

Laura Döring, Stefan Kellendonk, Marina Lemaire, Katrin Moeller, Stefan Büdenbender

Umfragebericht zur Data Literacy Bedarfserhebung für die
historisch arbeitenden Disziplinen

Autor:innenliste:

Auflistung der Autor:innen (gemäß [Credit-Taxonomie](#)) inkl. ORCID

Laura Döring | [Conceptualization](#), [Data curation](#), [Formal Analysis](#), [Methodology](#), [Validation](#), [Visualization](#), [Writing – original draft](#), [Writing – review & editing](#) | ORCID: 0009-0001-7129-2018

Stefan Kellendonk | [Data curation](#), [Formal Analysis](#), [Validation](#), [Visualization](#), [Writing – original draft](#), [Writing – review & editing](#) | ORCID: 0000-0003-1009-3311

Marina Lemaire | [Conceptualization](#), [Data curation](#), [Formal Analysis](#), [Methodology](#), [Validation](#), [Visualization](#), [Writing – original draft](#), [Writing – review & editing](#) | ORCID: 0000-0003-4726-2481

Katrin Moeller | [Formal Analysis](#) | ORCID: 0000-0003-4090-5667

Stefan Büdenbender | [Writing – original draft](#), [Writing – review & editing](#) | ORCID: 0000-0002-0478-0396

[Writing – review & editing](#): *Philipp Benra* | ORCID: 0009-0007-3572-7647

Susanne Buchmeier-Zisel

Ursula Lehmkuhl | ORCID: 0000-0003-3100-8809

Stefan Schmunk | ORCID: 0000-0001-9706-9757

im Namen des NFDI4Memory-Konsortiums.

Zusammenfassung

Im NFDI4Memory-Konsortium beschäftigt sich die Task Area „Data Literacy“ mit dem Thema Datenkompetenzen in den historisch arbeitenden Disziplinen. Ziel ist die Unterstützung von Forschenden und Lehrenden in allen Qualifikationsstufen durch die Entwicklung spezialisierter Angebote. Die in diesem Kontext durchgeführte Bedarfserhebung untersuchte die Selbsteinschätzung der Datenkompetenzen verschiedener Statusgruppen. Die Ergebnisse zeigen, dass die durchschnittlichen Zustimmungswerte in den sechzehn Data-Literacy-Kompetenzkategorien auf einer Likert-Skala von 1–5 zwischen 2,3 und 3,5 liegen. FDM-Mitarbeiter:innen mit Lehrtätigkeit bewerten sich in den meisten Data-Literacy-Kategorien am höchsten, wohingegen Bachelorstudierende sich am niedrigsten einschätzen. PostDocs, Professor:innen und Doktorand:innen liegen gemeinsam sehr nah am Durchschnittswert. Die Ergebnisse wurden nach Wissensleveln ausgewertet, wobei viele Bachelorstudierende als Lai:innen und FDM-Mitarbeiter:innen mit Lehre als Expert:innen eingestuft wurden. Die Verteilung der Wissenslevel ist über die Statusgruppen hinweg weitgehend ausgeglichen. Der Bericht bietet eine erste, beschreibende Sicht auf die Ergebnisse der Umfrage und zeigt auf, dass zunächst praxisorientierte Trainingskurse zu den Grundlagen des Forschungsdatenmanagements in den historischen Disziplinen entwickelt werden sollten.

Danksagung

Ein großer Dank geht an die NFDI4Memory-Community, die die Arbeiten nicht nur durch ihre überwältigende Teilnahme durch 581 Proband:innen an der Umfrage unterstützt hat, sondern bereits in der Phase der Entwicklung des Fragebogendesigns wertvolle Hinweise gab und das Umfragedesign testete. Ebenso bedanken wir uns bei allen Teilnehmenden an den drei Workshops¹, in denen wir (Zwischen-)Ergebnisse vorstellten und diskutierten. Die Einschätzungen und Interpretationen, die wir von den Teilnehmenden zu dem einen und anderen Phänomen in den Daten erhalten haben, haben uns bei der Auswertung der Daten sehr geholfen. Vielen Dank, dass Sie Ihre Expertise so bereitwillig mit uns geteilt haben. Ein besonderer Dank geht an Dr. Katrin Moeller vom Historischen Datenzentrum Sachsen-Anhalt. Sie unterstützte uns bei der Datenauswertung, insbesondere brachte sie die Idee der Kategorisierung nach Wissensleveln in die Auswertung ein, welche in der Gesamtschau auf die Ergebnisse eine wichtige Analyseperspektive für die Interpretation unserer Umfrageergebnisse darstellt.

Diese Arbeit ist im Rahmen des NFDI-Konsortiums 4Memory entstanden (www.4memory.de). Wir danken der Deutschen Forschungsgemeinschaft (DFG) für die finanzielle Unterstützung – Projektnummer 50 1609550.

¹ FORGE23 - Forschungsdaten in den Geisteswissenschaften - kritisch betrachtet (05.10.2023): Die Ermittlung von Data Literacy Bedarfen in Studium, Lehre und Qualifikation in den historisch arbeitenden; Wissenschaften. Ein kritischer Blick auf die Zwischenergebnisse der Erhebung; NFDI4Memory Community Forum, <https://4memory.de/2023/12/15/recap-erstes-nfdi4memory-community-forum/> (20.11.2023); Werkstattgespräch zu ersten Ergebnissen aus der Data Literacy Bedarfserhebung; NFDI4Memory Data Literacy Community Workshop (24.01.2024): Interner Community Workshop der Task Area „Data Literacy“ zur Interpretation und Diskussion der vorläufigen Ergebnisse der Bedarfserhebung

Inhaltsverzeichnis

Einleitung.....	1
Methodisches Vorgehen.....	4
Erstellung eines Data-Literacy-Kompetenzprofils.....	4
Erstellung des Fragebogendesigns	7
Berechnungen und Teilnehmendenzahl	10
Darstellung der Einzelergebnisse.....	11
Soziodemografische Fragen.....	11
Ich bin tätig in folgender Einrichtung: (Einfachauswahl)	11
Ich ordne mich folgender Gruppe zu: (Einfachauswahl)	12
Im folgenden Fachgebiet bin ich tätig:	14
Allgemeine Einschätzung	17
Ich fühle mich inhaltlich hinreichend sicher im Themenfeld Datenkompetenzen und Forschungsdatenmanagement in meinem Tätigkeitsbereich.....	17
Allgemeine FDM-Kenntnisse	19
Begriff „Forschungsdaten“.....	19
Begriff „Forschungsdatenmanagement“	20
Begriff „Forschungsdatenlebenszyklus“	21
Begriff „Datenmanagementplan“	22
Begriff „Metadaten“	22
Begriff „Open Science“	23
Begriffe „Open“ und „FAIR“	24
Kurzzusammenfassung der Kategorie Allgemeine FDM-Kenntnisse	25
Datenerschließung.....	26
Qualitative und quantitative Daten	26
Plattformen zur Datenrecherche	27
Recherchewerkzeuge und -strategien	28
Strategien zur Fehlervermeidung.....	28
Kurzzusammenfassung der Kategorie „Datenerschließung“.....	30
Datenorganisation.....	31
Strategien zum Ordnen von Daten.....	31
Forschungssoftware.....	32

Kurzzusammenfassung der Kategorie „Datenorganisation“	33
Datenevaluation.....	34
Digitale fachspezifische Forschungsmethoden.....	34
Visualisierungsarten.....	35
Zu folgenden digital gestützten Analysemethoden würde ich gerne mehr erfahren:	35
Kurzzusammenfassung Kategorie „Datenevaluation“	37
Datenaufbereitung	39
Begriff „Datendokumentation“	39
Begriff „Multimodalität“	40
Werkzeuge und Methoden zur Datendokumentation.....	41
Begriff „Ontologie“	42
Begriff „Metadatenstandards“.....	42
Zu folgenden Metadatenstandards wünsche ich mir mehr Informationen:	43
Zu folgenden Methoden zur Aufbereitung von Daten würde ich gerne mehr erfahren:	44
Kurzzusammenfassung der Kategorie „Datenaufbereitung“	47
Datenqualität.....	49
Bewertung von Datenqualität.....	49
Qualitätsmängel in Daten.....	50
Kurzzusammenfassung Kategorie „Datenqualität“.....	51
Datenkonvertierung	52
Begriffe „Datentyp / -format“ und „Dateityp / -format“.....	52
Begriff „Datenkonvertierung“	53
Kurzzusammenfassung in der „Datenkonvertierung“	54
Programmierung	55
Begriff „algorithmisches Denken“	55
Einfacher Programmcode	56
Zu folgenden Programmiersprachen benötige ich mehr Informationen:	57
Kurzzusammenfassung der Kategorie „Programmierung“	58
Recht	59
Fachspezifische rechtliche Aspekte	59
Begriff „Anonymisierung“.....	59
Begriff „Pseudonymisierung“.....	60
Begriff „Informierte Einwilligung“	61

Begriff „verschlüsselte Daten“	62
Zu folgenden ethischen und rechtlichen Aspekten benötige ich mehr Informationen: ..	63
Kurzzusammenfassung der Kategorie „Recht“	65
Ethik.....	67
Leitlinien zur Sicherung guter wissenschaftlicher Praxis	67
Begriff „Datenethik“	68
Begriff „CARE“	69
Kurzzusammenfassung in der Kategorie „Ethik“	70
Speicherung.....	71
Begriff „Speichersystem“	71
Grundprinzipien der Datensicherung	72
Begriff „Versionierungstool“	73
Speicherung und Langzeitarchivierung.....	74
Begriff „Datenbanksystem“	75
Kurzzusammenfassung der Kategorie „Speicherung“	76
Publikation	77
Begriff „Datenpublikation“	77
Fachspezifische Datenpublikationsorgane.....	78
Kurzzusammenfassung der Kategorie „Publikation“	79
Datennachnutzung.....	80
Nachnutzbarkeit von Daten und Datenformaten	80
Begriff „FAIR“	81
Methoden und Strategien, um Daten nachnutzbar(er) zu machen	82
Folgende Aspekte der Nachnutzung kenne ich/ wende ich an/ kann ich vermitteln:	83
Kurzzusammenfassung der Kategorie „Datennachnutzung“	85
Software/Tools.....	86
Zu folgenden Software/Tools hätte ich gerne mehr Informationen:	86
Didaktik.....	87
Anwendung didaktischer Methoden	87
Vermittlung didaktischer Methoden.....	87
Kurzzusammenfassung der Kategorie „Didaktik“	88
Vermittlung und Information.....	89

Über welche Formate möchte ich mich zu Datenkompetenz und Forschungsdatenmanagement weiterbilden: (Mehrfachauswahl).....	89
Weitere relevante Themen und Feedback.....	91
Ich erachte folgende weitere Themen als relevant, die im Fragebogen nicht vorgekommen sind:.....	91
Zusammenfassung.....	94
Interpretation(versuch) der Ergebnisse.....	100
Ausblick.....	108
Literaturverzeichnis	110

Einleitung

Im Rahmen des NFDI4Memory-Konsortiums arbeitet die Task Area 4 „Data Literacy“ an der Entwicklung von Services, Handreichungen und innovativen Lehr- und Lernformaten, um die Datenkompetenzen – auch Data Literacy (DL) genannt – und das aktive Forschungsdatenmanagement (FDM) von Forschenden und Lehrenden auf allen Qualifikationsstufen in den historisch arbeitenden Disziplinen zu verbessern. Um zielgruppenspezifische und dem Kompetenzniveau entsprechende DL-Angebote zu entwickeln, wurde in einem ersten Schritt eine Bedarfserhebung durchgeführt, deren Konzeption, die gewonnenen Ergebnisse und Erkenntnisse sowie die daraus resultierenden nächsten Schritte hier dargestellt werden.

Grundlage unserer Arbeit und der inhaltlichen Ausrichtung der durchgeführten Bedarfserhebung ist ein Konzept von „Data Literacy“, das wir im Rahmen der gemeinsamen Arbeit im Konsortium entwickelt haben und das für die Umsetzung des gesamten Arbeitsprogramms der Task Area „Data Literacy“ maßgeblich ist. Eine allgemeine von allen Akteur:innen verwendete Definition und Sichtweise auf Data Literacy existiert (bislang) nicht, auch deshalb, weil eine eindeutige Abgrenzung zu anderen digitalen Kompetenzen wie bspw. der Digital Literacy, Media Literacy, Statistical Literacy, Coding Literacy oder Information Literacy nicht möglich ist (Vgl. Grillenberger und Romeike 2019, 12). Daher haben wir uns innerhalb der Arbeitsgruppe bislang auf folgendes Verständnis von „Data Literacy“ geeinigt:

Data Literacy oder Datenkompetenz bezeichnet die Fähigkeit einer Person, Daten zu erfassen, zu sammeln, zu managen, zu kuratieren, zu annotieren, zu verstehen, zu interpretieren, zu analysieren, zu visualisieren, zu kontextualisieren, zu dokumentieren und daraus sinnvolle Schlussfolgerungen zu ziehen. Dies ist besonders wichtig im Kontext der Arbeit mit Forschungsdaten und der Anwendung digitaler Forschungsmethoden. Datenkompetenzen umfassen ein breites Spektrum an Fähigkeiten und Kenntnissen, die notwendig sind, um Forschungsdaten effektiv, qualitätsgesichert und unter Berücksichtigung der Regeln guter wissenschaftlicher Praxis im Einklang mit den FAIR- und Open Science-Prinzipien zu handhaben. Dies umfasst das Wissen und Verständnis um grundlegende Begriffe und Konzepte des FDM, die Fähigkeit zur Datenorganisation, -strukturierung und -verarbeitung, die Fertigkeiten im sicheren Umgang mit digitalen Werkzeugen und Technologien sowie die Motivation und Offenheit gegenüber datenethischen und -rechtlichen Aspekten. Ebenso wichtig ist die Bereitschaft, den notwendigen Kulturwandel hin zu einer etablierten Datenkultur in den historisch arbeitenden Disziplinen mitzutragen und gestalten zu wollen.

Der Bericht gliedert sich in fünf Kapitel: Im ersten Abschnitt wird das methodische Vorgehen dargelegt. Erläutert werden die methodischen Grundlagen für die Entwicklung des Frage-

bogendesigns und das Vorgehen bei der Datenauswertung. Da es keine vergleichbare Erhebung in den historisch arbeitenden Disziplinen im deutschsprachigen Raum gibt und auch nicht auf ein bereits entwickeltes und erprobtes Fragebogendesign zurückgegriffen werden konnte, welches das von uns entwickelte und definierte Datenkompetenzspektrum abbildet, musste zunächst das Umfragedesign aus verschiedenen DL-Kompetenzprofilen abgeleitet werden. Dieser Schritt wird im Kapitel „Methodisches Vorgehen“ dargelegt. Des Weiteren wird in diesem Kapitel für ein besseres Verständnis der Darstellung der Einzelergebnisse das Vorgehen bei der Auswertung der Daten erläutert. Danach folgen die Ergebnisdarstellungen zu jeder einzelnen Frage in einem vergleichenden Blick auf alle Zielgruppen. Wir haben uns in diesem Kapitel bewusst dazu entschieden, eine rein beschreibende Darstellung der Ergebnisse ohne Interpretationen zu präsentieren. Die Interpretation und die übergreifende Gesamtschau auf die Ergebnisse erfolgt im vorletzten Kapitel „Zusammenfassung“.

Bei der Datenauswertung wurden die Lehrer:innen und Lehramtsreferendar:innen aufgrund ihrer geringen Teilnahme von nur dreizehn Personen herausgenommen. Die geringe Anzahl ließ keine statistisch valide und repräsentative Auswertung zu. Über die Gründe der geringen Teilnahme kann nur gemutmaßt werden. Für die Kontextualisierung der geringen Beteiligung sind Informationen interessant, die wir in Feedback-Gesprächen mit Vertreter:innen dieser Zielgruppen erhalten haben. So wurde u. a. darauf hingewiesen, dass es bereits zahlreiche Initiativen zur Förderung der digitalen Kompetenzen von Lehrkräften gibt. Das BMBF fördert bspw. auf vielfältige Art und Weise die Digitalisierung der Schulen, insbesondere mit der „Qualitätsoffensive Lehrerbildung“ (Vgl. BMBF 2015), in der ca. 56 Projekte mit dem Schlagwort „Digitalisierung“ seit 2015 Mittel erhalten. Unterstützt werden darüber hinaus vier Kompetenzzentren für digitales und digital gestütztes Unterrichten (Vgl. BMBF 2022), in denen rund 200 Projekte Fort- und Weiterbildungsformate und -materialien für digital gestützten Unterricht in unterschiedlichen Fächern entwickeln und bereitstellen. Diese Initiativen werden im Kompetenzverbund lernen:digital (Vgl. Kompetenzverbund lernen:digital 2023) zusammengeführt, um den Dialog zwischen Wissenschaft und Praxis für die digitale Transformation von Schule und Lehrkräftebildung sicherzustellen und die Projektergebnisse in der Breite bekannt zu machen. Des Weiteren wird von der Kultusministerkonferenz (KMK) „das Lernen mit und über digitale Medien und Werkzeuge“ (KMK 2017, 11) als Ziel für die Schulen formuliert. Insgesamt fokussieren sich die Angebote und Projekte im Bereich Digitalität im schulischen Unterricht primär auf die kompetente Anwendung von digitalen Medien und Tools - „Digital Literacy“. Projekte und Angebote, die die Datenkompetenz und „Data Literacy“ von Lehrer:innen verbessern wollen und dabei explizit die Datenebene in den Blick nehmen, sind in diesem Bereich weniger stark vertreten. Die anwendungsorientierte Perspektive der bestehenden Förderprogramme im schulischen Bereich zeigt sich u. a. auch darin, dass die Begriffe „Datenkompetenz“ oder „Data Literacy“ in der KMK Digitalstrategie

2017 vermieden werden, an ihre Stelle treten „Medien-“, „Computer-“ oder „Informationskompetenz“. Dies deutet darauf hin, dass Geschichtslehrer:innen sehr wahrscheinlich ein gänzlich anderes digitales Kompetenzprofil benötigen als alle anderen Zielgruppen, die NFDI4Memory adressiert. Deshalb wird die Task Area „Data Literacy“ im weiteren Projektverlauf mit Akteur:innen aus dem Schulsektor folgende Fragen bearbeiten: Brauchen Geschichtslehrer:innen ergänzende fachspezifische Lehr- / Lernangebote zu den digitalen Kompetenzen, die in der Schule benötigt werden, oder reichen ihnen die vielfältigen bereits vorhandenen Angebote der oben genannten Kompetenzzentren aus? Wie ist es um die digitalen Kompetenzen bei den Geschichtslehrer:innen bestellt und welche Unterstützung würden sie sich noch von NFDI4Memory wünschen? Auf welche Art und Weise können Geschichtslehrer:innen kontaktiert werden, um sie einzubeziehen und die ggf. zu entwickelnden Materialien bekannt zu machen? In Abhängigkeit von den Antworten auf diese Fragen werden dann entsprechende Maßnahmen ergriffen.

Im vorletzten Kapitel „Zusammenfassung“ wird einerseits über alle DL-Kategorien hinweg ein vergleichender Blick auf die Ergebnisse in den verschiedenen Zielgruppen geworfen. Dabei wurden Quantilenberechnungen vorgenommen, um die Verteilung der Proband:innen in DL-Wissenslevel (Lai:innen, Starter:innen, Kenner:innen und Expert:innen) zu ermitteln. Andererseits werden Interpretationsansätze für die Ergebnisse der Umfrage dargestellt. An dieser Stelle werden auffällige Analyseergebnisse zu einzelnen Fragen benannt und Interpretationsansätze angeboten. Im letzten Kapitel wird ein Ausblick gegeben, wie diese Daten und daraus bereits gewonnene bzw. noch zu gewinnende Erkenntnisse in die weitere Arbeit der Task Area „Data Literacy“ einfließen werden.

Methodisches Vorgehen

Erstellung eines Data-Literacy-Kompetenzprofils

Für die Befragung konnte kaum auf vorhandene Fragebogendesigns für die Ermittlung von Datenkompetenzen zurückgegriffen werden, da bislang nur wenige Erhebungen dieser Art durchgeführt wurden. Die auffindbaren Studien ermitteln eher Digital Literacy (Vgl. Stürz u. a. 2023; Swiatek u. a. 2020; Bandtel, Kauz und Weißker 2021). Data-Literacy-Kompetenzen wurden dabei nur am Rande erhoben. Für den Bereich des FDM gibt es zwar diverse Bedarfserhebungen, aber sie wurden weniger auf den Kenntnisstand der Befragten ausgerichtet, sondern zielen vielmehr auf benötigte Infrastrukturen und Services ab (Vgl. Iglezakis und Schembera 2018; Gronwald 2019; Lemaire u. a. 2016; Langhanke und Stille 2015). Eine systematische Erhebung mit einem geistes- oder gar geschichtswissenschaftlichen Fokus liegt bislang nicht vor.

Dementsprechend war das Fragebogendesign eine besondere Herausforderung: Für dessen Entwicklung wurden FDM- und DL-Kompetenzmatrizen zu einem DL-Kompetenzprofil für die historisch arbeitenden Wissenschaften zusammengeführt. Als Grundlage sind die „Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards“ (Petersen u. a. 2023), die „FDM-Kompetenzmatrix“ des DIAMANT-Modells (Lemaire u. a. 2020, 30-40) und die „Taxonomy of Digital Research Activities in the Humanities (TaDiRAH)“ (Schöch, Perkins und Borek 2015) verwendet worden. Um die eher generischen Matrizen fachspezifisch zu ergänzen und sich so einem FDM- / DL-Kompetenzprofil für die historisch arbeitenden Disziplinen anzunähern, wurden Informationen aus Modulhandbüchern, Studiengangsbeschreibungen sowie Zusatz- und Zertifikatskursen von Digital-History-Studiengängen an deutschen Universitäten integriert.

Hierfür wurde unter Zuhilfenahme der Forschungssoftware FuD² ein Mapping der genannten Kompetenzmatrizen durchgeführt. In diesem Kontext verwenden wir den Begriff „Mapping“ als Überbegriff einer Analysemethode, welche dazu dient, Daten/Informationen aus unterschiedlichen Quellen aufeinander abzubilden, wenn sie die gleiche semantische Information enthalten. Dadurch wurden Beziehungen, Zusammenhänge und Überschneidungen zwischen den Kompetenzmatrizen identifiziert. Ziel des Mappings war es, ein vorläufiges „Data Literacy“-Kompetenzprofil für die historischen Disziplinen zu erstellen. Ausgangspunkt der Zusammenführung war die „FDM-Kompetenzmatrix“ des DIAMANT-Modells. Diese Matrix basiert auf dem FDM-Referenzprozess, welcher entlang des Forschungsprozesses die notwendigen FDM-Prozessschritte und die zugehörigen Akteur:innen darstellt (Vgl. Lemaire u. a. 2020, 11–29).

² www.fud.uni-trier.de (03.06.2024)

Ziel dieser Kompetenzmatrix ist es, außeruniversitären Forschungseinrichtungen und Hochschulen einen Referenzprozess als Benchmark für die Implementierung einer Informationsarchitektur bereitzustellen, um eine FDM-Service-Landschaft möglichst effizient und bedarfsgerecht aufzubauen und weiterzuentwickeln. Die Kompetenzbeschreibungen richten sich daher immer an eine bestimmte Akteursrolle zu einer bestimmten FDM-Funktion im Gesamtprozess³. Die DIAMANT-Kompetenzmatrix inkludiert das „Librarians’ competencies profile for research data management“ (Schmidt und Shearer 2016), „Data Stewardship on the map: A study of tasks and roles in Dutch research institutes“ (Verheul u. a. 2019), „A Pilot Competency Matrix for Data Management Skills: A Step toward the Development of Systematic Data Information Literacy Programs“ (Nelson 2017) sowie „Future Skills: Ansätze zur Vermittlung von Data-Literacy-Kompetenzen“ (Heidrich, Bauer und Krupka 2018). Für die strategischen Kompetenzen wurde auf „RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement“ (Hartmann, Jacob und Weiß 2019) zurückgegriffen. Weitere fehlende Kompetenzbeschreibungen wurden von den Autor:innen des DIAMANT-Modells hinzugefügt. Diese Matrix wurde im ersten Schritt mit der „Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards“ (Petersen u. a. 2023) gemappt, nachfolgend als LZM bezeichnet.

Die LZM orientiert sich nicht am Forschungsprozess, sondern ist nach unterschiedlichen FDM-Themenfeldern strukturiert und benennt generische Lernziele für die verschiedenen Qualifikationsstufen. Dabei orientiert sich diese an der Taxonomie von Bloom und Anderson-Krathwohl (Vgl. Krathwohl 2002), wobei die Kompetenzbereiche (Sach-, Methoden-, Selbst- und Sozialkompetenz) den Lernzielen zugeordnet sind. Diese Lernziele wurden für die Qualifikationsniveaus Bachelor, Master, PhD und Data Stewards formuliert. Sie unterscheidet sich strukturell in ihrem Aufbau stark von der DIAMANT-FDM-Kompetenzmatrix. Dennoch war ein inhaltliches Mapping zwischen den Kompetenzbeschreibungen möglich und es wurde eine hohe Deckung erreicht.⁴ Am Ende stand eine fusionierte DIAMANT-LZM-Matrix, die v. a. hinsichtlich allgemeiner FDM-Kenntnissen erweitert wurde, insbesondere was Kernkonzepte des FDM wie Open Science, FAIR-Prinzipien, Data Policies und Datenmanagementpläne sowie den Bereich Vermittlung von FDM- / DL-Kompetenzen betrifft. Ansonsten handelt es sich bei den Erweiterungen eher um Konkretisierungen vorhandener Kompetenzbeschreibungen der

³ Z. B. sind für die Rolle des Forschenden in der Phase der „Erhebung, Analyse und Auswertung“ bei der Aktivität der Datensammlung folgende Kompetenzbeschreibungen zu finden: „1. Kenntnisse über Datenerschließung und -sammlung; 2. Fähigkeit, die Qualität und Zuverlässigkeit von Daten und Datenquellen kritisch zu bewerten und sicherzustellen.“

⁴ Bis auf fünf Kompetenzen der LZM konnten alle anderen einer Kompetenz in der DIAMANT-Matrix zugeordnet werden. Dafür wurden allerdings nur knapp über 30 % der Kompetenzen der DIAMANT-Matrix benötigt. Die meisten Übereinstimmungen gab es im Bereich Datenanwendung, -evaluation, -management und -sammlung. Des Weiteren gab es Überlappungen im Bereich Datenaufbereitung für Publikation und Archivierung sowie bei der Auswahl eines Datenarchivs.

DIAMANT-Kompetenzmatrix. Lediglich die explizite Nennung von Programmierkenntnissen erforderte eine Ergänzung. Weil beide Matrizen jedoch einen generischen DL-Ansatz abbilden, und damit die Konkretisierung anhand von „disziplinüblichen“ geisteswissenschaftlichen Datentypen und Methoden fehlte, wurde im nächsten Mappingschritt die „Taxonomy of Digital Research Activities in the Humanities“ TaDiRAH (Schöch, Perkins und Borek 2015) mit der DIAMANT-LZM-Matrix gemappt. Die Taxonomie dient der Beschreibung von Forschungsaktivitäten in den Digital Humanities unter Berücksichtigung der verwendeten Theorien und Methoden in dieser Disziplin. TaDiRAH soll insbesondere für Vorhaben nützlich sein, die Informationen zu Werkzeugen, Methoden, Projekten oder Literatur im Bereich der DH sammeln (Vgl. Borek u. a. 2021). Wie erwartet, konnten die TaDiRAH-Kompetenzen mit den vorhandenen Kompetenzbeschreibungen zusammengeführt werden und dienten insbesondere der geisteswissenschaftlichen Konkretisierung im Bereich der anzuwendenden Methoden zur Datengewinnung, -aufbereitung, -analyse und -visualisierung. Dies zeigt sich in den 50 % der erfolgten Zuordnungen aus der TaDiRAH-Matrix in den Bereichen Datenanwendung, -evaluation, -management und -sammlung.⁵ Lediglich die TaDiRAH-Forschungsaktivitäten „Dissemination“ und „Rahmenaktivitäten“ ergänzten weitere Kompetenzen wie z. B. Community Building und Projektmanagement. Hinzu kam im Bereich „Programmierkenntnisse“ noch die Web- bzw. Toolentwicklung. Am Ende dieses Mappingschrittes stand eine fusionierte DIAMANT-LZM-TaDiRAH-Kompetenzmatrix.

Die Integration von TaDiRAH ermöglicht zwar eine geisteswissenschaftliche Annäherung an ein DL-Kompetenzprofil für die historisch arbeitenden Disziplinen, allerdings ist TaDiRAH stark sprachwissenschaftlich geprägt und repräsentiert noch nicht ausreichend die geschichtswissenschaftlichen Forschungsmethoden. Um diese Lücke zu schließen, wurden systematisch Modulhandbücher, Studiengangsbeschreibungen sowie Zusatz- und Zertifikatskurse von Digital History Studiengängen analysiert und mit der DIAMANT-LZM-TaDiRAH-Kompetenzmatrix gemappt. Dies führte zu einer deutlichen geschichtswissenschaftlichen Schärfung. Dadurch gelang die Annäherung v. a. im Bereich der digitalen Methoden und Tools, die für die historischen Fächer eine größere Relevanz zu haben scheinen, wie z. B. Netzwerkanalyse, geographische Visualisierungen, Topic Modeling, Transkriptionstools, insbesondere automatische Texterkennung von Druck- und Handschriften, Text-Mining, Datenbank, spezifische Datenformate / -standards (bspw. TEI, IIIF) und auch Programmierkenntnisse.⁶

⁵ Insgesamt wurden für das Mapping knapp 15 % der Kompetenzen des DIAMANT-LZM-Index benötigt.

⁶ Der Schwerpunkt der Zuweisungen war mit nahezu 70 % im Bereich Datenevaluation. Danach folgen mit großem Abstand Datenmanagement, -sammlung und -aufbereitung.

Diese Zusammenführung der unterschiedlichen DL-Kompetenzprofile erhebt keinen Anspruch auf Vollständigkeit und wird im Rahmen der Aufgaben der Measure 1 „Data Literacy Training Lab“ der 4Memory Task Area „Data Literacy“ weiterentwickelt zu einem Datenkompetenzprofil für die historisch arbeitenden Disziplinen.⁷ Es diente hier aber als Ausgangsbasis für die Erstellung eines Fragebogendesigns für die Bedarfserhebung, die nachfolgend beschrieben wird.

Erstellung des Fragebogendesigns

Für die Entwicklung des Fragebogendesigns wurden die Ergebnisse der Zusammenführung der DL- und FDM-Kompetenzprofile quantitativ ausgewertet, um festzustellen, welche DL-Aspekte bei einer Bedarfserhebung unbedingt abgefragt werden sollten. Daraus ergaben sich folgende 16 Kategorien, ergänzt um eine allgemeine Einschätzung:

Übersicht Kategorien

- Allgemeine Einschätzung
- Didaktik*
- Allgemeine FDM-Kenntnisse
- Datenerschließung
- Datenorganisation
- Datenevaluation
- Datenaufbereitung
- Datenqualität
- Datenkonvertierung
- Programmierung
- Recht
- Ethik
- Speicherung
- Publikation
- Datennachnutzung
- Software/Tools
- Vermittlung und Information

*nur im Fragebogen der Niveaustufe 3: „Vermitteln“ (s. u.) vorkommend

Zu jeder Kategorie wurden die am häufigsten genannten Aspekte aus dem Datenkompetenzprofil jeweils pro Niveaustufe als Fragen formuliert. Aufgrund des breiten Adressat:innenkreises der Umfrage und den damit einhergehenden Kompetenzunterschieden zwischen den Qualifikationsstufen in den Bereichen FDM und DL, wurde eine Unterteilung in die drei aufeinander aufbauenden Niveaustufen Wissen → Anwenden → Vermitteln vorgenommen:

⁷ Dies wird in Zusammenarbeit mit der NFDI4Memory-Community erfolgen.

Niveaustufe 1 bezieht sich auf die Kompetenz „Wissen“ und fragt nach dem Vorhandensein von DL- und FDM-Kenntnissen. Ihr wurden die Zielgruppen Studierende im Bachelor und Master zugeteilt.

Niveaustufe 2 meint die Kompetenz „Anwenden“ des vorhandenen Wissens zum Umgang mit (Forschungs-)Daten. Ihr wurden Promovierende sowie PostDocs, Data Stewards / FDM-Mitarbeitende / Datenkurator:innen und GLAM-Mitarbeitende, die angeben, keine Lehrtätigkeit auszuüben, zugeteilt.

Niveaustufe 3 umfasst die Kompetenz „Vermitteln“ und basiert auf den beiden vorhergehenden Kompetenzstufen. Hier wird die Frage adressiert, inwieweit die Proband:innen sich in der Lage fühlen, Data-Literacy-Kenntnisse und deren Anwendungen und Nutzungsmöglichkeiten zu vermitteln / zu lehren. Zugeordnet wurden dieser Stufe Personen mit einer Vermittlungs- und Lehrtätigkeit wie Professor:innen, Lehrer:innen und Lehramtsreferendar:innen sowie PostDocs, Data Stewards / FDM-Mitarbeitende / Datenkurator:innen und GLAM-Mitarbeitende, die angeben, Lehrtätigkeiten auszuüben.

Das bedeutet, dass die Kategorien und die ihnen zugeordneten Fragen in allen drei Fragebögen bis auf wenige Ausnahmen identisch sind und sich lediglich in der Formulierung und dem damit adressierten Kompetenzniveau voneinander unterscheiden.






Im Folgenden eine exemplarische Darstellung der für die unterschiedlichen Kompetenzniveaus formulierten Aussagen am Beispiel der Kategorie „Allgemeine FDM-Kenntnisse“ und der Frage nach dem Forschungsdaten-Begriff:

Niveaustufe 1: „Ich weiß, was unter dem Begriff ‚Forschungsdaten‘ zu verstehen ist.“

Niveaustufe 2: „Ich kann den Begriff ‚Forschungsdaten‘ eigenständig definieren und Beispiele nennen.“

Niveaustufe 3: „Ich kann vermitteln, was unter dem Begriff ‚Forschungsdaten‘ zu verstehen ist.“

Des Weiteren wurde entschieden, einen Selbsteinschätzungsfragebogen zu konzipieren, da der Aushandlungsprozess, welche Kompetenzen in welcher Qualifikationsstufe zu den einzelnen Aspekten in welcher Qualität/ Ausprägung erfüllt sein sollten, noch am Anfang steht und, wie zuvor bereits erläutert, ein eigenes Arbeitspaket der Task Area „Data Literacy“ ist. Die Proband:innen bekamen für die Selbsteinschätzung folgende Likert-Skala⁸:

„nicht beurteilbar“	„trifft überhaupt nicht zu“	„trifft überwiegend nicht zu“	„trifft teils-teils zu“	„trifft überwiegend zu“	„trifft völlig zu“
?					
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Alle drei Fragebögen beginnen mit soziodemographischen Fragen. Dieser Abschnitt besteht aus Fragen zur institutionellen Verortung, zu einer persönlichen Zuordnung zu einer Statusgruppe und zur Zugehörigkeit zu einem Fachgebiet. Aufgrund der zielgruppenspezifischen Adressierung richtet sich der soziodemographische Abschnitt zuvorderst an Personen, die an Hochschulen, außeruniversitären Forschungseinrichtungen, IT- / FDM-Serviceeinrichtungen oder GLAM-Institutionen beschäftigt sind. Es werden Personen adressiert, die sich den historisch arbeitenden Disziplinen zuordnen. Um hier für die spätere Datenauswertung möglichst mit standardisierten Angaben arbeiten zu können und den meisten Proband:innen ein bekanntes Vokabular zur Verfügung zu stellen, wurde die DFG-Fachsystematik als Grundlage ausgewählt und die Fächer mit historischem Bezug in eine Auswahlliste übernommen. Wegen der hohen Relevanz für die GLAM-Zielgruppe wurde zusätzlich die Disziplin „Informations-, Bibliotheks- und Archivwissenschaft“ ergänzt, weil es kein Äquivalent in der verwendeten Systematik gibt.⁹ Darüber hinaus wurde für GLAM-Mitarbeitende und Data Stewards / FDM-Mitarbeitende / Datenkurator:innen ein Freitextfeld angeboten, da in diesen Bereichen oftmals eine genaue Fachzuordnung nicht möglich ist und sich diese Gruppe in der Regel mehreren Disziplinen zuordnen kann. Für die korrekte Zuordnung der Befragten hinsichtlich ihrer Niveaustufe im Fragebogen erhielten PostDocs, Data Stewards / FDM-Mitarbeitende / Datenkurator:innen sowie GLAM-Mitarbeitende die zusätzliche Frage, ob Lehre zu ihrem Aufgabenprofil gehört.

⁸ Die Likert-Skala ist eine nach Rensis Likert 1932 veröffentlichte Methode, Verhaltenseinstellungen messbar zu machen (Vgl. Batterton und Hale 2017). Likert beschreibt das Konstrukt als Verhaltensskala, wobei eine Gruppe von zusammenhängenden Fragen die Einstellung zu einem gewissen Subjekt messbar macht. Die Skala besteht aus fünf Antwortmöglichkeiten, diese werden im Folgenden aufgelistet: „Stimme überhaupt nicht zu“, „Stimme überwiegend nicht zu“, „Unentschlossen“, „Stimme zu“, „Stimme voll und ganz zu“ (Vgl. Batterton und Hale 2017). In der Umfrage wurde die Auswahlmöglichkeit „Unentschlossen“ durch „Stimme teils-teils zu“ ersetzt. Außerdem wurde eine weitere Antwortoption hinzugefügt, die jedoch nicht Teil der Skala ist. Proband:innen konnten demnach auch die Option „Nicht beurteilbar“ auswählen.

⁹ In der Auswertung stellte sich bei den Disziplinen heraus, dass in der Auswahlliste die Literaturwissenschaften gefehlt hatten. Diese Disziplin wurde als weiteres Item in der Auswertung hinzugefügt.

Berechnungen und Teilnehmendenzahl

Wie im vorherigen Abschnitt erwähnt, wurden die Teilnehmer:innen entsprechend ihrer angegebenen Statusgruppe einer bestimmten Niveaustufe zugeordnet. Abhängig von dieser Niveaustufe wurden die zugehörigen Fragen gestellt. Den Auswahloptionen der Likert-Skala wurden die Werte von 1 bis 5 von „trifft überhaupt nicht zu“ bis „trifft völlig zu“ zugeordnet. Mit diesen Werten wurden die Mittelwerte berechnet.

Des Weiteren hatten die Proband:innen die Auswahloption „nicht beurteilbar“ zur Verfügung. Hat eine Person diese Option gewählt, wurde sie nicht in der Mittelwertberechnung berücksichtigt, weshalb die Proband:innenanzahl zwischen den Fragen variieren kann. Die Anzahl der Proband:innen mit einer gültigen Antwort pro Frage wird in der nachstehenden Auswertung jeweils durch „n“ angegeben.

Während die Beantwortung der meisten Fragen obligatorisch war, war die Beantwortung der Auswahllisten optional. Daher variiert die Anzahl der teilnehmenden Personen (n) bei diesem Fragetyp ebenfalls. Da hier mehr als eine Angabe möglich war, wird an dieser Stelle einmal die Anzahl der Teilnehmer:innen, die mindestens eine Angabe gemacht haben und die Anzahl der Nennungen angegeben.

Die Einzelfragen innerhalb jeder Kategorie bieten jeweils eine Übersicht über die Mittelwerte pro Statusgruppe. Um Trends auf Kategorieebene zu identifizieren, wurde ein Gesamtmittelwert für jede Kategorie aus den Mittelwerten der einzelnen Statusgruppen berechnet. Dazu wurde der Mittelwert über alle (nicht mit „nicht beurteilbar“) beantworteten Fragen der jeweiligen Kategorie pro Person berechnet und aus diesen Mittelwerten wiederum der Gesamtmittelwert pro Statusgruppe ermittelt.

Darstellung der Einzelergebnisse

Soziodemografische Fragen

Ich bin tätig in folgender Einrichtung: (Einfachauswahl)

Bitte ordnen Sie sich nur einer Einrichtung / einer Gruppe zu. Falls Sie in mehreren Einrichtungen tätig sind, wählen Sie die Einrichtung mit dem größten Stellenanteil.

- | | |
|---|---|
| <input type="checkbox"/> Universität | <input type="checkbox"/> Archiv |
| <input type="checkbox"/> Hochschule für angewandte Wissenschaften | <input type="checkbox"/> Bibliothek |
| <input type="checkbox"/> außeruniversitäre Forschungseinrichtung | <input type="checkbox"/> Museum |
| <input type="checkbox"/> IT- / FDM-Serviceeinrichtung | <input type="checkbox"/> Schule |
| | <input type="checkbox"/> Sonstiges: _____ |

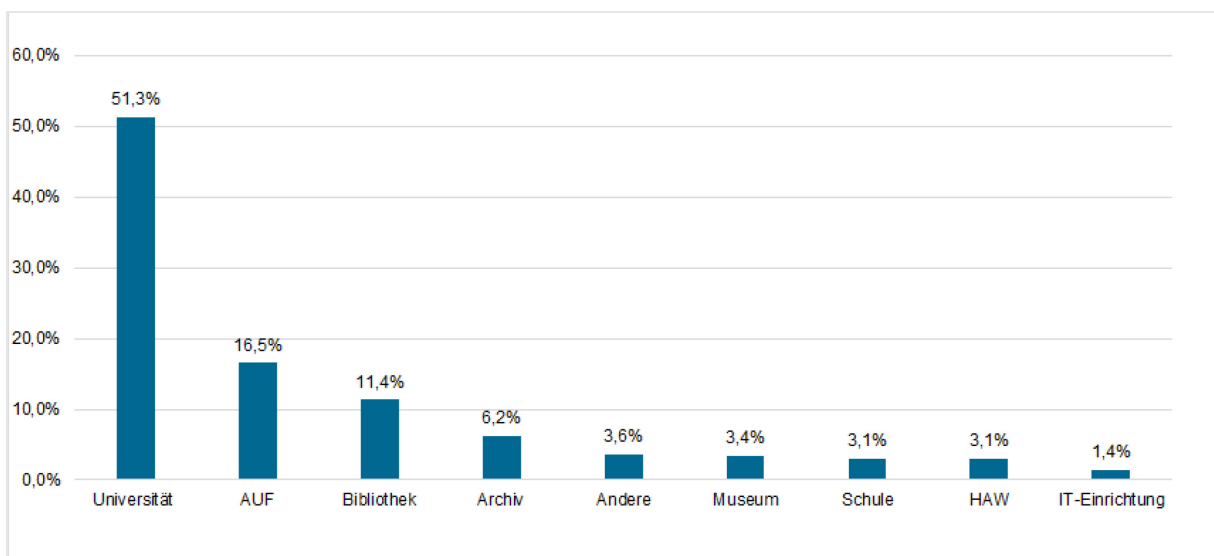


Abb. 1: Prozentuale Verteilung nach Einrichtungen | n = 581

Von den insgesamt 581 Teilnehmenden sind über die Hälfte an Universitäten tätig. Die zweitgrößte Gruppe, bestehend aus etwa 17 % der Teilnehmenden, arbeitet an außeruniversitären Einrichtungen, während Bibliotheken mit etwa 11 % an dritter Stelle stehen. Archive sind die viertstärkste Gruppe mit rund 6 %. Die restlichen Institutionen sowie die Kategorie „Andere“ verzeichnen jeweils Anteile von weniger als 4 %.

Ich ordne mich folgender Gruppe zu: (Einfachauswahl)

- Studierende (BA)
- Studierende (MA)
- Promovierende
- PostDoc
- Professor:in
- Data Steward / FDM-Mitarbeiter:in / Datenkurator:in
- GLAM-Mitarbeiter:in
- Lehrer:in
- Lehramtsreferendar:in

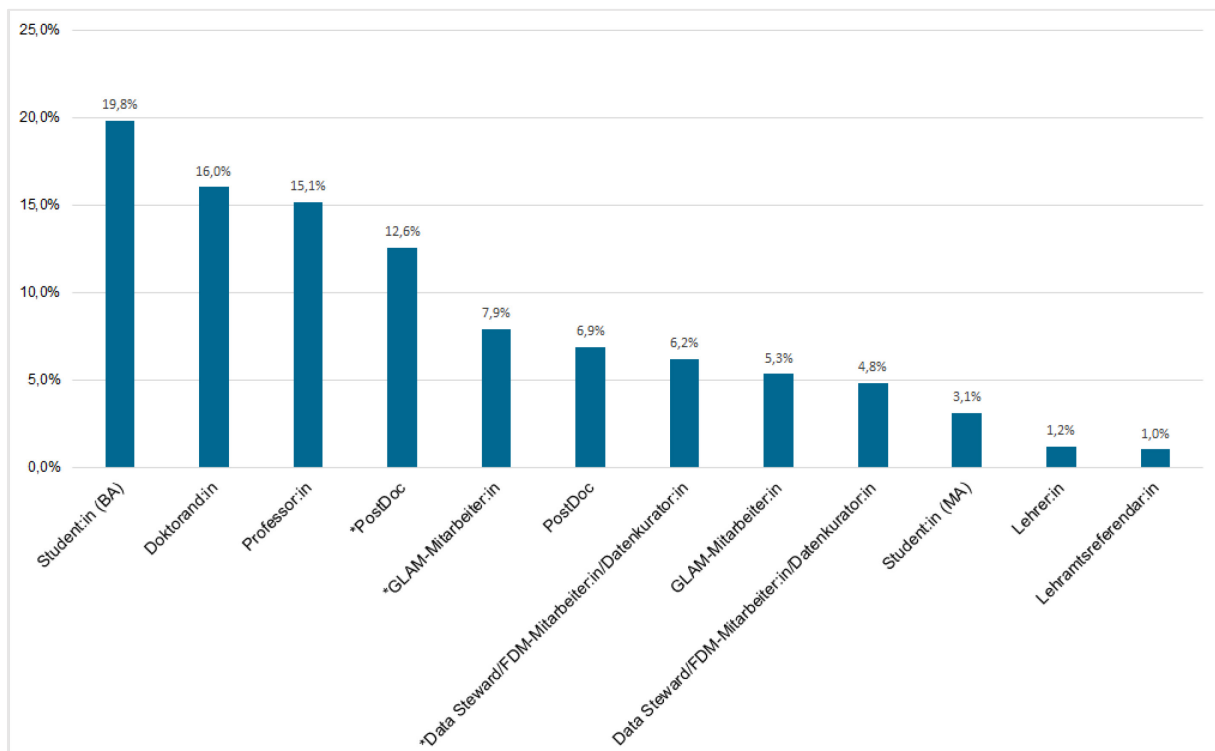


Abb. 2: Prozentuale Verteilung nach Statusgruppen (mit * markierte Statusgruppen = Lehrende) | n = 581

Statusgruppe	Häufigkeit	Anteil
Student:in (BA)	115	19,8 %
Student:in (MA)	18	3,1 %
Doktorand:in	93	16,0 %
PostDoc	40	6,9 %
GLAM-Mitarbeiter:in	31	5,3 %
Data Steward/ FDM-Mitarbeiter:in/ Datenkurator:in	28	4,8 %
*PostDoc	73	12,6 %
*GLAM-Mitarbeiter:in	46	7,9 %
*Data Steward/ FDM-Mitarbeiter:in/ Datenkurator:in	36	6,2 %
Professor:in	88	15,1 %
Lehrer:in	7	1,2 %
Lehramtsreferendar:in	6	1,0 %

Tab. 1: Verteilung nach Statusgruppen mit Häufigkeits- und Prozentangaben

Die größten Gruppen unter den Teilnehmenden sind Studierende in Bachelorstudiengängen sowie PostDocs (lehrende und nicht-lehrende) mit 19,8 % bzw. 19,4 %. Die Statusgruppen FDM-Mitarbeitende, GLAM-Mitarbeitende (jeweils lehrend und nicht-lehrend), Doktorand:innen, und Professor:innen sind mit Anteilen von 11–16 % vertreten. Hingegen sind die Lehramtsreferendar:innen, Lehrer:innen und Studierende in Masterstudiengängen mit 1–3 % unterrepräsentiert. Bis auf die drei letztgenannten ist die Verteilung der Gruppen so, dass sie miteinander vergleichbar sind.

