaften“, die Prof. Dr. Markus Strohmaier seit Januar 2022 innehat, kooperiert ebenfalls eng mit dem MCDS.

Die Mission des MCDS ist es, ein besseres Verständnis für Daten und ihren Wert für die Erforschung der digitalen Transformation von Individuum, Wirtschaft und Gesellschaft

zu schaffen. Dabei müssen Regeln für den verantwortungsvollen Umgang mit Daten und Analysemethoden zum Nutzen der Gesellschaft identifiziert werden. Die Ermöglichung innovativer, interdisziplinärer Forschung, die gesellschaftliche Prozesse und globale Herausforderungen in den Blick nimmt, die auch für lokale Institutionen und Unternehmen relevant sind, ist ein weiteres Anliegen des MCDS.

Das MCDS versteht sich als Plattform zur Vernetzung und Stärkung der Forschung zu und mit Data Science an der Universität Mannheim und darüber hinaus. In dieser Funktion organisiert es Lehrveranstaltungen, Vernetzungsevents und Vorträge und leistet einen Beitrag zur Wissenschaftskommunikation, indem es regelmäßig über neue Projekte und Fortschritte im Bereich Data Science an der Universität Mannheim informiert. Es ist dabei eng in den Struktur- und Entwicklungsplan der Universität eingebunden, der die Stärkung von Data Science in Lehre und Forschung verankert.

5 Workshop-Beiträge

Der Workshop umfasst drei geladene Keynotes zu den Themenbereichen (1) Data Science Einrichtungen an Hochschulen (Prof. Dr. Philipp Wieder), (2) Data Science in der Lehre (Prof. Dr. Peer Kröger) und (3) Data Science und Datenmanagement / NFDI (Dr. Sonja Schimmler). Die dazugehörigen Sessions werden mit thematisch passenden Impulsvorträgen ausgestaltet. Für den weiteren Austausch wird es eine Poster-Session geben.

Insgesamt wurden acht Beiträge zur Publikation und weitere Beiträge zur Beteiligung ohne Veröffentlichung eingereicht. Dabei ist eine Vielzahl von Data-Science-Einrichtungen vertreten, die wir hier in alphabetischer Reihenfolge auflisten:

- Bielefeld Center for Data Science, Universität Bielefeld
- Campus-Institut Data Science, Georg-August-Universität Göttingen
- Center for Interdisciplinary Digital Sciences, Technische Universität Dresden
- Data Science Center, Universität Bremen
- Digital Science Center, Universität Kiel
- House of Computing and Data Science, Universität Hamburg
- Mannheim Center for Data Science, Universität Mannheim
- Munich Data Science Institute, Technische Universität München

Darüber hinaus sind vielfältige Beiträge aus den Bereichen Lehre und Weiterbildung dabei, z. B. zu der BERD (*Business, Economic, and Related Data*) Academy, dem Projekt DALIA (*Data Literacy Alliance*), der Helmholtz School for Marine Data Science (*MarDATA*), dem Mannheim Master in Social Sciences, dem Zertifikatskurs

„Forschungsdatenmanagement für Studierende“, der NFDI4Earth Academy, dem Master of Science „Data Science“ der FU Berlin und der interdisziplinären Promovierenden-Ausbildung „Data Train – Training in Forschungsdatenmanagement und Data Science“ der U Bremen Research Alliance. Auch aus der Schnittmenge zwischen Data Science und Datenmanagement sind mit NFDI4DataScience, dem Projekt „Forschungsdatenmanagement an Hochschulen für angewandte Wissenschaften in Rheinland-Pfalz“ (FDM@HAW.rlp) und dem Projekt „institutionalisiertes und nachhaltiges Forschungsdatenmanagement in Brandenburg“ (IN-FDM-BB) Initiativen vertreten.

6 Literaturverzeichnis

- [Eu18] European Commission, Directorate-General for Research and Innovation, Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data, Brüssel, 2018.
- [Ki18] Kirchherr, J. et al.: Future Skills: Welche Kompetenzen in Deutschland fehlen, Stifterverband für die Deutsche Wissenschaft, Essen, 2018.
- [Ra19] Rat für Informationsinfrastrukturen: Digitale Kompetenzen - dringend gesucht! Empfehlungen zu Berufs- und Ausbildungsperspektiven für den Arbeitsmarkt Wissenschaft, Göttingen, 2019.
- [Ge19] Gesellschaft für Informatik e.V. (GI), Data Science: Lern- und Ausbildungsinhalte, Berlin, 2019.
- [Ge23] Gesellschaft für Informatik e.V. (GI), <https://gi.de/datascience>, Stand: 14.06.2023.
- [Wi20] Wissenschaftsrat: Zum Wandel in den Wissenschaften durch datenintensive Forschung, Köln, 2020.
- [Eu3] European Open Science Cloud (EOSC), <https://eosc-portal.eu/>, Stand: 14.06.2023.
- [Ga23] Gaia-X, <https://gaia-x.eu/>, Stand: 14.06.2023.
- [Na23] Nationale Forschungsdateninfrastruktur (NFDI), <https://www.nfdi.de/>, Stand: 14.06.2023.
- [Bu23a] Bundesministerium für Bildung und Forschung (BMBF), <https://www.bmbf.de/bmbf/shareddocs/kurzmeldungen/de/2022/06/50-millionen-foerderung-fuer-ki-kompetenzzentren.html>, Stand: 14.06.2023.
- [Bu23b] Bundesministerium für Bildung und Forschung (BMBF), https://www.bildung-forschung.digital/digitalezukunft/de/wissen/Datenkompetenzen/datenkompetenzzentren_f%C3%BCr_die_wissenschaft_ordner/datenkompetenzzentren_fuer_die_wissenschaft_node.html, Stand: 14.06.2023.
- [Wi16] Wilkinson, M.D. et al.: The FAIR Guiding Principles for scientific data management and stewardship, Scientific Data 3, 2016. DOI: 10.1038/sdata.2016.18.

- [SD21] Steinmann, L., Drechsler, R.: Verzahnung von Data Stewardship und Data Science – Wege und Perspektiven, Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern Nr. 3, S. 82-91, 2021. DOI: 10.17192/bfdm.2021.3.8342.
- [St23] Steinmann, L. et al.: Das Data Science Center an der Universität Bremen: Interdisziplinärer Knotenpunkt und Service-Infrastruktur für die datenintensive Forschung, in: Heuveline, Vincent und Bisheh, Nina (Hrsg.): E-Science-Tage 2023: Share Your Research Data, Heidelberg: heiBOOKS, (in review).
- [Pi21] Pigeot I. et al.: Etablierung eines kooperativen Forschungsdatenmanagements in der U Bremen Research Alliance, Zenodo, 2021, DOI: 10.5281/zenodo.4775371.