





Das Konzept des House of Computing and Data Science als Blaupause zur Digitalisierung der Forschung

Chris Biemann ¹, Tilo Böhmman², Sebastian Gerling ³, Jan Louis ⁴ und Martin Semmann ^{5 6}

Abstract: Forschung auf internationalem Niveau kommt nicht mehr ohne die strukturelle Nutzung der Potenziale der Digitalisierung, dabei insbesondere in den Bereichen Data Science und Künstlicher Intelligenz, aus. Gleichzeitig ist der Brückenschlag in die wissenschaftliche Anwendung aufgrund sehr heterogener Fächerkulturen, Vorerfahrungen und der hohen Methodenvielfalt nicht trivial. Ein vielversprechender Ansatz, mit diesen Herausforderungen umzugehen, ist das Konzept des House of Computing and Data Science. Als zentrale Einrichtung wird das Ziel eines reziproken, gemeinsamen Erkenntnisgewinns und Ausbau von Methodenkompetenz verfolgt. Dabei werden Impulse in die gesamte Universität gegeben, gleichzeitig aber auch Erfahrungen und Kompetenzen aus den dezentralen Bereichen gebündelt.

Keywords: Data Science, Künstliche Intelligenz, Digitalisierung, Organisationsstruktur, Research Software Engineering.


1 Einleitung


Die Digitalisierung ist ein zentraler Treiber von Veränderungen für Unternehmen, die Zivilgesellschaft, aber auch die Forschung. Damit einher gehen sich verändernde Anforderungen, um erfolgreiche, sichtbare und nicht zuletzt auch wirkungsvolle Forschung durchzuführen. So sind mittlerweile bspw. über 2/3 der Forschenden in ihrer Arbeit abhängig von Forschungssoftware [He18]. Zudem werden Forschungsaktivitäten zunehmend kollaborativer innerhalb der Disziplinen, was ebenfalls zu einer Dynamisierung von Wissenschaft geführt hat und immer noch führt [WJU07]. Die Digitalisierung erfordert vor dem Hintergrund heterogener Disziplinen eine Anpassung in

¹ Universität Hamburg, HCDS & Fachbereich Informatik, Vogt-Kölln-Str. 30, 22527 Hamburg,


chris.biemann@uni-hamburg.de,  <https://orcid.org/0000-0002-8449-9624>

² Universität Hamburg, Fachbereich Informatik, Vogt-Kölln-Str. 30, 22527 Hamburg, tilo.boehmann@uni-hamburg.de

³ Universität Hamburg, CDO, Mittelweg 177, 20148 Hamburg, sebastian.gerling@uni-hamburg.de,  <https://orcid.org/0000-0001-6047-3481>

⁴ Universität Hamburg, Fachbereich Physik, Luruper Chaussee 149, 22761 Hamburg, jan.louis@uni-hamburg.de  <https://orcid.org/0000-0001-7254-6641>

⁵ Universität Hamburg, HCDS, Vogt-Kölln-Str. 30, 22527 Hamburg, martin.semmann@uni-hamburg.de,

 <https://orcid.org/0000-0001-5316-3696>

⁶ Die Reihung der Autoren erfolgt alphabetisch.

drei zentralen Bereichen.

1. Die Adaption und Nutzung digitaler Methoden in den Disziplinen
2. Die konsequente Weiterentwicklung bestehender Methoden sowie Forschung an neuen Methoden, insbesondere auch durch Impulse aus der Anwendung
3. Der kritische Diskurs der Folgen von Digitalisierung auf die einzelnen Fachdisziplinen, Forschung generell, aber auch auf Gesellschaften

Dieses Spannungsfeld auszutarieren stellt eine Herausforderung, vor allem für Volluniversitäten dar, da ein gemeinsames Handeln und miteinander lernen erforderlich ist. Dies ist insbesondere der Breite der potenziellen Anwendungsdisziplinen und den damit einhergehenden umfassenden methodischen Herausforderungen und Anpassungsbedarfen deutlich. Neben solchen relevanten weichen Faktoren bedarf es aber auch eines Veränderungsprozesses in der Wahrnehmung digitaler Infrastrukturen [KMY21]. Wissenschaftliche Software stellt einen Treiber und Ermöglicher von Erkenntnisgewinn dar, sollte daher auch entsprechend als relevanter Forschungsbeitrag beurteilt werden [a]. Dieser Beitrag manifestiert sich in der technologischen Unterstützung von Verfahren innerhalb der Forschungskontexte, wodurch sowohl Effizienzgewinne, aber auch eine grundsätzliche Durchführbarkeit von Forschungstätigkeiten ermöglicht wird. So könnten bspw. durch automatisierte Analysemöglichkeiten großer Textkorpora Veränderungen im Zeitverlauf beforscht werden, die ohne die entsprechenden Technologien nicht bzw. nicht in diesem Umfang möglich wären. Zudem gelten insbesondere Anforderungen des wissenschaftlichen Arbeitens auch für wissenschaftliche Software [AIRe13, Ch22, Ha20, HePa15]. Die dazu nötige Professionalisierung der Entwicklung von Forschungssoftware kann allerdings nicht in allen Disziplinen gleichermaßen durch Promovierende, die die Treiber der (Weiter-)Entwicklung von Forschungssoftware sind, gesichert werden. Dazu bedarf es eines Brückenschlags in die Informatik, als (auch) ingenieurwissenschaftlich geprägte Fachdisziplin, die gerade in der systematischen Entwicklung von Software einen starken Ankerpunkt hat.