Bei den Bachelorstudierenden handelt es sich zum großen Teil um Studierende im ersten Studiensemester der Geschichtswissenschaften.¹⁰ Auch dies muss bei der Interpretation der Ergebnisse dieser Gruppen berücksichtigt werden.

¹⁰ Während der dreimonatigen Erhebungsphase wurde die Unterrepräsentanz von Studierenden festgestellt. Durch gezieltes Ansprechen von Lehrpersonen zum Semesterbeginn 2023/2024 sollte diese Zielgruppe noch erreicht werden. Von denjenigen Lehrpersonen, die sich daran beteiligt haben, wissen wir, dass es sich dabei um Erstsemesterveranstaltungen handelte.

Im folgenden Fachgebiet bin ich tätig:

- Alte Kulturen
- Ur- und Frühgeschichte
- Klassische Philologie
- Alte Geschichte
- Klassische Archäologie
- Ägyptische und Vorderasiatische Altertumswissenschaften
- Geschichtswissenschaften
- Mittelalterliche Geschichte
- Frühneuzeitliche Geschichte
- Neuere und Neueste Geschichte (einschließlich Europäische Geschichte der Neuzeit und Außereuropäische Geschichte)
- Wissenschaftsgeschichte
- Kunst-, Musik-, Theater- und Medienwissenschaften
- Kunstgeschichte
- Sprachwissenschaften
- Historische Linguistik
- Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft
- Ethnologie und Europäische Ethnologie
- Asienbezogene Wissenschaften
- Afrika-, Amerika- und Ozeanienbezogene Wissenschaft
- Islamwissenschaften, Arabistik, Semitistik
- Religionswissenschaft und Judaistik
- Theologie
- Evangelische Theologie
- Katholische Theologie
- Philosophie
- Geschichte der Philosophie
- Wirtschaftswissenschaften
- Wirtschaft- und Sozialgeschichte
- Rechtswissenschaften / Rechtsgeschichte
- Erziehungswissenschaft und Bildungsforschung
- Informationswissenschaft, Bibliotheks- und Archivwissenschaft (Ergänzung)

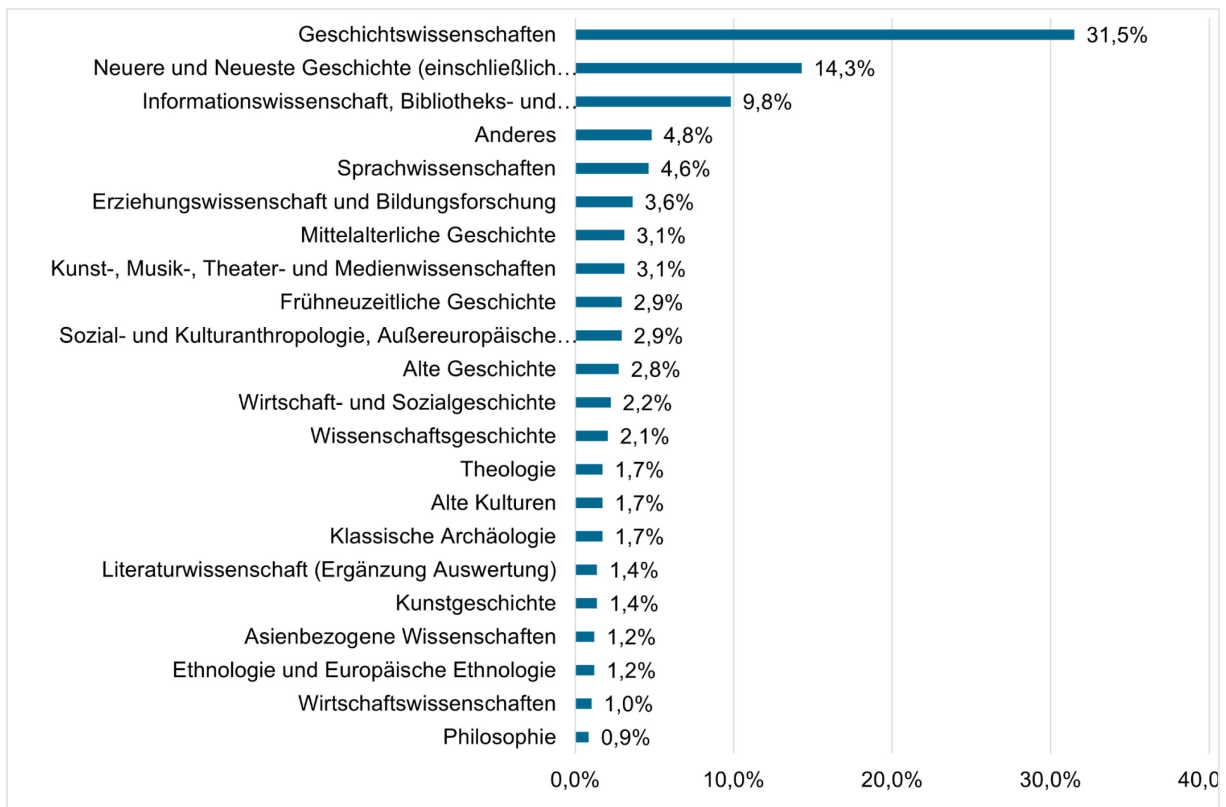


Abb. 3: Verteilung nach historischen Fachdisziplinen | n = 581

Die Liste der Fachdisziplinen beruht auf einem Auszug der DFG-Fachsystematik und beinhaltet speziell Fächer mit historischem Bezug. Die Fachdisziplin Informationswissenschaft, Bibliotheks- und Archivwissenschaft wurde aufgrund der hohen Relevanz der zielgruppenspezifischen Befragung ergänzt.¹¹ Da sich die Umfrage vorwiegend an historisch arbeitende Disziplinen richtet, ist die Verteilung nach Fachbereichen hier wenig überraschend. Über 30 % haben als zugehörige Fachdisziplin Geschichtswissenschaften ausgewählt. Der Fachbereich ist somit am stärksten in der Umfrage vertreten. Auf dem zweiten Platz mit knapp 15 % befindet sich Neuere und Neueste Geschichte und dahinter mit ca. 10 % Informationswissenschaft, Bibliotheks- und Archivwissenschaften. Für die spätere Quantilenauswertung war die Zusammenführung von Fachdisziplinen, die weniger als fünfmal ausgewählt wurden, in Fachkollegien notwendig, um eine ausreichende Anzahl an Proband:innen in verschiedenen disziplinären Gruppierungen zu erhalten.¹² Die Zusammenführung ist in der folgenden Übersicht dargestellt.

¹¹ Für GLAM- und FDM-Mitarbeiter:innen wurde ein Freitextfeld angeboten, da in diesen Bereichen eine genaue Fachzuordnung oftmals nicht möglich ist und sich diese Gruppe in der Regel mehreren Disziplinen zuordnen kann. Diese Auswahlmöglichkeit wurde mit „Anderes“ gekennzeichnet.

¹² Vgl. im Kapitel Zusammenfassung [Abschnitt zur Auswertung nach Wissensleveln in den Fachkollegien, Abb. 87, S. 99.](#)

Disziplinen	Fachkollegium
Ur- und Frühgeschichte	Alte Kulturen
Klassische Philologie	
Ägyptische und Vorderasiatische Altertumswissenschaften	
Historische Linguistik	Sprachwissenschaften
Afrika-, Amerika- und Ozeanienbezogene Wissenschaft	Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft
Islamwissenschaften, Arabistik, Semitistik	
Religionswissenschaft und Judaistik	
Evangelische Theologie	Theologie
Katholische Theologie	
Rechtswissenschaften	Anderes ¹³

Tab. 2: Zuordnung von Fachdisziplinen mit weniger als fünf Proband:innen zum übergeordneten Fachkollegium (außer Rechtswissenschaften)

¹³ Da es sich bei Rechtswissenschaften schon um das Fachkollegium handelt, wurden diese aufgrund der geringen Proband:innenzahl „Anderes“ zugeordnet.

Allgemeine Einschätzung

Ich fühle mich inhaltlich hinreichend sicher im Themenfeld Datenkompetenzen und Forschungsdatenmanagement in meinem Tätigkeitsbereich

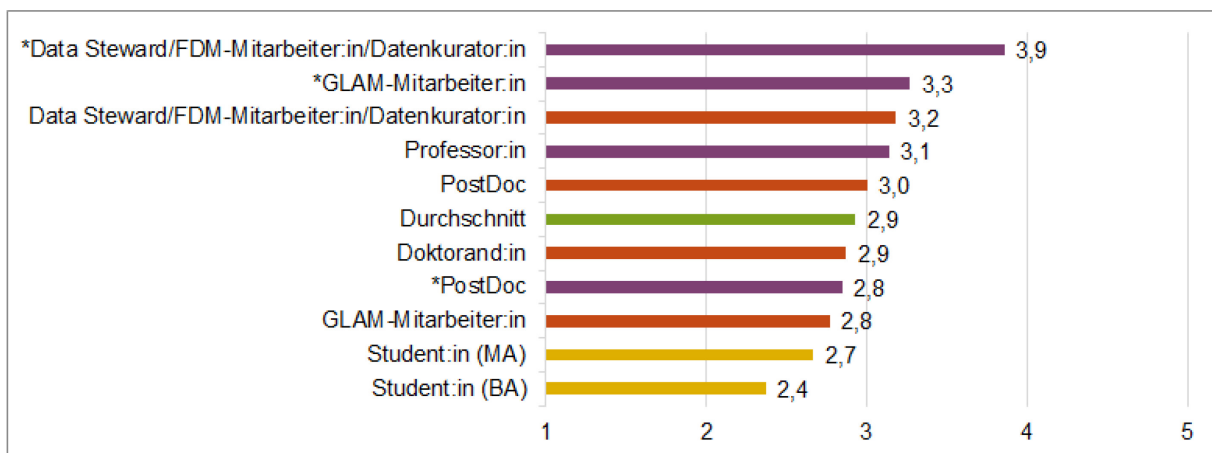
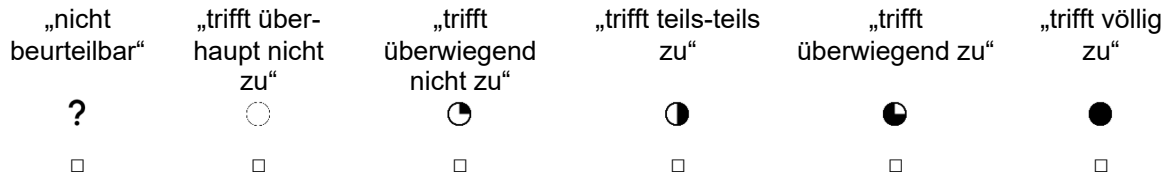


Abb. 4: Verteilung Statusgruppen, Allgemeine Einschätzungsfrage | n = 545

Die Frage nach inhaltlicher Sicherheit im Themenfeld Datenkompetenz und Forschungsdatenmanagement im eigenen Tätigkeitsbereich wird durchschnittlich mit einer 2,9 bewertet. An dieser Stelle ist anzumerken, dass es die einzige inhaltliche Frage ist, welche alle Gruppen aller Niveaustufen exakt gleich formuliert gestellt bekommen haben. Am kompetentesten haben sich die lehrenden FDM-Mitarbeitenden eingeschätzt (3,9) und am niedrigsten die Bachelorstudierenden (2,4). Auf den ersten Blick ist dies eine erwartbare Verteilung. Generell liegen die meisten Statusgruppen sehr nah beieinander, sodass signifikante Unterschiede nicht auszumachen sind. Die meisten Werte liegen zwischen 2,7 und 3,3.

Es folgt ein direkter Vergleich zwischen Erst- und Gesamteinschätzung über alle Kategorien hinweg:

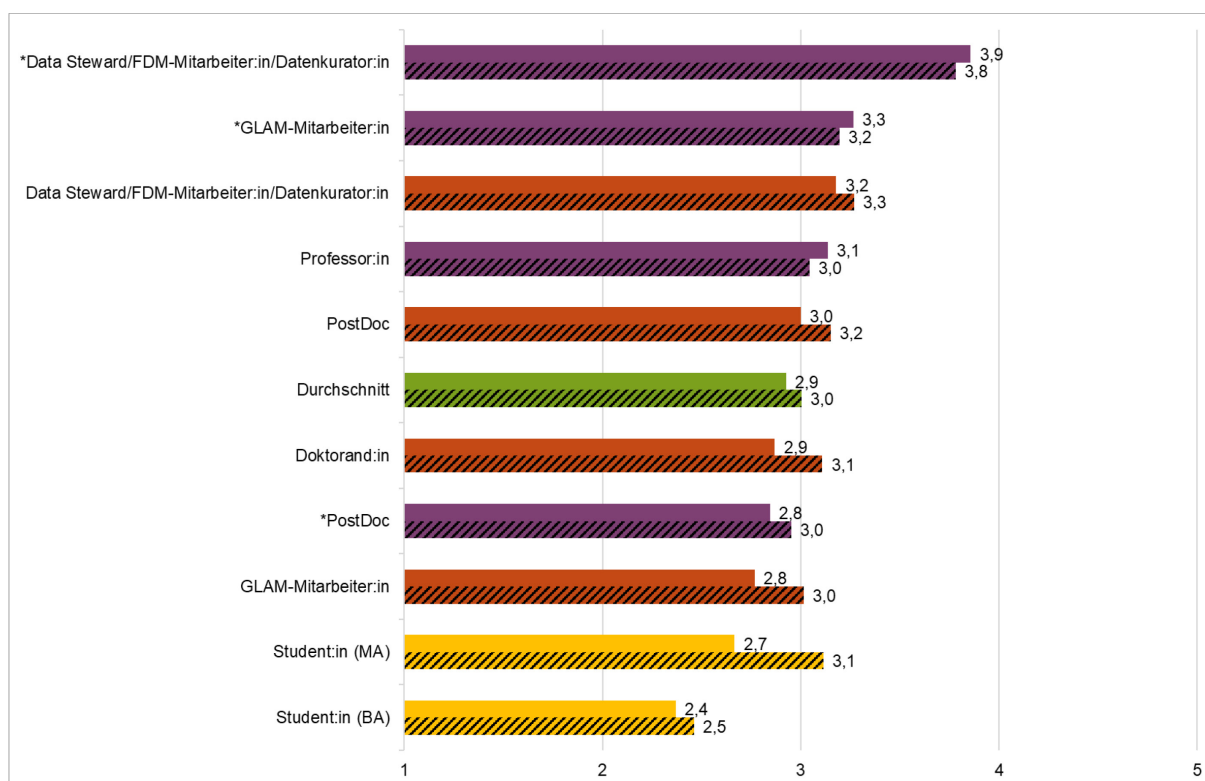


Abb. 5: Verteilung nach Statusgruppen, Vergleich zwischen Erst- und Gesamteinschätzung über alle Kategorien; Ersteinschätzung | n = 545, Gesamteinschätzung | n = 567

Die Balken mit dem schraffierten Muster zeigen die durchschnittliche Gesamtbewertung über alle Kategorien für die jeweilige Statusgruppe. Bei den meisten Gruppen ist zu erkennen, dass die anfängliche Einschätzung mit der Gesamtbewertung nahezu übereinstimmt, mit Unterschieden von $\Delta = 0,1-0,2$. Der überwiegende Teil der Statusgruppen hat sich zu Beginn schlechter eingeschätzt, als hinterher das Ergebnis der Gesamtbewertung ergab. Interessant ist, dass die „Über-“Schätzung der allgemeinen digitalen Kompetenzen nur in der Gruppe der Lehrenden (außer PostDocs) zu beobachten ist. Die größte Differenz besteht in der Statusgruppe der Masterstudierenden mit $\Delta = 0,4$.

Allgemeine FDM-Kenntnisse

Begriff „Forschungsdaten“

- Niveaustufe 1. Ich weiß, was unter dem Begriff „Forschungsdaten“ zu verstehen ist.
- Niveaustufe 2. Ich kann den Begriff „Forschungsdaten“ eigenständig definieren und Beispiele nennen.
- Niveaustufe 3. Ich kann vermitteln, was unter dem Begriff „Forschungsdaten“ zu verstehen ist.

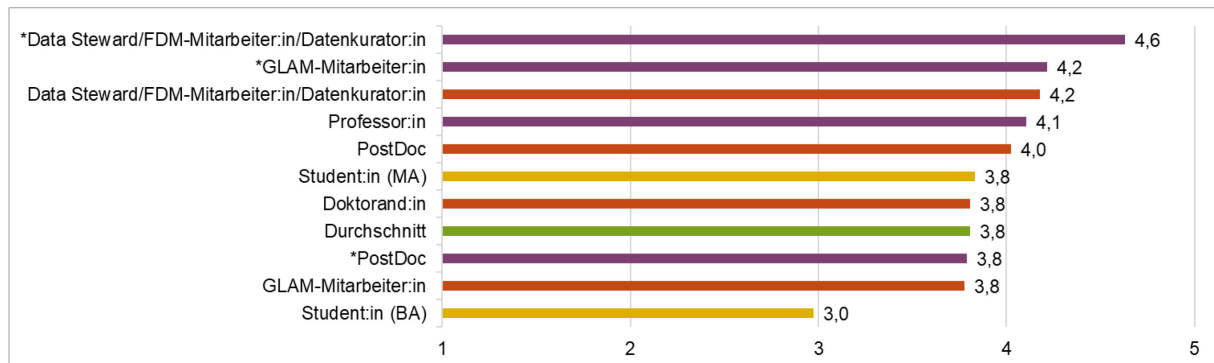


Abb. 6: Verteilung nach Statusgruppen, Abfrage nach Forschungsdatenbegriff | n = 558

Inhaltlich wurde in Frage 1 der Kategorie „Allgemeine FDM-Kenntnisse“ nach dem Begriff „Forschungsdaten“ gefragt. Mit einer durchschnittlichen Bewertung von 3,8 wurde die Frage im Allgemeinen mit einer eher positiven Selbsteinschätzung bewertet. Am kompetentesten schätzen sich die FDM-Mitarbeitenden mit Lehre ein (4,6), am niedrigsten die Bachelorstudierenden (3,0). Bis auf die Studierenden im Bachelor liegen alle Gruppen recht nah beieinander, mit einem Unterschied von maximal $\Delta = 0,4$. Auffallend ist die Einschätzung innerhalb der Niveaustufe 3 (lila). Drei der vier Statusgruppen schätzen sich mit erreichten Werten über 4,0 äußerst kompetent ein. Lediglich die Gruppe der lehrenden PostDocs fällt im Vergleich mit $\Delta = 0,3-0,8$ ab.

Begriff „Forschungsdatenmanagement“

- Niveaustufe 1. Ich kenne den Begriff „Forschungsdatenmanagement“.
- Niveaustufe 2. Ich bin mit dem Begriff „Forschungsdatenmanagement“ vertraut und kann Aspekte dessen benennen.
- Niveaustufe 3. Ich kann vermitteln, was Inhalte und Aspekte des „Forschungsdatenmanagements“ sind.

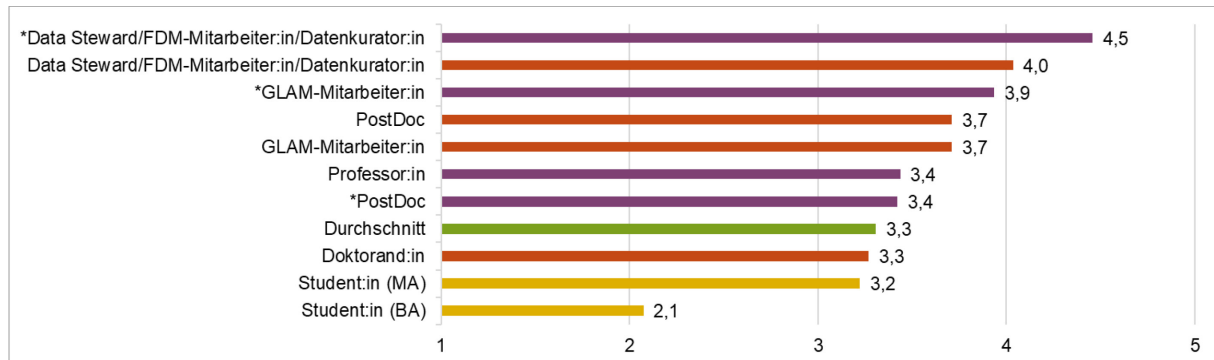


Abb. 7: Verteilung nach Statusgruppen, Abfrage des Begriffs „Forschungsdatenmanagement“ | n = 554

Die Frage nach (Aspekten des) Forschungsdatenmanagement(s) wurde durchschnittlich mit 3,3 (trifft „teils-teils zu“) bewertet. Data Stewards mit und ohne Lehre schätzen sich mit 4,0 und 4,5 am kompetentesten ein. Während die meisten Statusgruppen in der Bewertung relativ nah beieinander liegen, fallen die Studierenden im Bachelor deutlich ab. Mit 2,1 liegen sie $\Delta = 1,2$ unterhalb des Durchschnitts. Interessant ist auch die ähnliche Selbsteinschätzung der Gruppen mit und ohne Lehre. Wobei sich bis auf die Statusgruppe der PostDocs die Lehrenden, wenn auch nicht deutlich, etwas besser einschätzen.

Begriff „Forschungsdatenlebenszyklus“

- Niveaustufe 1. Die verschiedenen Stationen im „Forschungsdatenlebenszyklus“ kenne ich.
- Niveaustufe 2. Die verschiedenen Stationen in einem „Forschungsdatenlebenszyklus“ kann ich benennen und auf meine eigene Forschung übertragen.
- Niveaustufe 3. Ich kann vermitteln, was die verschiedenen Stationen eines „Forschungsdatenlebenszyklus“ sind.

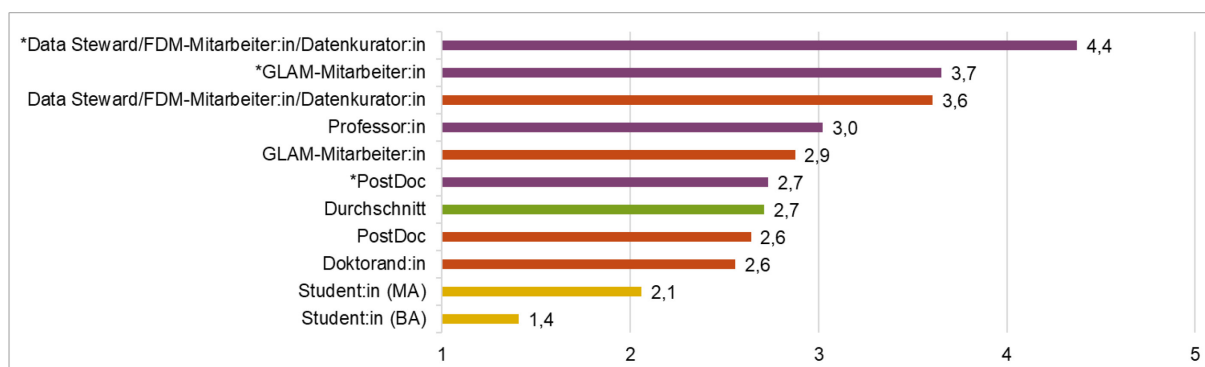


Abb. 8: Verteilung nach Statusgruppen, Abfrage des Begriffs „Forschungsdatenlebenszyklus“ | n = 540

Kenntnisse zu den verschiedenen Stationen des „Forschungsdatenlebenszyklus“ wurden durchschnittlich mit 2,7 bewertet. Im Gegensatz zu den vorherigen Auswertungen dieser Kategorie sieht die Verteilung hier etwas diverser aus. Am kompetentesten schätzen sich FDM-Mitarbeitende mit Lehre (4,4) ein, mit einem Unterschied von $\Delta = 0,7$ zur nächsten Gruppe. Deutlich abgeschlagen sind die Bachelorstudierenden. Mit einer durchschnittlichen Einschätzung von 1,4 liegen sie im unteren Bereich der Skala mit einem der niedrigsten Werte in der gesamten Befragung. Die durchschnittliche Bewertung liegt jedoch auch bei einem Großteil der Statusgruppen unter 3,0. Lehrende liegen mit Werten zwischen 2,7 und 4,4 geschlossen oberhalb des Durchschnitts und haben mitunter die höchsten Werte erreicht. Der Abstand zwischen den Data Stewards und den GLAM-Mitarbeitenden in der Lehre zu den anderen Statusgruppen, die überwiegend an Hochschulen arbeiten, ist mit $\Delta = 0,7-1,7$ bemerkenswert hoch.

Begriff „Datenmanagementplan“

- Niveaustufe 1. Ich weiß, was ein „Datenmanagementplan“ ist.
 Niveaustufe 2. Einen „Datenmanagementplan“ habe ich (mit Hilfe) bereits angefertigt.
 Niveaustufe 3. Ich kann vermitteln, was Inhalte und Aspekte eines „Datenmanagementplans“ sind.

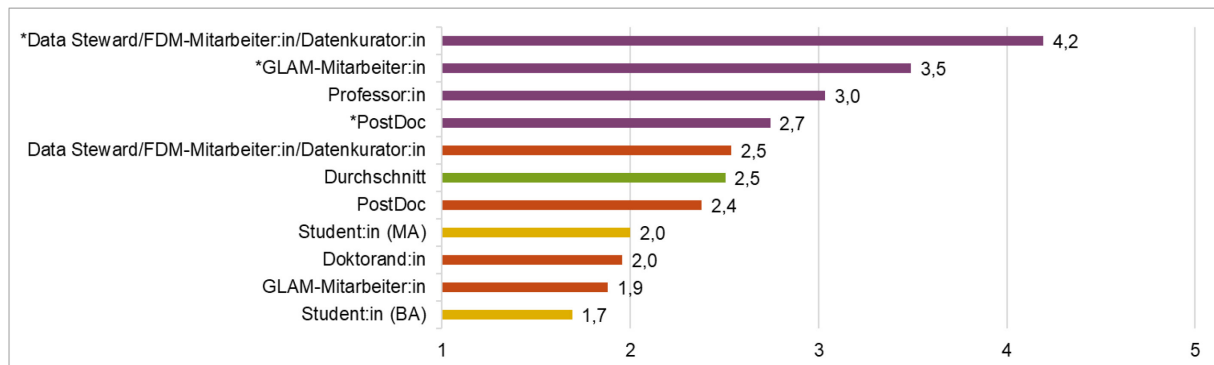


Abb. 9: Verteilung der Statusgruppen, Abfrage des Begriffs „Datenmanagementplan“ | (n = 545)

Bei der Frage nach Datenmanagementplänen liegt die Selbsteinschätzung der verschiedenen Statusgruppen auffällig auseinander. FDM-Mitarbeitende mit Lehre schätzen sich mit 4,2 deutlich am kompetentesten ein, Bachelorstudierende mit 1,7 am niedrigsten. Deutlich ist hier die Verteilung der Statusgruppen der Niveaustufe 3. Diese befindet sich geschlossen oberhalb des Durchschnitts. Eine etwas diversere Verteilung ist in der Niveaustufe 2 sichtbar. Allgemein schätzen sich die Gruppen niedriger ein, mit einer Bewertungsspanne von 1,9–2,5.

Begriff „Metadaten“

- Niveaustufe 1. Ich weiß, was unter dem Begriff „Metadaten“ zu verstehen ist.
 Niveaustufe 2. Ich habe bereits eigenständig Metadaten erstellt.
 Niveaustufe 3. Ich kann vermitteln, was Metadaten sind und welche Funktionen sie erfüllen.

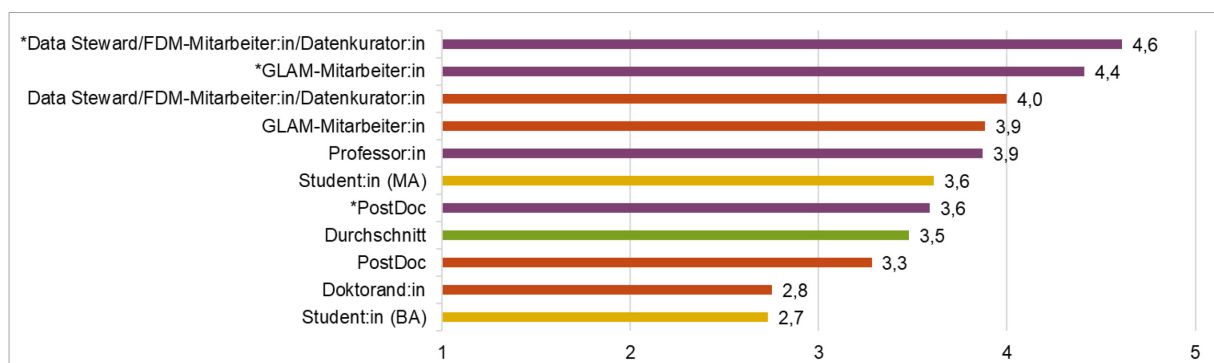


Abb. 10: Verteilung der Statusgruppe, Abfrage des Begriffs „Metadaten“ | (n = 549)

Die Frage nach dem Begriff (und der Funktion von) „Metadaten“ wurde durchschnittlich mit 3,5 bewertet. Mit 4,6 schätzen sich die FDM-Mitarbeitenden mit Lehre am kompetentesten ein, erreichen mit dieser Bewertung auch einen der höchsten Werte der Umfrage. Acht von zehn

Statusgruppen liegen im Durchschnitt über 3,0, lediglich Doktorand:innen und Studierende im Bachelor erreichen mit jeweils 2,8 bzw. 2,7 niedrigere Werte.

Begriff „Open Science“

Niveaustufe 1. Das Konzept „Open Science“ ist mir bekannt.

Niveaustufe 2. Ich kann das Konzept „Open Science“ anhand meiner Daten wiedergeben.

Niveaustufe 3. Ich kann vermitteln, was man im Allgemeinen unter dem Konzept „Open Science“ versteht.

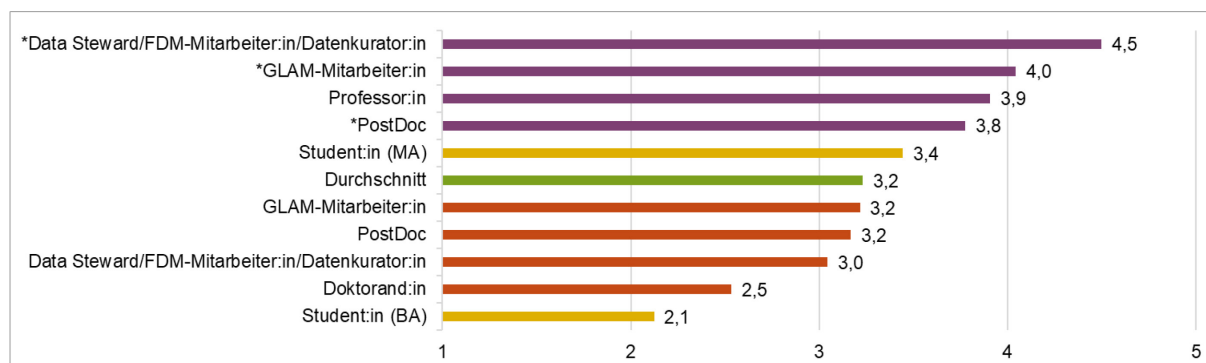


Abb. 11: Verteilung der Statusgruppen, Abfrage des Begriffs „Open Science“ | (n = 540)

Die Frage nach dem Konzept von „Open Science“ wurde im Durchschnitt mit 3,2 bewertet. Am kompetentesten hat sich die Niveaustufe 3 eingeschätzt. Mit Werten von 3,8–4,5 liegen die Statusgruppen geschlossen oberhalb des Durchschnitts. Ein gegenteiliges Bild zeichnet sich in Niveaustufe 2 ab. Alle Gruppen dieser Niveaustufe liegen entweder auf oder unter dem Durchschnitt. Am niedrigsten in dieser Gruppe haben sich die Doktorand:innen eingeschätzt. Interessant ist die unterschiedliche Einschätzung der Gruppe der Studierenden. Während Bachelorstudierende sich mit 2,1 insgesamt am niedrigsten eingeschätzt haben, befinden sich die Studierenden im Master mit einer Bewertung von 3,4 oberhalb des Durchschnitts über allen Gruppen der Niveaustufe 2.

Begriffe „Open“ und „FAIR“

- Niveaustufe 1. Ich kenne den Unterschied zwischen den Prinzipien „Open“ und „FAIR.“
- Niveaustufe 2. Was der Unterschied zwischen den Prinzipien „Open“ und „FAIR“ ist, kann ich anhand von Beispieldaten aufzeigen.
- Niveaustufe 3. Ich kann vermitteln, was der Unterschied zwischen den Prinzipien „Open“ und „FAIR“ ist.

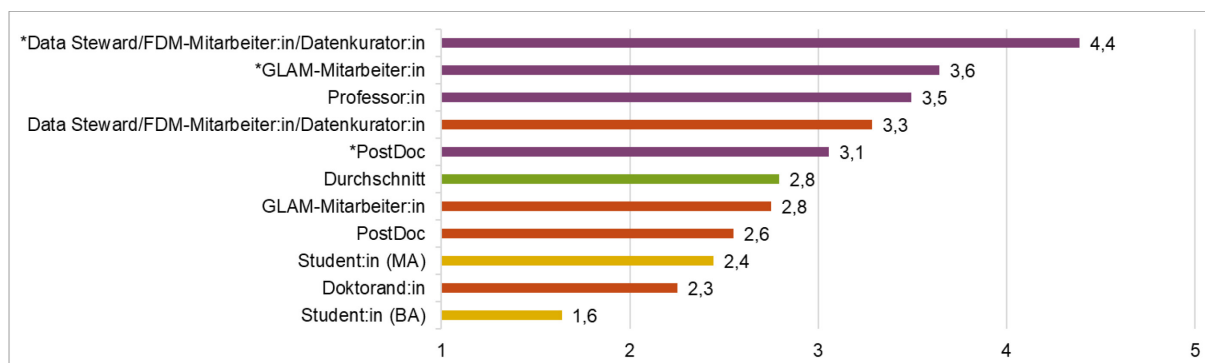


Abb. 12: Verteilung der Statusgruppe, Abfrage der Begriffe „Open“ und FAIR | (n = 548)

Die letzte Frage in der Kategorie „Allgemeine FDM-Kenntnisse“ behandelt die Unterschiede zwischen den „Open“- und „FAIR“-Prinzipien. Durchschnittlich wurde diese Frage mit 2,8 beantwortet. Mit 4,4 schätzen sich die FDM-Mitarbeitenden mit Lehre am kompetentesten und die Bachelorstudierenden mit 1,6 am niedrigsten ein. Allgemein zeigt sich auch eine eher diverse Verteilung der Statusgruppen, wobei auch hier ein deutlicher Unterschied zwischen Lehrenden (lila) und nicht-Lehrenden (orange) vorzufinden ist. Markant ist die geringe Selbsteinschätzung der PostDocs ohne Lehre v. a. mit einem Abstand von $\Delta = 0,5$ zu PostDocs mit Lehre.

Kurzzusammenfassung der Kategorie Allgemeine FDM-Kenntnisse

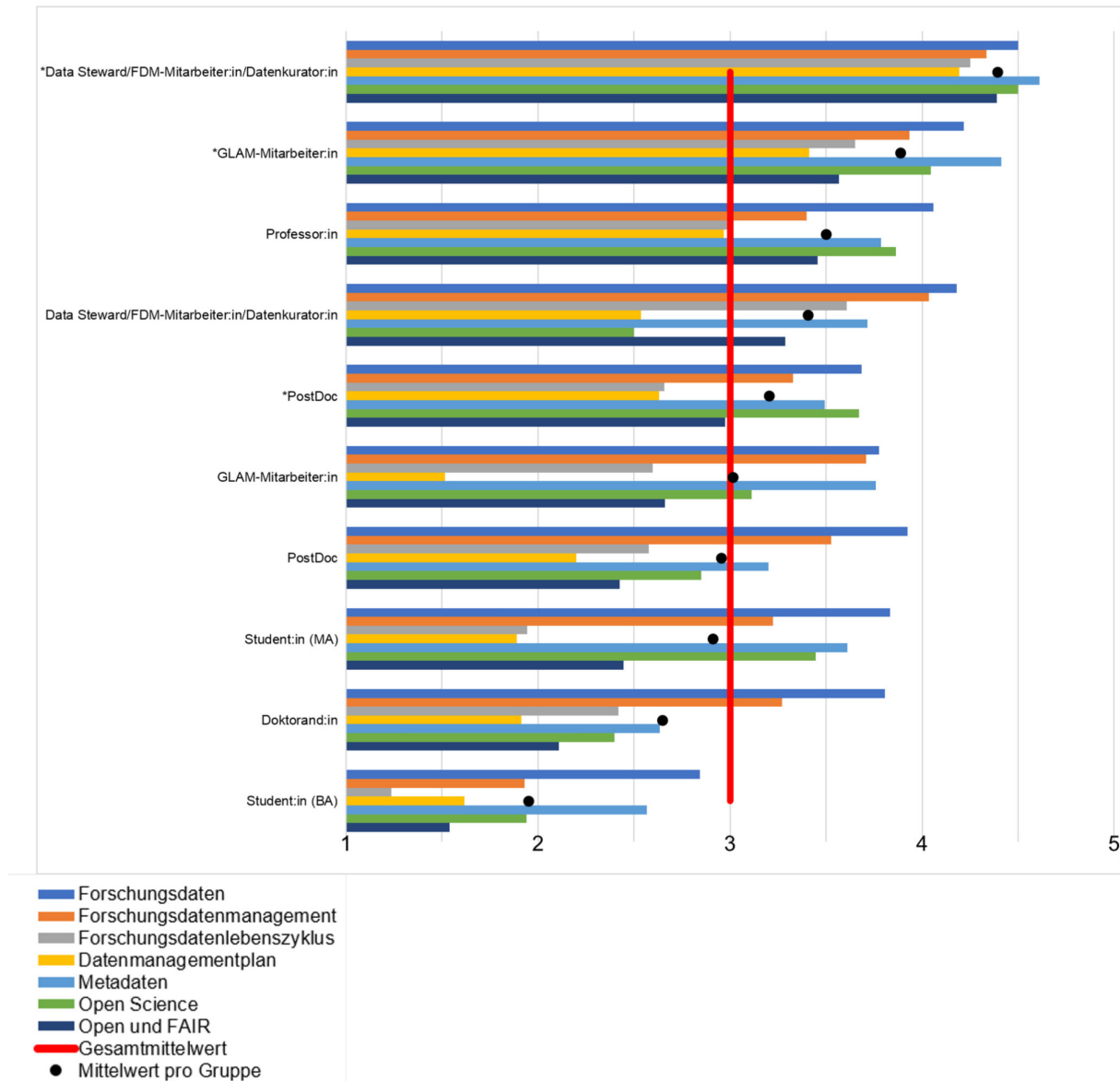


Abb. 13: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben der Kategorie „Allgemeine FDM-Kenntnisse“ | n = 568

In dieser Kategorie geht es um allgemeine Begrifflichkeiten, Konzepte und Prinzipien im Forschungsdatenmanagement, die noch keine fachspezifisch geschichtswissenschaftlichen Aspekte adressieren. Durchweg haben sich die FDM-Mitarbeitende mit Lehre am kompetentesten eingeschätzt. Allgemein ist deutlich zu erkennen, dass sich die Niveaustufe 3 geschlossen über dem Durchschnitt befindet und mitunter die höchsten Werte erreicht. Mit einer durchschnittlichen Bewertung von 3,0 haben sich die Proband:innen im Schnitt neutral eingeschätzt. Studierenden im Bachelor und Master liegen insgesamt unterhalb des Durchschnitts. Vor allem die Bachelorstudierenden verzeichnen mitunter die geringsten Werte in der gesamten Kategorie.

Datenerschließung

Qualitative und quantitative Daten

- Niveaustufe 1. Mir ist der Unterschied zwischen qualitativen und quantitativen Daten bekannt.
- Niveaustufe 2. Die Unterschiede zwischen qualitativen und quantitativen Daten sowie die unterschiedliche methodische Verarbeitungsweise sind mir bekannt.
- Niveaustufe 3. Ich kann vermitteln, worin der Unterschied zwischen qualitativen und quantitativen Daten liegt.

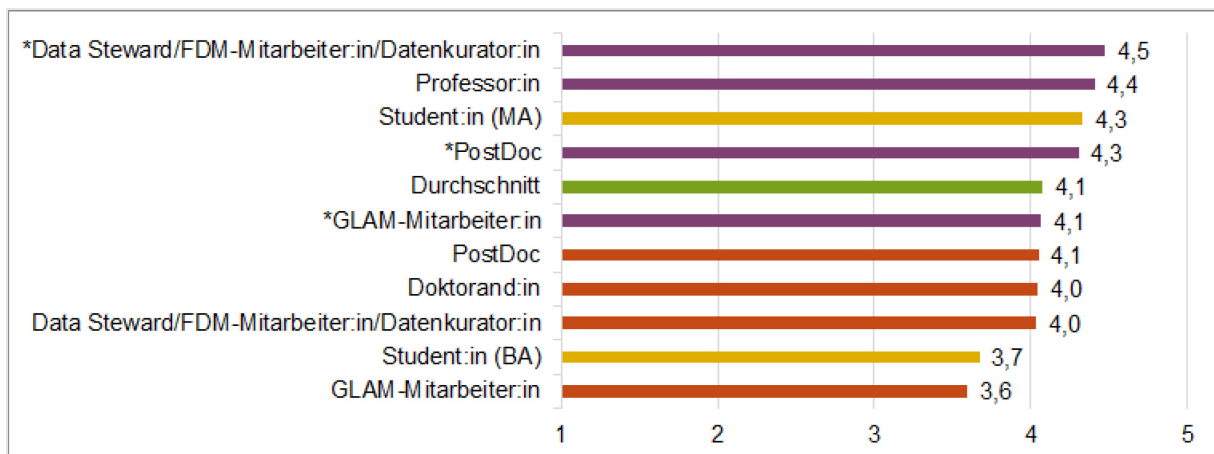


Abb. 14: Verteilung der Statusgruppen, Abfrage qualitative und quantitative Daten | n = 562

Mit einer Durchschnittsbewertung von 4,1 wird die Frage nach der Unterscheidung von qualitativen und quantitativen Daten von den Teilnehmenden relativ hoch in der Selbsteinschätzung bewertet. Die Gruppe der lehrenden FDM-Mitarbeitenden schätzt sich mit 4,5 am besten ein. Bemerkenswert ist hierbei die relativ hohe Selbsteinschätzung der Studierenden im Master und der Unterschied zwischen GLAM-Mitarbeitenden mit und ohne Lehre ($\Delta = 0,5$).

Plattformen zur Datenrecherche

- Niveaustufe 1. Ich kenne unterschiedliche Plattformen zur Datenrecherche in meinem Fachgebiet.
- Niveaustufe 2. Bei der Datenrecherche nutze ich unterschiedliche fachspezifische Plattformen.
- Niveaustufe 3. Ich kann vermitteln, welche Plattformen es zur Datenrecherche in meinem Fachgebiet gibt.

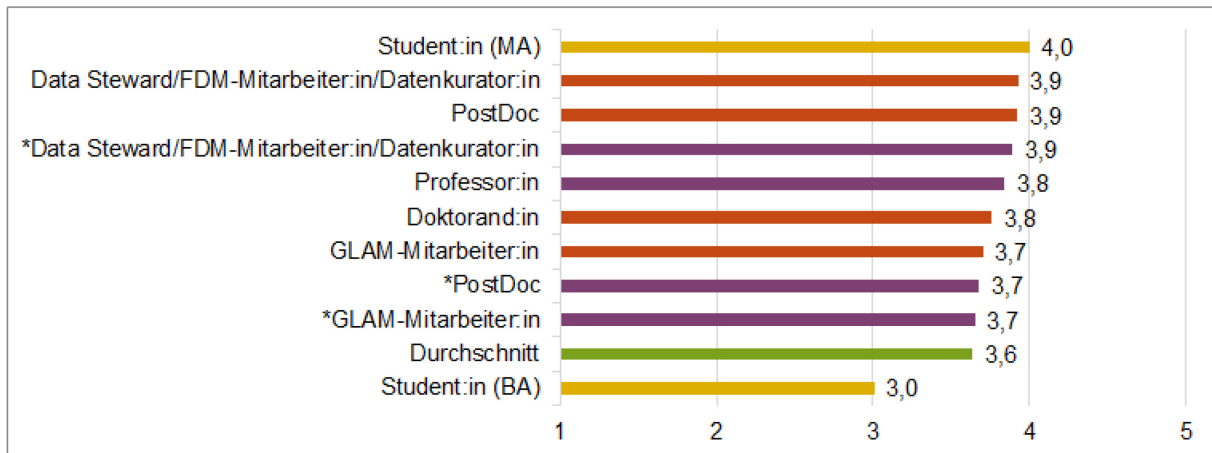


Abb. 15: Verteilung der Statusgruppen, Abfrage Plattformen zur Datenrecherche | n = 552

Bei der Frage nach Plattformen zur Datenrecherche liegt die durchschnittliche Selbsteinschätzung bei 3,6. Generell liegen alle Gruppen bis auf Studierende im Bachelor sehr eng beieinander. Auffällig ist hier der große Unterschied innerhalb der Niveaustufe 1 zwischen Bachelorstudierenden, die sich am geringsten mit 3,0 eingeschätzt haben und Masterstudierenden, die sich mit 4,0 am besten eingeschätzt haben.

Recherchewerkzeuge und -strategien

- Niveaustufe 1. Recherchewerkzeuge und -strategien sind mir bekannt.
- Niveaustufe 2. Ich wende spezielle Recherchewerkzeuge und -strategien an.
- Niveaustufe 3. Ich kann vermitteln, welchen Nutzen der Einsatz von Recherchewerkzeugen und -strategien hat.

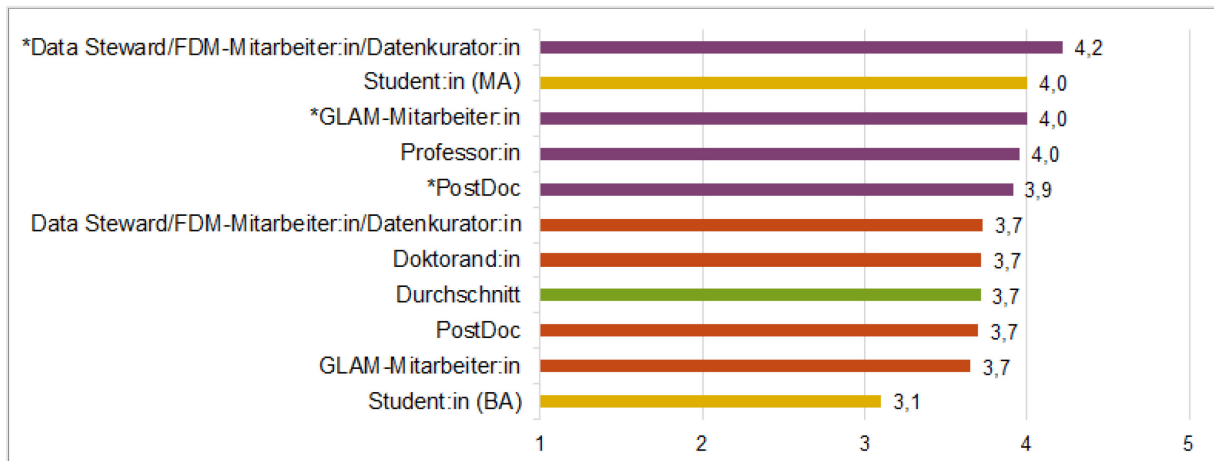


Abb. 16: Verteilung der Statusgruppen, Abfrage nach Recherchewerkzeugen und -strategien | n = 555

Die Frage nach Recherchewerkzeugen und -strategien wurde von den Teilnehmenden im Durchschnitt mit 3,7 bewertet. Die Niveaustufe 2 befindet sich genau auf dem Gesamtmittelwert der Frage. Niveaustufe 3 liegt mit $\Delta = 0,2-0,5$ geschlossen darüber. Die positive Selbsteinschätzung der Masterstudierenden (4,0) und die zurückhaltende Einschätzung der Bachelorstudierenden (3,1) sind am auffälligsten.