2 Der Lösungsansatz: Das House of Computing and Data Science

Ausgehend von diesen Rahmenbedingungen, wurde das House of Computing and Data Science (HCDS) an der Universität Hamburg Ende 2021 als eine zentrale Einrichtung der Hochschule gegründet. Die Initialisierungsphase wurde mit insgesamt knapp 8 VZÄ bestritten, die neben der wissenschaftlichen Leitung, der Geschäftsführung auch (Senior) Softwareentwickler und die Koordination von Digitalkompetenzen umfassen. Das HCDS dient als Blaupause für den Ausbau, Weiterentwicklung und Transfer digitaler Methodenkompetenz (Ebene 1) und gleichermaßen für interdisziplinäre Forschung mit deutlichem Beitrag auch in den Bereichen Data Science und künstlicher Intelligenz (Ebene 2). Neben diesen fachlichen Schwerpunkten, zielt das Leitbild des House of Computing and Data Science auf eine Open Science-Kultur ab, wobei parallel insbesondere die

Grundsätze einer professionellen Softwareentwicklung, im Sinne eines wissenschaftlichen Artefaktes, dem Aufbau und der Pflege einer Community von Forschenden mit Affinität für digitale Methoden und Digitalisierung, aber auch der Weiterbildung in den Disziplinen [Co21, HePa15].

Das Konzept des HCDS verbindet diese Aspekte über die Ebenenstruktur und übergreifende Elemente. Aufgrund der Etablierung als eine zentrale Hochschuleinheit, stellt eine wichtige Rolle die Vernetzung und Etablierung als erster Kontaktpunkt dar. Dies geschieht auf Ebene 1, dem Methodenkompetenzzentrum. Hier werden Anfragen aus der Breite der Hochschule aufgenommen, Beratungsformate etabliert und bei Bedarf Weitervermittlung zu spezifischeren Expertisen. Damit wird ein niedrighschwelliges Angebot realisiert, dass es Forschenden ermöglicht Unterstützung in der Digitalisierung und Data Science zu erhalten. Dabei wird die Digitalisierung als Treiber exzellenter Forschung sehr umfassend verstanden, weshalb verschiedene Kompetenzprofile innerhalb des HCDS vorgehalten werden müssen. Zentrale Anforderung ist dabei eine ausgeprägte Kompetenz in der Informatik, insb. Methoden und Technologien der künstlichen Intelligenz und Digitalisierung, gleichzeitig aber auch einschlägige Erfahrungen in weiteren Fachdisziplinen. Hiermit kann ein schneller Transfer und Adaption von digitalen Innovationen gesichert werden. Insbesondere Anfragen, die Standardisierungspotentiale bieten oder mit bereits bestehenden Methoden und Lösungen realisiert werden können, werden durch Mitarbeitende der zentralen Einheit direkt umgesetzt. Beispielsweise zeigen sich bereits Überschneidungen bei den Anforderungen aus den Geisteswissenschaften. Hier bietet sich die Chance eine Toolsuite zu entwickeln, die insbesondere für Zeitreihenanalysen verschiedener Textkorpora dient. Eine Weiterentwicklung kann dabei dann zu potenziellen Mehrwerten bei allen Anwendenden führen. Komplexere Anfragen, die auch ein Potential für gemeinsame Forschungstätigkeiten bergen, werden auf Ebene 2 weiter spezifiziert. Hierfür wurde ein Förderprogramm initiiert, dass sogenannte Cross-Disciplinary Labs (CDL) fördert. Zielsetzung ist, interdisziplinäre Verbünde zu fördern, die durch die Kooperation wissenschaftliche Beiträge in den beteiligten Disziplinen erarbeiten. Also Forschungsvorhaben, die als Tandem zwischen Informatik und Anwendung positioniert werden. Hiermit kann die interdisziplinäre Kooperation initiiert und ausgeweitet werden, um eine Stärkung der digitalen Methodenwissenschaft und der Fachwissenschaft zu erreichen.

Das Scharnier der beiden Ebenen bildet der systematische Austausch von Methodenwissen, damit einerseits Wissen über die Einführung und Adaption in Fachwissenschaften vorgehalten und nachgenutzt werden kann. Gleichmaßen werden aber auch methodische Weiterentwicklungen aus den CDLs zurückbezogen und für weitere Nutzungskontexte aufbereitet.

Für das 2-Ebenenmodell des House of Computing and Data Science stehen mittel- und langfristige Forschungsk Kooperationen im Vordergrund, wobei als Ankerdisziplin Informatik bzw. informatische Methoden dienen. Solche interdisziplinären Verbünde bieten ein großes Potential zur Digitalisierung in der Breite einer Volluniversität beizutragen, da unterschiedliche Disziplinen mit ihren spezifischen Anforderungen und

Methoden, aber insbesondere auch Vorerfahrungen gezielt unterstützt werden können. Gleichzeitig können Impulse in die Weiterentwicklung von Methoden der Informatik geleistet werden.

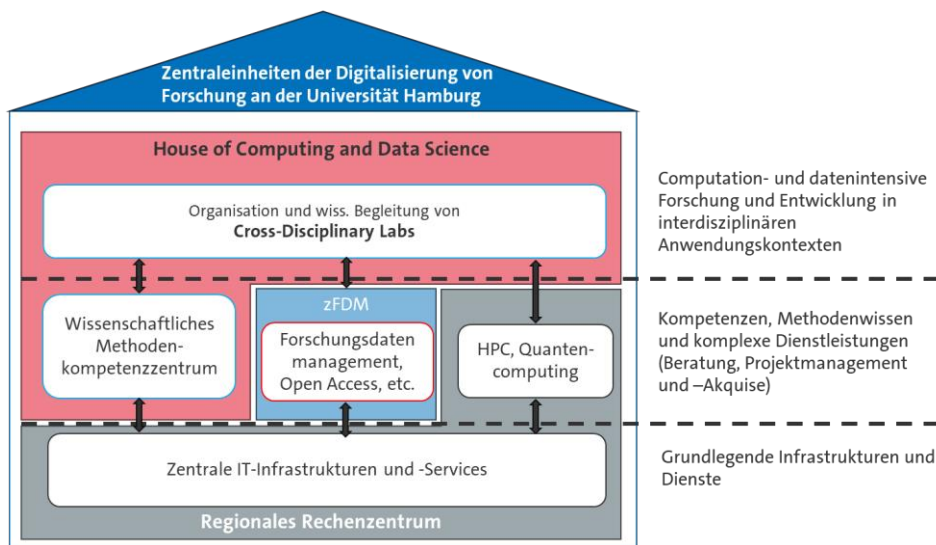


Abb. 1: 2-Ebenenmodell des House of Computing and Data Science

Über beide Ebenen hinaus besteht der generelle Bedarf nach Austausch, Professionalisierung der Softwareentwicklung im Forschungskontext und nach Bildungsangeboten. Hier kann das House of Computing and Data Science durch die Positionierung als zentrale Einheit, verortet zwischen Chief Digital Officer und Vizepräsident für Forschung, sowohl aus sich heraus, als auch aus den an Verbänden beteiligten Disziplinen ein Angebot schaffen, dass in die Breite der Hochschule wirkt und spezifische Lücken der Curricula schließt. Konkret werden bereits vereinzelt Lehrveranstaltungen positioniert, es ist aber angestrebt, dies perspektivisch und unter der Bereitstellung weiterer Ressourcen systematisch auszubauen. Dies sorgt in Konsequenz zu weiteren Potentialen für die Bildung, Forschung, aber auch den generellen Transfer in die Gesellschaft [Ho23, Pi22].

3 Beispielhafte Anwendungsfälle

Analog zur Bandbreite der Fachdisziplinen, ist das Portfolio an Unterstützungsleistungen auch umfassend und je in, sofern kapazitär möglich, maßgeschneiderter Form. Zur

Verdeutlichung werden drei beispielhafte Anwendungsfälle vorgestellt.