Strategien zur Fehlervermeidung

- Niveaustufe 1. Mir sind Strategien zur Fehlervermeidung beim Umgang mit Daten bekannt.
- Niveaustufe 2. Ich nutze aktiv Strategien zur Fehlervermeidung beim Umgang mit Daten.
- Niveaustufe 3. Ich kann vermitteln, welche Strategien es zur Fehlervermeidung beim Umgang mit Daten gibt.

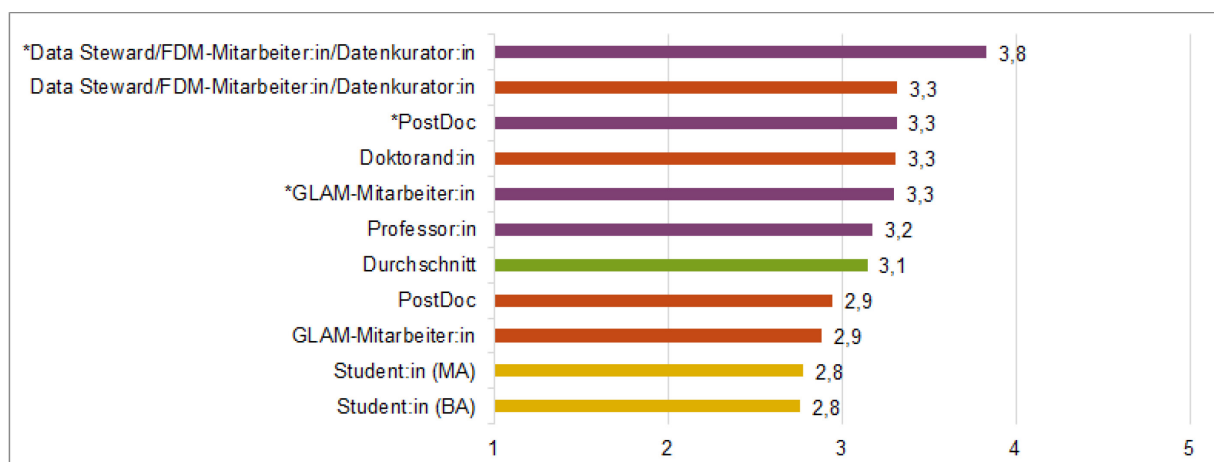


Abb. 17: Verteilung der Statusgruppen, Abfrage nach Strategien zur Fehlervermeidung | n = 536

Mit einer durchschnittlichen Bewertung von 3,1 („trifft teils-teils zu“) bewegt sich bei der Frage nach Strategien zur Fehlervermeidung beim Umgang mit Daten die Selbsteinschätzung der Teilnehmenden nach Gruppen zwischen 2,8 und 3,8. Allgemein eine durchschnittliche Einschätzung, bei der sich die einzelnen Gruppen nur um $\Delta = 0,1-0,2$ unterscheiden. Die FDM-Mitarbeitenden mit Lehre stechen hierbei hervor, mit einem Unterschied von $\Delta = 0,5$ zur nächsten Gruppe. Die Gruppe der Lehrenden (lila) hat sich am kompetentesten im Vergleich zu den anderen Gruppen eingeschätzt und befindet sich oberhalb des Durchschnitts. Vor allem der Unterschied zwischen Lehrenden und Nicht-Lehrenden der Gruppen PostDocs und GLAM-Mitarbeitende ($\Delta = 0,4$) ist bemerkenswert.

Kurzzusammenfassung der Kategorie „Datenerschließung“

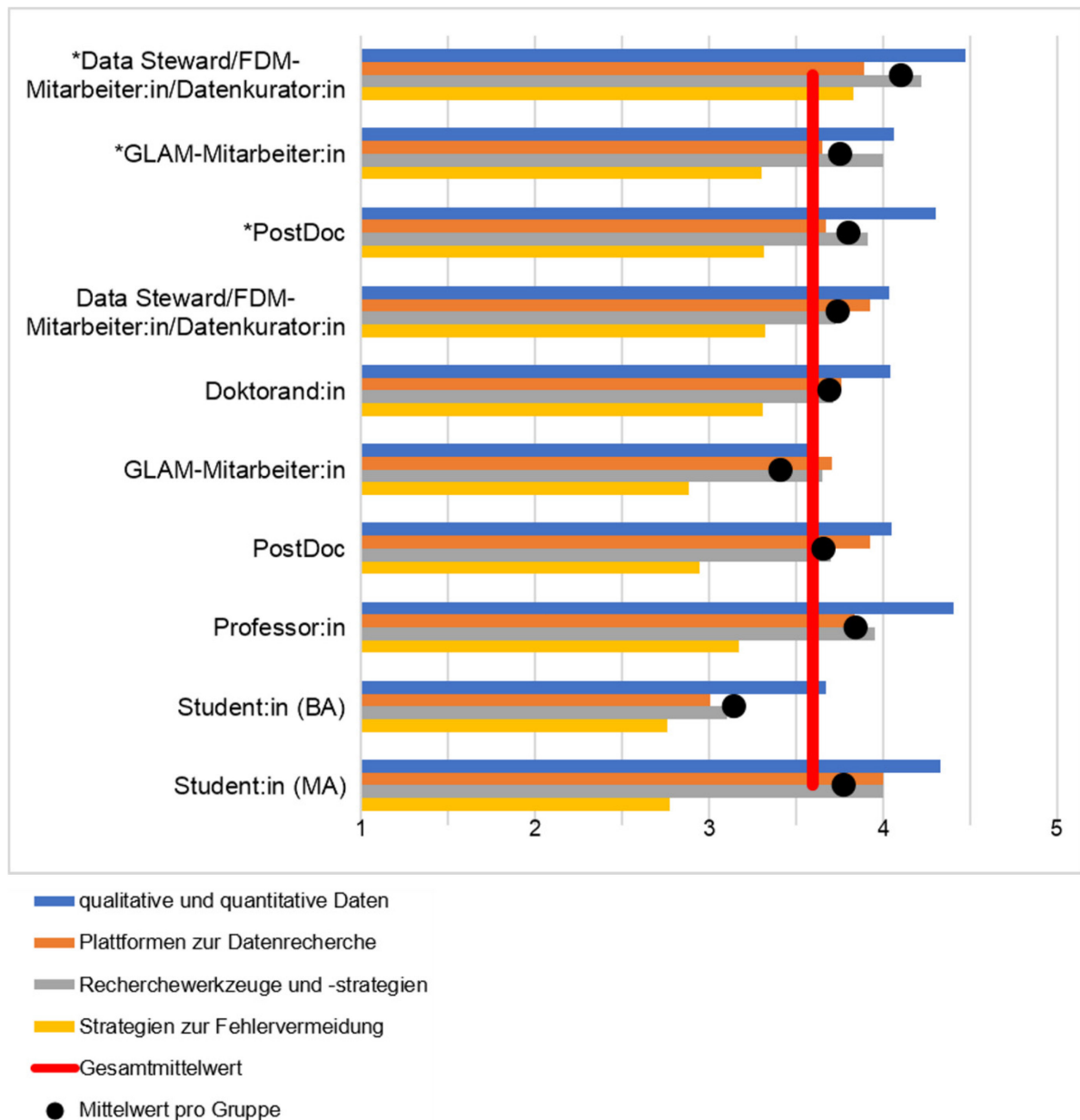


Abb. 18: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben der Kategorie „Datenerschließung“ | n = 566

Die durchschnittliche Bewertung über alle Fragen in der Kategorie „Datenerschließung“ liegt bei 3,6. Durchweg am kompetentesten haben sich FDM-Mitarbeitende mit Lehre eingeschätzt. Allgemein liegen alle Gruppen recht nah beieinander, sodass ein signifikanter Unterschied nicht feststellbar ist. Die Bachelorstudierenden und GLAM-Mitarbeitenden ohne Lehre ($\bar{x} = 3,1$ und $\bar{x} = 3,4$) haben sich am niedrigsten eingeschätzt.

Datenorganisation

Strategien zum Ordnen von Daten

- Niveaustufe 1. Ich kenne Strategien, um Daten zu ordnen und zu strukturieren.
- Niveaustufe 2. Ich verfolge Strategien, um meine Daten zu ordnen und zu strukturieren.
- Niveaustufe 3. Ich kann vermitteln, welche Strategien es gibt, um Daten zu ordnen und zu strukturieren.

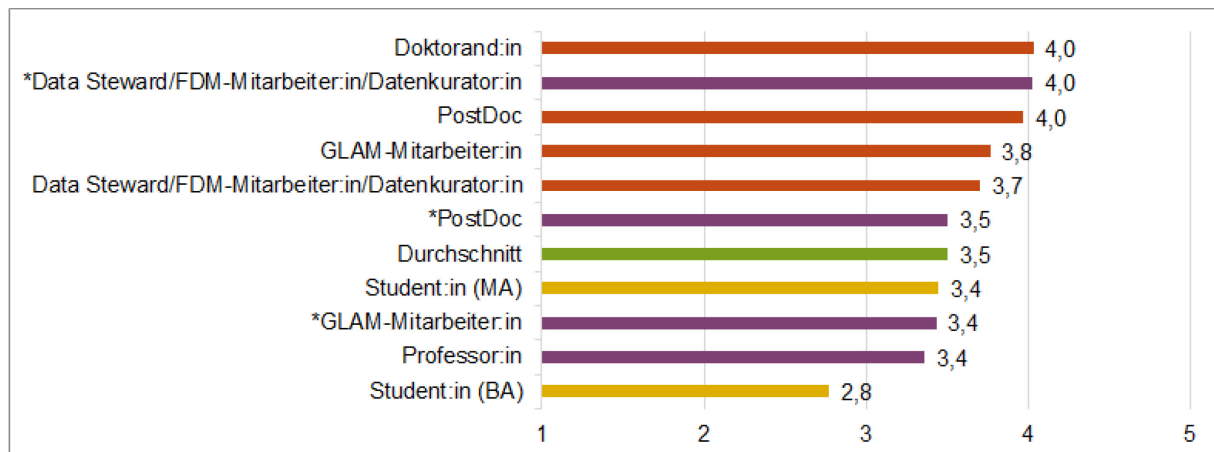


Abb. 19: Verteilung der Statusgruppen, Abfrage nach Strategien zum Ordnen von Daten | n = 555

Der Durchschnitt bei der Frage nach Strategien zum Ordnen und Strukturieren von Daten liegt bei 3,5 (zwischen „trifft teils-teils zu“ und „trifft überwiegend zu“). Während sich die Studierenden im Bachelor mit 2,8 am niedrigsten eingeschätzt haben, teilen sich die Gruppen Doktorand:innen, FDM-Mitarbeitende mit Lehre und PostDocs ohne Lehre mit 4,0 den ersten Platz. Im überwiegenden Teil der Kategorien liegt die Gruppe der Lehrenden (lila) meist geschlossen oberhalb des Durchschnitts. Hier zeigt sich eine andere Verteilung. Lediglich FDM-Mitarbeitende mit Lehre haben sich besser als der Durchschnitt eingeschätzt. Die anderen Statusgruppen liegen knapp darüber oder darunter. GLAM-Mitarbeitende und Professor:innen haben sich am niedrigsten eingeschätzt. Ein gegenteiliges Bild zeichnet sich bei der Gruppe der Niveaustufe 2 ab: sie liegt geschlossen über dem Durchschnitt mit $\Delta = 0,2-0,5$.

Forschungssoftware

- Niveaustufe 1. Ich kenne (fachspezifische) Forschungssoftware (z. B. Analysewerkzeuge, Datenmanagementsysteme, Visualisierungstools).
- Niveaustufe 2. Ich wende (fachspezifische) Forschungssoftware (z. B. Analysewerkzeuge, Datenmanagementsysteme, Visualisierungstools) aktiv an.
- Niveaustufe 3. Ich kann vermitteln, welchen Nutzen (fachspezifische) Forschungssoftware (z. B. Analysewerkzeuge, Datenmanagementsysteme, Visualisierungstools) haben.

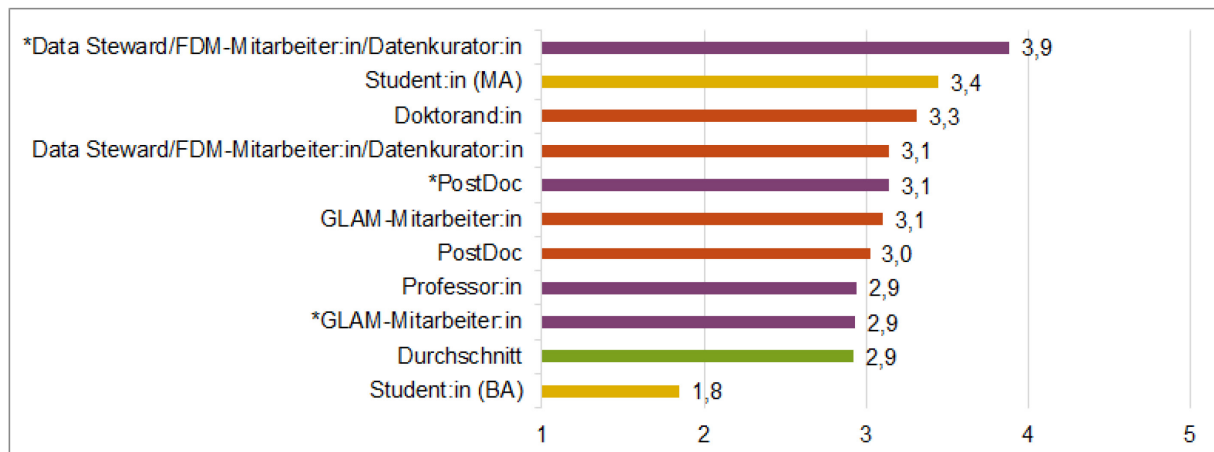


Abb. 20: Verteilung der Statusgruppen, Abfrage nach Forschungssoftware | n = 553

Die Teilnehmenden haben sich hier eher zurückhaltend eingeschätzt. Mit 2,9 liegt der Durchschnitt im Mittelfeld. Während sich die meisten Gruppen zwischen 2,9 und 3,4 eingeordnet haben, gibt es zwei Ausreißer. Mit 1,8 haben sich die Studierenden im Bachelor mit einem Unterschied von $\Delta = 1,1$ zur nächsten Gruppe am niedrigsten eingeschätzt. Die lehrenden FDM-Mitarbeitenden haben sich am kompetentesten eingeschätzt und heben sich mit $\Delta = 0,5$ von den nächstgelegenen Masterstudierenden deutlich ab.

Kurzzusammenfassung der Kategorie „Datenorganisation“

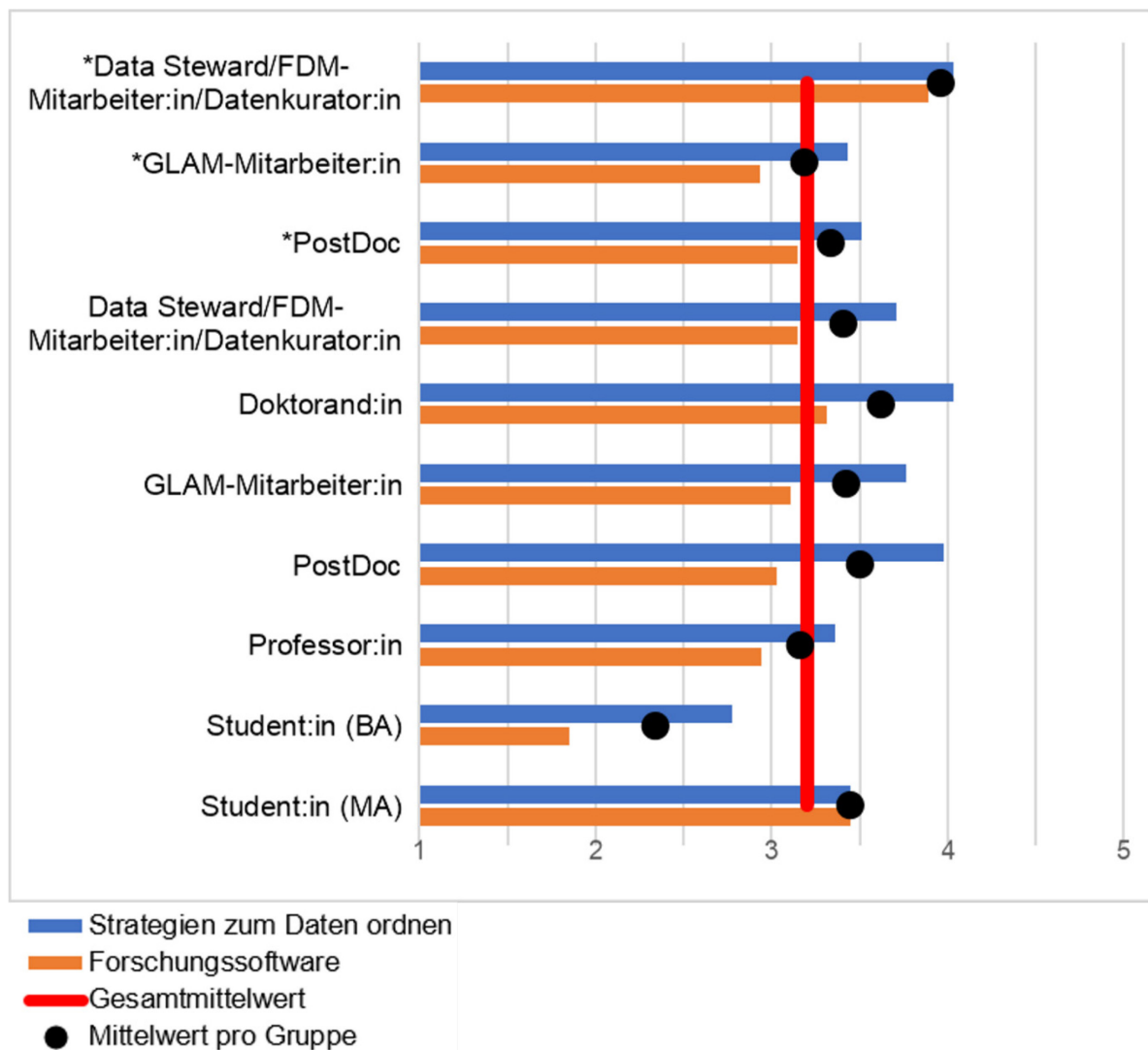


Abb. 21: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben der Kategorie „Datenorganisation“ | n = 563

Die Kategorie „Datenorganisation“ wurde im Durchschnitt mit 3,2 („trifft teils-teils zu“) beantwortet. Gefragt wurde nach Strategien zum Ordnen von Daten und (fachspezifischer) Forschungssoftware. Mit 4,0 haben sich lehrende FDM-Mitarbeitende am besten eingeschätzt und liegen mit $\Delta = 0,8$ deutlich über dem Durchschnitt. Wirklich signifikante Unterschiede liegen zwischen den einzelnen Statusgruppen nicht vor. Interessant ist die hohe Selbsteinschätzung der Masterstudierenden. Mit 3,4 liegen sie $\Delta = 0,2$ oberhalb des Durchschnitts und vor dem Großteil der anderen Statusgruppen.

Datenevaluation

Digitale fachspezifische Forschungsmethoden

- Niveaustufe 1. Mir sind digitale fachspezifische Forschungsmethoden bekannt.
- Niveaustufe 2. Ich integriere digitale fachspezifische Forschungsmethoden in meine Arbeit.
- Niveaustufe 3. Ich kann vermitteln, was der Nutzen von digitalen fachspezifischen Forschungsmethoden ist.

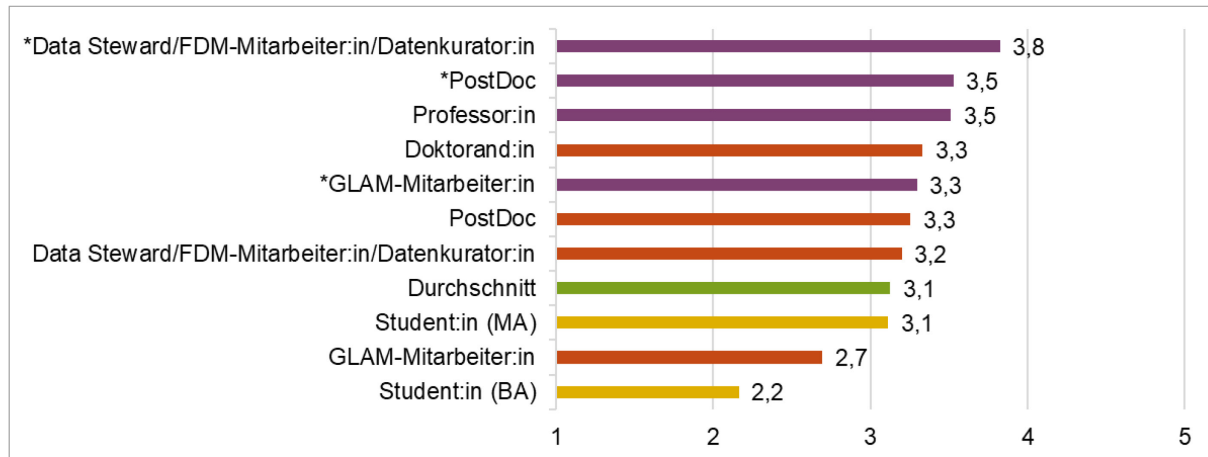


Abb. 22: Verteilung der Statusgruppen, Abfrage nach digitalen, fachspezifischen Forschungsmethoden | n = 546

Die Frage nach digitalen fachspezifischen Forschungsmethoden wird im Durchschnitt mit 3,1 („trifft teils-teils zu“) bewertet. Die Selbsteinschätzung der Gruppen bewegt sich zwischen 2,2 (Studierende im Bachelor) und 3,8 (FDM-Mitarbeitende mit Lehre). Es ist ein deutlicher Unterschied zwischen lehrenden und nicht-lehrenden Statusgruppen erkennbar. Die Gruppe der Lehrenden befindet sich geschlossen über dem Durchschnitt und erreicht mit Werten von 3,3 - 3,8 die höchste Selbsteinschätzung. Bei der Gruppe der Nicht-Lehrenden sieht die Verteilung diverser aus. Während Doktorand:innen, PostDocs und FDM-Mitarbeitende vergleichbar hohe Werte erzielen, fallen GLAM-Mitarbeitende ohne Lehre eher ab.

Visualisierungsarten

Niveaustufe 1. Ich kenne unterschiedliche Visualisierungsarten.

Niveaustufe 2. Ich nutze aktiv unterschiedliche Visualisierungsarten und kenne die Vor- und Nachteile dieser.

Niveaustufe 3. Ich kann vermitteln, was die Vor- und Nachteile von unterschiedlichen Visualisierungsarten sind.

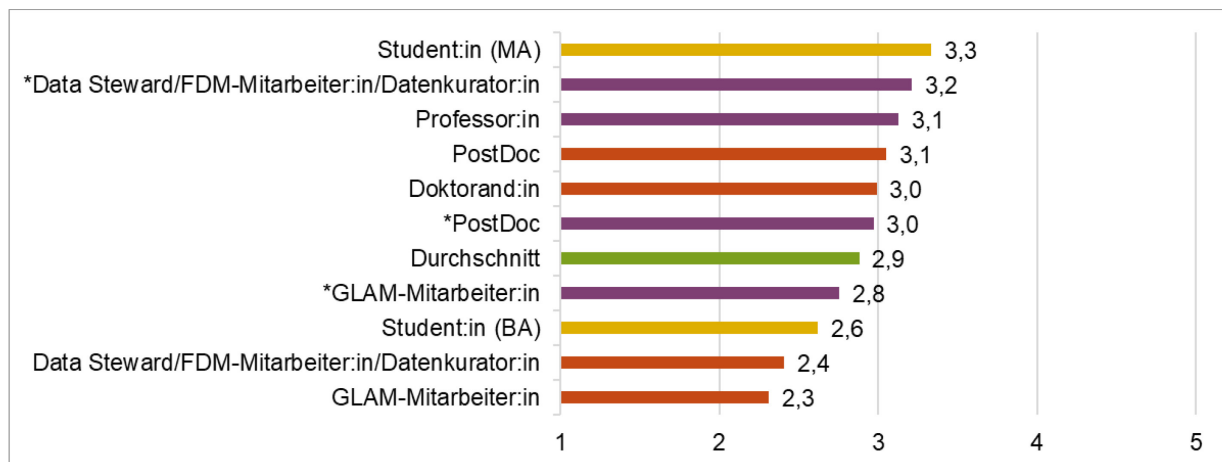


Abb. 23: Verteilung nach Statusgruppen, Abfrage nach Visualisierungsarten | n = 547

Bei der Frage nach Visualisierungsarten zeigt sich im Vergleich zu den meisten anderen Fragen eine ungewöhnliche Verteilung. Allgemein liegt die Verteilung zwischen 2,3 (GLAM-Mitarbeitende ohne Lehre) und 3,3 (Masterstudierende) eher eng beieinander. Vor allem die FDM- und GLAM-Mitarbeitenden ohne Lehre liegen mit jeweils $\Delta = 0,5$ und $\Delta = 0,6$ deutlich unter dem Durchschnitt. Der Durchschnittswert fällt mit 2,9 eher niedrig aus.

Zu folgenden digital gestützten Analysemethoden würde ich gerne mehr erfahren:

- Strukturanalyse
- Räumliche Analyse
- Analyse von Relationen
- Netzwerkanalyse
- Inhaltsanalyse
- Stilistische Analyse
- Statistik
- Simulation
- Diskursanalyse
- Topic modeling
- Distant reading
- Natural Language Processing
- Photogrammetrie
- Sonstige _____

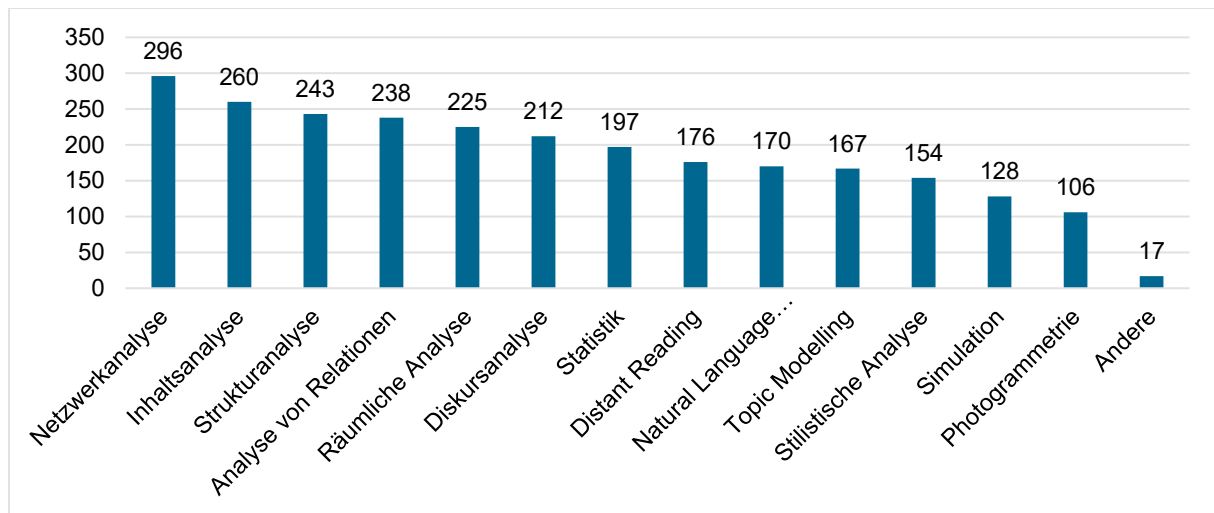


Abb. 24: Abfrage nach Analysemethoden, Antworten: 2589 | n = 529

Bei der Abfrage, zu welchen Analysemethoden mehr Informationen gewünscht werden, war es möglich, keine, eine oder mehrere Auswahlmöglichkeiten auszuwählen. Insgesamt haben 529 von 568 Teilnehmenden abgestimmt. Die Methode der Netzwerkanalyse wurde mit 296 Stimmen am häufigsten ausgewählt, gefolgt von der Inhaltsanalyse (260) und Strukturanalyse mit jeweils 243 am dritthäufigsten.

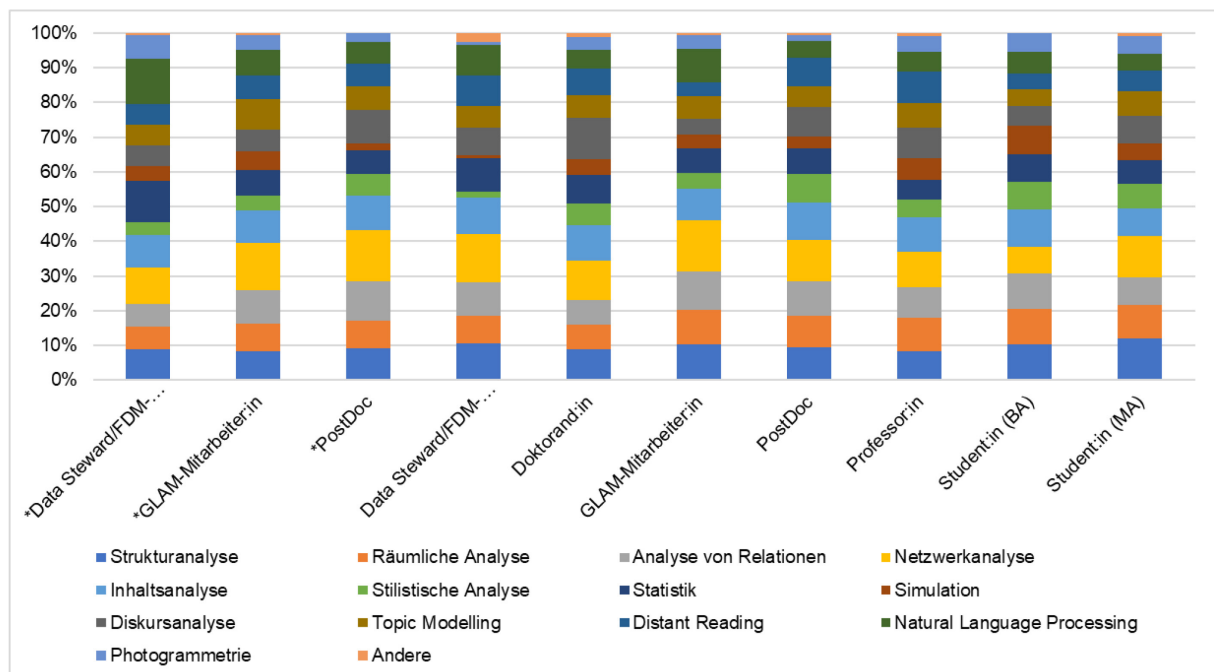


Abb. 25: Verteilung nach Statusgruppen und Auswahloption von Analysemethoden, Antworten: 2589 | n = 529

Mit Blick auf die Verteilung der Statusgruppe wird deutlich, dass acht von zehn Gruppen sich am meisten für die Netzwerkanalyse interessieren. Insgesamt handelt es sich um eine recht homogene Verteilung: Bei den Bachelorstudierenden sind mehr Informationen zur Inhaltsanalyse und bei Doktorand:innen zur Diskursanalyse gewünscht. Topic Modeling und Stilistische

Analyse ist bei den FDM-Mitarbeitenden ohne Lehre wenig gewünscht. PostDocs mit Lehre haben ebenfalls ein geringeres Interesse an Topic Modelling. Siebzehn Personen haben die Auswahlmöglichkeit „Andere“ angewählt. Während sich zwei Proband:innen mehr Informationen zu Stematologie wünschen, handelt es sich bei den verbleibenden fünfzehn Angaben um Einmalnennungen. Darunter vertreten sind u. a. Bildanalyse, Videografie und Eye Tracking.

Kurzzusammenfassung Kategorie „Datenevaluation“

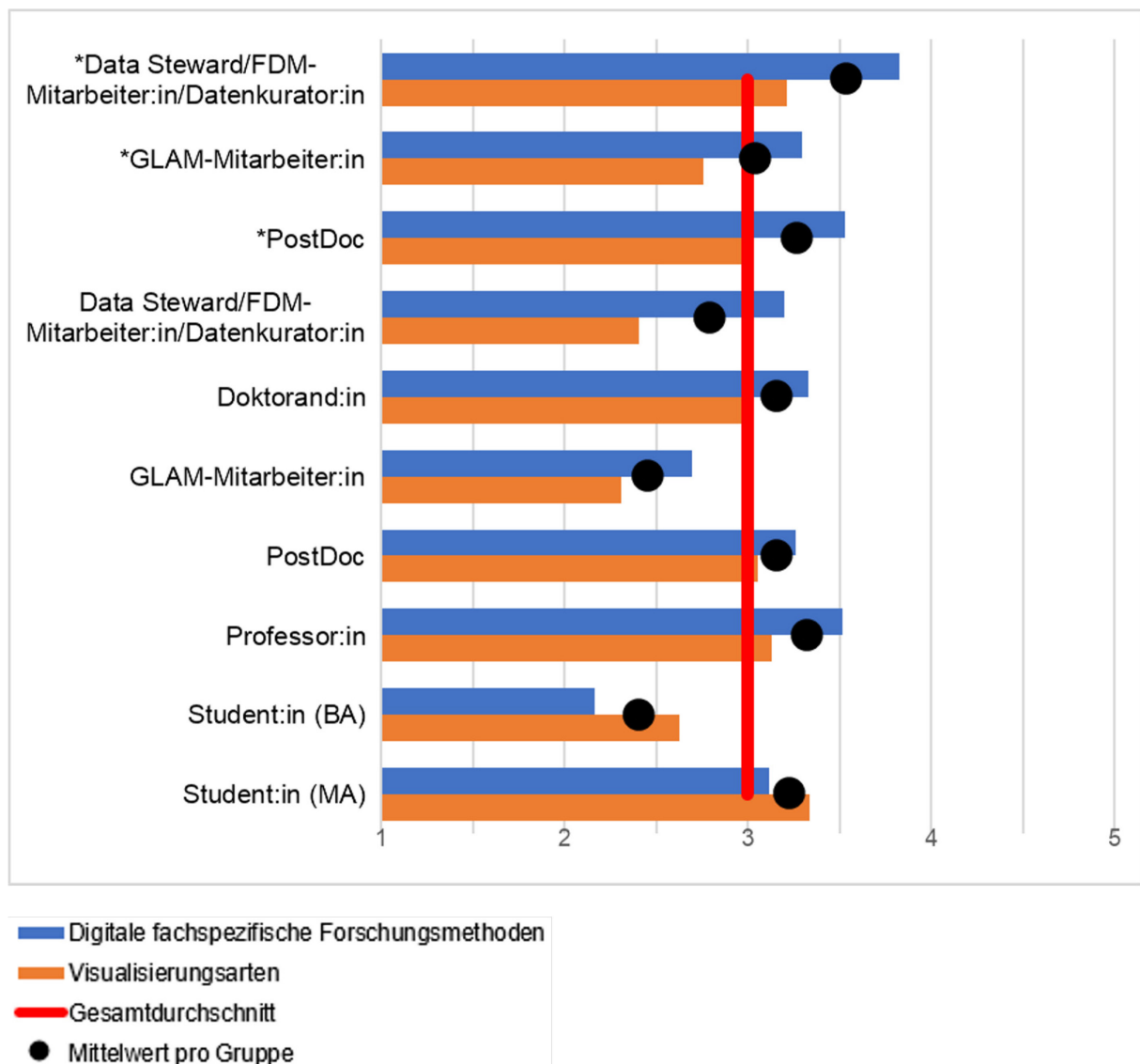


Abb. 26: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Datenevaluation“ | n = 557

Die Kategorie „Datenevaluation“ wurde im Durchschnitt mit 3,0 bewertet. Gefragt wurde nach digitalen, fachspezifischen Forschungsmethoden und Visualisierungsarten. Die Gesamtansicht zeigt ein eher verhaltenes Bild. Am kompetentesten schätzen sich die FDM-Mitarbeitenden mit Lehre (3,5) und am niedrigsten die GLAM-Mitarbeitenden ohne Lehre und die Ba-

chelorstudierenden (2,4) ein. Das Ergebnis der Auswahllisten zeigt ein deutliches Interesse an verschiedenen Datenanalysemethoden. 529 von 568 Personen haben an der fakultativen Abfrage teilgenommen. Netzwerk-, Inhalts- und Strukturanalyse wurden mit jeweils über 245 Stimmen am häufigsten ausgewählt.

Datenaufbereitung

Begriff „Datendokumentation“

- Niveaustufe 1. Ich weiß, was unter dem Begriff „Datendokumentation“ zu verstehen ist.
- Niveaustufe 2. Die Inhalte und Aspekte einer „Datendokumentation“ kann ich aufzählen und umsetzen.
- Niveaustufe 3. Ich kann vermitteln, was die Inhalte und Aspekte einer „Datendokumentation“ sind.

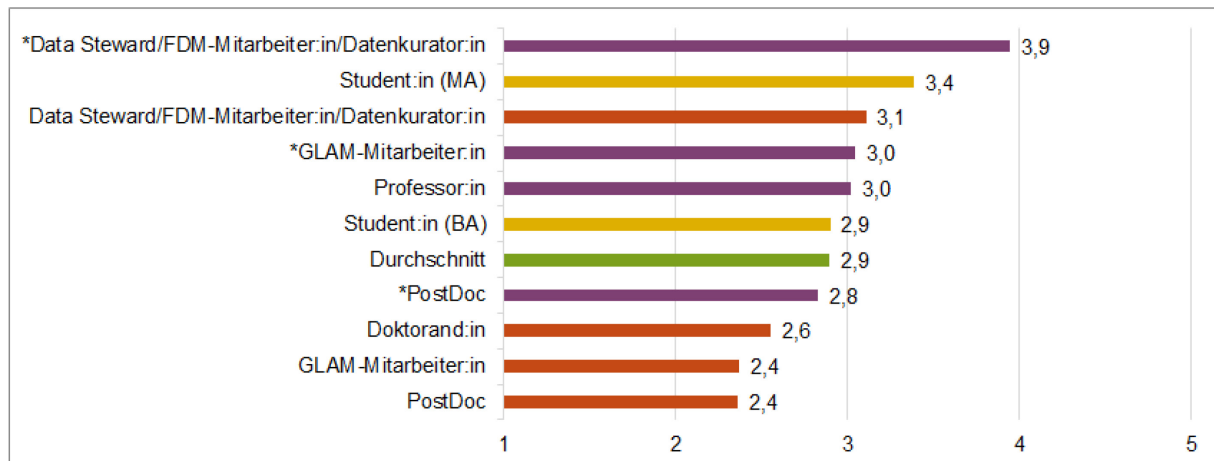


Abb. 27: Verteilung der Statusgruppen, Abfrage des Begriffs „Datendokumentation“ | n = 549

Bei der Frage nach der „Datendokumentation“ liegt der Durchschnitt bei 2,9. Am kompetentesten haben sich FDM-Mitarbeitende mit Lehre (3,9) eingeschätzt und am niedrigsten PostDocs und GLAM-Mitarbeiter:innen ohne Lehre (je 2,4). Auffällig ist hier die Selbsteinschätzung der Niveaustufe 2 (orange): Bis auf die FDM-Mitarbeitenden befindet sich die Gruppe mit Werten von 2,4 bis 2,6 geschlossen unterhalb des Durchschnitts. Masterstudierende haben sich mit 3,4 kompetenter eingeschätzt.

Begriff „Multimodalität“

- Niveaustufe 1. Den Begriff „Multimodalität“ im Kontext von Daten kenne ich.
- Niveaustufe 2. Ich kann vermitteln, was der Begriff „Multimodalität“ im Kontext von Daten bedeutet.
- Niveaustufe 3. Mir ist die Multimodalität von Daten bewusst und somit auch die unterschiedlichen Aufbereitungsmethoden.

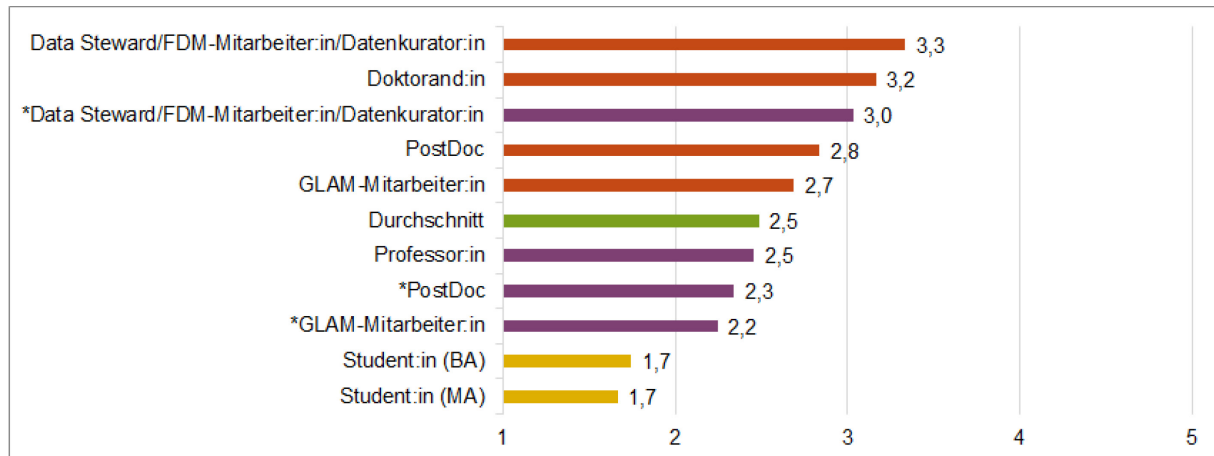


Abb. 28: Verteilung der Statusgruppen, Abfrage des Begriffs „Multimodalität“ | n = 541

Die Frage nach „Multimodalität“ im Kontext von Daten wurde durchschnittlich mit 2,5 bewertet. Am kompetentesten schätzen sich FDM-Mitarbeitende ohne Lehre (3,3) und am niedrigsten Bachelor- und Masterstudierende mit jeweils 1,7 ein. Hier zeichnet sich im Vergleich zur ersten Frage der Kategorie „Datenaufbereitung“ ein anderes Bild ab. Während bei der ersten Frage die Statusgruppen der Niveaustufe 2 vermehrt unterhalb des Durchschnitts angesiedelt waren, zeigt sich hier eine gegenteilige Verteilung. Alle vier Gruppen der Niveaustufe 2 befinden sich geschlossen oberhalb des Durchschnitts und erzielen die höchsten Werte bei dieser Frage. Auffällig ist die deutlich geringere Selbsteinschätzung der Studierenden. Mit Blick auf die Gruppe der Lehrenden fällt auf, dass bis auf die FDM-Mitarbeitenden alle Lehrenden unterhalb des Durchschnitts mit Werten zwischen 2,2 und 2,5 liegen.

Werkzeuge und Methoden zur Datendokumentation

- Niveaustufe 1. Ich kenne Werkzeuge und Methoden zur Datendokumentation.
 Niveaustufe 2. Ich kann Werkzeuge und Methoden zur Datendokumentation anwenden.
 Niveaustufe 3. Ich kann vermitteln, welche Werkzeuge und Methoden es zur Datendokumentation gibt.

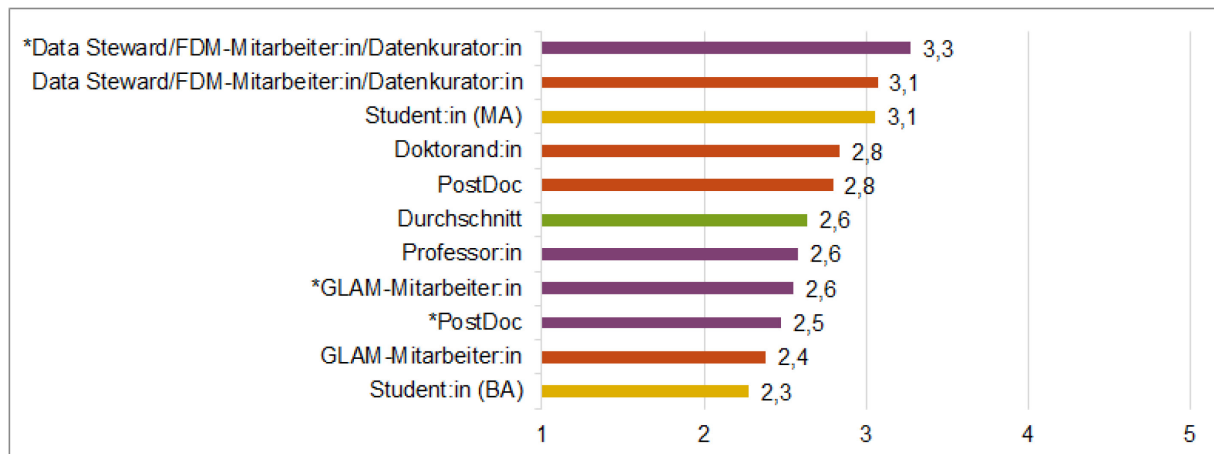


Abb. 29: Verteilung der Statusgruppen, Abfrage nach Werkzeugen und Methoden zur Datendokumentation | n = 550

Bei der Frage nach Werkzeugen und Methoden zur Datendokumentation liegt der Durchschnitt bei 2,6. Am kompetentesten mit 3,3 schätzen sich FDM-Mitarbeitende mit Lehre und am niedrigsten Bachelorstudierende mit 2,3 ein. Während bis auf die FDM-Mitarbeitenden mit Lehre die Niveaustufe 3 geschlossen unterhalb des Durchschnitts vorzufinden ist, zeigt sich ein gegenteiliges Bild bei den Statusgruppen der Niveaustufe 2. Bis auf die GLAM-Mitarbeitenden ohne Lehre befinden sich hier alle Gruppen oberhalb des Durchschnitts. Allgemein liegen jedoch alle Gruppen nah beieinander, mit Unterschieden von $\Delta = 0,1$ – $0,3$. Auch zwischen dem höchsten und dem niedrigsten Wert liegt lediglich ein Unterschied von $\Delta = 1,0$. Die wohl auffälligste Differenz in der Einschätzung innerhalb einer Statusgruppe zeigt sich bei den Bachelor- und Masterstudierenden. Obwohl sich die Bachelorstudierenden hier am niedrigsten eingeschätzt haben, liegen sie nur $\Delta = 0,1$ unterhalb der Gruppe der GLAM-Mitarbeitenden ohne Lehre und $\Delta = 0,2$ unterhalb der Einschätzung der PostDocs mit Lehre.

Begriff „Ontologie“

- Niveaustufe 1. Ich weiß, was eine Ontologie ist.
- Niveaustufe 2. Ich habe bereits mit fachspezifischen Ontologien gearbeitet.
- Niveaustufe 3. Ich kann vermitteln, wie eine Ontologie aufgebaut ist und genutzt wird.