Die Beratung (1) durch das wissenschaftliche Methodenkompetenzzentrum ist ein niedrigschwelliges Angebot das zunächst mit einer fokussierten Klärung des Forschungsgegenstandes und Reflektion des methodischen Status quo beginnt. Dabei werden die Gespräche von Mitarbeitenden des HCDS geführt, die einschlägige, relevante Vorerfahrungen bzw. potenziell übertragbare Vorerfahrungen haben. Basierend auf den initialen Informationen können dann Potenziale digitaler Methoden bzw. vorhandener Lösungen vorgestellt werden. Ziel ist es, effizient zu einer hinreichend nutzbaren und nutzenstiftenden Lösung zu kommen. Gleichmaßen können aber nicht alle Anfragen in diesem Schritt gelöst werden. Insbesondere komplexere, zeitkritische oder entwicklungsintensive Vorhaben werden dann systematisch dahingehend geprüft, ob eine co-Finanzierung oder Akquise von Drittmitteln realistisch und zweckmäßig ist. Es resultieren Projekte (2), die als klassische Forschungsprojekte kooperativ initiiert und bearbeitet werden. Beispielhaft kann die Kooperation im Rahmen des D-WISE Projektes zur digitalen Wissenssoziologischen Diskursanalyse benannt werden [Fi23]. Abschließend gibt es das Konstrukt der Cross-Disciplinary Labs (3) als interdisziplinärer Rahmen für langfristige Kooperationen mit dem Ziel sowohl einen wissenschaftlichen Beitrag in der Methodenwissenschaft, aber auch in der Anwendungsdisziplin zu schaffen. Zeitgleich erfolgt ein systematischer Rückbezug auf das Methodenkompetenzzentrum, um die Resultate nachzuhalten. Dadurch wird sichergestellt, dass die involvierten Parteien jeweils forschend eingebunden sind und keine Anbieter-Kunde-Relation entsteht.

4 Fazit

Das vorgestellte Konzept des HCDS stellt eine gut übertragbare Blaupause dar, um dem steigenden Bedarf nach digital- und Data Science-Kompetenzen in einer Volluniversität umzusetzen. Die besondere Herausforderung sind dabei die stark dezentrale Organisation und gleichzeitig heterogenen Vorerfahrungen. Durch die zentrale Positionierung des HCDS und dem Anerkennen und Ausbauen der dezentralen Kompetenzen, kann eine Digitalisierung der Forschung zügig gesichert bzw. initiiert werden. Ziel ist damit perspektivisch ein sich erweiterndes Portfolio an Methoden und Tools bereitzustellen, die flexibel in verschiedene Anwendungskontexte übertragen werden können. Um dies zu sichern ist ein weiterer Ausbau der Schnittstellen zwischen der zentralen Einheit und dezentralen Einheiten zu Themen der Digitalisierung, aber auch dezentralen Forschungseinheiten angestrebt. Zudem sollte perspektivisch ein Angebot entwickelt werden, dass auch Master- bzw. Promotionsniveau Lehrangebote schafft, die systematisch eine data- und digital-literacy unterstützen.

Referenzen

[AIRe13] Alden, K.; Read, M.: Scientific software needs quality control. Nature 448/2013, S. 448–

448, 2013.

- [Co21] Cohen, J. et al: The Four Pillars of Research Software Engineering. IEEE Software 1/2021, S. 97–105, 2021.
- [Ch22] Chue Hong, N. et al.: FAIR Principles for Research Software (FAIR4RS Principles). <https://doi.org/10.15497/RDA00068>, Stand Version 1 24.05.2022, 2022.
- [Fi23] Fischer, T. et al.: D-WISE - Digitale Wissenssoziologische Diskursanalyse. In (Busch, A.; Trilcke, P. Hrsg.): DHd 2023 Open Humanities Open Culture. 9. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum", Trier, S. 388-389, 2023.
- [Ha20] Hasselbring, W. et al.: From FAIR research data toward FAIR and open research software. it - Information Technology 1/2020, , S. 39–47, 2020.
- [Ho23] Horsfall, D. et al.: Research software engineering accelerates the translation of biomedical research for health. Nature Medicine, 6/2023, S. 1313–1316, 2023.
- [HePa15] Hey, Tony ; Payne, Mike C.: Open science decoded. Nature Physics 5/2015, S. 367–369, 2015.
- [He18] Hettrick, Simon: Software in Research Survey. <https://doi.org/10.5281/zenodo.1183562>, Stand Version 1, 23.02.2018.
- [Ka21] Katz, D. et al.: Recognizing the value of software: a software citation guide, F1000Research 9/2021, 2021.
- [KMY21] Knowles, R.; Mateen, B.; Yehudi, Y.: We need to talk about the lack of investment in digital research infrastructure. Nature Computational Science 3/2021, S. 169–171, 2021.
- [Pi22] Pierce, M. et al.: The science gateway community institute’s consulting services program: Lessons for research software engineering organizations. arXiv preprint, arXiv:2209.03958, 2022.
- [WJU07] Wuchty, Stefan ; Jones, Benjamin F. ; Uzzi, Brian: The Increasing Dominance of Teams in Production of Knowledge. Science 5827/2007, S. 1036–1039, 2007.