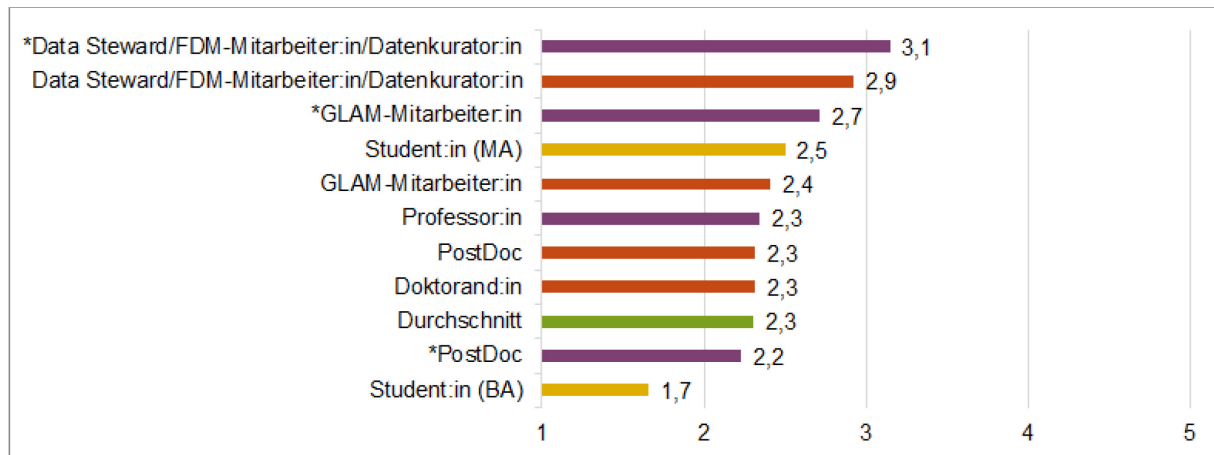


Abb. 30: Verteilung der Statusgruppen, Abfrage des Begriffs „Ontologie“ | n = 531

Die Frage nach Ontologien wird durchschnittlich zurückhaltend mit 2,3 bewertet. Den höchsten Wert erreicht die Gruppe der FDM-Mitarbeitenden mit Lehre mit 3,1, den niedrigsten die Bachelorstudierenden mit 1,7. Im Vergleich zu den anderen Fragen der Kategorie „Datenaufbereitung“ wird die Frage nach Ontologien am niedrigsten bewertet. Ein Großteil der befragten Gruppen liegt zwischen 2 und 3. Bis auf die Bachelorstudierenden schätzen sich die anderen Statusgruppen recht ähnlich ein, mit Unterschieden von $\Delta = 0,1-0,2$.

Begriff „Metadatenstandards“

- Niveaustufe 1. Ich kenne „Metadatenstandards“ für mein Fachgebiet.
- Niveaustufe 2. Ich achte bei der Erstellung und Bereitstellung von Metadaten auf „Metadatenstandards“ für mein Fach.
- Niveaustufe 3. Ich kann vermitteln, welche Metadatenstandards für mein/e Fachgebiet/e relevant sind.

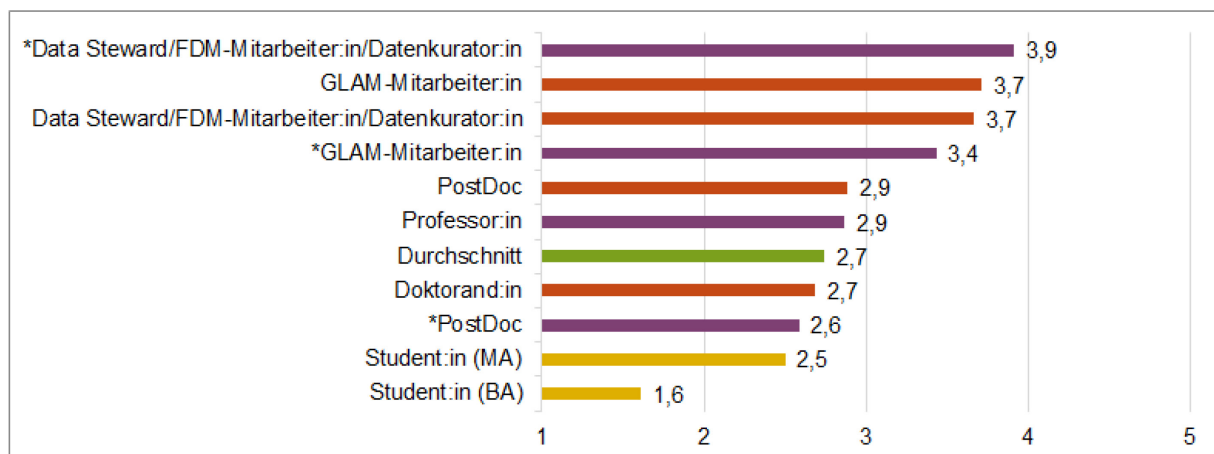


Abb. 31: Verteilung der Statusgruppen, Abfrage des Begriffs „Metadatenstandards“ | n = 526

Die letzte inhaltliche Frage des Themenblocks „Datenaufbereitung“ fragt nach Metadatenstandards für das eigene Fach. Auch wenn die durchschnittliche Einschätzung von 2,7 eher im unteren Bereich der Skala liegt, wird deutlich, dass sich vor allem die Gruppen GLAM und FDM mit und ohne Lehre mit Werten zwischen 3,4 und 3,9 deutlich kompetenter im Vergleich zu den anderen Statusgruppen einschätzen. Studierende schätzen sich deutlich niedriger ein, mit Werten zwischen 1,6 und 2,5.

Zu folgenden Metadatenstandards wünsche ich mir mehr Informationen:

- TEI (Texte)
- CEI (Urkunden)
- EAD (Archivdaten)
- DDI (Interviews)
- IIF (Bild-, Audio-, Videomaterial)
- CMIF (Briefe)
- METS/MODS, DC (bibliographische Informationen)
- CIDOC CRM (Kulturgüter und Museumsdokumentation)
- Sonstige _____

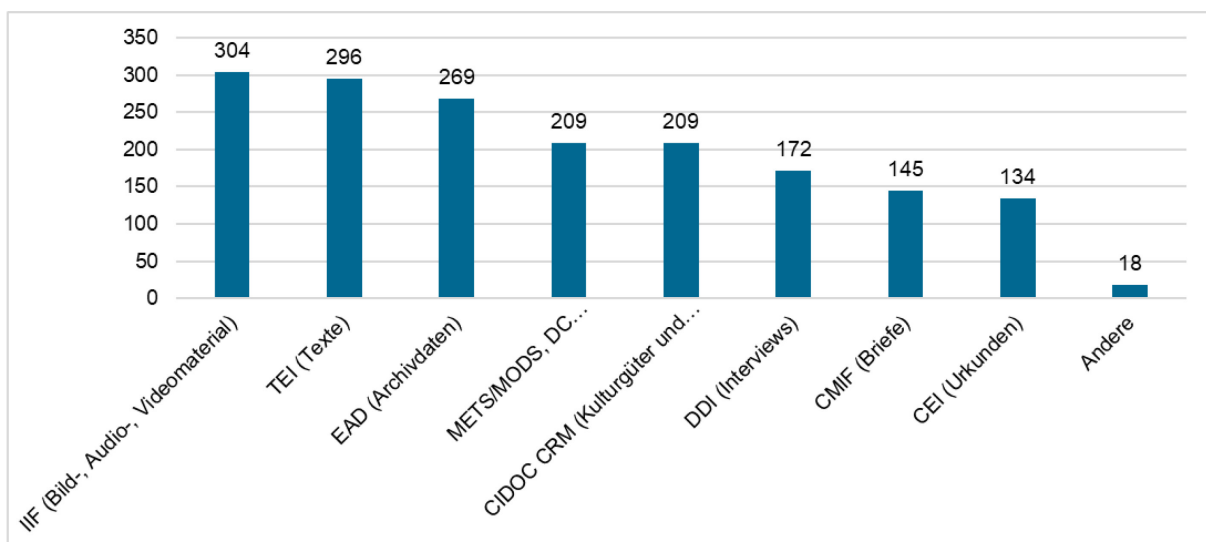


Abb. 32: Abfrage nach Metadatenstandards; Antworten: 1756 | n = 497

Insgesamt gaben 497 Personen mindestens ein Votum ab. Mit 304 Stimmen wurde der Metadatenstandard IIF am häufigsten ausgewählt. Deutliches Interesse besteht auch an den Standards TEI und EAD. Mit 134 Stimmen wurde an dem Metadatenstandard für Urkunden (CEI) das geringste Interesse bekundet.

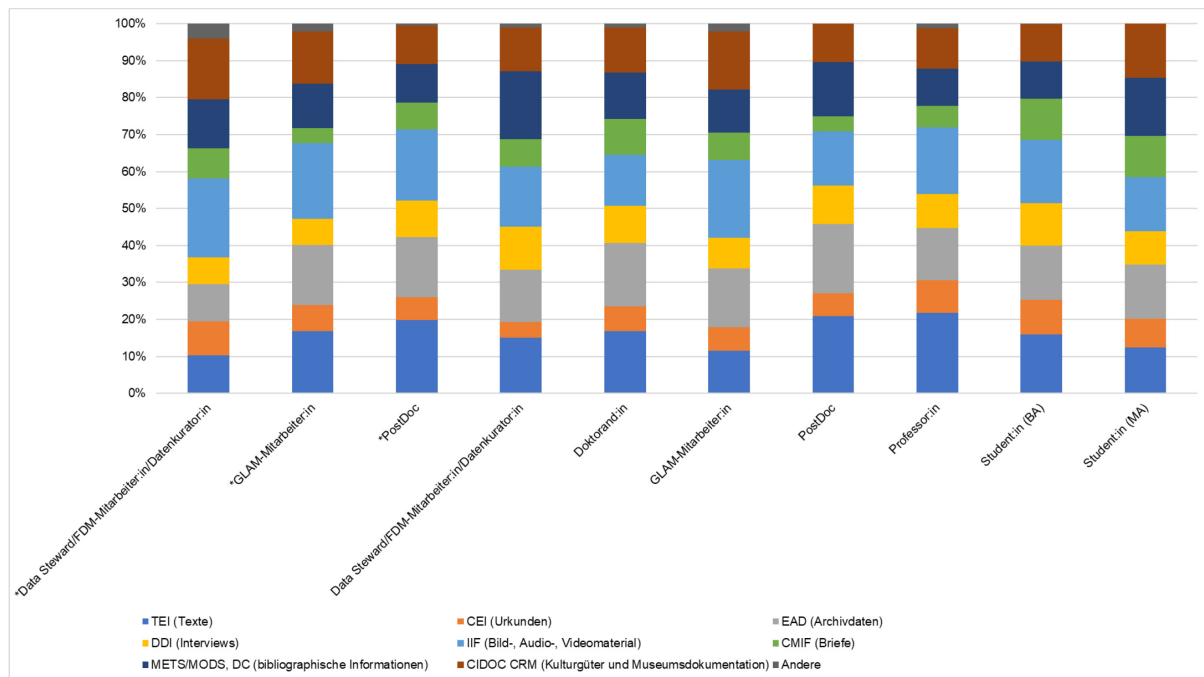


Abb. 33: Abfrage nach Metadatenstandards, Verteilung nach Statusgruppe und Auswahloption, Nennungen: 1756 | n = 497

Mit Blick auf die Verteilung innerhalb der Statusgruppen zeigt sich, dass FDM-Mitarbeitende mit Lehre, GLAM-Mitarbeitende mit und ohne Lehre und Bachelorstudierende das größte Interesse am Metadatenstandard IIF bekunden. TEI wurde am häufigsten innerhalb der Gruppen PostDocs mit und ohne Lehre und Professor:innen ausgewählt. Doktorand:innen bekunden als einzige Gruppe überwiegend Interesse an dem Standard EAD. Es lassen sich zwar durchaus Tendenzen erkennen, die Verteilung innerhalb der Statusgruppen ist aufgrund der Gruppengröße aber so divers, dass ein Statusgruppen-übergreifender Vergleich schwierig ist.

Zu folgenden Methoden zur Aufbereitung von Daten würde ich gerne mehr erfahren:

- Transkription
- Annotation
- Automatische Texterkennung (OCR, HTR)
- Linked Open Data
- Datenbereinigung/-konsolidierung
- Named entity recognition
- Lemmatisierung
- PoS-Tagging (grammatische Annotation)
- Graphentechnologie
- Literaturverwaltung
- Textmining
- Datenmodellierung
- Visualisierungsverfahren
- Sonstige _____

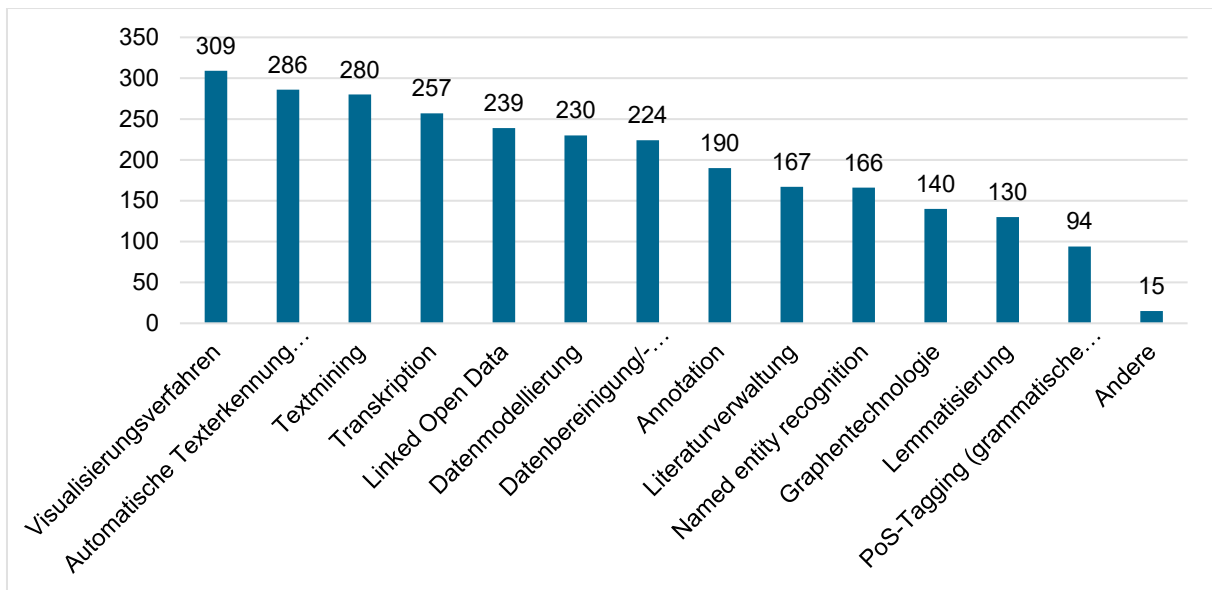


Abb. 34: Abfrage nach Methoden zur Aufbereitung von Daten, Nennungen: 2727 | n = 530

530 von 568 Proband:innen haben mindestens eine Option ausgewählt. Das meiste Interesse wurde mit insgesamt 309 Stimmen an Visualisierungsverfahren geäußert. Zusätzlich dazu wurde vermehrt das Verfahren der automatischen Texterkennung, Textmining, Datenbereinigung /-konsolidierung und Transkription ausgewählt. Mit 94 Stimmen stößt die Methode PoS-Tagging auf das geringste Interesse.

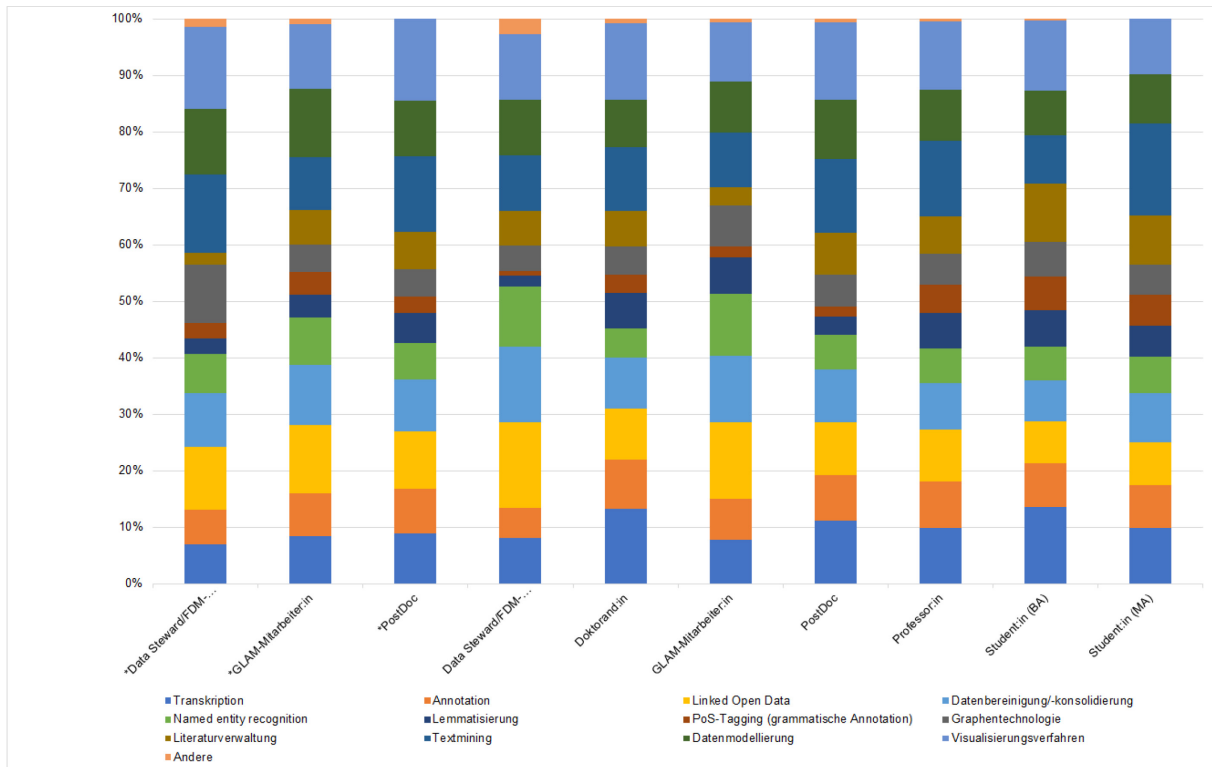


Abb. 35: Abfrage nach Methoden zur Aufbereitung von Daten, Verteilung nach Statusgruppe und Auswahloption, Nennungen: 2778 | n = 543

Innerhalb der Statusgruppen zeigt sich insgesamt ein erhöhtes Interesse an Visualisierungsverfahren und Linked Open Data, bei den Gruppen Professor:innen und Studierende an Textmining und Transkription. Das geringste Interesse durch fast alle Statusgruppen hinweg besteht an dem Verfahren PoS-Tagging. Lediglich bei den lehrenden FDM-Mitarbeitenden wurde Literaturverwaltung am seltensten ausgewählt.

Insgesamt ist die Verteilung zwischen den Statusgruppen sehr gleichmäßig. Lediglich kleine Nuancen zeigen sich z. B. bei der größeren Bedeutsamkeit der Annotation für Doktorand:innen und BA-Studierende, der geringeren Nachfrage nach Literaturverwaltung bei FDM-Mitarbeitenden mit Lehre und GLAM-Mitarbeitenden ohne Lehre. Die Graphentechnologien stoßen auf ein größeres Interesse bei den DataStewards ohne Lehre.

Kurzzusammenfassung der Kategorie „Datenaufbereitung“

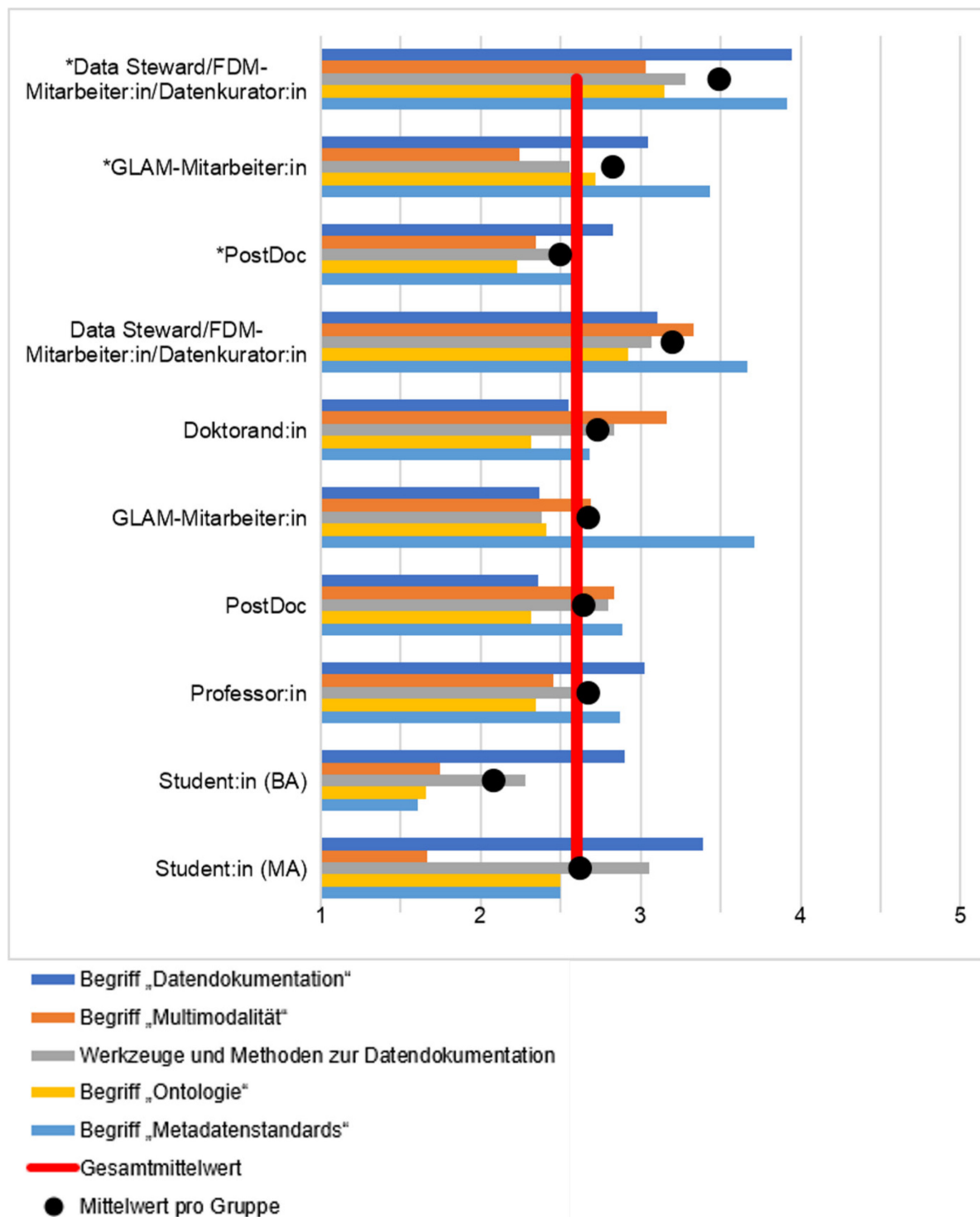


Abb. 36: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Datenaufbereitung“ | n=562

Die Kategorie „Datenaufbereitung“ wurde im Durchschnitt mit 2,6 bewertet. Am kompetentesten haben sich lehrende FDM-Mitarbeitende mit 3,5 und am niedrigsten Bachelorstudierende mit 2,1 eingeschätzt. Allgemein ist eine eher zurückhaltende Einschätzung vorzufinden. Die meisten Statusgruppen bewegen sich zwischen den Werten 2,5 und 2,8 mit nicht signifikanten Unterschieden von $\Delta = 0,1$. Die Verteilung nach Institutionen zeigt einen ähnlichen Ausgang.

Personen, die in IT-/FDM-Serviceeinrichtungen tätig sind, schätzen sich durchschnittlich mit 4,1 ein und liegen damit deutlich an der Spitze, mit einem Unterschied zum allgemeinen Durchschnitt von $\Delta = 1,5$. Universitäten befinden sich im Gegensatz dazu mit 2,5 unterhalb des Durchschnitts. Im Hinblick auf die Einschätzung nach Fachbereichen lässt sich feststellen, dass diese relativ nah beieinander liegen, mit Werten zwischen 2,1 und 3,3. Am kompetentesten haben sich Personen eingeschätzt, die sich dem Fachbereich Ethnologie und Europäische Ethnologie zugeschrieben haben, am niedrigsten Personen aus dem Fachbereich Neuere und Neueste Geschichte. Die im eher unteren Bereich liegende Durchschnittsbewertung zeigt, dass bei dem Thema „Datenaufbereitung“ bei mehreren Gruppen Schulungsbedarf besteht.

Datenqualität

Bewertung von Datenqualität

Niveaustufe 1. Kriterien zur Bewertung von Datenqualität kenne ich.

Niveaustufe 2. Ich kann die Qualität von Daten durch Bewertungskriterien beurteilen.

Niveaustufe 3. Ich kann vermitteln, welchen Nutzen Kriterien zur Bewertung von Datenqualität haben.

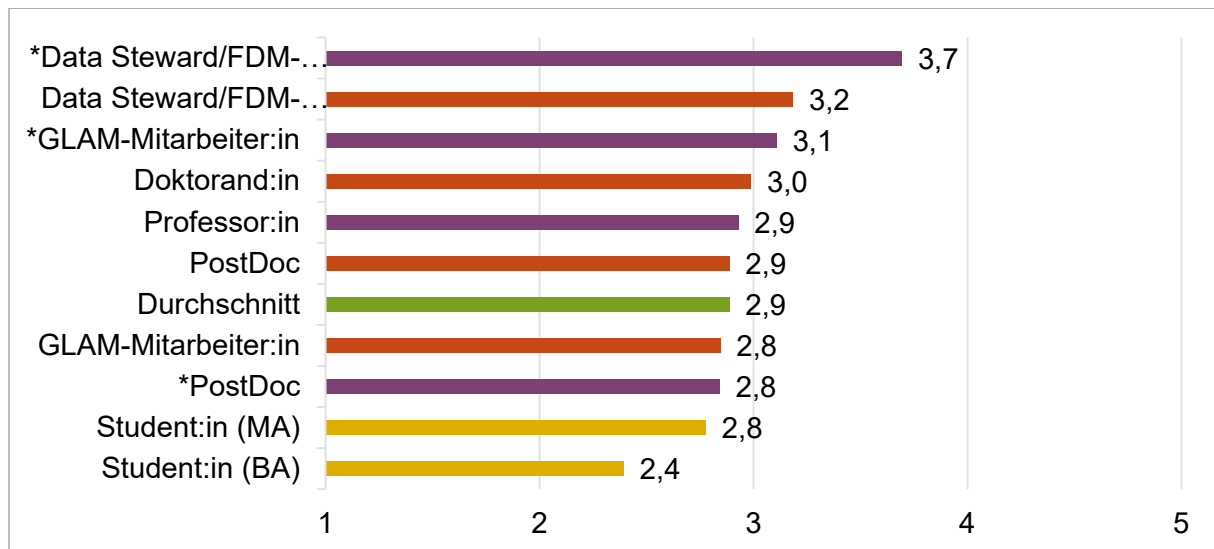


Abb. 37: Verteilung der Statusgruppen, Abfrage nach der Bewertung von Datenqualität | n = 544

Die Frage nach den Kriterien zur Bewertung von Datenqualität wurde im Durchschnitt mit 2,9 bewertet. Am besten schätzen sich FDM-Mitarbeitende mit Lehre ein, mit einer durchschnittlichen Selbsteinschätzung von 3,7, am niedrigsten Studierende im Bachelor mit 2,4. Die einzelnen Statusgruppen liegen in der Regel mit einem Unterschied von $\Delta = 0,1-0,4$ relativ nah beieinander. Lediglich die FDM-Mitarbeitenden mit Lehre schätzen sich im Vergleich zu den anderen Gruppen ($\Delta = 0,5$ zu GLAM-Mitarbeitende mit Lehre) deutlich kompetenter ein. Der überwiegende Teil der Mittelwerte liegt um 3.

Qualitätsmängel in Daten

Niveaustufe 1. Ich kann Qualitätsmängel in Daten erkennen.

Niveaustufe 2. Qualitätsmängel können in meinen eigenen Daten von mir identifiziert und korrigiert werden.

Niveaustufe 3. Ich kann vermitteln, wie man Qualitätsmängel in Daten erkennt.

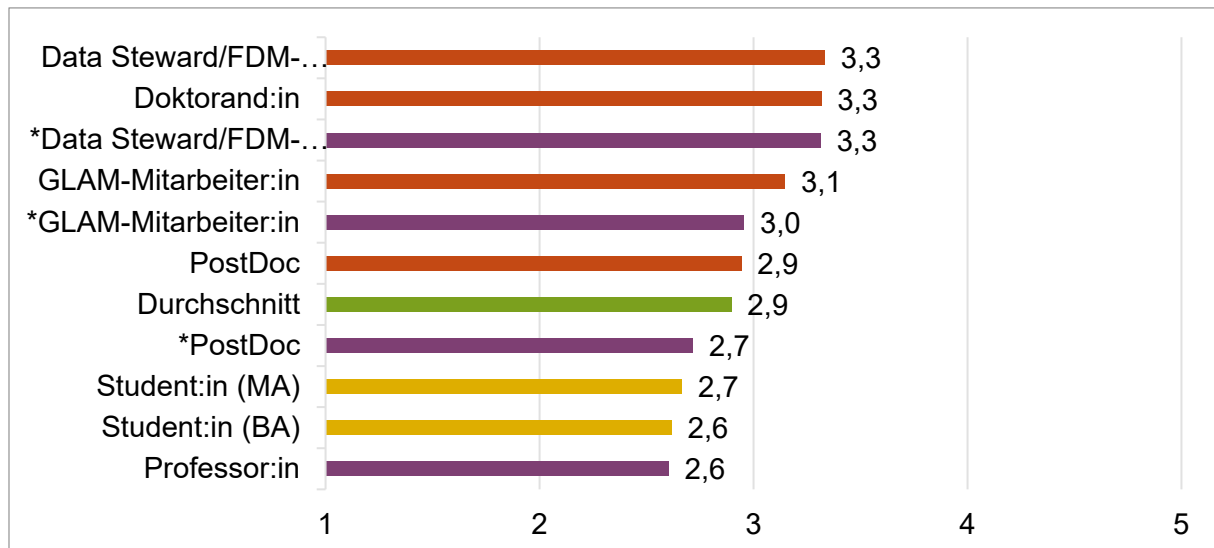


Abb. 38: Verteilung der Statusgruppen, Abfrage nach Qualitätsmängel in Daten | n = 541

Bei der Frage nach Qualitätsmängeln in Daten liegt der Durchschnittswert bei 2,9. FDM-Mitarbeitende und Doktorand:innen schätzen sich am höchsten (3,3) und Professor:innen und Bachelorstudierende am niedrigsten (2,6) ein. Allgemein ist zu erkennen, dass FDM- und GLAM-Mitarbeitende mit Werten ≥ 3 über dem Durchschnitt liegen. PostDocs, Professor:innen sowie Studierende liegen auf oder unter dem Durchschnitt. Insgesamt gruppieren sich alle um den Wert „trifft teils-teils zu“ mit $\Delta = 0,1-0,4$.

Kurzzusammenfassung Kategorie „Datenqualität“

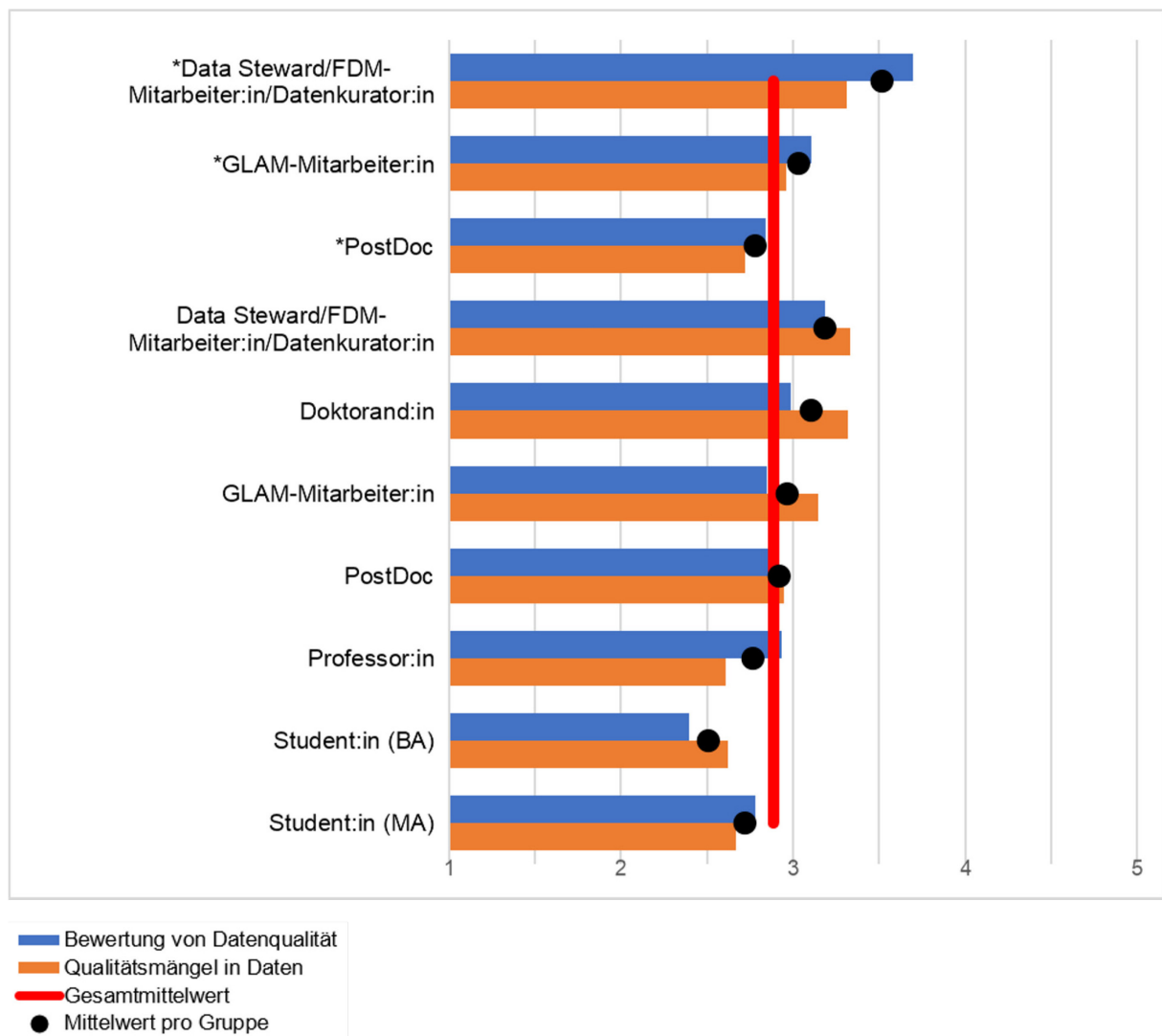


Abb. 39: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Datenqualität“ | n = 552

Die durchschnittliche Selbsteinschätzung der Teilnehmenden in der Kategorie „Datenqualität“ liegt bei 2,9. Allgemein ist die Einschätzung eher verhalten. Tendenziell haben sich die Befragten bei der Frage danach, Qualitätsmängel (in den eigenen Daten) erkennen zu können, etwas besser eingeschätzt als bei der Frage nach Qualitätskriterien zur Bewertung von Daten.

Datenkonvertierung

Begriffe „Datentyp / -format“ und „Dateityp / -format“

- Niveaustufe 1. Die Begriffe Datentyp / -format und Dateityp / -format kenne ich.
- Niveaustufe 2. Ich kann die Begriffe Datentyp / -format und Dateityp / -format sicher voneinander abgrenzen und anhand von Beispielen erklären.
- Niveaustufe 3. Ich kann vermitteln, wo genau die Unterschiede zwischen Datentyp / -format und Dateityp / -format liegen.

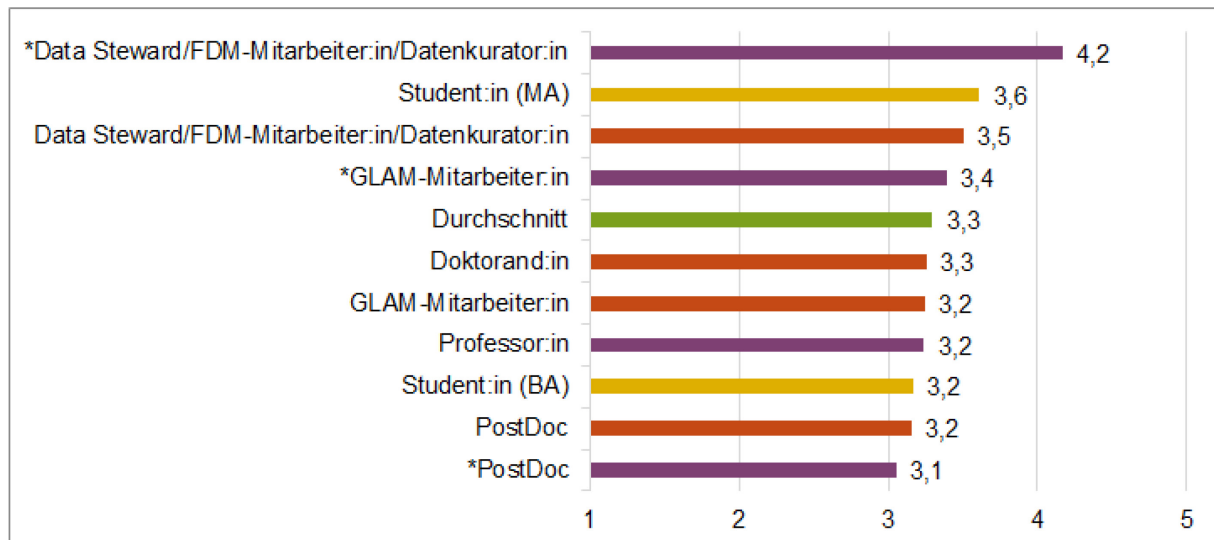


Abb. 40: Verteilung der Statusgruppen, Abfrage der Begriffe „Datentyp / -format“ und „Dateityp / -format“ | n = 550

Die Frage nach Daten- und Dateitypen und -formaten wurde durchschnittlich mit 3,3 beantwortet. Die FDM-Mitarbeitenden mit Lehre schätzen sich mit 4,2 am kompetentesten ein, die lehrenden PostDocs mit 3,1 am niedrigsten. Der Großteil der Statusgruppen liegt unterhalb des Durchschnitts von 3,3, wenn auch nur mit geringen Abständen von $\Delta = 0,1-0,2$. Auffällig ist die hohe Selbsteinschätzung der Masterstudierenden, die sich mit 3,6 am zweitbesten eingeschätzt haben.

Begriff „Datenkonvertierung“

Niveaustufe 1. Ich weiß, was Datenkonvertierung bedeutet.

Niveaustufe 2. Ich kann Daten von einem Format in ein anderes konvertieren.

Niveaustufe 3. Ich kann vermitteln, welche Aspekte der Begriff „Datenkonvertierung“ umfasst.

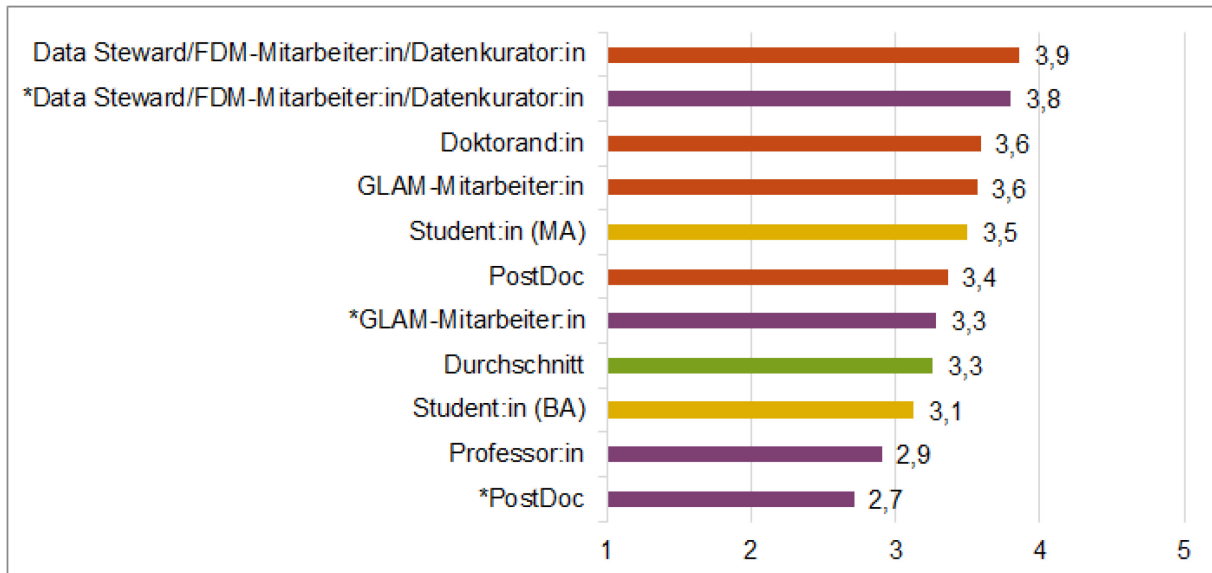


Abb. 41: Verteilung der Statusgruppen, Abfrage des Begriffs „Datenkonvertierung“ | n = 556

Bei der Frage nach Datenkonvertierung liegt der Durchschnitt bei 3,3. Am kompetentesten haben sich die FDM-Mitarbeitenden ohne Lehre (3,9) und am geringsten PostDocs mit Lehre (2,7) eingeschätzt. Die sehr diverse Verteilung der Gruppe der Lehrenden ist hier besonders auffällig. Während sich Professor:innen und PostDocs eher niedrig einschätzen, befinden sich GLAM-Mitarbeitende knapp oberhalb des Durchschnitts und FDM-Mitarbeitende an zweiter Position. Auffällig ist auch die im Vergleich relativ hohe Selbsteinschätzung der Doktorand:innen. Die Gruppen der Niveaustufe 2 liegen geschlossen über dem Durchschnitt mit $\Delta = 0,1-0,6$

Kurzzusammenfassung in der „Datenkonvertierung“

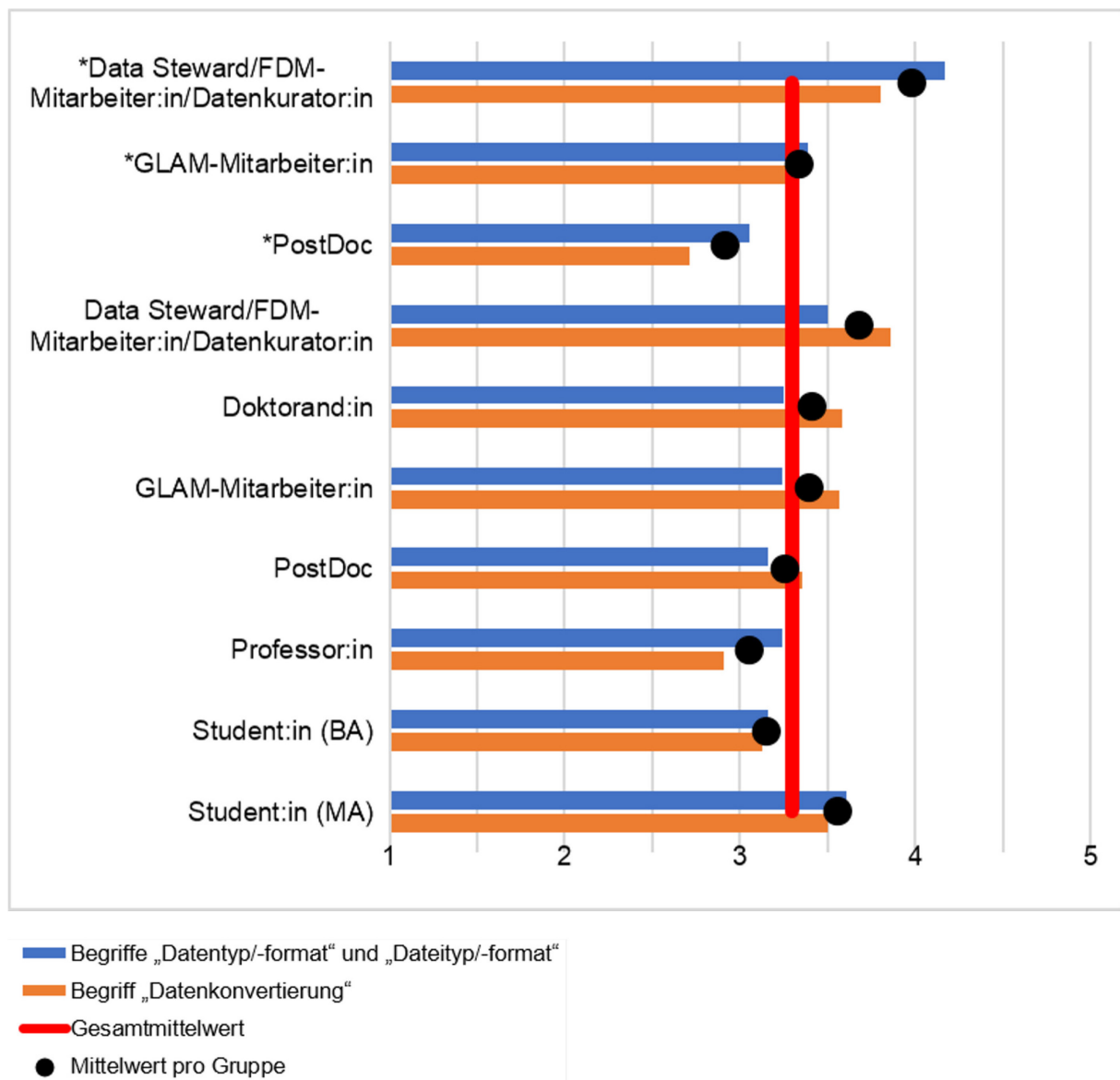


Abb. 42 Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Datenkonvertierung“ | n = 559

In der Kategorie „Datenkonvertierung“ wurde durchschnittlich ein Wert von 3,3 erreicht. FDM-Mitarbeitende mit Lehre schätzen sich mit 4,0 am kompetentesten und PostDocs mit Lehre mit einer Durchschnittsbewertung von 2,9 am niedrigsten ein. Ein klares Bild zeichnet sich im Hinblick auf die Gesamtverteilung der Kategorie „Datenkonvertierung“ nicht ab. Die Verteilung der Statusgruppen innerhalb der Niveaustufen zeigt kein aussagekräftiges Muster. So befinden sich einzelne Gruppenmitglieder aus allen drei Niveaustufen sowohl oberhalb als auch unterhalb des Durchschnitts.

Programmierung

Begriff „algorithmisches Denken“

- Niveaustufe 1. Ich weiß, was unter algorithmischem Denken verstanden wird.
- Niveaustufe 2. Ich setze in meiner Arbeit Methoden ein, die algorithmisches Denken voraussetzen.
- Niveaustufe 3. Ich kann vermitteln, bei welchen Methoden und Techniken algorithmisches Denken in meinem Fach relevant ist.

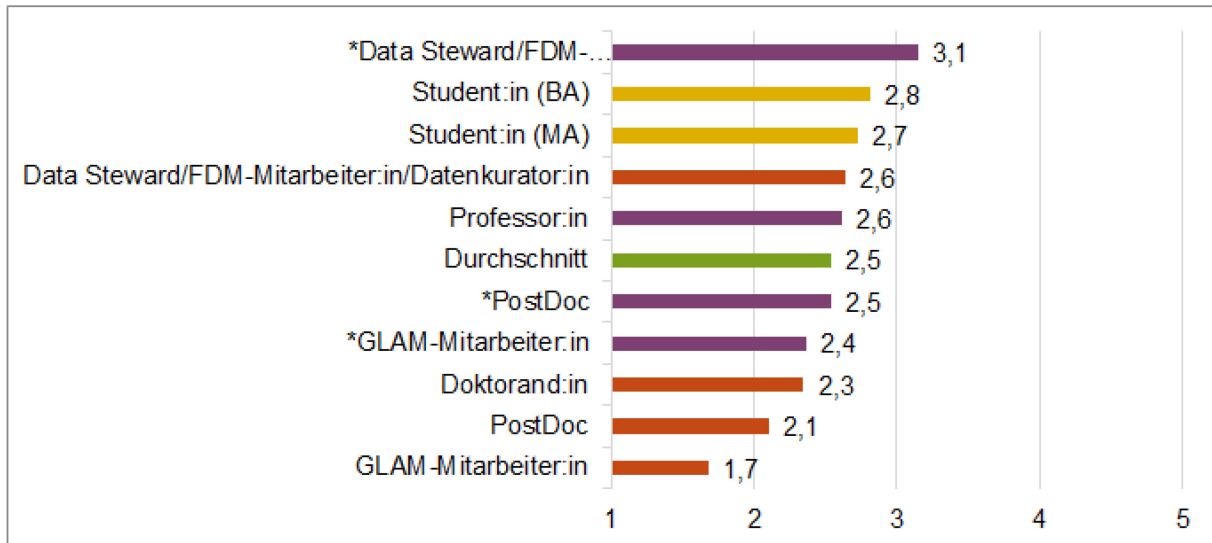


Abb. 43 Verteilung der Statusgruppen, Abfrage des Begriffs „algorithmisches Denken“ | n = 538

Bei der Frage nach algorithmischem Denken liegt die durchschnittliche Bewertung bei 2,5. Während sich die FDM-Mitarbeitenden mit Lehre am kompetentesten (3,1) einschätzen, fallen die GLAM-Mitarbeitenden ohne Lehre mit einer Selbsteinschätzung von durchschnittlich 1,7 stark ab. Die Gruppen der Niveaustufe 1 schätzen sich mitunter am besten ein.

Einfacher Programmcode

- Niveaustufe 1. Ich kann einen einfachen Programmcode in einer mir bekannten Programmiersprache nachvollziehen.
- Niveaustufe 2. Ich kann einen (einfachen) Programmcode in einer mir bekannten Programmiersprache schreiben.
- Niveaustufe 3. Ich kann vermitteln, wie man einen einfachen Programmcode in einer ausgewählten Programmiersprache schreibt.

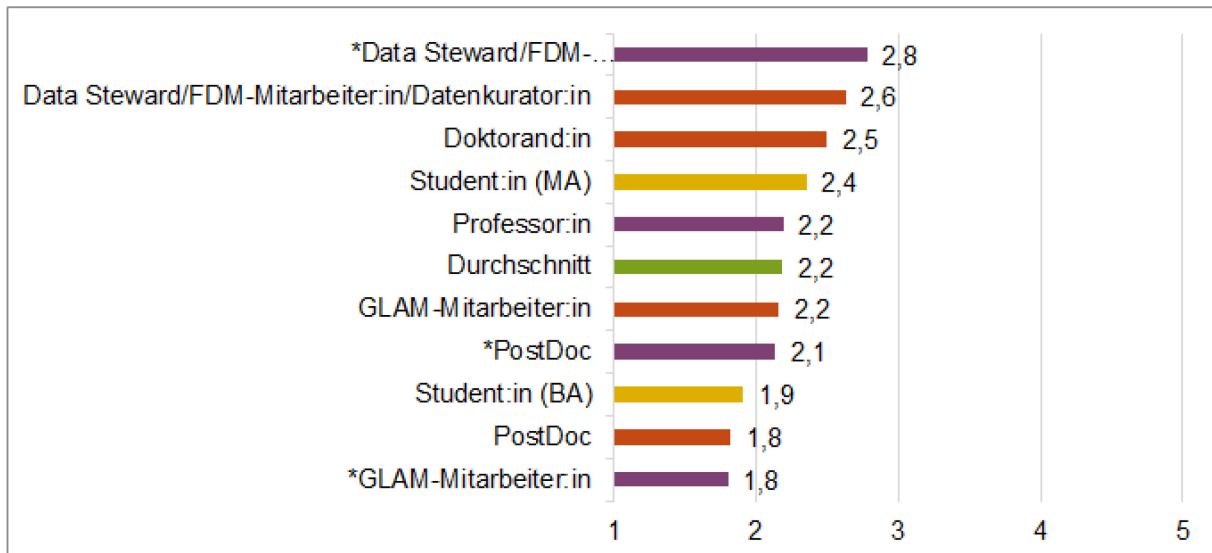


Abb. 44: Verteilung der Statusgruppen, Abfrage nach einfachem Programmcode | n = 546

Die Frage, ob man einen Programmcode verstehen bzw. schreiben und nachvollziehen kann, wird mit einer durchschnittlichen Bewertung von 2,2 bewertet. Mit einer Selbsteinschätzung von 2,8 haben sich die FDM-Mitarbeitenden mit Lehre am kompetentesten eingeschätzt. Auffallend ist hier die geringe Selbsteinschätzung der PostDocs und GLAM-Mitarbeitenden ohne Lehre, mit jeweils 1,8. Allgemein liegen die Statusgruppen nah beieinander ($\Delta = 0,1-0,3$), so dass keine signifikanten Kompetenzunterschiede auszumachen sind.

Zu folgenden Programmiersprachen benötige ich mehr Informationen:

- Python
- Java
- PHP
- R
- Sonstige _____

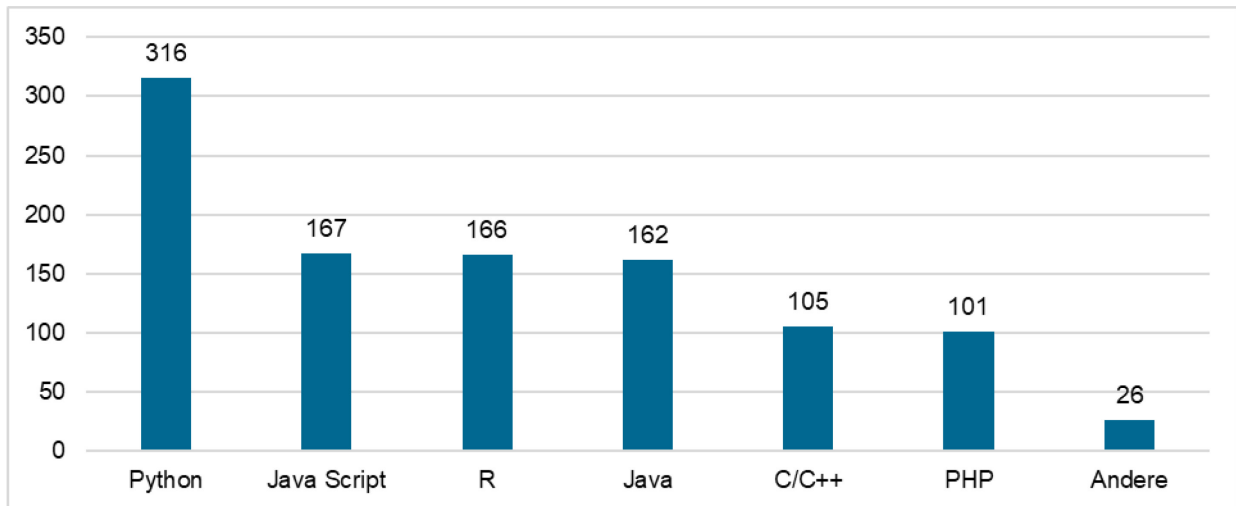


Abb. 45: Abfrage nach Programmiersprachen, Nennungen: 104 | n = 425

Neben inhaltlichen Fragen zum Thema Programmierung hatten die Proband:innen die Möglichkeit anzukreuzen, zu welchen Programmiersprachen sie gerne weitere Informationen hätten. Über 73 % der Personen haben mindestens einen Wunsch zu mehr Informationen geäußert. Mit insgesamt 316 Stimmen wünscht sich ein Großteil der Befragten mehr Informationen zu der Programmiersprache Python. Interesse an Java, JavaScript und R wurde mit einer relativ ausgeglichenen Verteilung ausgewählt.

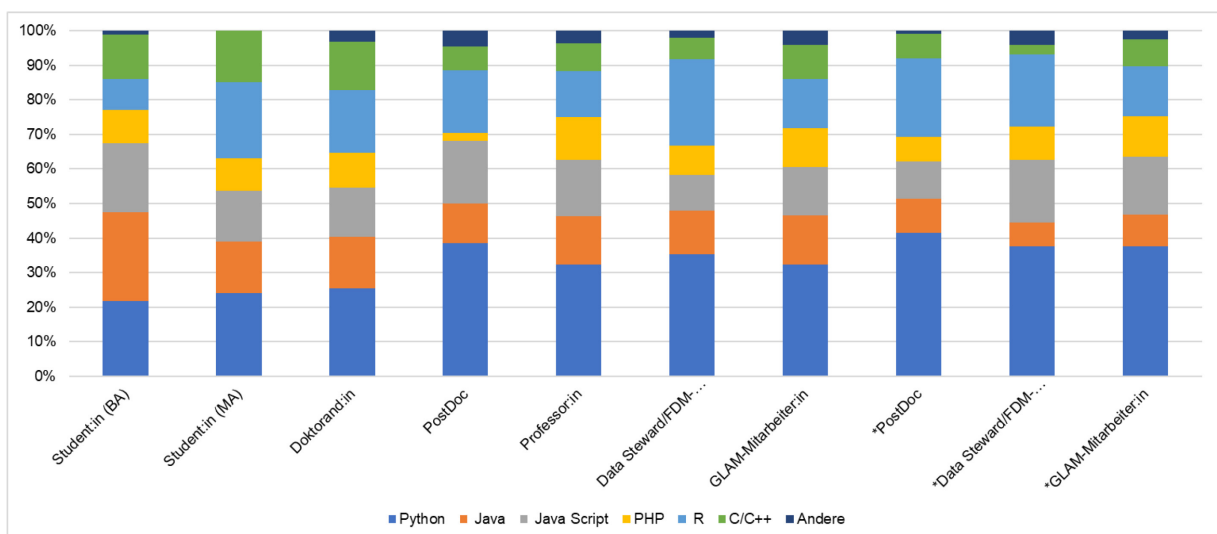
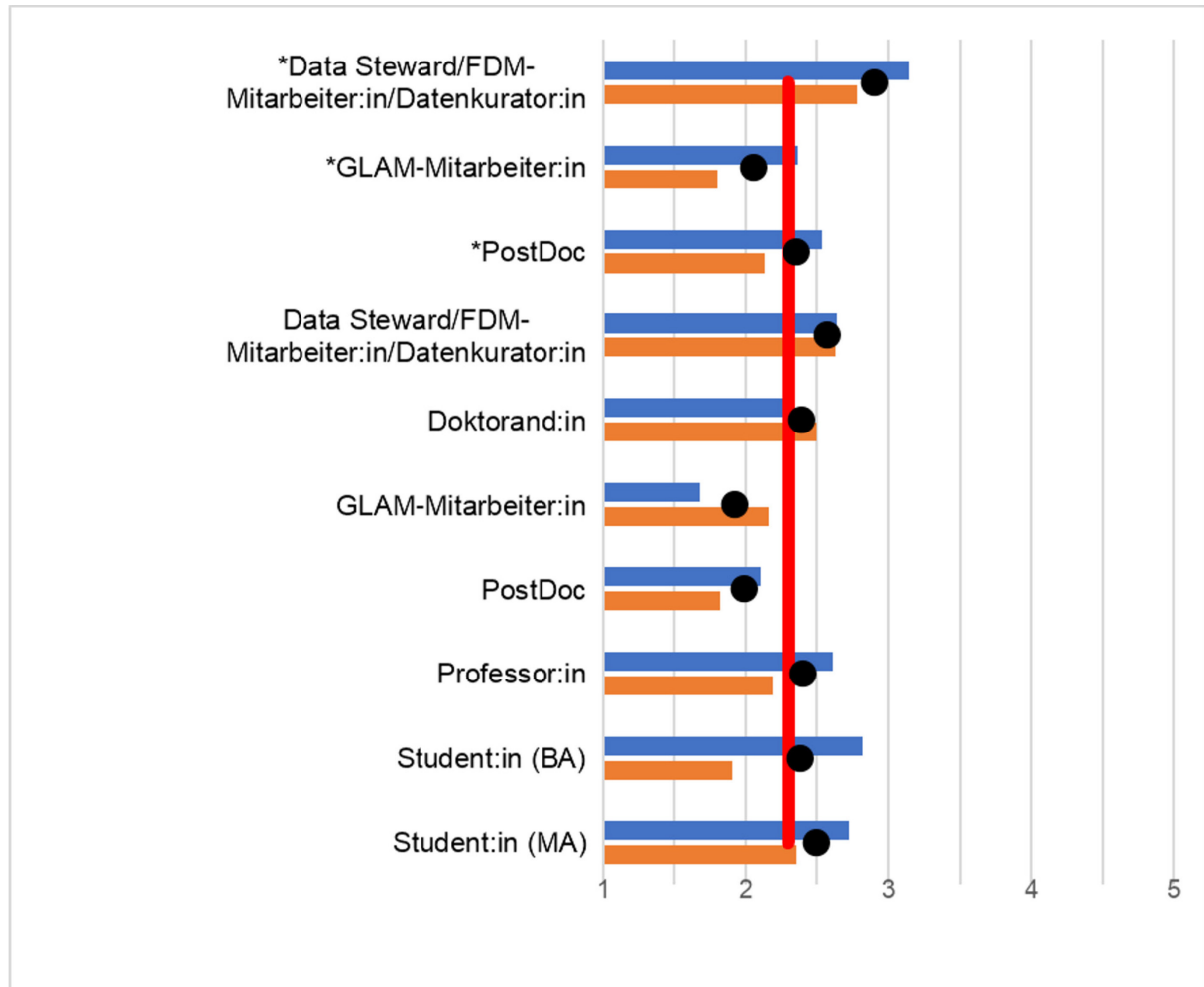


Abb. 46: Abfrage nach Programmiersprachen, Verteilung nach Statusgruppen und Auswahloptionen, Nennungen: 1043 | n = 425

Das erhöhte Interesse an der Programmiersprache Python bestätigt sich auch innerhalb der Statusgruppen. Bis auf Studierende im Bachelor, die gerne mehr Informationen zu Java hät-

ten, haben alle Teilnehmenden mit teilweise deutlichem Abstand zu den anderen Auswahlmöglichkeiten Python am häufigsten ausgewählt. Das geringste Interesse besteht an den Programmiersprachen C/C++ und PHP; bei der Gruppe der Bachelorstudierenden an R.

Kurzzusammenfassung der Kategorie „Programmierung“



- Begriff „algorithmisches Denken“
- Einfacher Programmcode
- Gesamtmittelwert
- Mittelwert pro Gruppe

Abb. 47: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Programmierung“ | n = 557

Im Vergleich zu anderen Kategorien liegt die durchschnittliche Bewertung der Kategorie „Programmierung“ mit 2,3 im unteren Bereich. Dies ist nochmal deutlicher an der Einzelverteilung der Gruppen zu erkennen. FDM-Mitarbeitende mit Lehre schätzen sich am kompetentesten ein, erreichen jedoch auch nur einen Wert von 2,9. Am niedrigsten schätzen sich GLAM-Mitarbeitende und PostDocs ohne Lehre ein, mit jeweils 1,8 und 1,9 liegen sie deutlich unter dem Durchschnitt. Tendenziell bewerten sich die Proband:innen besser bei der Frage nach algorithmischem Denken als bei der Frage nach einem Programmiercode.

Recht

Fachspezifische rechtliche Aspekte

- Niveaustufe 1. Ich kenne fachspezifische, rechtliche Aspekte, die in meinem Fachgebiet zu berücksichtigen sind.
- Niveaustufe 2. Die fachspezifischen rechtlichen Aspekte beachte ich während meiner Arbeit.
- Niveaustufe 3. Ich kann vermitteln, welche fachspezifischen, rechtlichen Aspekte für mein/e Fachgebiet/e zu berücksichtigen sind.

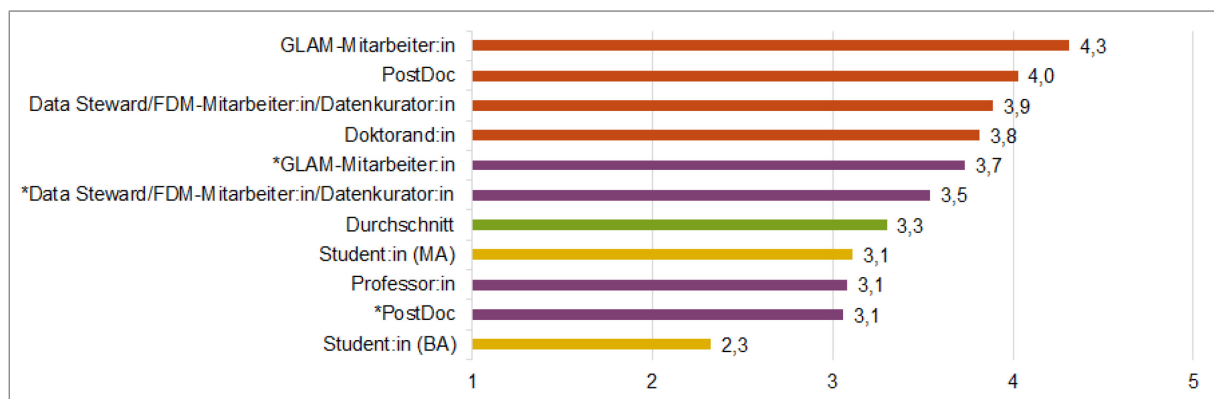


Abb. 48: Verteilung der Statusgruppen, Abfrage nach fachspezifischen, rechtlichen Aspekten | n = 546

Bei der Frage nach fachspezifischen, rechtlichen Aspekten liegt die durchschnittliche Bewertung bei 3,3. Deutlich zu erkennen ist hier die Verteilung der Niveaustufe 2. Diese befindet sich geschlossen oberhalb des Durchschnitts und erreicht mit 3,8 bis 4,3 die höchsten Werte. Teilweise deutlich geringer schätzt sich die Statusgruppen der Niveaustufe 3 ein. Vor allem Post-Docs mit Lehre und Professor:innen liegen unterhalb des allgemeinen Durchschnitts.

Begriff „Anonymisierung“

- Niveaustufe 1. Ich kenne den Begriff Anonymisierung im Hinblick auf Daten.
- Niveaustufe 2. Ich habe bereits Daten anonymisiert.
- Niveaustufe 3. Ich kann vermitteln, wie man Daten anonymisiert.

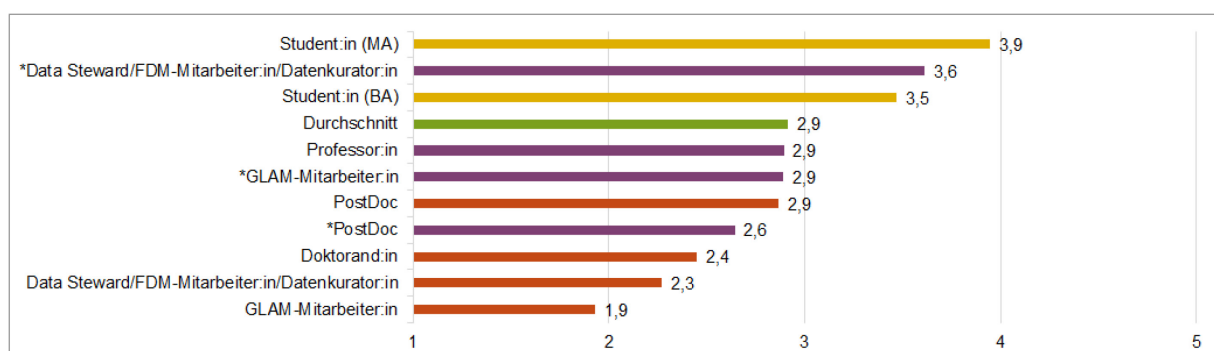


Abb. 49: Verteilung der Statusgruppen, Abfrage des Begriffs „Anonymisierung“ | n = 545

Die Frage nach der Anonymisierung von Daten wurde durchschnittlich mit 2,9 bewertet. Am kompetentesten schätzen sich hier Masterstudierende mit 3,9 und am niedrigsten GLAM-Mit-

arbeitende mit 1,9 ein. Gegenteilig zur Frage zuvor haben sich hier die Gruppen der Niveaustufe 3 im Vergleich zur Niveaustufe 2 teilweise deutlich besser eingeschätzt. Besonders interessant ist die Diskrepanz zwischen GLAM-Mitarbeitenden mit und ohne Lehre ($\Delta = 1,0$) und FDM-Mitarbeitenden mit und ohne Lehre ($\Delta = 1,3$). Ebenfalls augenfällig ist die hohe Einschätzung der Studierenden. Sowohl Master- als auch Bachelorstudierende schätzen sich mit 3,5 und 3,9 am kompetentesten ein.

Begriff „Pseudonymisierung“

Niveaustufe 1. Ich kenne den Begriff Pseudonymisierung im Hinblick auf Daten.

Niveaustufe 2. Ich habe bereits Daten pseudonymisiert.

Niveaustufe 3. Ich kann vermitteln, wie man Daten pseudonymisiert.

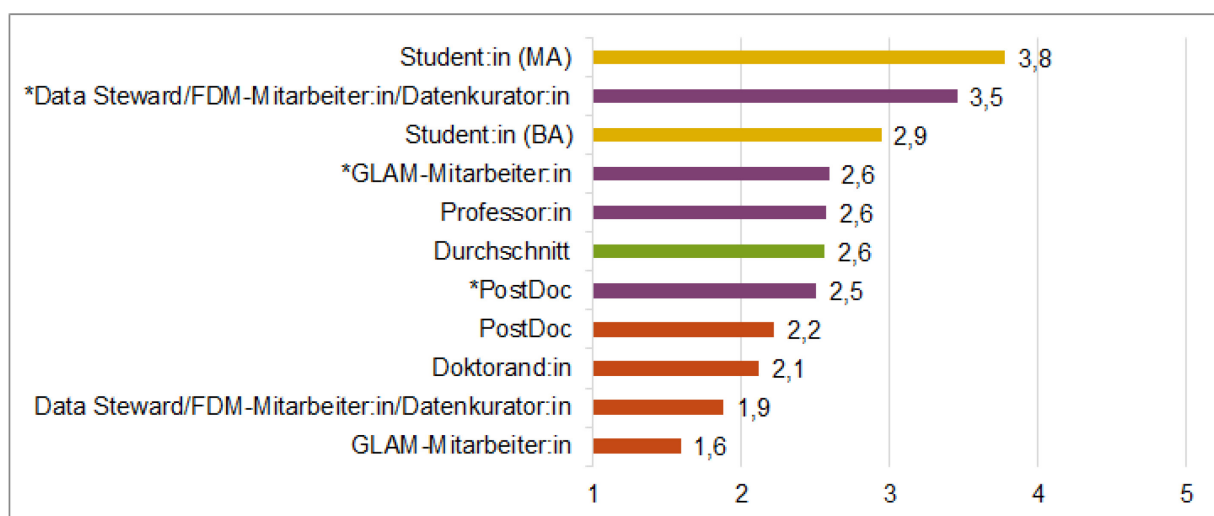


Abb. 50: Verteilung der Statusgruppen, Abfrage des Begriffs „Pseudonymisierung“ | n = 535

Die Frage nach der Pseudonymisierung von Daten wurde durchschnittlich mit 2,6 bewertet. Am kompetentesten schätzen sich die Masterstudierenden mit 3,8 und am niedrigsten die GLAM-Mitarbeitenden mit 1,6 ein. Wie bei der Frage zuvor erreichen die Studierenden im Vergleich zu den restlichen Statusgruppen mitunter die höchsten Werte. Alle Statusgruppen der Niveaustufe 2 befinden sich teils deutlich mit einem Unterschied von $\Delta = 0,4$ – $1,0$ unterhalb des Durchschnitts. Die Lehrenden erreichen zwar tendenziell bessere Werte, befinden sich aber auch knapp unter bzw. auf dem Durchschnitt. Lediglich die FDM-Mitarbeitenden mit Lehre erreichen im Gesamten einen höheren Wert von 3,5.

Begriff „Informierte Einwilligung“

Niveaustufe 1. Ich weiß, was eine Informierte Einwilligung ist.

Niveaustufe 2. Ich habe schon einmal eine Informierte Einwilligung eigenständig entworfen.

Niveaustufe 3. Ich kann vermitteln, wie man eine Informierte Einwilligung erstellt.

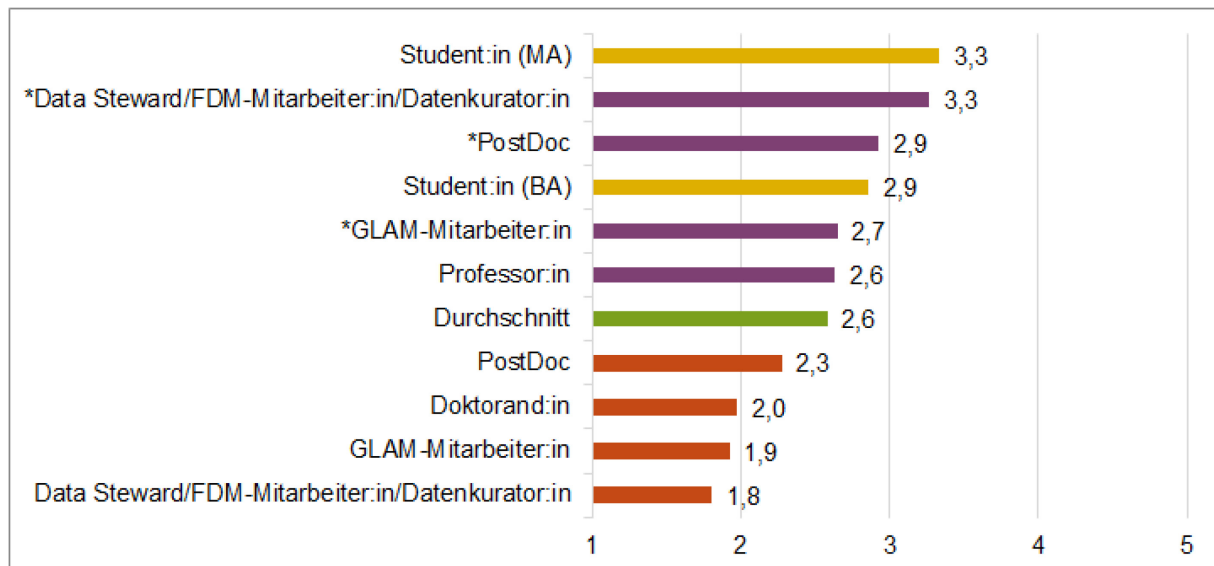


Abb. 51: Verteilung der Statusgruppen, Abfrage des Begriffs „Informierte Einwilligung“ | n = 529

Die Frage nach einer Informierten Einwilligung wurde durchschnittlich mit 2,6 bewertet. Am kompetentesten schätzen sich die Masterstudierenden mit 3,3 und am geringsten die nicht-lehrenden FDM-Mitarbeitenden ein. Auch hier zeigt sich, wie in den beiden Fragen zuvor, die geringere Relevanz von Verfahren im Umgang mit personenbezogenen Daten in der Breite der Geschichtswissenschaften. Daher haben sich aufgrund der Fragestellung die Studierenden am besten eingeschätzt. Der Unterschied zwischen Statusgruppen mit und ohne Lehre ist deutlich zu erkennen. Zwischen FDM-Mitarbeitenden aus beiden Gruppen besteht ein Unterschied von $\Delta = 1,7$, zwischen GLAM-Mitarbeitenden ein Unterschied von $\Delta = 0,8$ und zwischen PostDocs ein Unterschied von $\Delta = 0,6$.

Begriff „verschlüsselte Daten“

- Niveaustufe 1. Ich weiß, was verschlüsselte Daten sind.
 Niveaustufe 2. Ich habe bereits Daten verschlüsselt.
 Niveaustufe 3. Ich kann vermitteln, wie man Daten verschlüsselt.

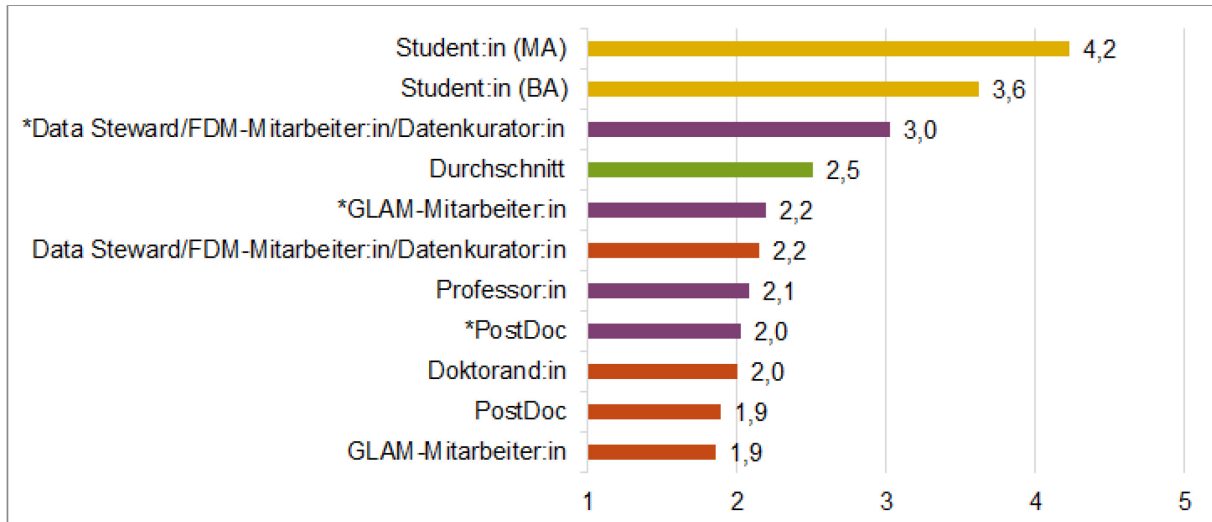


Abb. 52: Verteilung der Statusgruppen, Abfrage des Begriffs „Verschlüsselte Daten“ | n = 543

Die Frage nach verschlüsselten Daten wurde im Durchschnitt mit 2,5 bewertet. Während sich die Masterstudierenden mit 4,2 am kompetentesten einschätzen, erreichen GLAM-Mitarbeitende und PostDocs ohne Lehre den niedrigsten Wert mit 1,9. Ähnlich wie die vorherigen Auswertungen in dieser Kategorie haben sich die Studierenden mit jeweils 3,6 und 4,2 am kompetentesten eingeschätzt und liegen mit $\Delta = 1,1$ – $1,7$ oberhalb des Durchschnitts. Auffällig ist die geringere Bewertung der Niveaustufen 2 und 3. Lediglich lehrende FDM-Mitarbeitende liegen mit $\Delta = 0,5$ oberhalb des Durchschnitts. Alle anderen Statusgruppen befinden sich darunter und erreichen Werte von 1,9 bis 2,2.

Zu folgenden ethischen und rechtlichen Aspekten benötige ich mehr Informationen:

- Datenschutz
- Archivrecht
- Lizenzrecht
- Urheberrecht
- Datensicherheit
- Persönlichkeitsrecht
- Zugriffsschutz und -rechte
- Sonstige _____

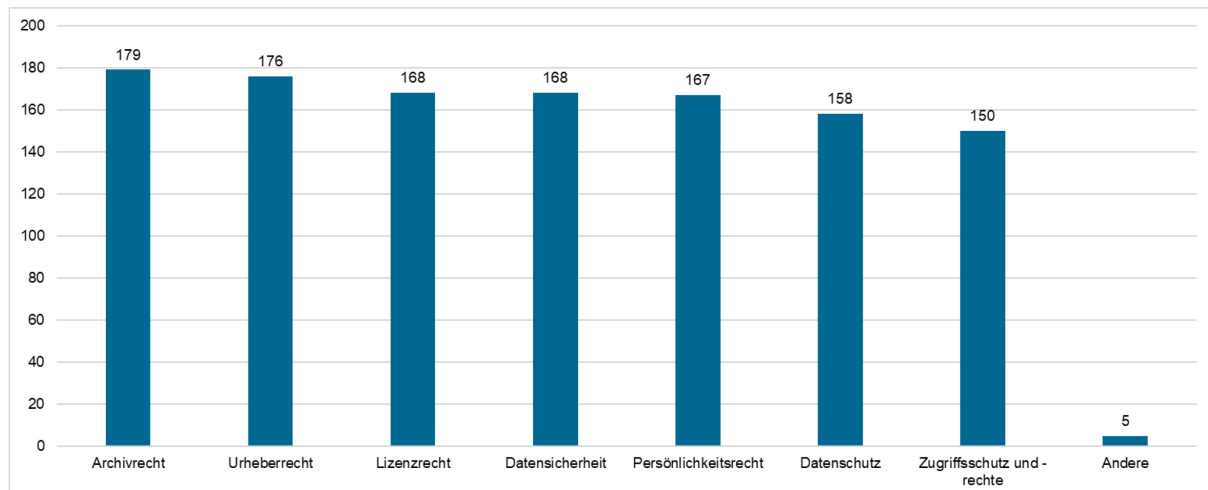


Abb. 53: Abfrage nach ethischen und rechtlichen Aspekten, Nennungen: 1171 | n = 291

Insgesamt haben 291 von 325¹⁴ Personen angegeben, dass sie zusätzliche Informationen zu ethischen und rechtlichen Aspekten benötigen. Das Interesse an diesen Themen scheint bei gut 90 % Beteiligung hoch zu sein. Im Vergleich zu den vorherigen Auswahlmöglichkeiten lässt sich hier eine vergleichsweise ausgeglichene Verteilung der Optionen feststellen. Die Option „Archivrecht“ wurde mit knapp 179 Stimmen am häufigsten gewählt, während das geringste Interesse mit 150 Stimmen für Zugriffsschutz und -rechte besteht.

¹⁴ Aufgrund einer falschen Filterzuweisung innerhalb des Fragebogentools SoSciSurvey wurde die Auswahlliste zu der Kategorie „Recht“ den Gruppen der Niveaustufe 3 nicht angezeigt. Alle Werte beziehen sich daher ausschließlich auf die Angaben der Gruppen der Niveaustufe 1 und 2.

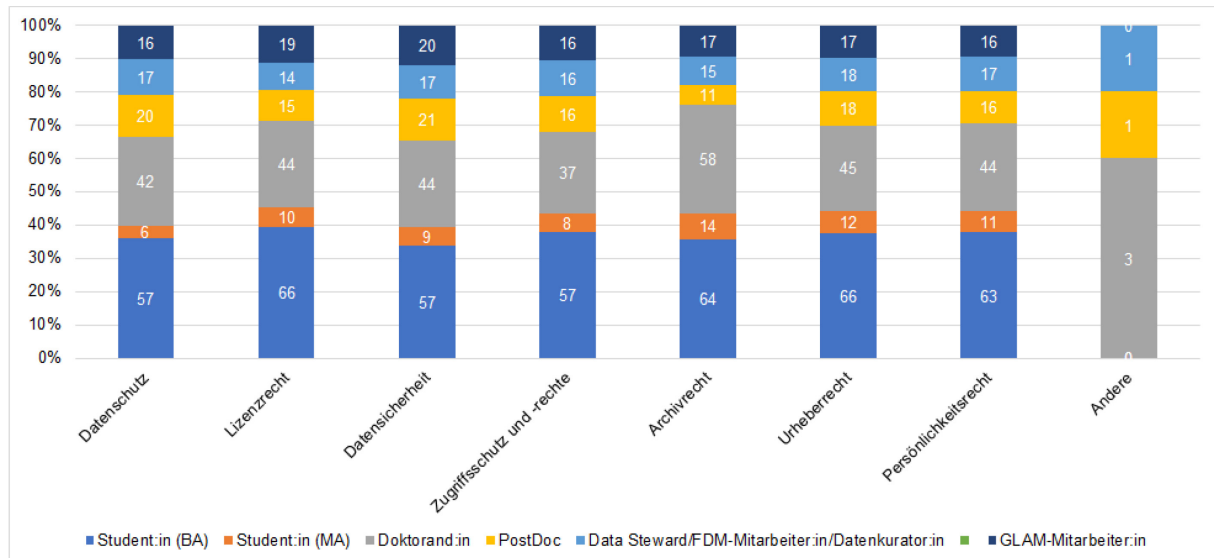


Abb. 54: Abfrage nach ethischen und rechtlichen Aspekten, Verteilung nach Statusgruppen und Auswahloptionen, Nennungen: 1171, | n = 291

Der Blick auf die Verteilung innerhalb der Statusgruppen zeigt hier eine relativ gleichmäßige Verteilung. Auffällig ist lediglich das geringere Interesse im Vergleich zu den anderen Statusgruppen am Datenschutz bei den MA-Studierenden und am Archivrecht bei den PostDocs. Vermutlich haben viele PostDocs bereits einschlägige Erfahrung in ihrer Karriere in Archiven gesammelt.

Kurzzusammenfassung der Kategorie „Recht“

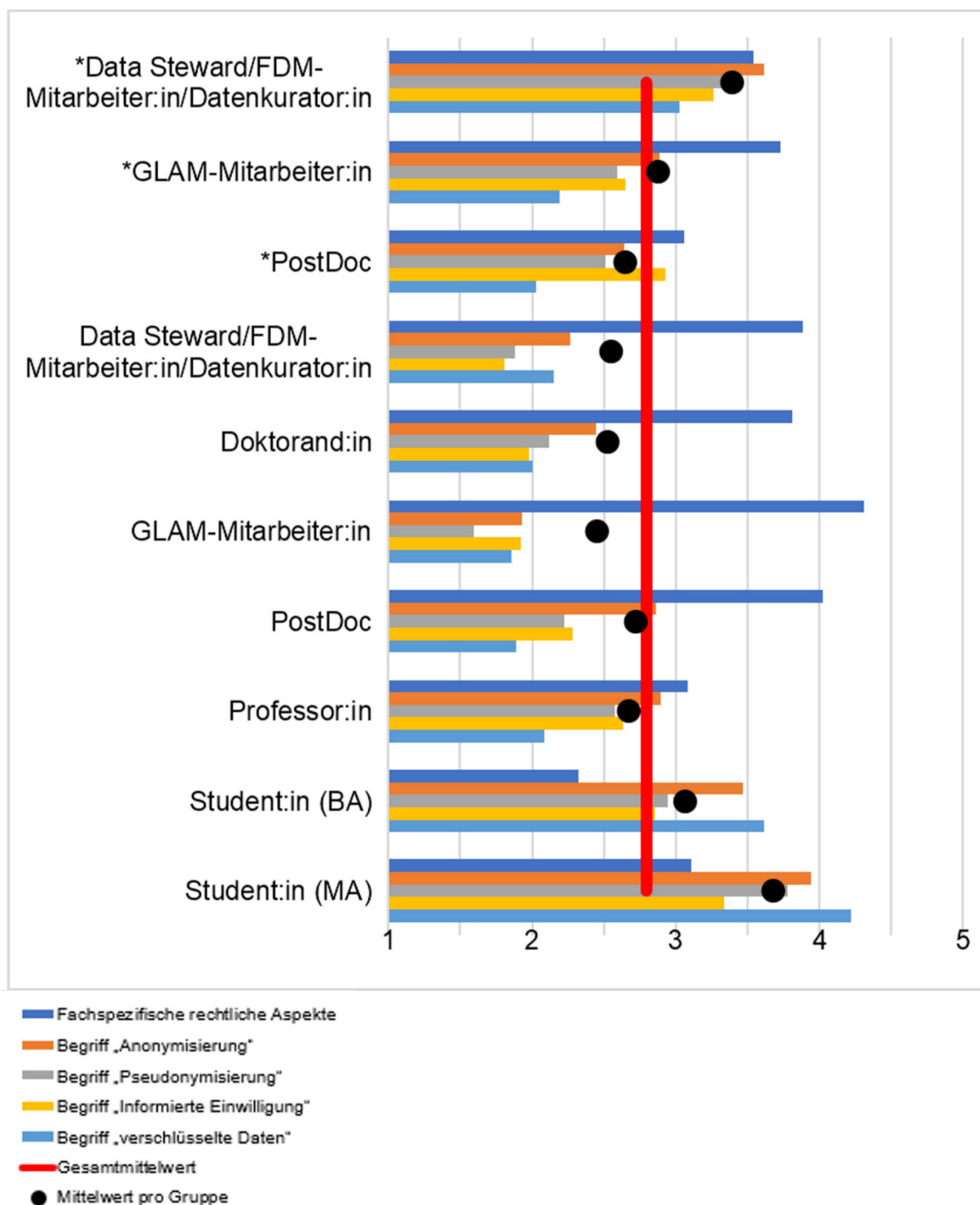


Abb. 55: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Recht“ | n = 562

Die Kategorie „Recht“ wurde im Durchschnitt von den Proband:innen mit 2,8 bewertet, was an der Fragestellung für die Niveaustufen 2 und 3 lag und der relativ geringen Relevanz des Datenschutzes in den meisten historischen Disziplinen. Daher haben sich die Masterstudierenden mit einem Durchschnittswert von 3,7 durchweg am besten eingeschätzt, die Bachelor-

studierenden mit 3,1 etwas schlechter. Lediglich die lehrenden FDM- und GLAM-Mitarbeitenden positionieren sich, abgesehen von den Studierenden, oberhalb des Gesamtdurchschnitts der Kategorie „Recht“. Die restlichen Statusgruppen liegen mit Durchschnittswerten von 2,5–2,7 alle nah beieinander darunter.

Ethik

Leitlinien zur Sicherung guter wissenschaftlicher Praxis

Niveaustufe 1. Ich kenne die Leitlinien zur Sicherung guter wissenschaftlicher Praxis.

Niveaustufe 2. Ich verfolge die Leitlinien zur Sicherung guter wissenschaftlicher Praxis.

Niveaustufe 3. Ich kann vermitteln, was die Leitlinien zur Sicherung guter wissenschaftlicher Praxis zum Forschungsdatenmanagement sagen.

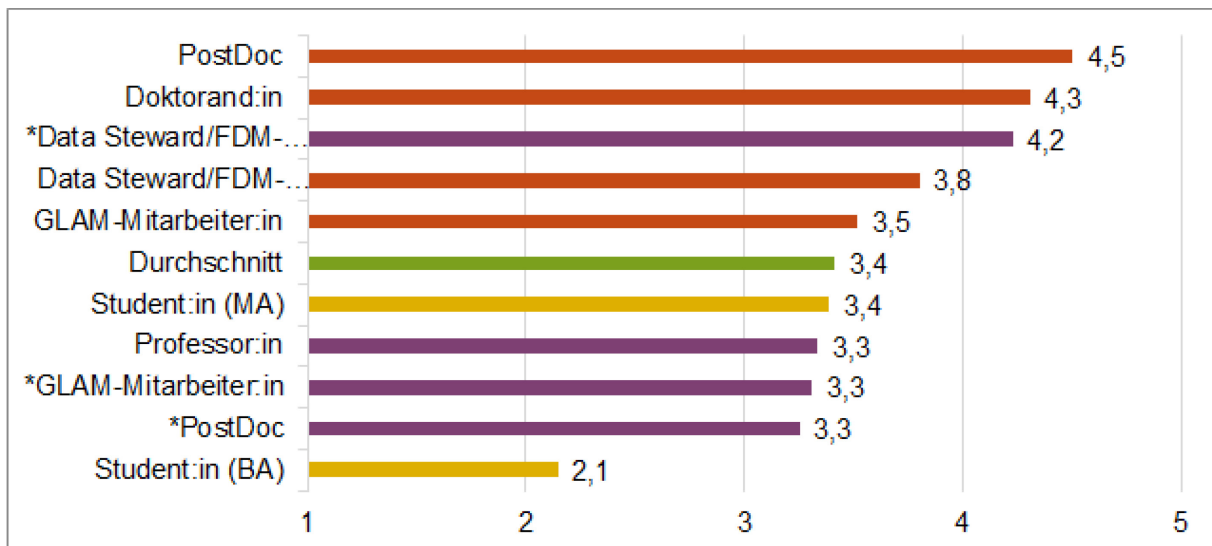


Abb. 56: Verteilung der Statusgruppen, Abfrage nach Leitlinien zur Sicherung guter wissenschaftlicher Praxis | n = 541

Bei der Frage nach Leitlinien zur Sicherung guter wissenschaftlicher Praxis liegt der Durchschnitt bei 3,4. Am besten hat sich die Gruppe der PostDocs ohne Lehre mit 4,5 und am geringsten die Bachelorstudierenden mit 2,1 eingeschätzt. Mit Ausnahme der GLAM-Mitarbeitenden ohne Lehre positioniert sich die Niveaustufe 2 deutlich im oberen Bereich. Interessant ist die, wenn auch nicht deutlich, niedrigere Bewertung der Gruppen GLAM-Mitarbeitende und PostDocs mit Lehre.

Begriff „Datenethik“

- Niveaustufe 1. Ich weiß, was unter dem Begriff „Datenethik“ zu verstehen ist.
- Niveaustufe 2. Die Überprüfung von ethischen Dimensionen ist in meiner eigenen Forschungspraxis ein fester Bestandteil.
- Niveaustufe 3. Ich kann vermitteln, welche Inhalte und Aspekte unter dem Begriff „Datenethik“ gefasst werden.

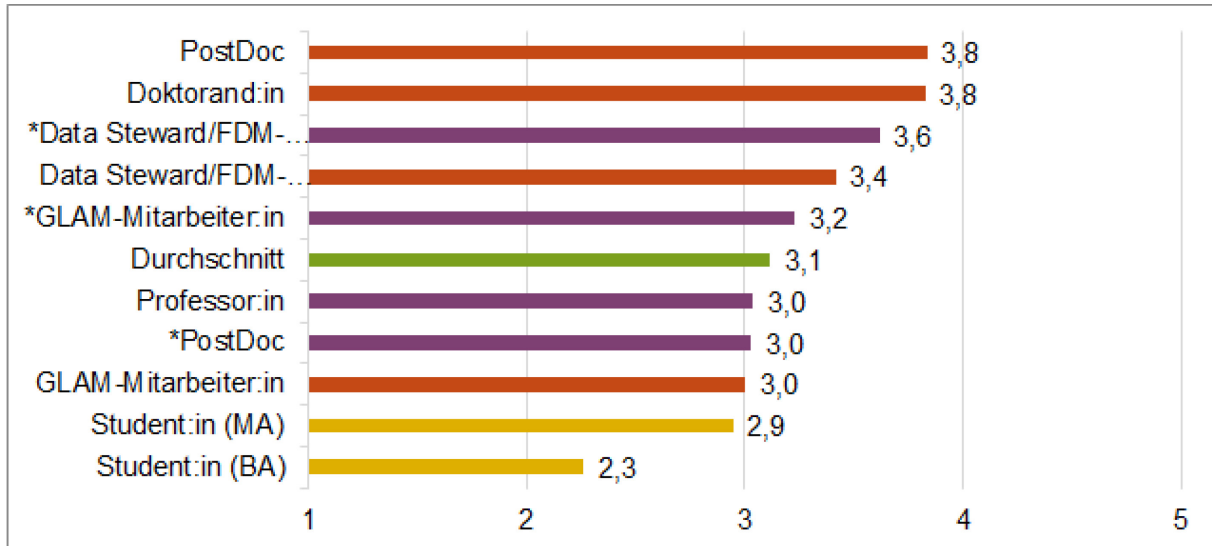


Abb. 57: Verteilung der Statusgruppen, Abfrage des Begriffs „Datenethik“ | n = 527

Die Frage nach „Datenethik“ und ethischen Dimensionen in der Forschungspraxis wird durchschnittlich mit 3,1 bewertet. Am besten schätzen sich PostDocs ohne Lehre (3,8) und am niedrigsten Studierende im Bachelor (2,3) ein. Der Abstand zwischen lehrenden und nicht-lehrenden PostDocs ist mit $\Delta = 0,8$ bemerkenswert. Dabei liegen die lehrenden PostDocs mit den Professor:innen bei 3,0 gleichauf und die Doktorand:innen ziehen bei 3,8 mit den nicht-lehrenden PostDocs gleich.

Begriff „CARE“

Niveaustufe 1. Ich kenne das Akronym CARE.

Niveaustufe 2. Die CARE-Prinzipien integriere ich in meine Arbeit.

Niveaustufe 3. Ich kann vermitteln, wofür das Akronym CARE steht.

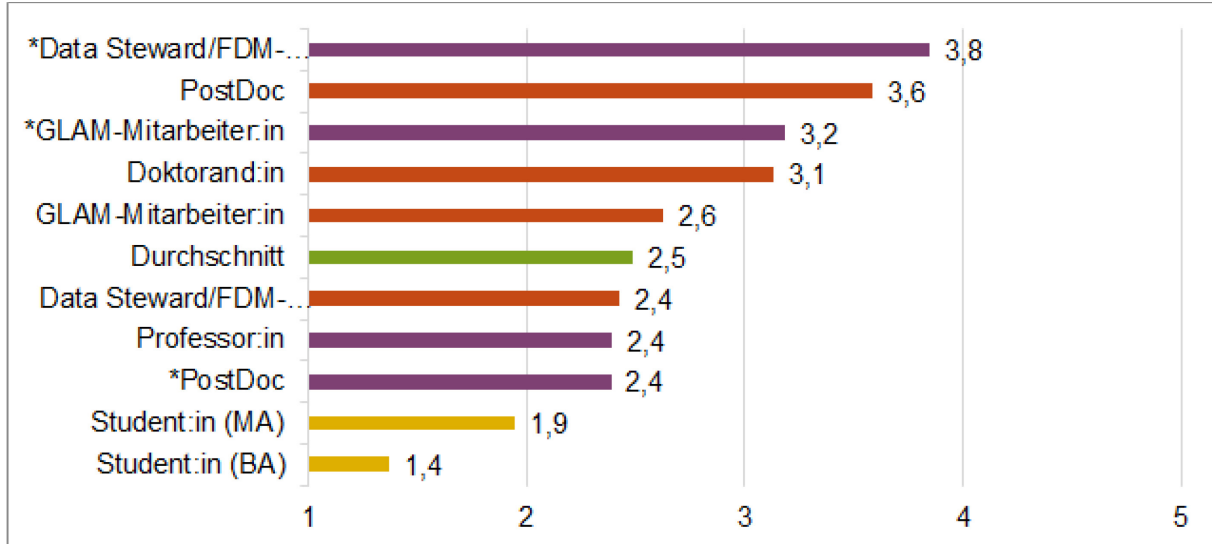


Abb. 58: Verteilung der Statusgruppen, Abfrage des Begriffs „CARE“ | n = 469

Bis auf die Gruppe der nicht-lehrenden FDM-Mitarbeitenden liegen alle oberhalb des Durchschnitts, der mit 2,5 zwischen „teils-teils“ und „trifft eher nicht zu“ liegt. Professor:innen und PostDocs mit Lehre haben sich mit 2,4 eher niedrig eingeschätzt. Auffällig ist auch, dass 99 von 568 Teilnehmenden die Frage mit „nicht beurteilbar“ bewertet haben.

Kurzzusammenfassung in der Kategorie „Ethik“

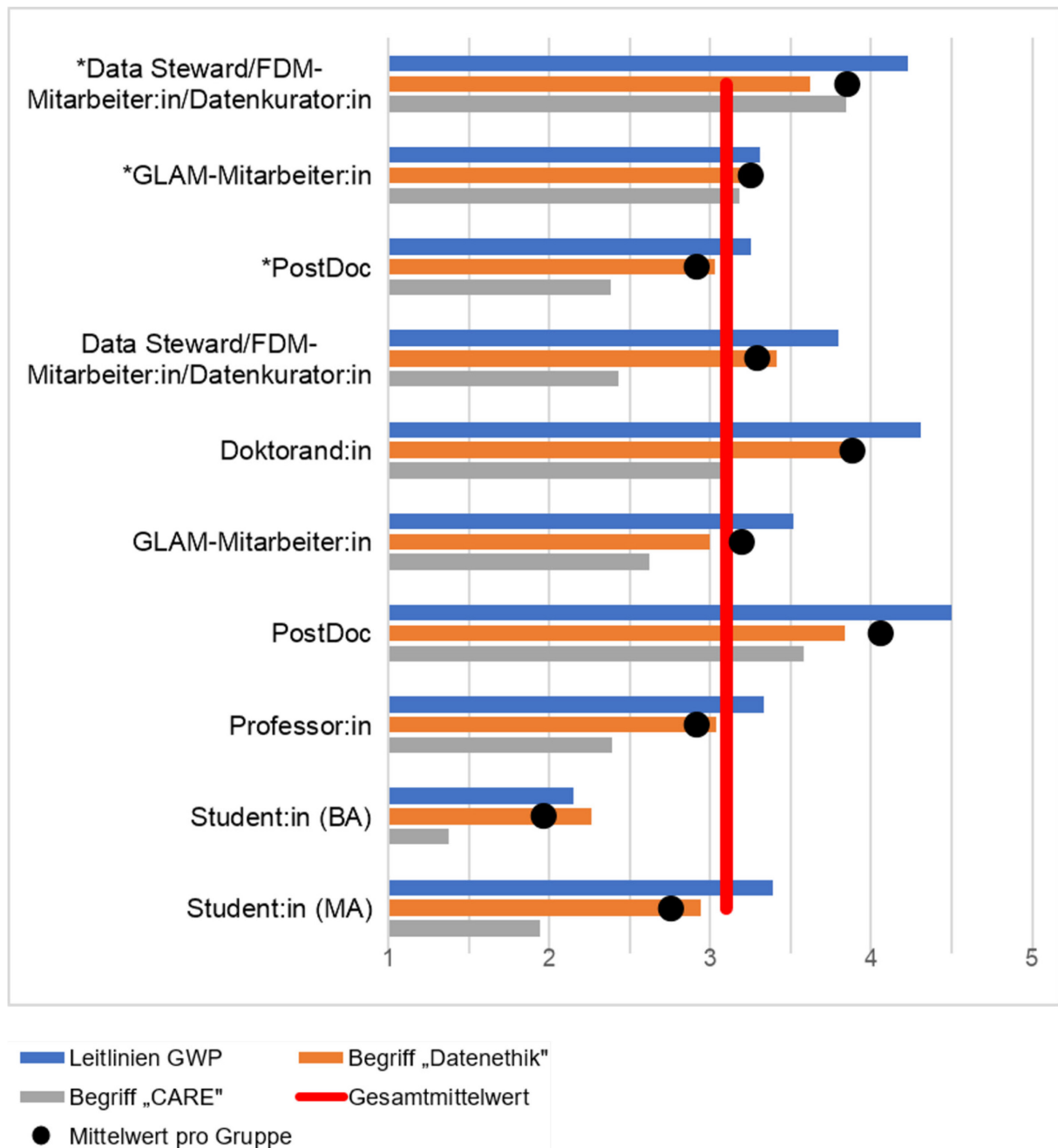


Abb. 59: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Ethik“ | n = 554

Die Kategorie „Ethik“ wurde durchschnittlich mit 3,1 bewertet. Am kompetentesten schätzen sich PostDocs ohne Lehre (4,1) ein und am niedrigsten die Bachelorstudierenden (2,0). Die meisten Gruppen liegen recht nah beieinander. Niveaustufe 2 erreicht durchschnittlich die höchsten Werte. Bei allen drei Fragen liegen Professor:innen und PostDocs mit Lehre unterhalb des Durchschnitts.

Speicherung

Begriff „Speichersystem“

Niveaustufe 1. Ich weiß, was man unter „Speichersystemen“ im Allgemeinen versteht.

Niveaustufe 2. Ich verwende ein Speichersystem für meine Daten.

Niveaustufe 3. Ich kann vermitteln, was Speichersysteme sind und welchen Nutzen sie haben.

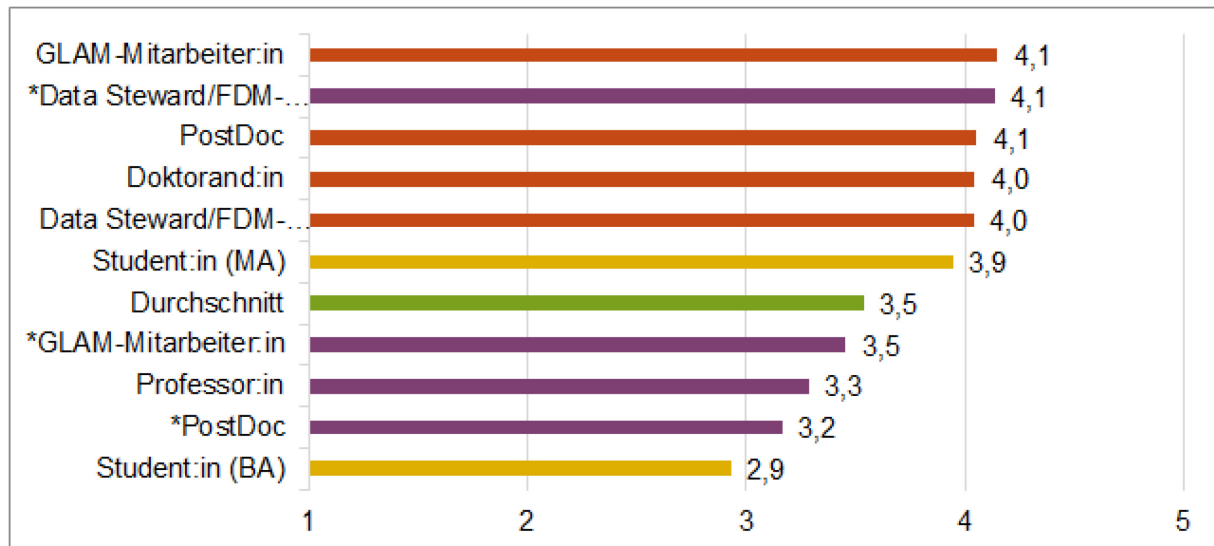


Abb. 60: Verteilung der Statusgruppen, Abfrage des Begriffs „Speichersystem“ | n = 550

Durchschnittlich wurde die Frage nach Speichersystemen mit 3,5 bewertet. Am kompetentesten schätzen sich GLAM-Mitarbeitende ohne Lehre und lehrende FDM-Mitarbeitende mit jeweils 4,1 ein. Den niedrigsten Wert hat die Statusgruppe Studierende im Bachelor mit 2,9 erreicht. Hier zeigt sich deutlich die Verteilung nach Niveaustufen. Alle Statusgruppen der Niveaustufe 2 haben Werte über 4,0 erreicht und befinden sich damit geschlossen über dem Durchschnitt. Gegenteiliges lässt sich bei der Niveaustufe 3 feststellen. Bis auf FDM-Mitarbeitende befinden sich die restlichen Statusgruppen unterhalb des Durchschnitts. Der Unterschied zwischen lehrend und nicht-lehrend liegt bei den Statusgruppen PostDocs bei $\Delta = 0,9$, GLAM-Mitarbeitende bei $\Delta = 0,6$ und FDM-Mitarbeitende bei $\Delta = 0,1$.

Grundprinzipien der Datensicherung

Niveaustufe 1. Ich weiß, was Grundprinzipien der Datensicherung sind.

Niveaustufe 2. Die Grundprinzipien der Datensicherung integriere ich in meine Arbeit mit Daten.

Niveaustufe 3. Ich kann vermitteln, was Grundprinzipien der Datensicherung sind.

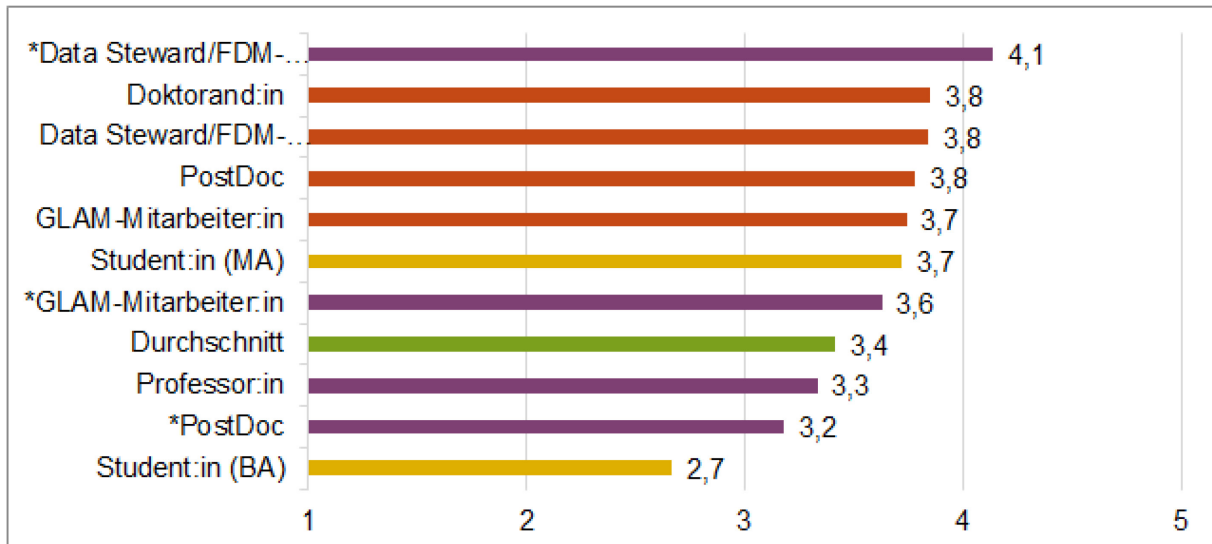


Abb. 61: Verteilung der Statusgruppen, Abfrage nach Grundprinzipien der Datensicherung | n = 543

Die Frage nach Grundprinzipien der Datensicherung wurde durchschnittlich mit 3,4 bewertet. Am kompetentesten schätzen sich FDM-Mitarbeitende mit Lehre (4,1) ein und am niedrigsten Bachelorstudierende (2,7). Ähnlich wie bei der Frage nach Speichersystemen ordnet sich die Niveaustufe 2 auch hier geschlossen oberhalb des Durchschnitts ein, mit Werten zwischen 3,7 und 3,8. Ein diverseres Bild zeichnet sich bei der Niveaustufe 3 ab. Die einzelnen Statusgruppen erreichen Werte zwischen 3,2 und 4,1. Lediglich Professor:innen und PostDocs mit Lehre liegen, wenn auch nicht signifikant, unterhalb des Durchschnitts.

Begriff „Versionierungstool“

Niveaustufe 1. Ich weiß, was ein Versionierungstool ist.

Niveaustufe 2. Ich integriere Versionierungstools in meine Arbeit.

Niveaustufe 3. Ich kann vermitteln, wie Versionierungstools verwendet werden und welchen Nutzen sie haben.

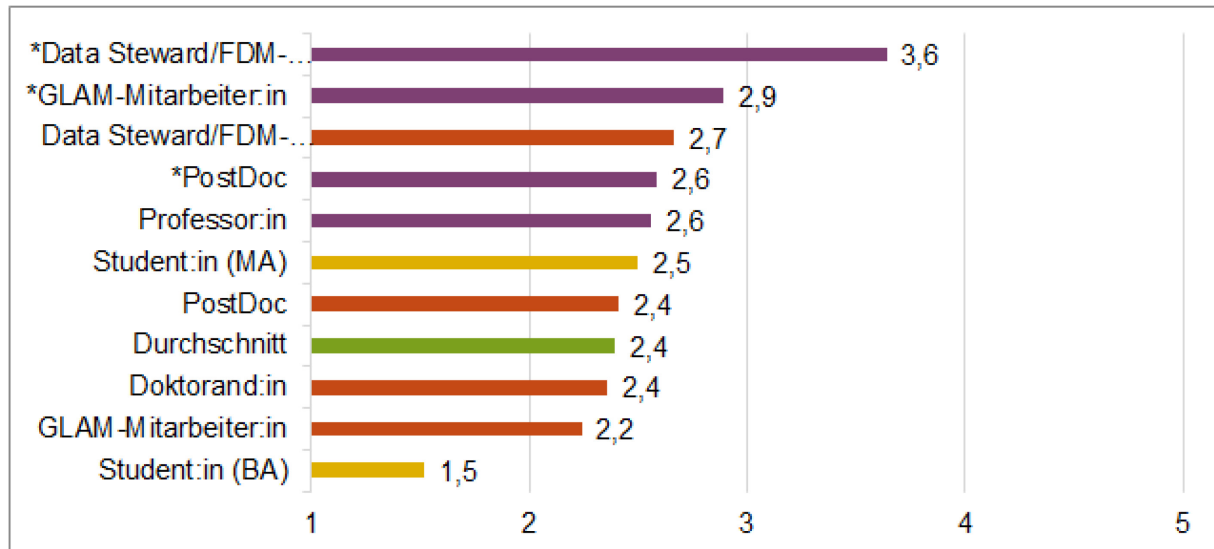


Abb. 62: Verteilung der Statusgruppen, Abfrage des Begriffs „Versionierungstool“ | n = 529

Mit einer durchschnittlichen Bewertung von 2,4 haben sich die Teilnehmenden bei der Frage nach Versionierungstools eher zurückhaltend eingeschätzt. Lediglich der höchste und niedrigste Wert stechen sehr hervor. Ein Großteil der Gruppen liegt im unteren Mittelfeld zwischen 2,2 und 2,9. FDM-Mitarbeitende mit Lehre schätzen sich bei dieser Frage mit 3,6 am kompetentesten ein, Studierende im Bachelor mit 1,5 am niedrigsten. Im Gegensatz zu den zwei bisherigen Fragen der Kategorie schätzen sich die Statusgruppen der Niveaustufe 3 mitunter am besten ein und liegen somit geschlossen oberhalb des Durchschnitts. Die Gruppen der Niveaustufe 2 erzielen Werte zwischen 2,2 und 2,7. Vergleichsweise geringe Werte sind in den Gruppen GLAM-Mitarbeitende und PostDocs ohne Lehre sowie Doktorand:innen vorzufinden.

Speicherung und Langzeitarchivierung

- Niveaustufe 1. Ich weiß, was der Unterschied zwischen Speicherung und Langzeitarchivierung ist.
- Niveaustufe 2. Ich kann die Begriffe Speicherung und Langzeitarchivierung sicher voneinander abgrenzen.
- Niveaustufe 3. Ich kann vermitteln, was der Unterschied zwischen Speicherung und Langzeitarchivierung ist.

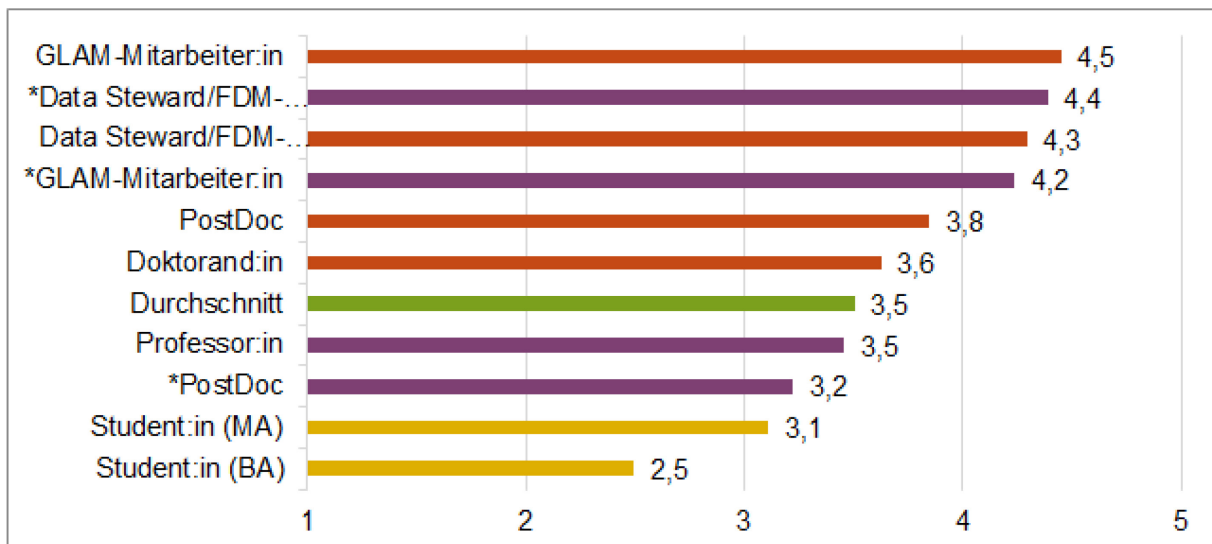


Abb. 63: Verteilung der Statusgruppen, Abfrage nach Speicherung und Langzeitarchivierung | n = 552

Wesentlich kompetenter schätzen sich die Personen bei der Frage nach dem Unterschied zwischen Speicherung und Langzeitarchivierung ein. Der Durchschnittswert liegt bei 3,5. Am kompetentesten schätzen sich GLAM-Mitarbeitende ohne Lehre (4,5) und am niedrigsten Bachelorstudierende (2,5) ein. Auffällig ist die hohe Einschätzung der Gruppen GLAM- und FDM-Mitarbeitende mit und ohne Lehre. Diese erreichen alle Werte über 4,0.

Begriff „Datenbanksystem“

Niveaustufe 1. Ich weiß, was ein Datenbanksystem ist.

Niveaustufe 2. Ich kenne fachspezifische Datenbanksysteme und kann deren Funktionalität nutzen.

Niveaustufe 3. Ich kann vermitteln, wie Datenbanksysteme aufgebaut sind und welchen fachspezifischen Nutzen sie erfüllen.

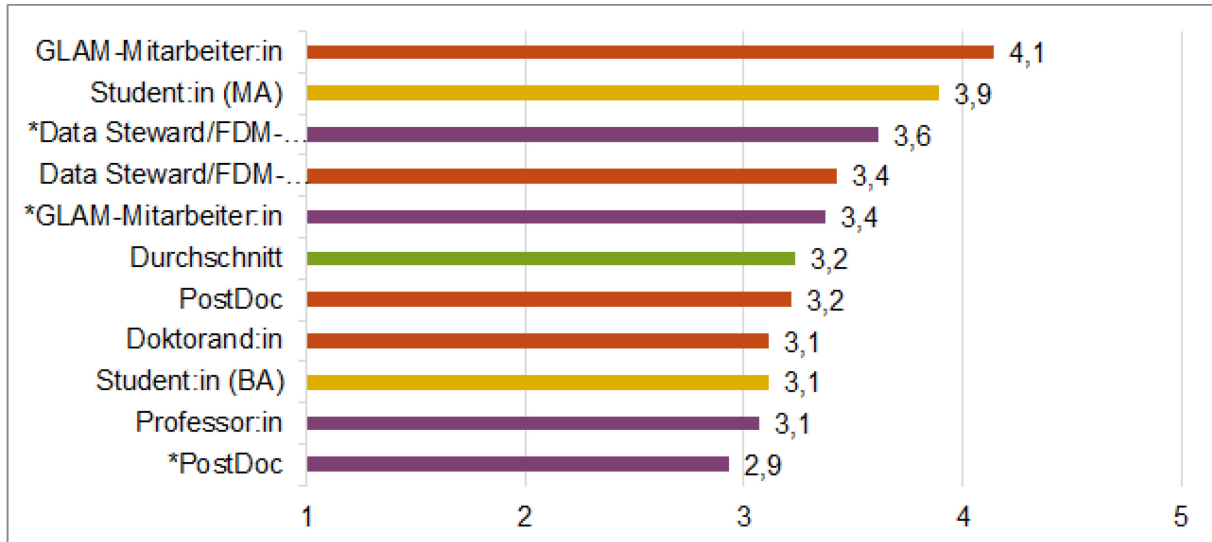
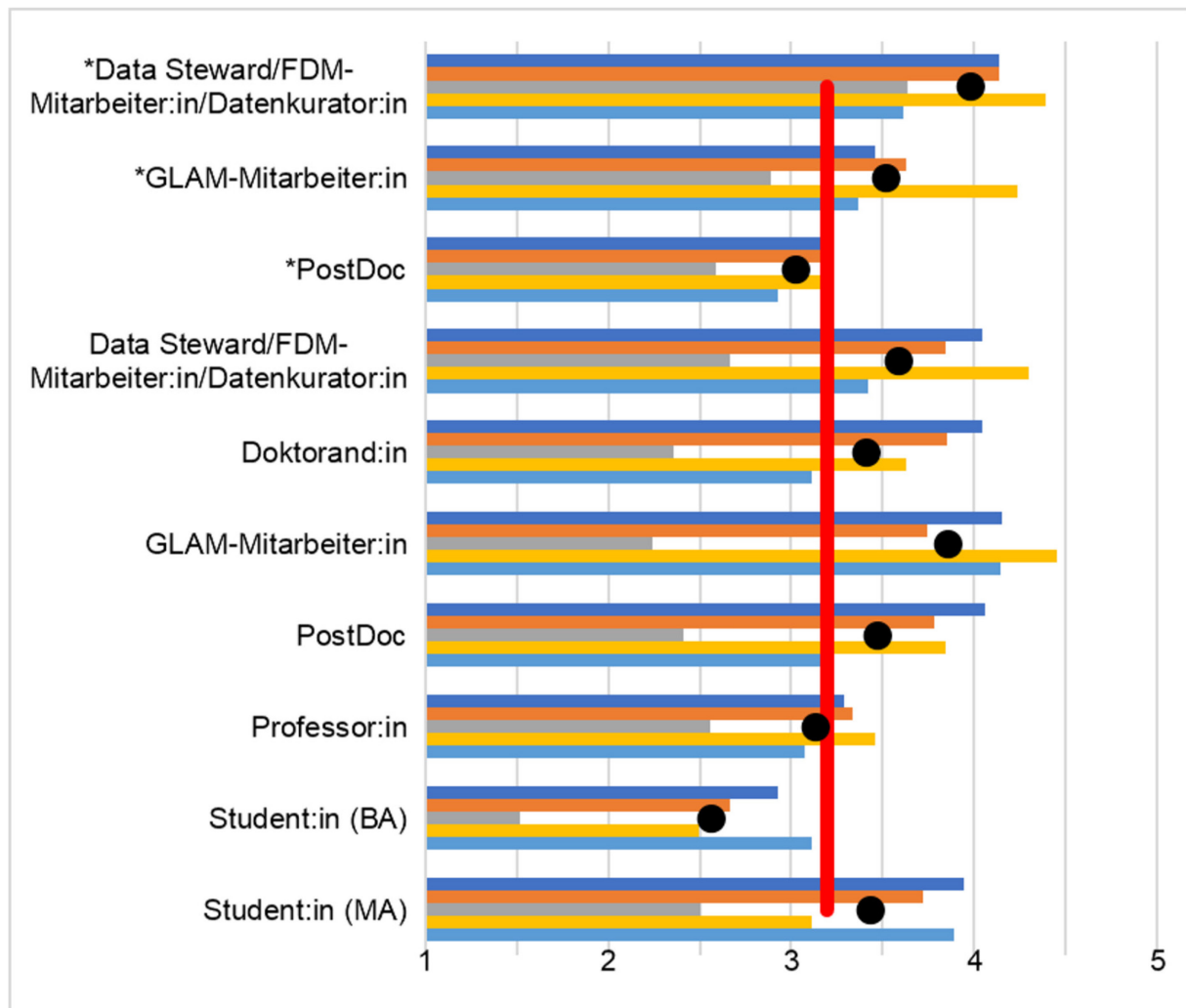


Abb. 64: Verteilung der Statusgruppen, Abfrage nach Speicherung und Langzeitarchivierung | n = 552

Die Frage nach Datenbanksystemen wird im Durchschnitt mit 3,2 bewertet. Am kompetentesten schätzen sich GLAM-Mitarbeitende ohne Lehre mit 4,1 und am niedrigsten PostDocs mit Lehre mit 2,9 ein. Allgemein liegen die meisten Gruppen eng beieinander, mit einem Unterschied von $\Delta = 0,1$ – $0,2$. Auffällig ist hier die hohe Einschätzung der Masterstudierenden. Mit 3,9 erreichen sie den zweithöchsten Wert und liegen damit nicht nur mit $\Delta = 0,7$ oberhalb des Durchschnitts, sondern auch mit einem Unterschied von $\Delta = 0,8$ über den Bachelorstudierenden.

Kurzzusammenfassung der Kategorie „Speicherung“



- Begriff „Speichersystem“
- Grundprinzipien der Datensicherung
- Begriff „Versionierungstool“
- Speicherung und Langzeitarchivierung
- Begriff „Datenbanksystem“
- Gesamtmittelwert
- Mittelwert pro Gruppe

Abb. 65: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Speicherung“ | n = 564

Die Kategorie „Speicherung“ wird im Gesamten eher neutral mit durchschnittlich 3,2 bewertet. FDM-Mitarbeitende mit Lehre erreichen mit 4,0 den höchsten und Bachelorstudierende mit 2,6 den geringsten Wert. Im Allgemeinen liegt die Einschätzung der Gruppen zwischen 3 und 4, mit Unterschieden von $\Delta = 0,1-0,3$. Auffällig ist die durchweg geringere Selbsteinschätzung der Professor:innen und PostDocs mit Lehre, die in vier von fünf Fragen knapp unterhalb des Gesamtdurchschnitts liegen.

Publikation

Begriff „Datenpublikation“

Niveaustufe 1. Ich weiß, was unter Datenpublikation zu verstehen ist.

Niveaustufe 2. Ich weiß, wie man Daten publizieren kann.

Niveaustufe 3. Ich kann vermitteln, was Inhalte und Nutzen einer Datenpublikation sind.

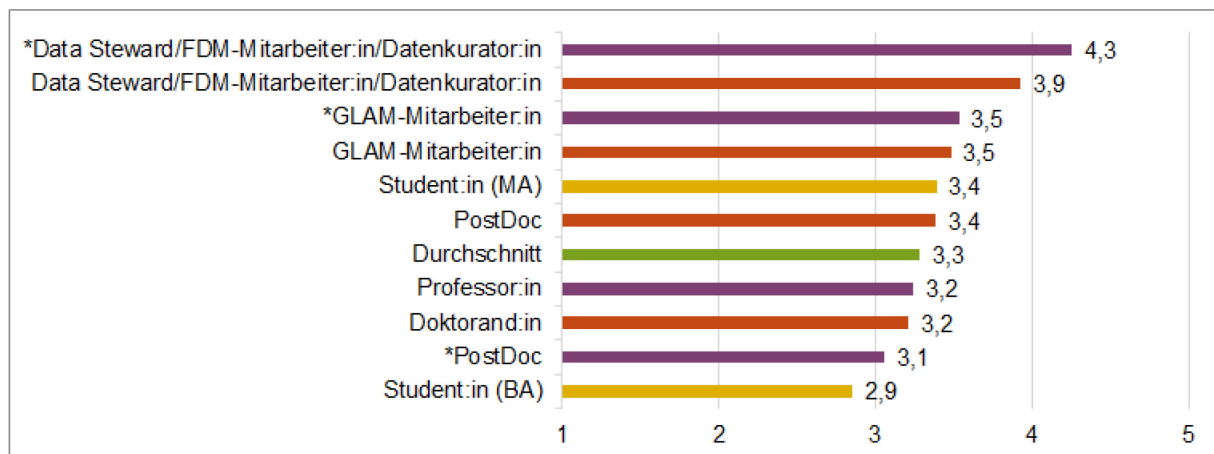


Abb. 66: Verteilung der Statusgruppen, Abfrage des Begriffs „Datenpublikation“ | n = 557

Bei der Frage nach „Datenpublikation“ erreichen die Proband:innen eine durchschnittliche Bewertung von 3,3. Mit 4,3 schätzen sich die FDM-Mitarbeitenden mit Lehre am kompetentesten und die Bachelorstudierenden mit 2,9 am niedrigsten ein. Generell liegen die einzelnen Statusgruppen nah beieinander ($\Delta = 0,1-0,2$). Anhand der Verteilung lässt sich keine Tendenz auf Ebene der Niveaustufen erkennen. Interessant ist die geringe Selbsteinschätzung der Gruppen Professor:innen und PostDocs mit Lehre. Beide befinden sich, wenn auch nicht signifikant, unterhalb des Durchschnitts.

Fachspezifische Datenpublikationsorgane

Niveaustufe 1. Ich kenne fachspezifische Datenpublikationsorgane

Niveaustufe 2. Ich nutze fachspezifische Publikationsorgane.

Niveaustufe 3. Ich kann vermitteln, welche fachspezifischen Datenpublikationsorgane es gibt.

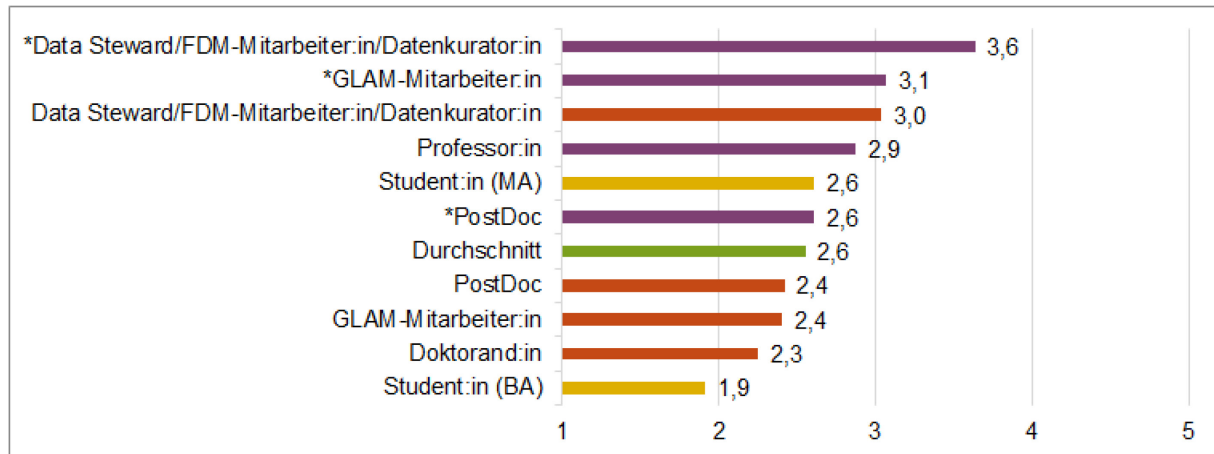


Abb. 67: Verteilung der Statusgruppen, Abfrage des Begriffs „Datenpublikation“ | n = 557

Die Frage nach fachspezifischen Datenpublikationsorganen wurde durchschnittlich mit 2,6 bewertet. FDM-Mitarbeitende mit Lehre schätzen sich mit 3,6 am kompetentesten und Bachelorstudierende mit 1,9 am niedrigsten ein. Allgemein zeigt sich hier eine erwartbare Verteilung. Auffällig ist jedoch die geringe Selbsteinschätzung der Niveaustufe 2. Bis auf FDM-Mitarbeitende liegen die Statusgruppen unterhalb des Durchschnitts. Die Ergebnisse bei den Gruppen der Professor:innen, Doktorand:innen und PostDocs von 2,3 bis 2,9 zeigen, dass die Praxis der Datenpublikation noch nicht in der Breite etabliert ist.

Kurzzusammenfassung der Kategorie „Publikation“

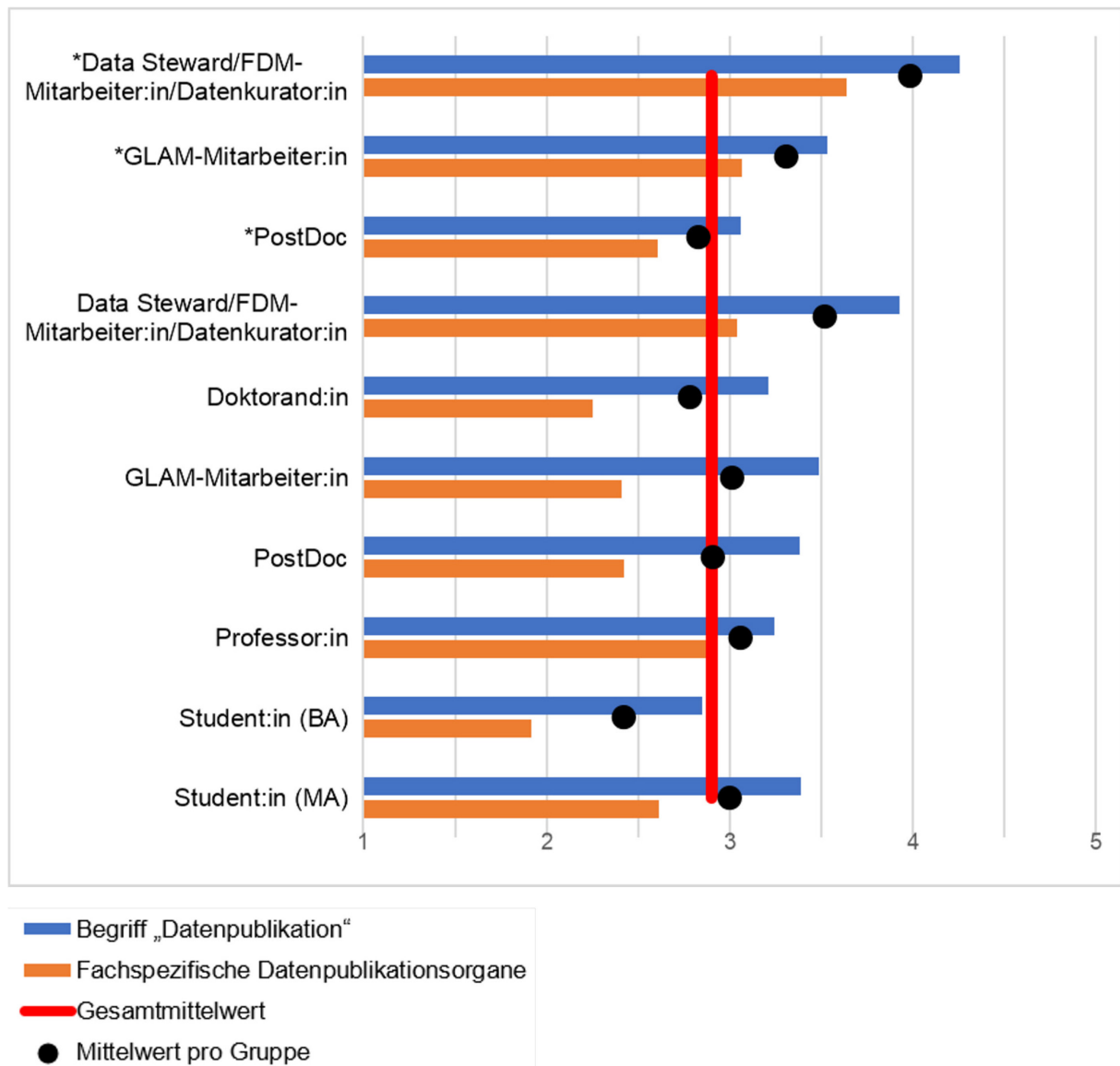


Abb. 68: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Publikation“ | n = 558

Mit einer durchschnittlichen Bewertung von 2,9 in der Kategorie „Publikation“ schätzen sich die Proband:innen eher vorsichtig ein. Die höchste Selbsteinschätzung im Mittel der Kategorie haben durchweg FDM-Mitarbeitende mit 3,5 ohne Lehrtätigkeit und 4,0 mit Lehraufgaben. Mit $\Delta = 0,6$ und $\Delta = 1,1$ liegen sie deutlich oberhalb des Durchschnitts.

Datennachnutzung

Nachnutzbarkeit von Daten und Datenformaten

- Niveaustufe 1. Über die Nachnutzbarkeit von Daten und Datenformaten mache ich mir während des Forschungsprozesses Gedanken.
- Niveaustufe 2. Ich weiß, welche Voraussetzungen erfüllt werden müssen, um (meine) Daten nachnutzbar zu machen.
- Niveaustufe 3. Ich kann vermitteln, warum die Nachnutzbarkeit von Daten und Datenformaten relevant ist.

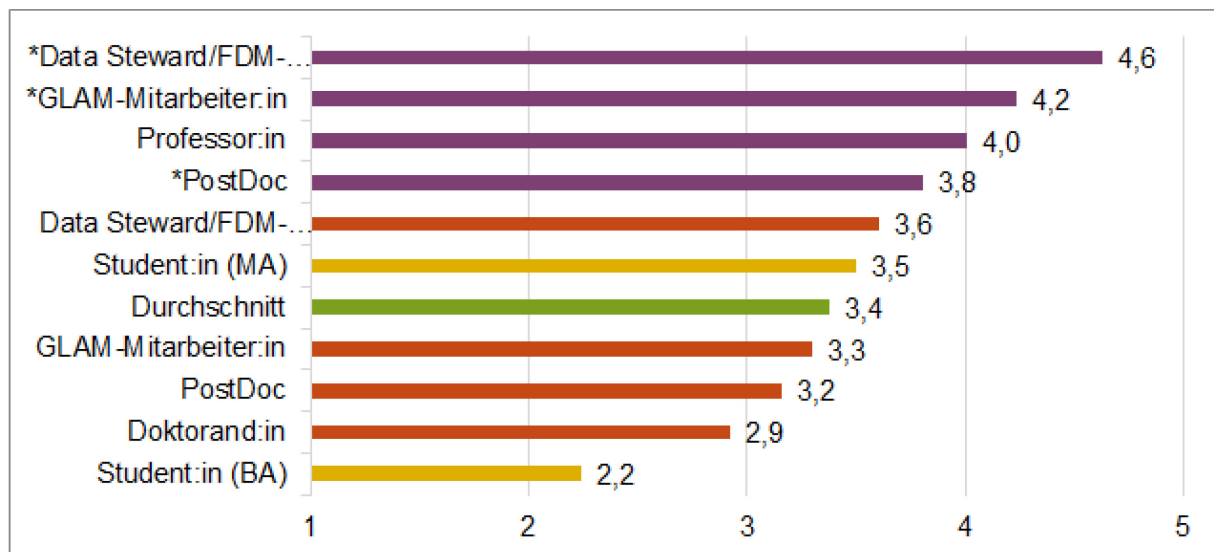


Abb. 69: Verteilung der Statusgruppen, Abfrage der Nachnutzbarkeit von Daten und Datenformaten | n = 555

Die Frage nach der Nachnutzbarkeit von Daten und Datenformaten wurde durchschnittlich mit 3,4 bewertet. Am kompetentesten schätzen sich FDM-Mitarbeitende mit Lehre (4,6) und am niedrigsten Bachelorstudierenden (2,2) ein. Die höchsten vier Werte von 3,8 bis 4,6 nehmen die Gruppen der Niveaustufe 3 ein. Gegensätzlich dazu befinden sich die Gruppen der Niveaustufe 2 bis auf FDM-Mitarbeitende ohne Lehre geschlossen unterhalb des Durchschnitts, mit Werten zwischen 2,9 und 3,6. Im Hinblick auf Statusgruppen mit und ohne Lehre sind signifikante Unterschiede vorzufinden. Zwischen PostDocs mit und ohne Lehre besteht ein Unterschied von $\Delta = 0,6$, zwischen GLAM-Mitarbeitenden von $\Delta = 0,9$ und zwischen FDM-Mitarbeitenden ein Unterschied von $\Delta = 1,0$. Diese Unterschiede lassen sich möglicherweise durch die unterschiedliche Fragestellung zwischen den Niveaustufen 2 und 3 erklären.

Begriff „FAIR“

Niveaustufe 1. Ich kenne das Akronym FAIR.

Niveaustufe 2. Ich achte darauf, die FAIR-Prinzipien im Hinblick auf meine Daten einzuhalten.

Niveaustufe 3. Ich kann vermitteln, was hinter dem Akronym FAIR steht.

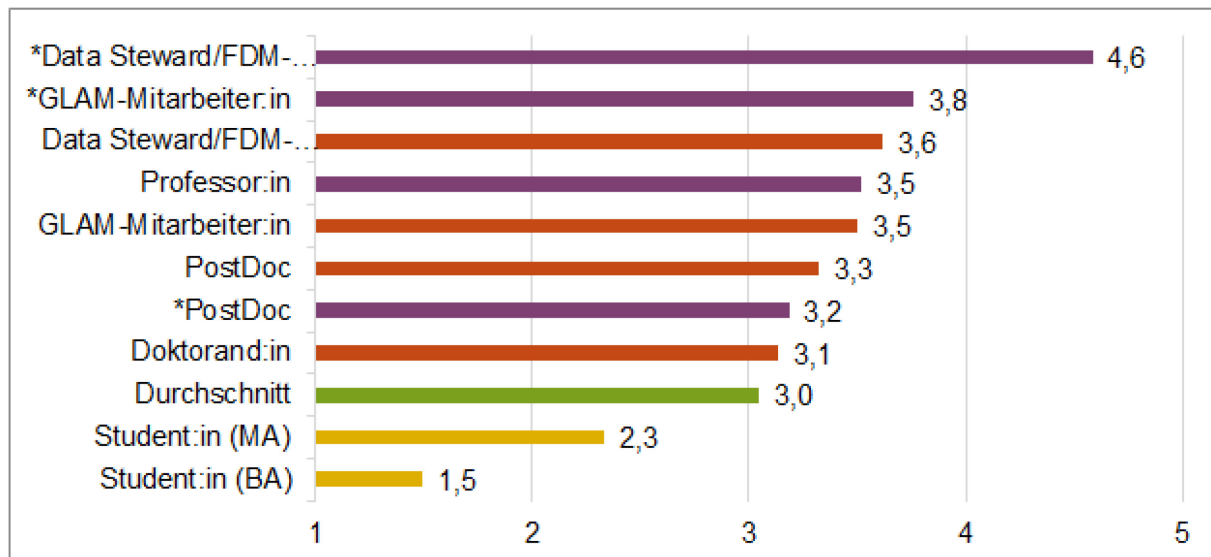


Abb. 70: Verteilung der Statusgruppen, Abfrage des Begriffs „FAIR“ | n = 499

Die Frage nach dem Akronym FAIR wurde eher disparat beantwortet. Während FDM-Mitarbeitende mit Lehre sich mit 4,6 am kompetentesten eingeschätzt haben, erreichen Studierende im Bachelor mit 1,5 einen der geringsten Werte der Umfrage. Der Unterschied zwischen diesen beiden Gruppen liegt bei $\Delta = 3,1$. Die meisten Gruppen schätzen sich zwischen 3 und 4 ein, bewegen sich also eher im positiven Mittelfeld. Der allgemeine Durchschnitt liegt bei 3,0. Niveaustufe 1 befindet sich als einzige geschlossen unterhalb des Durchschnitts. Bis auf FDM-Mitarbeitende mit und ohne Lehre erreichen die Statusgruppen der Niveaustufe 2 und 3 nah beieinanderliegende Werte mit Unterschieden von $\Delta = 0,1$ – $0,2$. Die Bewertungen der Masterstudierenden zeigen zwar einen Anstieg im Vergleich zu den Bachelorstudierenden, jedoch befinden sich diese mit $\Delta = 0,7$ dennoch deutlich unterhalb des allgemeinen Durchschnitts. Auffällig ist außerdem der hohe Anteil der Antwort „nicht beurteilbar“, die von 69 Teilnehmenden gewählt wurde.

Methoden und Strategien, um Daten nachnutzbar(er) zu machen

- Niveaustufe 1. Mir sind Methoden und Strategien bekannt, um Daten nachnutzbar(er) zu machen.
- Niveaustufe 2. Ich wende spezifische Methoden und Strategien an, um meine Daten nachnutzbar(er) zu machen.
- Niveaustufe 3. Ich kann vermitteln, wie man Methoden und Strategien anwendet, um Daten nachnutzbar(er) zu machen.

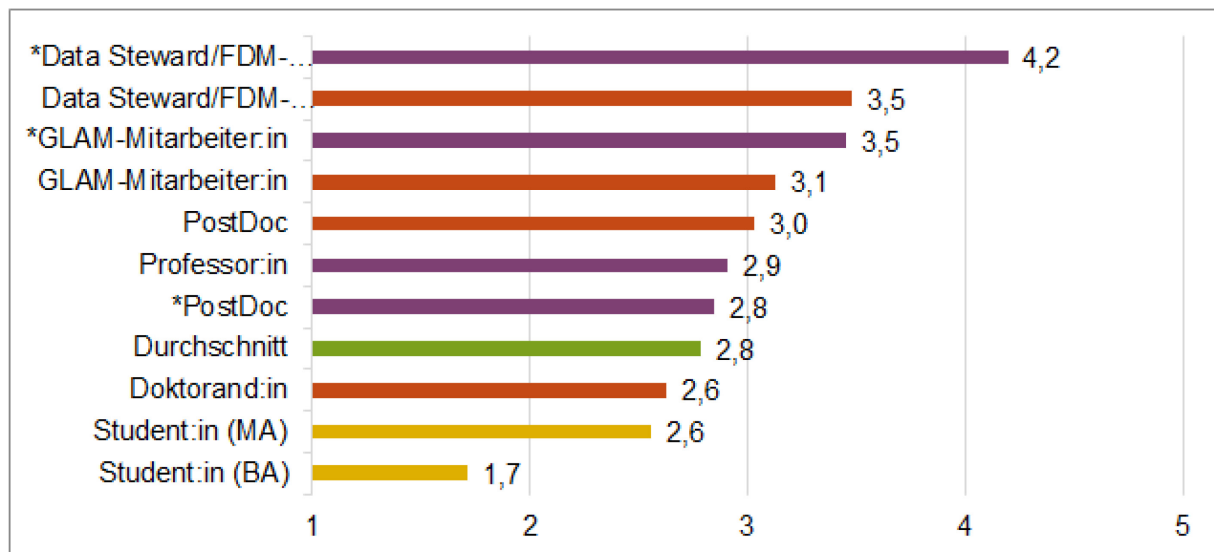


Abb. 71: Verteilung der Statusgruppen, Abfrage nach Methoden und Strategien um Daten nachnutzbar(er) zu machen | n = 531

Die Frage nach Methoden und Strategien, um Daten nachnutzbar(er) zu machen, wurde im Durchschnitt mit 2,8 bewertet und erreicht damit die geringste Durchschnittsbewertung der gesamten Kategorie „Datennachnutzung“. Auch bei dieser Frage ist eine große Diskrepanz zwischen dem niedrigsten (1,7) und dem höchsten (4,2) Wert vorzufinden, mit einem Unterschied von $\Delta = 2,5$. Bei den Gruppen mit und ohne Lehre schätzen sich die lehrenden Personen im Vergleich zu den nicht-lehrenden derselben Gruppe bis auf PostDocs besser ein. Zwischen den FDM-Mitarbeitenden mit und ohne Lehre ist ein Unterschied von $\Delta = 0,7$ vorhanden, zwischen GLAM-Mitarbeitenden ein Unterschied von $\Delta = 0,4$ und zwischen PostDocs ein Unterschied von $\Delta = 0,2$. Innerhalb der Niveaustufe 2 fallen die Doktorand:innen mit 2,6 deutlich ab und liegen damit $\Delta = 0,4$ unterhalb der nächsten zugehörigen Statusgruppe der Niveaustufe. Hier schätzen sich Doktorand:innen erstmals im Vergleich zu PostDocs deutlich geringer ein. Ein Kompetenzzuwachs von Masterstudierenden zu Doktorand:innen ist bei dem Thema nicht erkennbar.

Folgende Aspekte der Nachnutzung kenne ich/ wende ich an/ kann ich vermitteln:

- Versionierung
- Metadatenstandards
- Interoperabilitätsstandards
- Persistente Identifikatoren
- Andere

Aufgrund eines Konzeptionsfehlers bei der Bereitstellung der Auswahllisten ist die Vergleichbarkeit zwischen den einzelnen Niveaustufen bei dieser Frage nicht möglich. Aus diesem Grund haben wir uns dazu entschieden, die Ergebnisse in Form von Einzelgrafiken zu den Fragen darzustellen, um somit vollständige Transparenz zu gewährleisten, diese jedoch nicht weiter zu beschreiben.

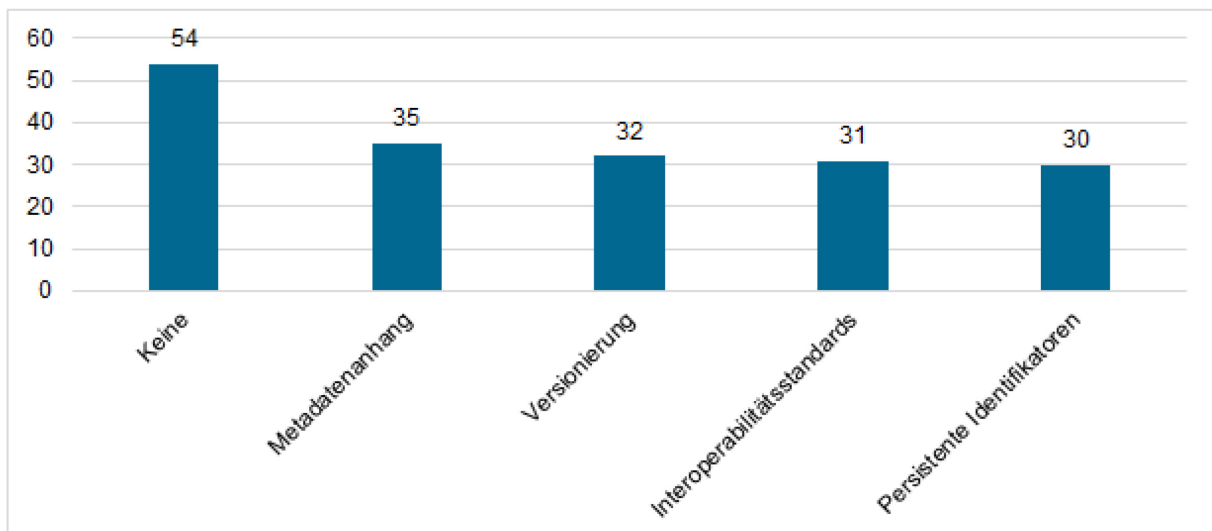


Abb. 72: „Folgende Aspekte der Nachnutzung kenne ich“ (Niveaustufe 1); Nennungen: 182 | n = 133

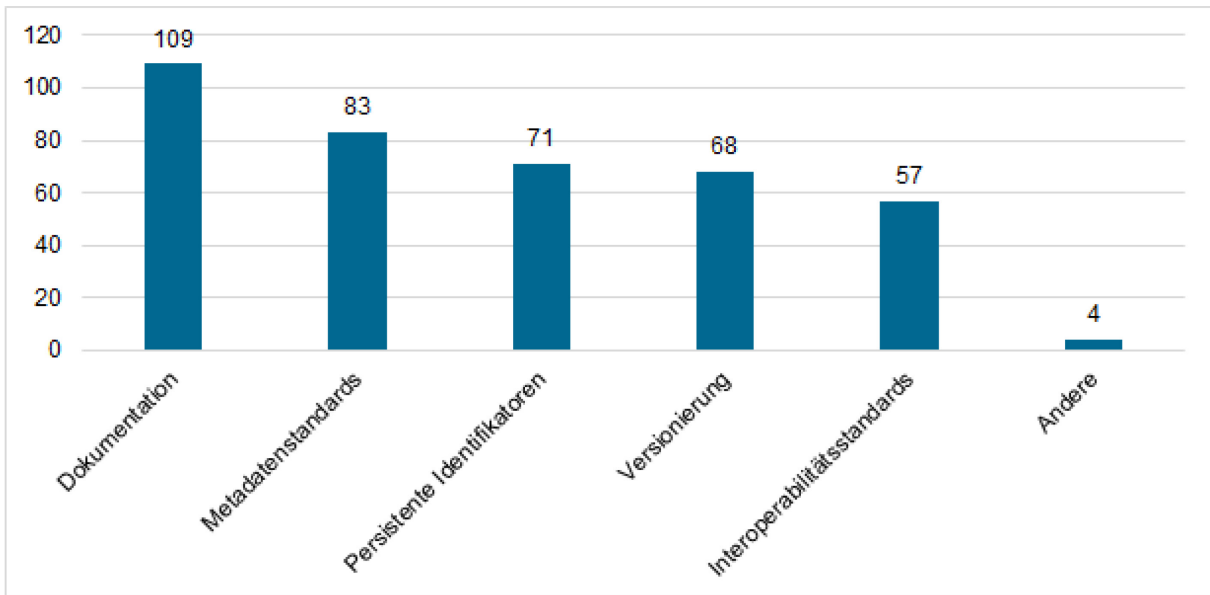


Abb. 73: „Folgende Aspekte der Nachnutzung wende ich an“ (Niveaustufe 2); Nennungen: 392 | n = 192

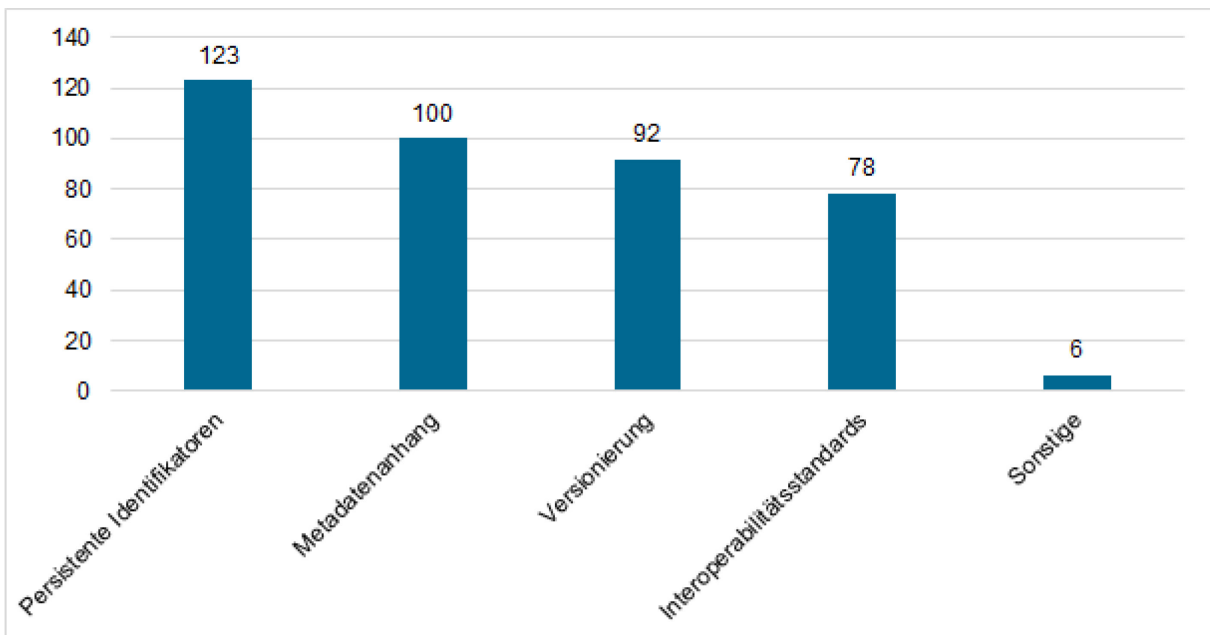


Abb. 74: „Folgende Aspekte der Nachnutzung kann ich vermitteln“ (Niveaustufe 3); Nennungen: 399 | n = 243

Kurzzusammenfassung der Kategorie „Datennachnutzung“

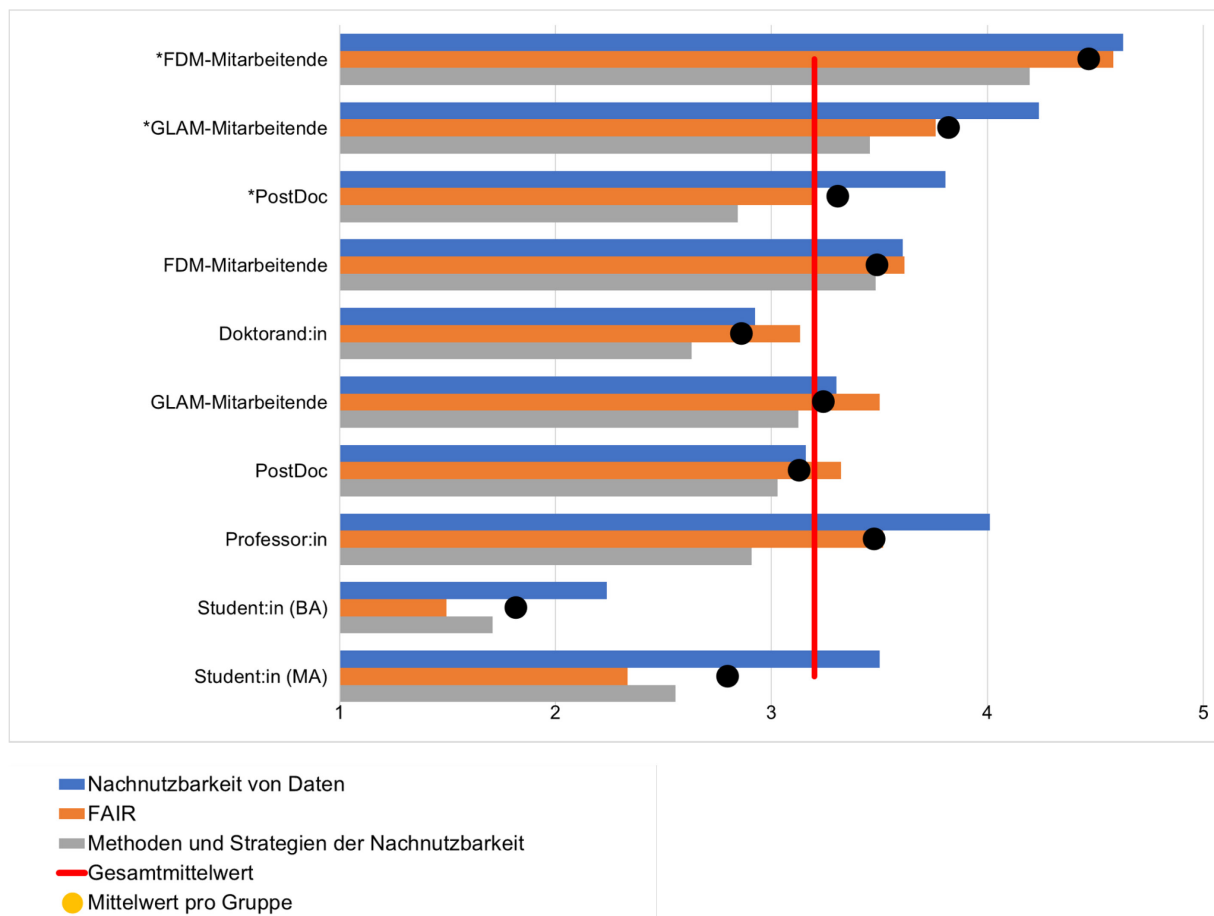


Abb. 75: Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Datennachnutzung“ | n = 560

Durchschnittlich wurde die Kategorie „Datennachnutzung“ mit 3,1 bewertet. Durchweg am kompetentesten schätzen sich FDM-Mitarbeitende mit Lehre und am geringsten Bachelorstudierende ein. Die durchschnittlich geringste Selbsteinschätzung ist bei der Frage nach Methoden und Strategien der Nachnutzbarkeit von Daten vorzufinden.

Didaktik

Anwendung didaktischer Methoden

Niveaustufe 3. Ich fühle mich hinreichend sicher in der Anwendung didaktischer Methoden zur Vermittlung von Datenkompetenzen und Forschungsdatenmanagement in meiner Lehrtätigkeit.

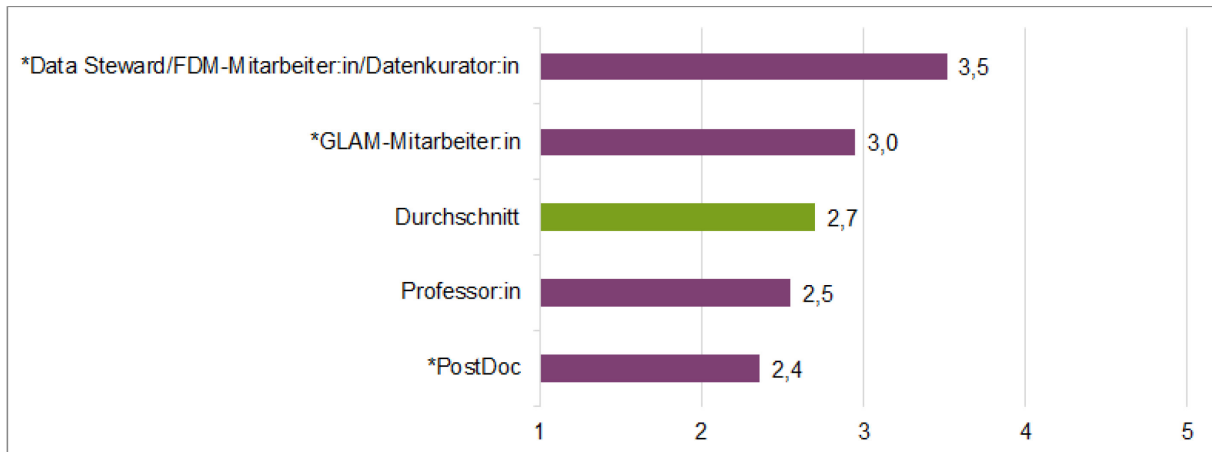


Abb. 77: Verteilung nach Statusgruppen mit Lehre, Abfrage nach der Anwendung didaktischer Methoden | n = 225

Bei der Frage nach der Anwendung didaktischer Methoden zur Vermittlung von Datenkompetenzen und Forschungsdatenmanagement liegt der Durchschnitt bei 2,7. Am kompetentesten haben sich FDM-Mitarbeitende eingeschätzt. Diese liegen mit 3,5 deutlich oberhalb des Durchschnitts. Die Werte der restlichen Gruppen bewegen sich zwischen 2,4 und 3,0. Allgemein ist eine eher zurückhaltende Einschätzung zu beobachten.

Vermittlung didaktischer Methoden

Niveaustufe 3. Ich kann vermitteln, wie didaktische Methoden eingesetzt werden, um Daten- und Forschungsdatenmanagement-Kompetenzen zu vermitteln.

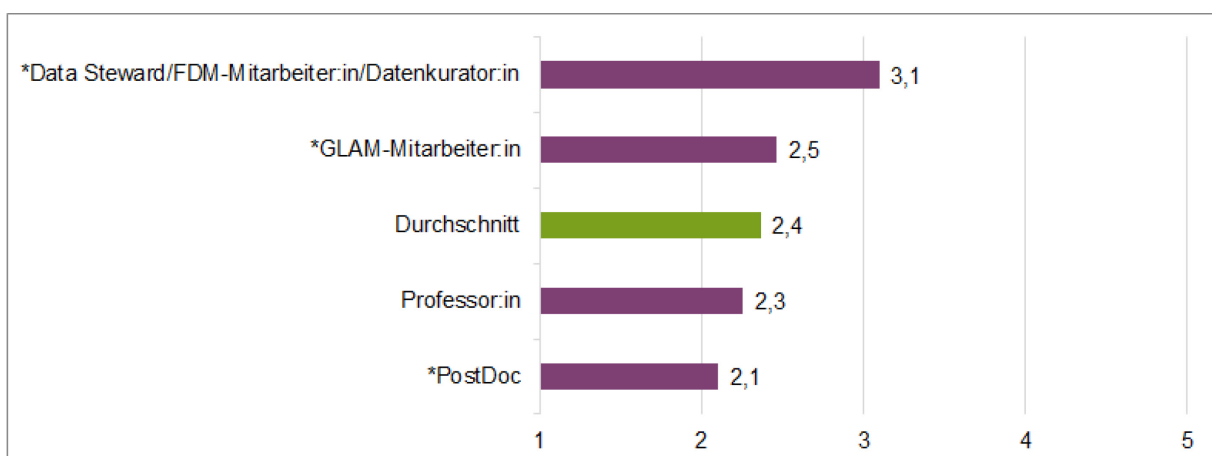


Abb. 78: Verteilung der Statusgruppen mit Lehre, Abfrage nach der Vermittlung von didaktischen Methoden | n = 218

Die Frage nach dem aktiven Einsatz von didaktischen Methoden zur Vermittlung von Daten- und Forschungsdatenmanagement-Kompetenzen wird im Durchschnitt mit 2,4 bewertet.

Hier sticht deutlich die Statusgruppe FDM-Mitarbeitende hervor. Mit 3,1 haben sich diese am besten bewertet. GLAM-Mitarbeitende, Professor:innen und PostDocs schätzen sich mit Werten zwischen 2,1 und 2,5 zurückhaltender ein. Insgesamt ist eine Unsicherheit innerhalb der Gruppen ablesbar.

Kurzzusammenfassung der Kategorie „Didaktik“

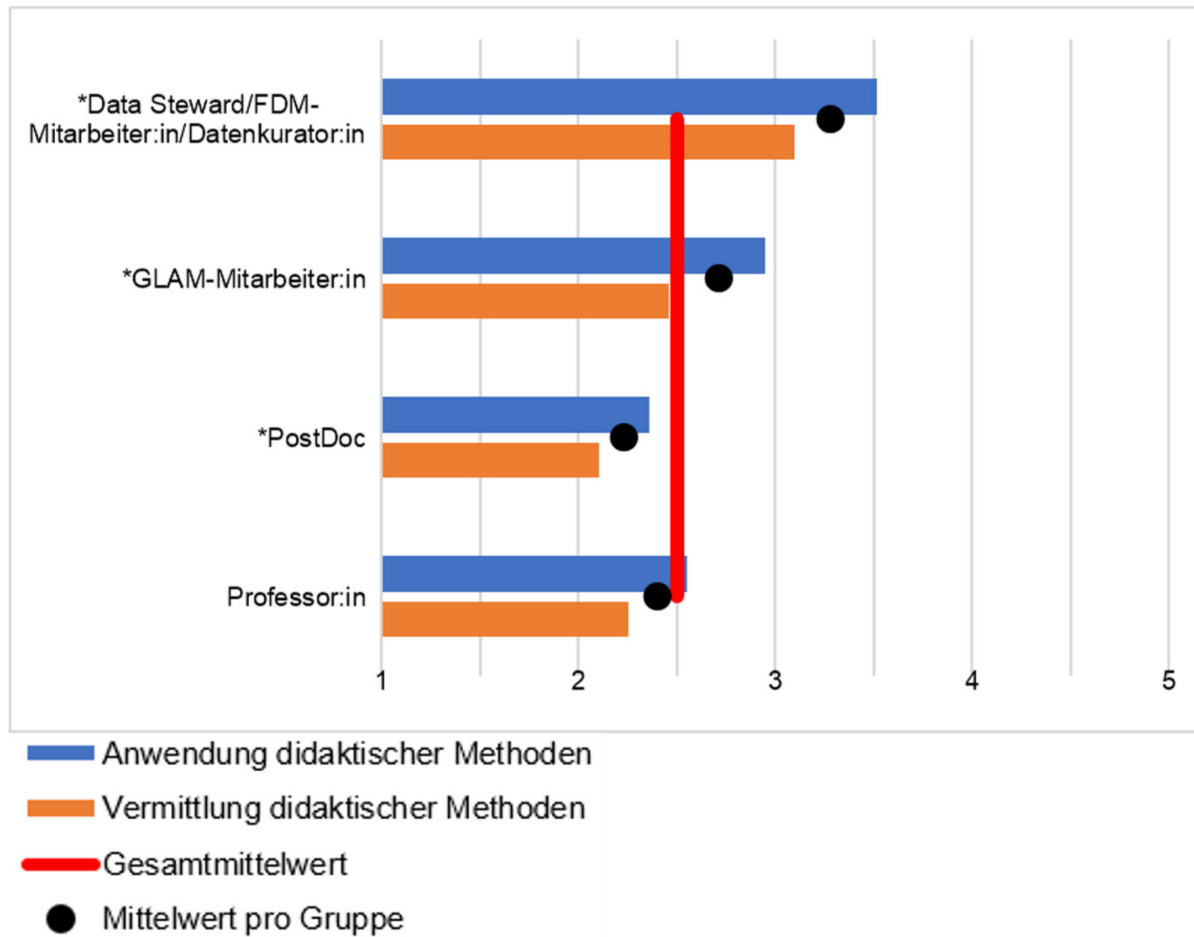


Abb. 79 Gesamtüberblick der Verteilungen der Einzelfragen pro Statusgruppe und Mittelwertangaben innerhalb der Kategorie „Didaktik“ | n =228

Die Kategorie „Didaktik“ besteht aus zwei Fragen, welche nach dem individuellen Empfinden bei der Anwendung von didaktischen Methoden zur Vermittlung von Daten- und Forschungsdatenmanagement-Kompetenzen und nach deren Vermittlung fragen. Die Reihenfolge der Einschätzung der Statusgruppen ist bei beiden Fragen identisch. So schätzen sich FDM-Mitarbeitende bei beiden Fragen am kompetentesten ein, danach folgen GLAM-Mitarbeitende, Professor:innen und PostDocs. Gut erkennbar ist hier der Unterschied zwischen der eigenen Anwendung und der Vermittlung der didaktischen Methoden an andere Lehrende.

Vermittlung und Information

Diese beiden Fragen wurden in allen drei Niveaustufen gestellt.

Über welche Formate möchte ich mich zu Datenkompetenz und Forschungsdatenmanagement weiterbilden: (Mehrfachauswahl)

- Video-Tutorial
- Beratung online
- Beratung persönlich
- Workshop online
- Workshop Präsenz
- OER (E-Learning)
- Integriert ins Studium
- Zusatzzertifikatskurs
- _____

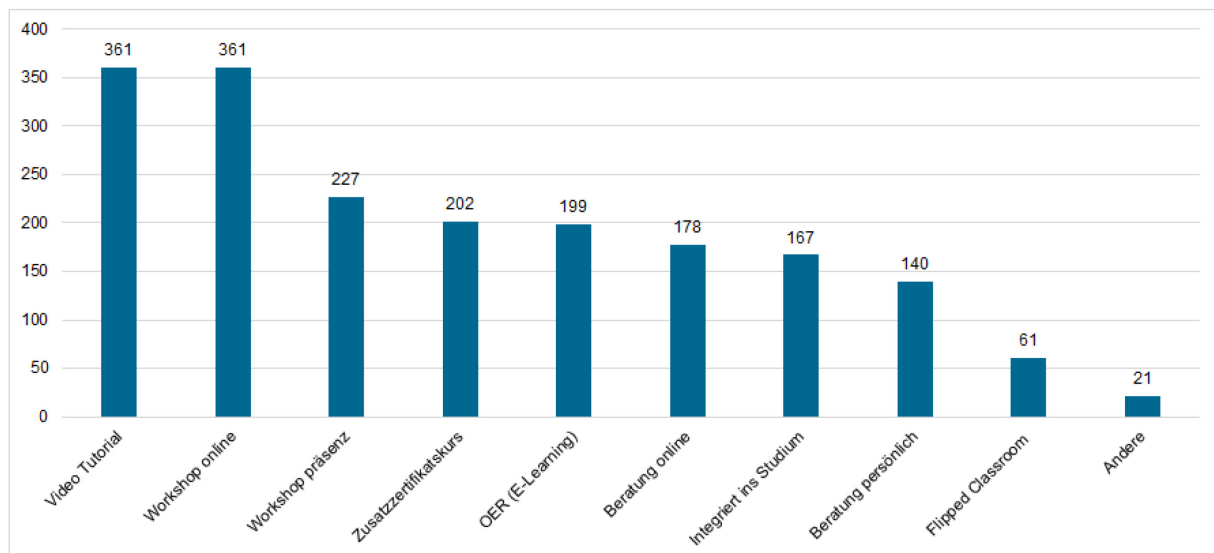


Abb. 80: Angaben zu Weiterbildungsformaten zu Datenkompetenz und Forschungsdatenmanagement, Nennungen: 1917 | n = 568

Die Frage nach Vermittlungsformaten zu Datenkompetenz und Forschungsdatenmanagement wurde von 568 Proband:innen beantwortet. Die beiden am häufigsten angeklickten Auswahlmöglichkeiten sind Video-Tutorials und Online-Workshops (jeweils 367 Stimmen). Im Mittelfeld bewegen sich Workshops in Präsenz, OER-Materialien und Zusatzzertifikatskurse. Eher geringes Interesse besteht an dem Konzept des Flipped Classrooms. Auch Beratungen fallen sowohl online als auch persönlich tendenziell eher ab.

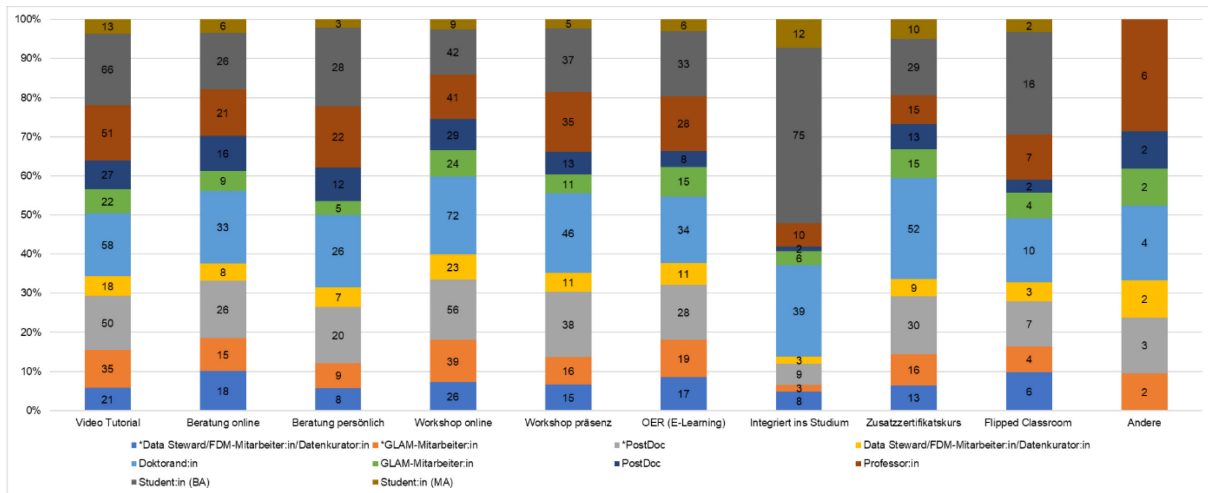


Abb. 81: Abfrage nach Weiterbildungsformaten, Verteilung nach Statusgruppe und Auswahloption | n = 568

Ein Blick auf die Verteilung innerhalb der Statusgruppen zeigt, dass die meisten Gruppen am häufigsten die Option „Workshop online“ ausgewählt haben. Das geringste Interesse besteht an „Flipped Classroom“. Interessant ist die deutliche Tendenz zu Online-Angeboten. Das Online-Äquivalent bestimmter Auswahloptionen wurde immer häufiger ausgewählt als die Varianten in Präsenz.

Neben den vorgegebenen Auswahlmöglichkeiten (konnten die Proband:innen in einem Freitextfeld unter dem Punkt „Andere“ weitere Vermittlungsformate angeben. Überwiegend wurde nach Verschriftlichungen gefragt, ob Dokumentationen, OER-Materialien oder Tutorials. Kurze (einführende) Texte scheinen gefragter zu sein als „langatmige“ Videos.

Weitere relevante Themen und Feedback

Ich erachte folgende weitere Themen als relevant, die im Fragebogen nicht vorgekommen sind:

Insgesamt nutzten diese Gelegenheit weitere Themen zu benennen 125 Personen und gaben mindestens einen Kommentar bei der Frage nach weiteren relevanten Themen ab.

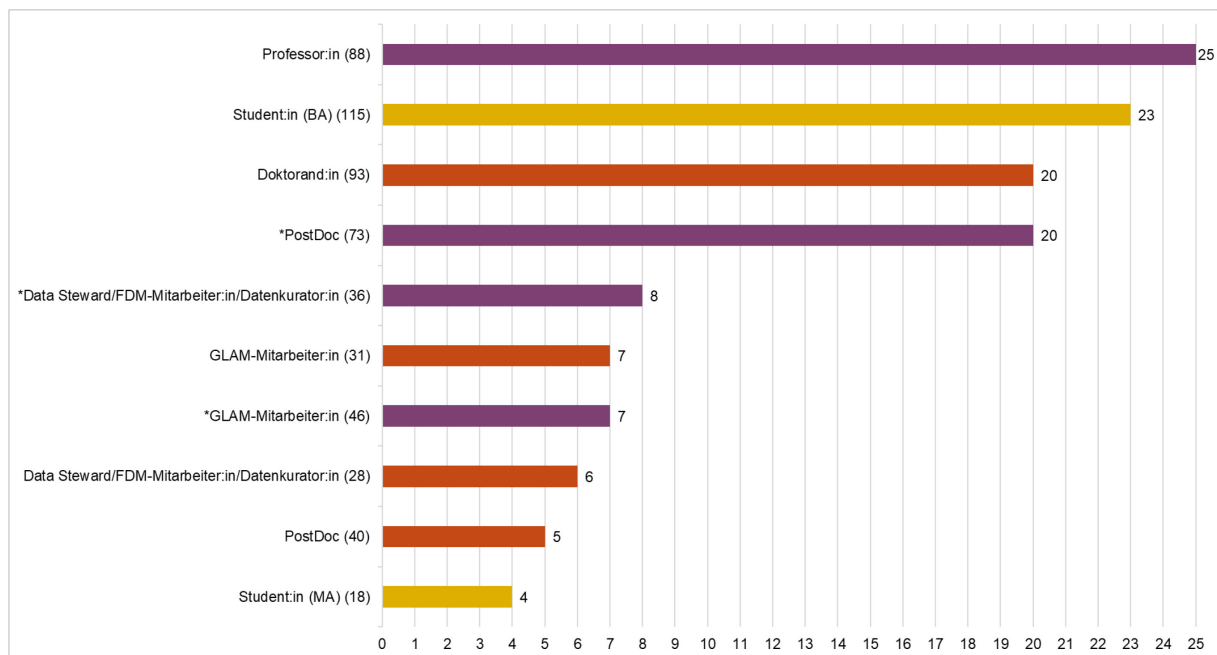


Abb. 82: Anzahl der Teilnehmer:innen mit mindestens einer Antwort pro Statusgruppe (in Klammern die Grundgesamtheit der Statusgruppen) | n = 568

Die Gruppen Professor:innen und Bachelorstudierende haben mit 25 bzw. 23 Stimmen am häufigsten teilgenommen. Dabei wurden nicht nur weitere Themen genannt, sondern auch Feedback und Kommentare zum Fragebogendesign und der Umfrage im Allgemeinen gegeben. Überwiegend wurden zu bereits abgefragten Themenfeldern weitere Aspekte gewünscht. Eine Reihe von Kommentaren können unter dem Begriff (geschichtswissenschaftliche) Forschungsdaten zusammengeführt werden. Inhaltlich zeigt sich eine grundlegende Unsicherheit darüber, ob es überhaupt Forschungsdaten in den Geschichtswissenschaften gibt und falls ja, wie genau diese aussehen.

„[R]eine - ‚Forschungsdaten‘ existieren in der Geschichtswissenschaft nicht. Es gibt – Quellen‘, aufbewahrt und gesichert und ggfs. überprüfbar in Archiven.“

„Fragebogen auf mein Wissenschaftsfeld nicht recht anwendbar; Fragen treffen nicht zu, da in der Geschichtswissenschaft idR keine Forschungsdaten wie etwa in der Biologie produziert werden.“

Ein weiterer identifizierter Themenblock befasst sich mit der Vermittlung und Lehre von DL- und FDM-Inhalten. Wiederholt wurde auf die Notwendigkeit der Integration dieser Inhalte ins Studium hingewiesen. Neben allgemeinen Aussagen wie: „Die Integration in Studiengänge ist ein relevantes Thema, das zu wenig Berücksichtigung findet.“ wurden auch sehr konkrete Themen und Einbindungsoptionen beschrieben: „Integration von Grundlagen zu Date[n]banken (Normalisierung etc.!) und Datenmodellierung in die geschichtswissenschaftlichen Fächer (ggf. als Wahlmodul o. ä.)“. Es wurde bspw. vorgeschlagen, Anfänger:innen durch kleine Projekte mit sofortigem Erfolg an die Thematiken heranzuführen.

An vielen Stellen wurden spezifische Aspekte zu verschiedenen Themenfeldern gewünscht. Im Bereich Programmierung wünschen sich demnach Proband:innen mehr Informationen zu Künstlicher Intelligenz und Large Language Models. Beim Umgang mit Daten empfanden die Teilnehmenden es als relevant, mehr über das Thema Quellen- / Daten- und Methodenkritik zu erfahren sowie den Umgang mit unterschiedlichen Daten (qualitativ, quantitativ, primär oder sekundär). Des Weiteren wurden die Sicherheit und der Schutz von Daten, einschließlich der Einhaltung gesetzlicher Bestimmungen wie z. B. der DSGVO genannt. Ebenso wichtig sind den Proband:innen die Vermeidung von Datenmissbrauch und die Gewährleistung einer sicheren Verarbeitung, um Fehler und Lücken in Daten zu vermeiden oder diese zumindest transparent zu machen. Die Teilnehmer:innen betonten auch die Bedeutung der Verschlüsselung und Langzeitarchivierung von Daten sowie die Zusammenarbeit zwischen der FDM-Community und Archivar:innen. Weiterhin wurde darauf hingewiesen, dass rechtliche Aspekte, wie der Umgang mit Urheberrechten sowie verschiedene rechtliche Verpflichtungen im Forschungsdatenmanagement nicht vernachlässigt werden sollten. Auch die Vermittlung von Methoden, Tools und Software wurde als wichtig erachtet, um grundlegende Statistikkennntnisse zu vermitteln und den Umgang mit Legacy-Software¹⁶ und Datenbanken zu erleichtern. Allgemein zeigt sich ein sehr diverses Spektrum an Themenfeldern und Inhalten. Vieles lässt sich in die Kategorien der Umfrage einordnen, andere Kommentare öffnen neue Bereiche wie z. B.

¹⁶ Legacy-Software bezieht sich auf ältere Software, die bereits seit einiger Zeit im Einsatz ist und oft nicht mehr aktiv weiterentwickelt wird. Diese Software kann in vielen Fällen immer noch wichtige Funktionen erfüllen, aber sie entspricht möglicherweise nicht mehr den aktuellen Standards und Anforderungen. Legacy-Software kann verschiedene Probleme mit sich bringen, darunter Sicherheitsrisiken, Inkompatibilität mit modernen Systemen und Schwierigkeiten bei der Wartung und Weiterentwicklung. Oftmals wird versucht, Legacy-Software durch neuere Lösungen zu ersetzen, die effizienter, sicherer und besser an die aktuellen Anforderungen angepasst sind.

Projektmanagement oder die Definition von Rollen und den damit einhergehenden Kenntnisständen und Kompetenzvorstellungen unterschiedlicher Akteur:innen. Schließlich wurde auch das Thema institutionelle Ressourcen adressiert, um ein umfassendes Verständnis und Know-how im Bereich Data Literacy überhaupt fördern zu können.

Zusammenfassung

Die Erhebung zielte darauf zu erfassen, wie es um die Datenkompetenzen bei den verschiedenen Statusgruppen in den historisch arbeitenden Disziplinen gemäß ihrer Selbsteinschätzung bestellt ist. Die Ergebnisse sind nicht überraschend, sondern bestätigen im Großen und Ganzen die Wahrnehmung vieler Akteur:innen in diesem Feld. Das folgende Diagramm Abb. 83 zeigt eine Gesamtauswertung der Mittelwerte für alle Kategorien für jede Statusgruppe.

Die Abbildung 83 zeigt, dass die Durchschnittswerte der Kategorien zwischen 2,3 und 3,5 liegen. Die Hälfte der Statusgruppen erzielt in der Kategorie „Datenerschließung“ die höchsten Werte. Fast durchgehend schätzen sich die Teilnehmenden in der Kategorie „Programmierung“ am wenigsten kompetent ein. Es zeigt sich, dass die Gruppen der FDM-Mitarbeiter:innen, v. a. jene mit Lehrtätigkeit, an der Spitze liegen. Sie erzielten in fast allen Kategorien höhere Selbsteinschätzungswerte als die anderen Gruppen. Insbesondere in den Bereichen „Datenerschließung“ und „Datennachnutzung“ schätzen sie sich besser ein als alle anderen Gruppen, was durch ihre Tätigkeit im Datenmanagement begründbar ist. Überraschend ist hingegen ihr geringerer Abstand zu den anderen Gruppen in der Kategorie „Datenevaluation“. Ebenfalls interessant ist der Umstand, dass die GLAM-Mitarbeitenden mit Lehre ähnlich abschneiden wie die Gruppe der FDM-Mitarbeiter:innen ohne Lehre. Dieses Ergebnis ist vermutlich auf ein ähnlich gelagertes Tätigkeitsfeld der beiden Gruppen zurückzuführen. In Bezug auf die beiden Gruppen ist weiterhin auffällig, dass sie sich im Bereich „Programmierung“ am dritt schlechtesten einschätzen und auch in der Kategorie „Recht“ recht niedrige Werte angeben.

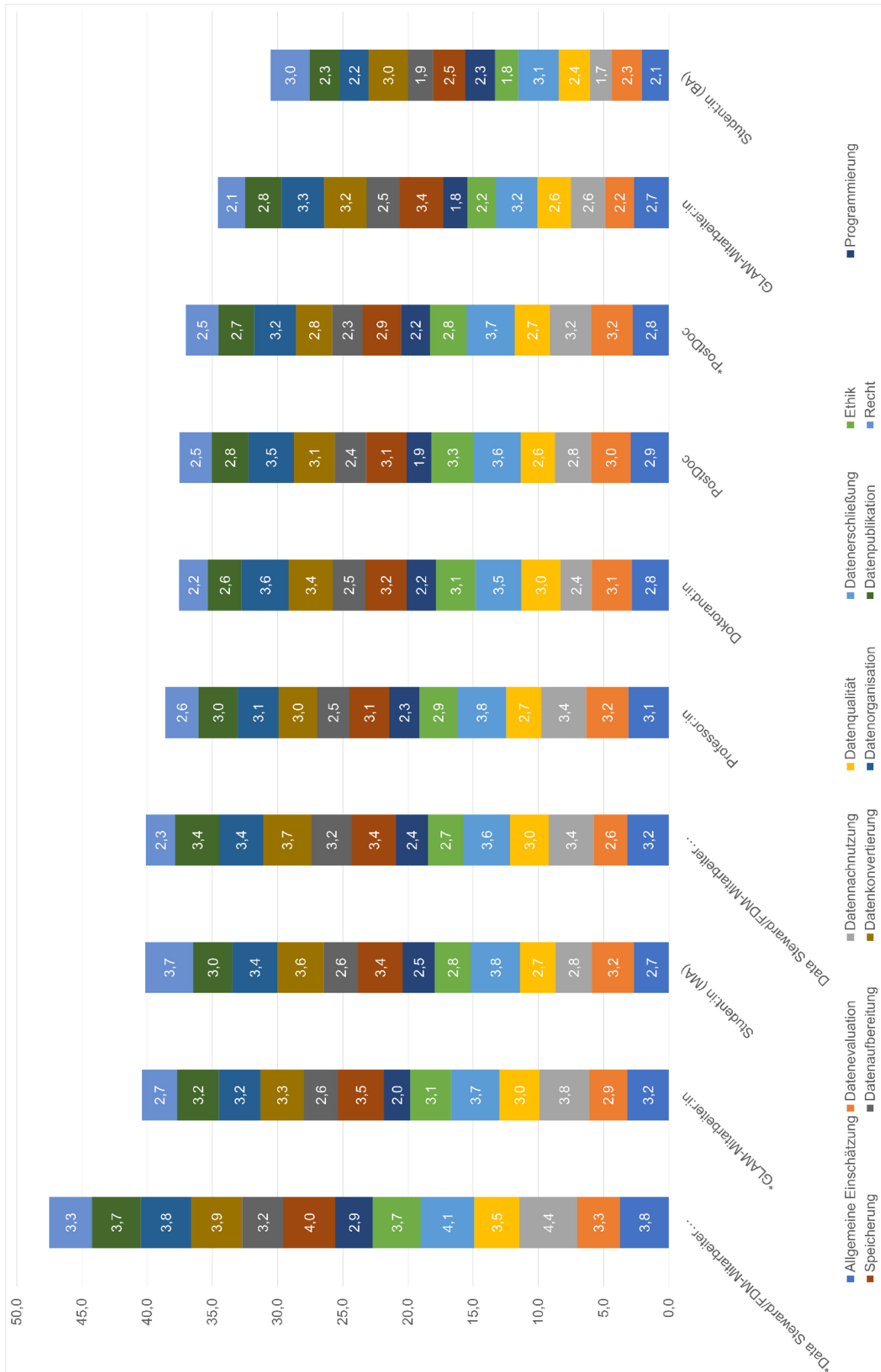


Abb. 83: Gesamtübersicht der Verteilung pro Statusgruppe über alle Kategorien | n = 567

Die Werte der Masterstudierenden müssen mit Vorsicht rezipiert werden, da sie mit nur 18 Proband:innen nicht repräsentativ sind. Am unteren Ende der Skala befinden sich die Bachelorstudierenden. Insbesondere in den Bereichen „Ethik“ und „Datennachnutzung“ zeigen sie deutlich niedrigere Werte. Offensichtlich verfügen Bachelorstudierende in diesen beiden Bereichen über weniger Erfahrung sowie weniger spezialisierte Kenntnisse. Zudem war ein großer Teil der Bachelorstudierenden erst am Anfang des ersten Semesters, als sie an der Umfrage teilnahmen. Den Bachelorstudierenden folgen die GLAM-Mitarbeitenden ohne Lehre. Sie schätzen sich in den Kategorien „Allgemeine Einschätzung“, „Datennachnutzung“, „Speicherung“, „Datenaufbereitung“ und „Datenorganisation“ deutlich besser ein als die Studierenden im Bachelor. Die Gruppen der PostDocs mit und ohne Lehre, Professor:innen und Doktorand:innen liegen im Gesamtdurchschnitt nah beieinander. Die Abweichungen zwischen diesen Gruppen liegen oft zwischen $\Delta = 0,1$ bis $\Delta = 0,4$. Aber z. B. bei der Kategorie „Datennachnutzung“ schätzen sich die Promovierenden deutlich schlechter ein als die Professor:innen und PostDocs mit Lehre, wohingegen sie sich bei der Datenkonvertierung mit $\Delta = 0,6$ besser einschätzen als PostDocs mit Lehre.

Das folgende Diagramm in Abb. 84 zeigt die Mittelwerte pro Statusgruppe über alle Kategorien hinweg.

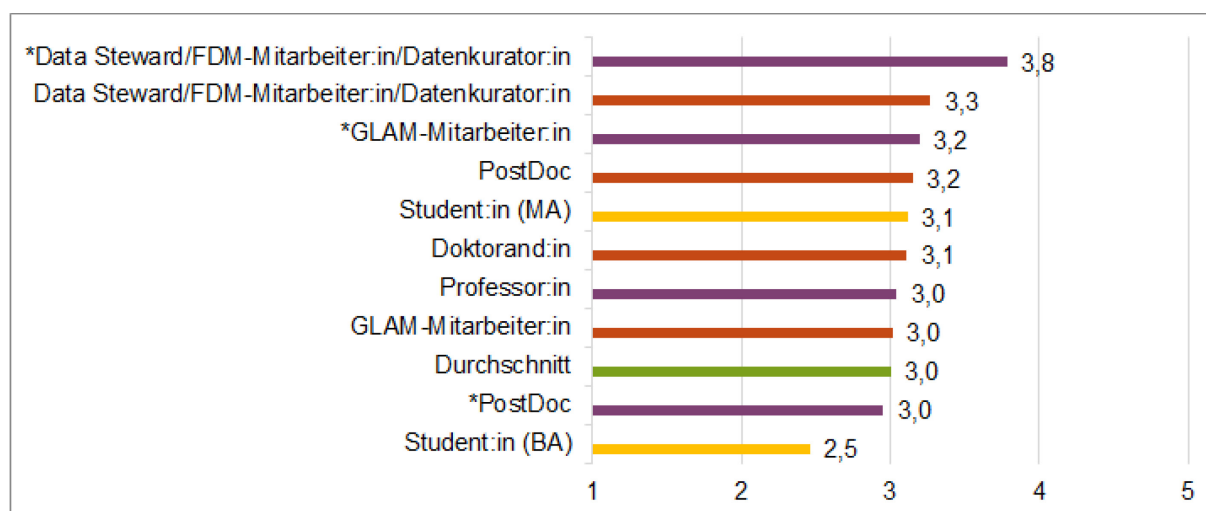


Abb. 84: Angabe der Mittelwerte pro Statusgruppe über alle Kategorien | n = 567

Die Grafik verdeutlicht nochmal den Eindruck, dass ein Großteil der Gruppen sich mit Werten zwischen 3,0 und 3,3 im Gesamtdurchschnitt dieser Umfrage im Mittelfeld bewegt. Die beiden statistischen Ausreißer sind lehrende FDM-Mitarbeitende mit 3,8 und Bachelorstudierende mit 2,5. Die durchschnittliche Einschätzung aller Statusgruppen zusammen liegt bei 3,0 und damit genau in der Mitte der Skala.

Darüber hinaus wurde eine Auswertung nach Wissensleveln vorgenommen, um eine komprimierte Übersicht über die einzelnen Gruppen zu erhalten. Der Fokus lag dabei auf

Statusgruppen, Institutionen und Fachkollegien. Die folgende Abbildung soll detaillierter aufzeigen, inwieweit sich das Kompetenzniveau innerhalb einer Gruppe unterscheidet. Um die Teilnehmenden in die Wissenslevel Lai:innen, Starter:innen, Kenner:innen und Expert:innen einzuteilen, wurde ermittelt, wie viele der Fragen von den einzelnen Teilnehmenden mit „trifft größtenteils zu“ oder „trifft völlig zu“ beantwortet wurden. Hat ein:e Teilnehmer:in bis zu 25 % der Fragen mit einer dieser beiden Optionen beantwortet, wurde sie als Lai:in eingestuft, bei 25–50 % als Starter:in, bei 50–75 % als Kenner:in und bei mehr als 75 % als Expert:in.

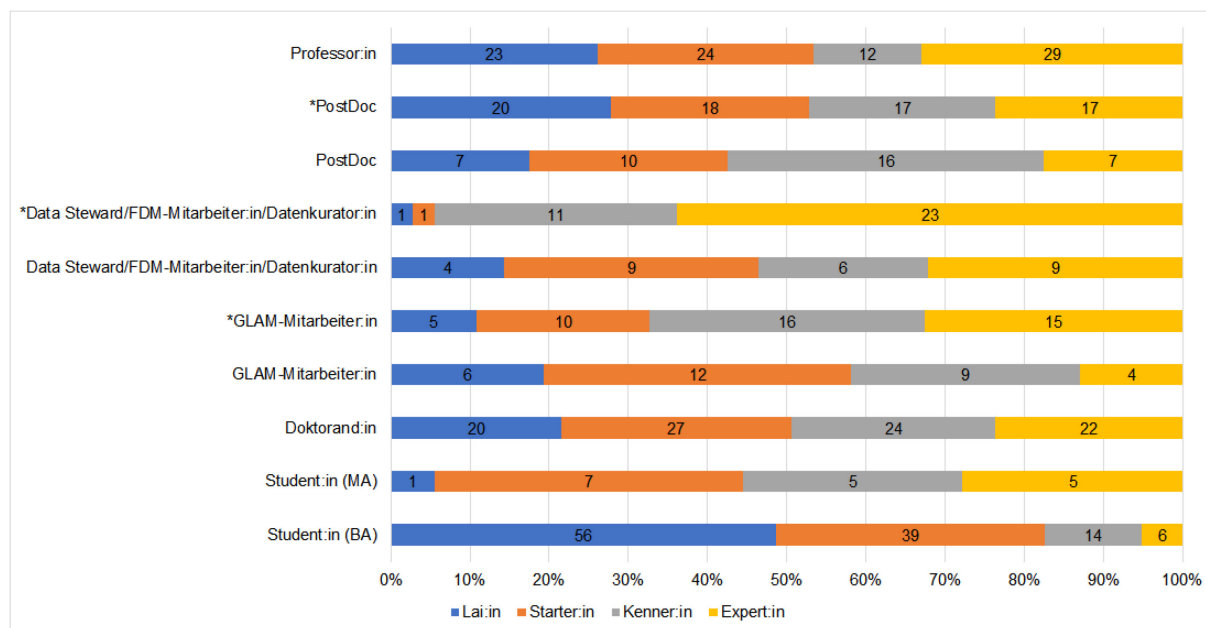


Abb. 85: Auswertung nach Wissensleveln pro Statusgruppe über alle Kategorien | n = 567

Wie bereits in den Einzelauswertungen der Fragen ersichtlich, fällt ein Großteil der Bachelorstudierenden unter das Wissenslevel Lai:in. Der Anteil an Expert:innen ist in der Gruppe der FDM-Mitarbeitenden mit Lehre am höchsten. Eine eher ausgeglichene Verteilung über alle Wissenslevel hinweg ist bei den Gruppen Doktorand:innen und PostDocs mit Lehre zu erkennen. Auffällig ist der hohe Anteil im Wissenslevel Kenner:in in der Gruppe der PostDocs ohne Lehre im Vergleich zu den PostDocs mit Lehre. In der Gruppe der Professor:innen ist der deutlich größere Anteil im Expert:in-Level im Verhältnis zu dem Kenner:in-Level bemerkenswert. Hier zeigt sich möglicherweise, dass sich einige bereits diesem Feld zugewandt haben und andere noch am Anfang stehen.

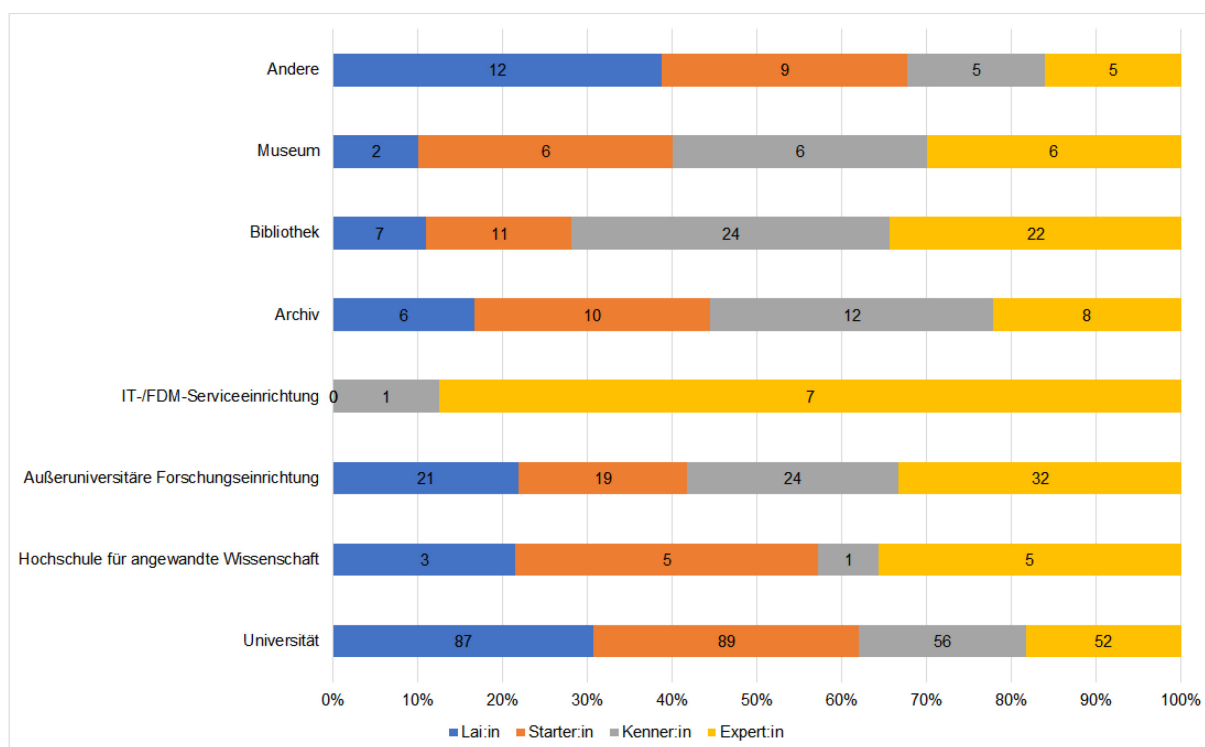


Abb. 86: Verteilung der Wissenslevel innerhalb der Einrichtungen | n = 567

Die Auswertung der Wissenslevel der Proband:innen nach zugehöriger Institution zeigt, dass der Anteil an Expert:innen in IT-/FDM-Serviceeinrichtungen am größten ist. Anzumerken ist jedoch, dass die Gruppe aus lediglich acht Personen besteht. Der geringste Anteil an Expert:innen ist an Universitäten zu finden.¹⁷ Von den 290 Personen, die angeben, an einer Universität tätig zu sein, fallen 90 unter das Level Lai:in und 91 unter Starter:in. Dies könnte daran liegen, dass die Gruppe der Bachelorstudierenden hauptsächlich an einer Universität verortet ist. Beim Herausrechnen der Studierenden in dieser Auswertung sinkt der Lai:in-Anteil von 31 % auf 24 %, der Starter:in-Anteil sinkt um 2 % auf 22 %, Kenner:in steigt um 3 % auf 23 % an und der Anteil an Expert:innen wächst von 18 % auf 25 %. Es liegt also nur bedingt an den Studierenden.

Wenn auch teilweise sehr knapp (mit Ausnahme der Hochschulen und Andere) zeigt die Verteilung bei allen anderen Institutionen, dass die Proband:innen vorwiegend den Wissensleveln Kenner:in und Expert:in zugeteilt werden können.

¹⁷ Mit Ausnahme von „Andere“.

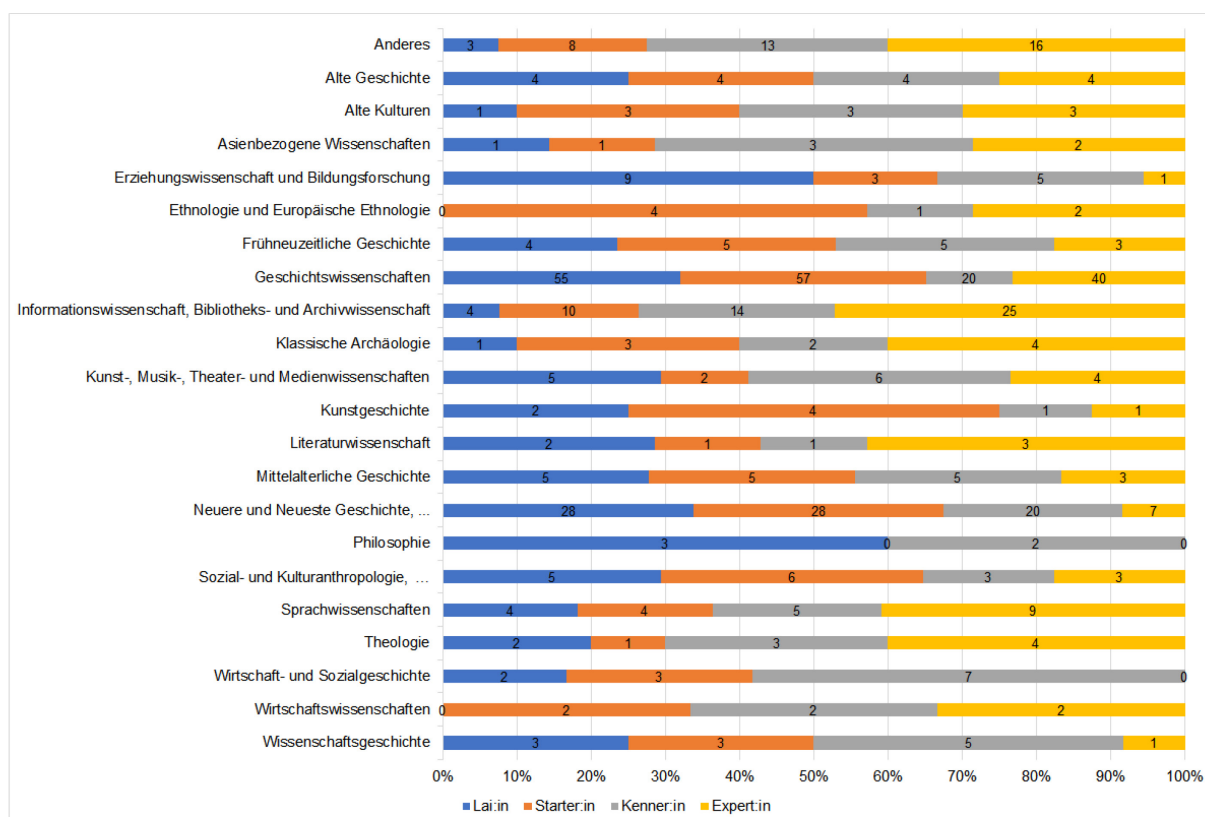


Abb. 87: Verteilung der Wissenslevel pro Fachkollegium | n = 567

Das dritte Diagramm (Abb. 87) zeigt die Verteilung der Wissenslevel in den Fachkollegien. Zur übersichtlicheren Darstellung wurden die einzelnen Disziplinen Fachkollegien zugeteilt, wenn sich weniger als fünf Teilnehmende einer Disziplin zugeordnet haben.¹⁸ Die Grundgesamtheit der jeweiligen Fachkollegien variiert zum Teil sehr stark, dies gilt es bei der nachfolgenden Interpretation zu berücksichtigen. Innerhalb der Fachkollegien Sprach-, Literaturwissenschaft sowie Informations-, Bibliotheks- und Archivwissenschaft befinden sich mehr als 40 % auf Expert:in-Level. In der Theologie, den asienbezogenen Wissenschaften, den Wirtschafts-, Sprach- sowie den Informations-, Bibliotheks- und Archivwissenschaften liegen über 60 % im Bereich Kenner:in und Expert:in. Das Fachkollegium Geschichtswissenschaften ist mit 172 zugehörigen Personen die größte Gruppe. 112 Personen wurden den Wissensleveln Lai:in und Starter:in zugeteilt, was 65 % dieser Gruppe ausmacht. Werden die Studierenden aus den gleichen Gründen wie bei den Universitäten herausgenommen, um einen klareren Blick auf die Forschenden zu bekommen, so sinkt der Anteil der Lai:innen und Starter:innen auf 45 % in den Geschichtswissenschaften und v. a. der Anteil an Expert:innen steigt von 23 % auf 40 %. Es zeigt sich, abgesehen von den Studierenden, durchaus eine Lücke zwischen Personen in den Geschichtswissenschaften, die gerade die ersten Schritte machen und Ex-

¹⁸ Vgl. zu den Fachdisziplinen die Ausführungen im Abschnitt [Soziodemographische Fragen, S. 13](#).

pert:innen.¹⁹ Fachkollegien mit einem Anteil von über 60 % an Lai:innen und Starter:innen sind Sozial- und Kulturanthropologie, Außereuropäische Kulturen, Judaistik und Religionswissenschaft; Neuere und Neueste Geschichte (einschließlich Europäischer Geschichte der Neuzeit und Außereuropäische Geschichte); Kunstgeschichte; Geschichtswissenschaften, Erziehungswissenschaft und Bildungsforschung.

Wie durch die unterschiedlichen Auswertungen dargestellt, gibt es, je nachdem welche Gruppe betrachtet wird, Ausreißer nach oben und unten. Insgesamt ist die Verteilung zwischen den Wissensleveln jedoch sehr gleichmäßig und liegt bei durchschnittlich etwa 25 %.

Lai:in	Starter:in	Kenner:in	Expert:in
25%	28%	23%	24%

Tab. 3: Gesamtverteilung der Proband:innen in den Wissensleveln | n = 567

Interpretation(sversuch) der Ergebnisse

Bei der Darstellung der Ergebnisse der Einzelauswertungen der Kategorien²⁰ haben wir uns bewusst für einen ausschließlich beschreibenden Ansatz entschieden. Zum einen können so Ergebnisse ohne vorheriges Framing beobachtet werden, zum anderen ist es nicht immer möglich, auf Grundlage der uns vorliegenden Datenlage fundierte Erklärungen abzugeben. Oftmals spiegeln die Ergebnisse zwar die langjährige Erfahrung im jeweiligen Feld wider, jedoch bräuchte es für viele Vermutungen weitere Erhebungen, im besten Fall qualitative Interviews, um die Hypothesen zu verifizieren. Zudem sind manche Interpretationsansätze für viele Fragen sehr ähnlich und würden sich fortlaufend wiederholen.

Um Redundanzen zu vermeiden, werden im Folgenden insbesondere Auffälligkeiten in den Daten der Einzelauswertung dargelegt. Die vorgestellten Erklärungsansätze für diese Auffälligkeiten wurden u. a. in einem Online-Workshop mit den Participants der Task Area „Data Literacy“ diskutiert und so in einem gewissen Maße validiert. Die im Austausch mit unseren Participants gewonnenen Interpretationsansätze werden in Teilen des nächsten Abschnitts präsentiert. Es konnten insgesamt sechs Aspekte herausgearbeitet werden, die die Auffälligkeiten in den Daten der Einzelauswertung zu erklären helfen: (1) (Fehlender) Praxisbezug und individuelle Arbeitspraxis, (2) Kompetenzstufen der verschiedenen Zielgruppen, (3) Beginn des Kulturwandels, (4) Diskrepanz zwischen Theorie und Praxis, (5) unterschiedliche Frage-

¹⁹ An dieser Stelle muss jedoch auch berücksichtigt werden, dass vermutlich der größere Anteil an Proband:innen grundsätzlich dieser Thematik offen(er) gegenüberstand und damit bereit war an der Umfrage teilzunehmen, weshalb möglicherweise Personen mit bereits hoher Expertise häufiger teilgenommen haben als Personen, die sich dieser Thematik bislang nicht zuwenden.

²⁰ Vgl. [Darstellung der Einzelergebnisse, S.11.](#)

stellungen in den Niveaustufen sowie (6) Begriffsverständnis und Interpretation der Fragen. Diese sechs Erklärungsmuster bedingen sich teilweise gegenseitig:

Das Fehlen eines direkten *Praxisbezugs und die individuelle Arbeitspraxis* können einen erheblichen Einfluss auf die Selbstbewertung von Kompetenzen haben. In Arbeitsfeldern, in denen die täglichen Aufgaben weniger auf eigener Forschung basieren und eher generische oder wissenschaftsunterstützende Tätigkeiten umfassen, neigen Personen dazu, ihre Fähigkeiten in fachspezifischen Bereichen aufgrund fehlender Alltagspraxis geringer einzuschätzen. In den Daten lässt sich dies innerhalb der Kategorien „Datenevaluation“ und „Datenkonvertierung“ erkennen. Z. B. bei der Frage nach digitalen, fachspezifischen Forschungsmethoden ist die Diskrepanz der Selbsteinschätzung von GLAM-Mitarbeiter:innen mit und ohne Lehre mit $\Delta = 0,6$ auffällig. Dies erklärt sich möglicherweise daraus, dass GLAM-Mitarbeitende ohne Lehre zu einem großen Teil in Bibliotheken und Archiven tätig sind und somit der Schwerpunkt ihrer Arbeit nicht auf der Durchführung von eigenen Forschungsprojekten liegt. Auch wenn die Personen einen fachbezogenen Hintergrund haben und fachspezifische Beratung anbieten, erfordern ihre Aufgaben weniger fachbezogene, methodische Tätigkeiten.

Neben der Praxis und dem individuellen Tätigkeitsprofil spielt das *Kompetenzlevel* der einzelnen Statusgruppen eine relevante Rolle bei der Interpretation der Daten. Die Selbsteinschätzung der Proband:innen unterliegt in besonderem Maße den Dynamiken der Gruppenzugehörigkeit, bspw. in der Statusgruppe der Bachelorstudierenden. Bei ihnen zeigt sich eine tendenzielle Zurückhaltung in der Einschätzung, die auf Unwissenheit, aber auch auf mögliche Unsicherheiten oder mangelnde praktische Erfahrungen zurückgeführt werden kann. So ist es nicht verwunderlich, dass sich diese Gruppe in so gut wie allen Fragen am niedrigsten einschätzt. Neben der geringen individuellen Erfahrung spielt die nicht ausreichende Integration der Vermittlung von digitalen Kompetenzen in der Schule²¹ sowie die (meist) noch nicht implementierten Data Literacy- und FDM-Inhalte im Studium eine bedeutende Rolle.

²¹ Im internationalen Vergleich liegen auch die computer- und informationsbezogenen Kompetenzen der Schüler:innen nur im Mittelfeld. Wenige Schüler:innen erreichen hohe digitale Kompetenzstufen, während knapp ein Drittel nur die niedrigsten erreicht. Insgesamt ist die Nutzung digitaler Medien im deutschen Unterricht gering und die Medienkompetenzen der Schüler:innen sind durchschnittlich (Vgl. Lachner, Scheiter und Stürmer 2020).

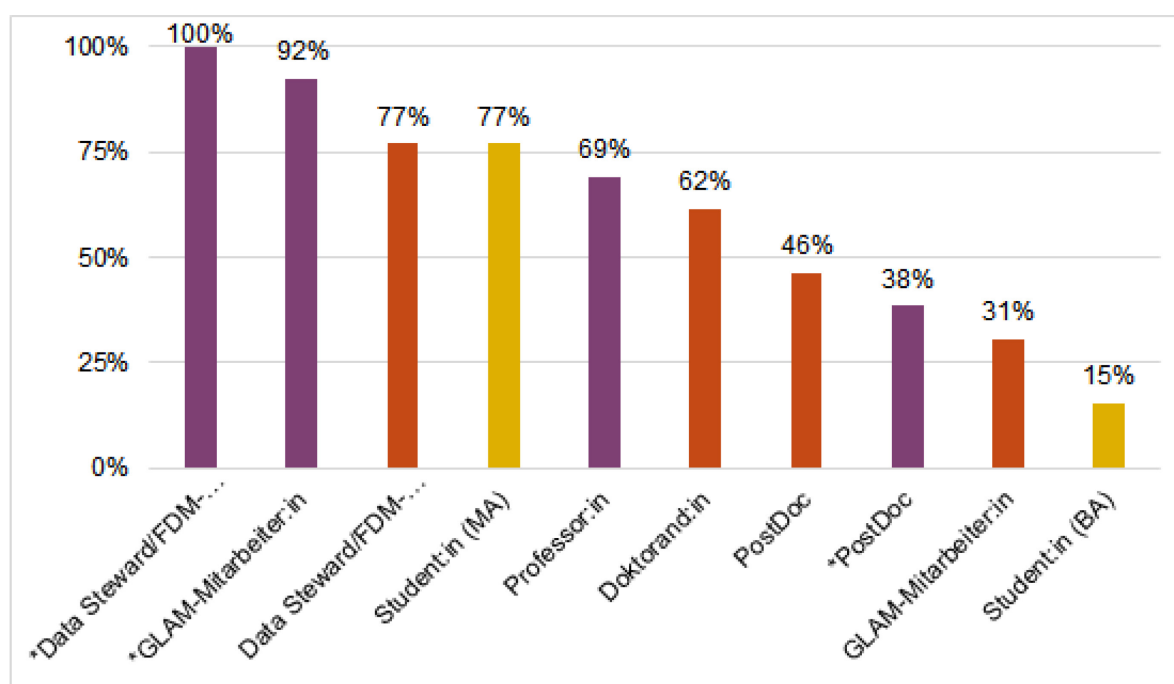


Abb. 88: Prozentualer Anteil, wie oft eine Gruppe über dem Kategorienmittelwert liegt | n = 568

Der Konzeption der drei Kompetenzniveaus (Wissen, Anwenden, Vermitteln) lag die Annahme zugrunde, dass Lehrende die höchsten Kompetenzen besitzen. Nun zeigt sich jedoch in Abbildung 88 dass diese These für die Gruppen der PostDocs und der GLAM-Mitarbeitenden ohne Lehre überhaupt nicht zutrifft, wohingegen sie für die lehrenden FDM- und GLAM-Mitarbeitenden, die nicht-lehrenden Data Stewards und durchaus auch noch auf die Professor:innen sowie die Doktorand:innen zutrifft.²² Insbesondere bleibt erklärungsbedürftig, wieso die lehrenden PostDocs sogar schlechter abschneiden als die nicht-lehrenden PostDocs. Eine Erklärung könnte sein, dass die lehrenden PostDocs sich zunächst noch Kompetenzen auf dem Niveau „Anwenden“ aneignen müssen, damit sie dieses Wissen anschließend an die Studierenden weitergeben können. Bei den GLAM-Mitarbeitenden ohne Lehre könnte dies, wie oben bereits erwähnt, an zu wenigen Berührungspunkten im Umgang mit Forschungsdaten in ihrem Arbeitsalltag liegen. Dass Promovierende bereits ein höheres Level aufweisen, könnte daran liegen, dass sie am Beginn ihrer Forschungstätigkeit stehen und durch die Schulungsangebote in den Graduiertenzentren bereits mehr Wissen in diesen Bereichen sammeln konnten. Zudem wird der jüngeren Generation auch eine höhere Affinität zu digitalen Medien und Werkzeugen nachgesagt, weshalb sie eventuell offener sind, sich mit diesen Themen auseinanderzusetzen. Dieser Aspekt leitet zu einem weiteren Erklärungsansatz über.

Die Geschichtswissenschaften stehen erst am *Beginn eines grundlegenden Kulturwandels*, wenn es um den qualitätsgesicherten Umgang mit Forschungsdaten geht. Obwohl die Digitali-

²² Die Anzahl der Masterstudierenden ist mit 18 Proband:innen zu klein, weshalb sie hier nicht in der Interpretation berücksichtigt werden.

sierung das Potenzial hat, neue Methoden und Ansätze in die Geschichtsforschung einzubringen, wird deutlich, dass der Aufbau einer soliden geschichtswissenschaftlichen Datenkultur noch erheblicher Anstrengungen bedarf. Bisher waren Historiker:innen oft auf traditionelle Methoden der Datensammlung und -analyse angewiesen, ohne eine breite Diskussion über die Sicherung der Datenqualität führen zu müssen. Doch mit dem Aufkommen digitaler Ressourcen und Werkzeuge wird die Nachfrage nach strukturierten Daten, Metadaten und nachvollziehbaren, gar reproduzierbaren Forschungsergebnissen größer. Um in diese Aushandlungsprozesse hineingehen sowie solche Daten produzieren zu können, sind Datenkompetenzen in den hier unterschiedlich abgefragten Facetten notwendig, jedoch sind sie weder ein etablierter Bestandteil der Wissenschaftspraxis noch der Ausbildung. Dies spiegelt sich in der Umfrage deutlich wider, bspw. in den Aspekten „Werkzeuge und Methoden zur Datendokumentation“²³, „Qualitätsmängel in Daten“²⁴ oder „Bewertung von Datenqualität“²⁵. Hier sind durchweg geringere Werte zu konstatieren. Da die Praxis der Datendokumentation und allein der Begriff „(Forschungs-)Daten“²⁶ in den historisch arbeitenden Disziplinen (noch) nicht in der Forschungspraxis etabliert ist, verwundern die Ergebnisse insgesamt nicht. Die weitestgehend geringe Selbsteinschätzung unterhalb von „teils-teils“ bei nahezu allen Statusgruppen zeigt, dass Datendokumentation noch keine gelebte Forschungspraxis ist. Das Thema Datenqualität ist in der Geschichtswissenschaft ebenfalls noch nicht umfassend diskutiert und ausgehandelt, was dazu führt, dass viele Historiker:innen zurückhaltend in ihrer Selbsteinschätzung sind. Die geringeren Bewertungen bei der Frage nach „Qualitätsmängeln in Daten erkennen“ könnten auch darauf zurückzuführen sein, dass viele Proband:innen zwar in der Lage sind, Fehler in den (eigenen) Daten zu erkennen, aber dennoch unsicher sind, weil sie nicht die nötigen Kriterien kennen, um diese Fehler v. a. in fremden Daten korrekt zu identifizieren und zu bewerten. Obwohl sicherlich offensichtliche Mängel häufig bemerkt werden, ist die Beurteilung der Datenqualität ohne tiefgehendes Verständnis der zugrundeliegenden Daten (v. a. digitaler Daten) und ohne theoretische Bewertungsmethoden sehr schwierig oder gar unmöglich. Die Unsicherheit/Vorsicht der Proband:innen ist also nur allzu verständlich. Das deutet einerseits darauf hin, dass das Wissen über Datenqualität in der Geschichtswissenschaft noch gering ist und andererseits Konzepte sowie Standards zur Beurteilung erst noch entwickelt werden müssen.²⁷ Ebenso zeigt sich an der am niedrigsten bewerteten Kategorie „Programmierung“, dass die Diskussion um die Notwendigkeit von Programmierkenntnissen in den historischen Disziplinen noch intensiviert werden muss. Denn obwohl die Werte sehr gering ausfallen, scheint

²³ Vgl. [Werkzeuge und Methoden zur Datendokumentation, S. 41.](#)

²⁴ Vgl. [Qualitätsmängel in Daten, S. 50.](#)

²⁵ Vgl. [Bewertung von Datenqualität, S. 49.](#)

²⁶ Vgl. [Begriff „Forschungsdaten“, S. 19.](#)

²⁷ Dies ist ein zentrales Arbeitsfeld der Task Area „Data Quality“ von NFDI4Memory. <https://4memory.de/task-areas/task-area-1-data-quality/> (zuletzt 21.06.2024).

die Motivation, diese Kompetenzen zu erwerben, hoch. Dies zeigt sich an der hohen Beteiligung der optional zu beantwortenden Frage nach mehr Informationen zu Programmiersprachen.²⁸ Dennoch sind die Ergebnisse wenig überraschend, da Programmierung und algorithmisches Denken weder Bestandteile der Curricula in den historischen Disziplinen sind, noch in der Breite der Community in der Forschungspraxis angewandt werden. Hinzu kommt, dass viele Personen in den geschichtswissenschaftlichen Fächern ein schlechtes Selbstbild hinsichtlich ihrer mathematisch-naturwissenschaftlichen Fähigkeiten haben, denen sie informatische Kenntnisse wie das einfache algorithmische Denken zuordnen. In dem Workshop mit den Participants der Task Area „Data Literacy“ wurden zu diesen Beobachtungen folgende Erklärungsansätze vorgeschlagen: Zum einen wurde argumentiert, dass algorithmisches Denken und Programmieren nicht das Gleiche sind und daher auch nicht unbedingt aufeinander aufbauen müssen, wie es die Fragen suggerieren. Zum anderen kann das Verständnis von algorithmischem Denken unterschiedlich sein. Denn Grundkenntnisse im Programmieren können vermittelt werden, ohne vorab algorithmisches Denken in Gänze zu beherrschen. Allgemein liegt die geringe Selbsteinschätzung sicherlich an der fachlichen Ausrichtung der Proband:innen: So wird logisches und analytisches Denken nicht als dezidiertes Bestandteil der klassischen Geschichtswissenschaften wahrgenommen, weil es eher den mathematisch-naturwissenschaftlichen Fächern zugeschrieben wird. Das ist eigentlich ein Trugschluss, denn ein Kern geschichtswissenschaftlicher Forschung besteht darin, anhand von Quellen aus der Vergangenheit, historische Geschehnisse, ihren Ablauf sowie ihre Ursachen und Wirkungen zu rekonstruieren. Hierzu bedarf es sehr wohl analytischer und auf Logik basierender Denk- und Argumentationsweisen, die im Unterschied zu den mathematisch-naturwissenschaftlichen Fächern allerdings seltener einer numerischen Logik folgen, sondern vielmehr auf hermeneutischen Prinzipien basieren.

Ein weiterer Erklärungsansatz betrifft die *Diskrepanz zwischen Theorie und Praxis*, v. a. die Auffälligkeiten in den Gruppen, in denen optional auch gelehrt wird (FDM- und GLAM-Mitarbeiter:innen, PostDocs), bzw. die allgemeine Unterscheidung der Niveaustufen 2 und 3. Die Unterschiede in den Ergebnissen zwischen Personen, die Kompetenzen theoretisch vermitteln können und jenen, die sie in der Praxis ausüben, sind teils signifikant. Während Theoretiker:innen möglicherweise ein fundiertes Verständnis für bestimmte Konzepte haben, fehlt ihnen oft die praktische Erfahrung, die für eine realistische Selbsteinschätzung ihrer Fähigkeiten entscheidend ist. Bei der Auswertung ist dies vor allem bei der Frage nach „Recherchewerkzeugen und -strategien“ sowie „Strategien zum Ordnen von Daten“ deutlich geworden. Die Verteilung bei der Frage nach „Recherchewerkzeugen und -strategien“ ist erwartbar.²⁹ Auf dem ersten Platz befinden sich FDM-Mitarbeitende mit Lehre und auf dem letzten Studierende im Bachelor. Markant ist jedoch die Gruppe der GLAM-Mitarbeitenden mit und ohne Lehre.

²⁸ Vgl. [Programmierung, S. 55](#).

²⁹ Vgl. [Recherchewerkzeuge und -strategien, S. 28](#).

Während sich GLAM-Mitarbeitende mit Lehre mit 4,0 als kompetent in diesem Bereich einschätzen, ist die Beurteilung bei den Mitarbeitenden ohne Lehre mit 3,7 leicht verhaltener. Theoretische Konzepte spielen in der Lehre eine größere Rolle, da es dabei nicht (immer) um das aktive Anwenden von Werkzeugen und Strategien geht, sondern lediglich um ihre Vermittlung. Des Weiteren wäre es denkbar, dass Recherchewerkzeuge und -strategien unbewusst eingesetzt werden und das Vorgehen womöglich nicht bewusst als Strategie begriffen wird. Anhand dieser Auswertung ist auch ein deutlicher Unterschied zwischen Lehrenden und Nicht-Lehrenden zu erkennen. Lehrende Personen schätzen sich als Gruppe geschlossen zunehmend kompetent ein (zwischen 3,9 und 4,2). Die Gruppe der Nicht-Lehrenden schätzt sich geschlossen mit 3,7 etwas verhaltener ein.³⁰ Anzumerken ist jedoch, dass zwischen den meisten Statusgruppen keine signifikanten Unterschiede bestehen ($\Delta = 0,2$).

Bei der Frage nach Strategien zum Ordnen von Daten ist die Verteilung der Niveaustufe 2 und 3 auffällig.³¹ Personen, die nicht lehren, arbeiten eventuell projektgebundener und aktiver mit und an Daten und sind in deren Management bereits geübter. Aus diesem Grund ist das Ordnen und Strukturieren näher am eigenen Arbeitsalltag und dadurch schätzen sich diese Gruppen möglicherweise kompetenter ein.

Ein weiterer nicht zu unterschätzender Aspekt sind die unterschiedlichen Fragestellungen in den Niveaustufen, weshalb der vergleichende Blick im Auswertungs- und Interpretationsprozess an manchen Stellen schwierig ist. Wie bereits im Kapitel zum methodischen Vorgehen beschrieben, wurde aufgrund des breiten Adressat:innenkreises und den damit einhergehenden anzunehmenden Kompetenzunterschieden in den Bereichen FDM und DL eine Unterteilung in 3 Niveaustufen vorgenommen: Wissen, Anwenden und Vermitteln, die als aufeinander aufbauend verstanden wurden. So zeigt sich bspw. bei der Frage nach dem Begriff „Anonymisierung“³², dass in erster Linie theoretische Kenntnisse bei der Fragestellung von Niveaustufe 1 abgefragt wurden, jedoch für Niveaustufe 2 und 3 eine aktive Komponente erfragt wurde. Es ist davon auszugehen, dass vermutlich einem Großteil der Zielgruppe der Begriff der „Anonymisierung“ schon einmal begegnet ist und sie wissen, was er bedeutet. Daten jedoch selbst zu anonymisieren, ist nur in zeithistorischen Forschungskontexten nötig,³³ womit für einen Großteil der Proband:innen keine Notwendigkeit besteht, diese Fähigkeiten zu besitzen. Die fehlende forschungsbasierte Motivation in den älteren geschichtswissenschaftlichen Fächern erklärt vermutlich auch, dass bis heute in den meisten historischen Studiengänge auf die Vermittlung von Anonymisierungsverfahren verzichtet wird. Das gleiche Phäno-

³⁰ Vgl. [Abb. 16 Recherchewerkzeuge, S. 28](#).

³¹ Vgl. [Datenorganisation, S. 31](#).

³² Vgl. [Begriff „Anonymisierung“, S. 59](#).

³³ Für die Proband:innen, die sich der Neueren und Neuesten Geschichte zugeordnet haben, ist der Mittelwert in der Kategorie „Recht“ 2,6.

men findet sich bspw. auch bei den Fragen zu „Forschungssoftware“³⁴, „Visualisierungsarten“³⁵ und „Metadatenstandards“³⁶.

Des Weiteren muss als möglicher Erklärungsansatz das *individuelle Begriffsverständnis und die Interpretation der Fragestellung* in Betracht gezogen werden. Die Umfrageteilnehmenden hatten vermutlich sehr unterschiedliche Vorstellungen und Verständnisse von den im Rahmen unserer Umfrage adressierten digitalen Kompetenzen und den damit verbundenen Konzepten und Instrumenten. Divergenzen im individuellen Begriffsverständnis reflektieren den aktuellen Stand der Fachdiskussion im Bereich von Data Literacy. Sie spiegeln die Vielgestaltigkeit und Fluidität des Verständnisses von Datenkompetenz und Forschungsdatenmanagement wider. Unterschiede in der Interpretation der Fragestellung resultieren schließlich auch aus der noch offenen Diskussion darüber, welche digitalen Kompetenzen in den historischen Disziplinen notwendig sind. Die in unseren Daten zum Ausdruck kommende Variabilität in den Selbstbewertungen scheint zu einem großen Teil darauf zurückgeführt werden zu können, dass die Befragten ihre Kompetenzen basierend auf ihrem eigenen Verständnis der Begriffe bewerteten. Als Beispiel kann hier die Frage nach den „Grundprinzipien der Datensicherung“³⁷ dienen. Eine mögliche Erklärung für die Zurückhaltung in der Niveaustufe 3 könnte die Überinterpretation des Begriffes „Datensicherung“ sein. Es ist davon auszugehen, dass die meisten Proband:innen irgendeine Form der Datensicherung, wie z. B. das manuelle Erstellen von Kopien auf externen Datenträgern praktizieren, ihnen die Sinnhaftigkeit von Datensicherung bewusst ist und sie einfache Backup-Verfahren kennen. Dazu gehört bei der Nutzung von Cloudspeichern z. B. auch die Verwendung von sicheren Passwörtern. All diese Aspekte fallen unter die Kategorie „Datensicherung“. Angesichts der teilweise niedrigen Bewertungen ist jedoch zu bezweifeln, dass die Proband:innen diese alltägliche Praxis der Datensicherung als Bestandteil der unter der Rubrik „Grundprinzipien der Datensicherung“ abgefragten Kompetenzen wahrnehmen und einordnen. Auch bei dem Begriff „Metadaten“³⁸ sind Auffälligkeiten in den Daten erkennbar. Vor allem bei den Doktorand:innen überrascht die geringe Selbsteinschätzung in Anbetracht der Tatsache, dass Doktorand:innen an unterschiedlichen Stellen in ihrer Tätigkeit mit der Erstellung von Metadaten in Berührung gekommen sein müssen, allein schon durch die Erstellung von Literatur- und Quellenverzeichnissen. Es ist davon auszugehen, dass die geringe Einschätzung weniger von der fehlenden Praxis als von der Unkenntnis sowie der Miss- oder Fehlinterpretation des Begriffes „Metadaten“ rührt. Als letztes Beispiel seien noch die Ergebnisse der Kategorie „Datenkonvertierung“ genannt. Da es sich bei den Fragen um grundlegende Kenntnisse der Datenkonvertierung handelt, wie den Unterschied zwischen Datentyp / -format und Dateityp / -format zu kennen sowie Daten von einem Format in ein anderes zu konvertieren, ist die zögerliche Einschätzung mit den vorliegenden Erkenntnissen nicht vollkommen zu erklären. Möglicherweise sind hier die Fehl- oder Überinterpretation der Begrifflichkeiten bzw. das mangelnde Bewusstsein über deren alltägliche Präsenz im eigenen

³⁴ Vgl. [Forschungssoftware](#), S. 32.

³⁵ Vgl. [Visualisierungsarten](#), S. 35.

³⁶ Vgl. [Begriff „Metadatenstandards“](#), S. 42.

³⁷ Vgl. [Grundprinzipien der Datensicherung](#), S. 72.

³⁸ Vgl. [Begriff „Metadaten“](#), S. 22.

Arbeitskontext eine Ursache. Zudem verengen viele historisch arbeitende Personen den Forschungsdatenbegriff auf digitale Daten und hier vor allem auf digitale quantitative Daten. Diese Verengung des Datenbegriffs führt auch zu der unter historisch arbeitenden Wissenschaftler:innen weit verbreiteten Einschätzung, dass sie in ihrer eigenen Forschung gar nicht mit Forschungsdaten arbeiten würden.³⁹

³⁹ Vgl. auch im [Abschnitt Weitere relevante Themen und Feedback, S. 91](#). Hier sei auf den in Publikationsvorbereitung befindlichen Artikel zu Forschungsdaten im Clio-Guide verwiesen. <https://guides.clio-online.de/guides/arbeitsformen-und-techniken> (14.05.2024).

Ausblick

Bei diesem Bericht handelt es sich um eine erste beschreibende Sicht auf die erhobenen Daten. Er dient als erste Publikation von weiteren, weil die Daten in vielerlei Hinsicht für verschiedene Arbeitspakete der Task Area „Data Literacy“ relevant sind. Hiermit wird ein erster Überblick gegeben, wie es in den historischen Disziplinen um die Datenkompetenzen aus der Perspektive der verschiedenen Akteursgruppen bestellt ist und um einen Ansatzpunkt zu haben, in welche Richtung die nächsten Schritte bei der Entwicklung der virtuellen Trainingskurse (Vgl. Paulmann u. a. 2022, 94; TA4-M4-T2) gehen sollen. Aus den Ergebnissen geht hervor, dass in den Grundlagen des (Forschungs-)Datenmanagement in allen Zielgruppen Weiterbildungsbedarfe bestehen, somit werden die ersten 4Memory-Lernangebote diese Aspekte adressieren. Hierfür werden generische FDM-Schulungskonzepte⁴⁰ berücksichtigt und an die Bedarfe der historisch arbeitenden Disziplinen angepasst: Einerseits sollen die Inhalte möglichst anwendungsnah sein und andererseits sollen die Lernformate den Wünschen entsprechend und den Gegenständen angemessen gestaltet werden. D. h., es werden zunächst primär die Personen aus den Wissensleveln Lai:in und Starter:in adressiert sowie die Kompetenzniveaus Wissen und Anwenden. Durch eine modulare Bereitstellung einzelner Aspekte der DataLiteracy zu den verschiedenen Kategorien werden auch Personen aus den Wissensleveln Kenner:in und Expert:in zu einzelnen Themen Angebote finden, mit denen sie sich erst noch vertraut machen müssen. Schließlich gab es wenige Proband:innen, die in allen DL-Kategorien ausgesprochen hohe Selbsteinschätzungswerte aufzeigten und keinerlei Schulungsbedarf haben.

Ein weiteres Arbeitspaket ist die Entwicklung des Data-Literacy-Kompetenzprofils. Aufbauend auf den Ergebnissen der Umfrage und dem entwickelten Fragebogendesign wird das Profil nun weiterentwickelt. Dabei werden die Umfragedaten Hinweise geben, wie die Kompetenzprofile in den verschiedenen Zielgruppen justiert werden sollten. Das Kompetenzprofil ist dann auch die Grundlage für die Entwicklung des Modellcurriculums (Vgl. Paulmann u. a. 2022, 91, TA4-M1-T3) sowie für die epochenspezifischen Beispiellehrpläne (Vgl. Paulmann u. a. 2022, 91, TA4-M2-T3). Das Kompetenzprofil wird darüber hinaus auch ein Vokabular bieten, um die Lern- und Lehrressourcen mit Metadaten für den Trainingskatalog (Vgl. Paulmann u. a. 2022, 90, TA4-M1-T1) anzureichern, um den Suchenden das Auffinden der für sie passenden Ressourcen zu erleichtern.

Ein weiterer Anwendungsfall für die Umfragedaten ist die Erstellung von Personae, um eine bessere Vorstellung von den Bedarfen unterschiedlicher Zielgruppen zu erhalten (Vgl.

⁴⁰ Z. B. das Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement (Biernacka u. a. 2023), die HeFDI Data Learning Materials (Krähwinkel u. a. 2022) sowie das Konzept des Zertifikatskurs Forschungsdatenmanagement in NRW (Slowig u. a. 2022)

Schweibenz 2004). Diese Profile helfen dabei, die von 4Memory entwickelten Dienste an verschiedene Nutzungsbedürfnisse anzupassen. Diese Arbeiten gehen auch in die NFDI-Sektion „Training & Education“ und das DALIA-Projekt⁴¹ ein.

Des Weiteren hat sich im Verlauf der Auswertung und Interpretation der Ergebnisse gezeigt, dass das entwickelte Fragebogendesign an manchen Stellen angepasst werden muss, weshalb eine überarbeitete Version des bereits veröffentlichten Designs (Vgl. Döring und Lemaire 2023) ebenfalls bereitgestellt werden wird.

Neben den hier genannten und von uns bislang geplanten Verwertungen der Erhebungsdaten, sind sicherlich aufgrund der Vielschichtigkeit der Umfragedaten weitere Nachnutzungsszenarien denkbar. So ist z. B anzunehmen, dass mit weitergehenden statistischen Methoden möglicherweise Korrelationen in den Daten gefunden werden, die tiefere Erkenntnisse oder Hinweise für eine bessere Passfähigkeit von 4Memory-Angeboten liefern. Da nicht nur die Task Area „Data Literacy“ diese Umfragedaten nachnutzen können soll und bereits mehrfach Interesse an den Daten geäußert wurde, sind die Umfragedaten veröffentlicht worden (Vgl. Döring, Kellendonk und Lemaire 2024).

⁴¹ <https://dalia.education/> (17.05.2024)

Literaturverzeichnis

- Bandtel, Matthias, Leonie Kauz und Natalia Weißker. 2021. Data Literacy Education für Studierende aller Fächer. Kompetenzziele, curriculare Integration und didaktische Ausgestaltung interdisziplinärer Lehr-Lern-Angebote. In: *Digitalisierung in Studium und Lehre gemeinsam gestalten: Innovative Formate, Strategien und Netzwerke*, 395–412. Wiesbaden: Springer Fachmedien. https://doi.org/10.1007/978-3-658-32849-8_23.
- Batterton, Katherine A. und Kimberly N. Hale. 2017. The Likert Scale What It Is and How To Use It. *Phalanx* 50, Nr. 2: 32–39. <https://www.jstor.org/stable/26296382> (zugegriffen: 23. Mai 2024)
- Biernacka, Katarzyna, Ron Dockhorn, Claudia Engelhardt, Kerstin Helbig, Juliane Jacob, Tereza Kalová, Adienne Karsten, u. a. 2023. *Train-the-Trainer-Konzept zum Thema Forschungsdatenmanagement*. Version 5. <https://doi.org/10.5281/zenodo.10122153>.
- BMBF, Bundesministerium für Bildung und Forschung. 2015. Qualitätsoffensive Lehrerbildung. *Bundesministerium für Bildung und Forschung - BMBF Qualitätsoffensive Lehrerbildung*. https://www.qualitaetsoffensive-lehrerbildung.de/SiteGlobals/Forms/lehrerbildung/projektarte/projektarte_formular.html?queryResultId=null&pageNo=0 (zugegriffen: 23. Mai 2024).
- . 2022. Kompetenzzentren für digitales und digital gestütztes Unterrichten. *BMBF-Homepage*. <https://www.bmbf.de/bmbf/de/bildung/digitalisierung-und-mint-bildung/digitaler-unterricht/kompetenzzentren-fuer-digitales-unterrichten.html> (zugegriffen: 28. Mai 2024).
- Borek, Luise, Canan Hastik, Vera Khramova und Jonathan Geiger. 2021. TaDiRAH. Taxonomy of Digital Research Activities in the Humanities. Version 2.0. <https://vocabs.dariah.eu/tadirah/en/> (zugegriffen: 21. Juni 2024).
- Döring, Laura, Stefan Kellendonk und Marina Lemaire. 2024. Daten der Data Literacy Bedarfserhebung für die historisch arbeitenden Disziplinen (Erhebungszeitraum: August-Oktober 2023). <https://doi.org/10.5281/zenodo.12200702>.
- Döring, Laura und Marina Lemaire. 2023. Fragebogendesign der Data Literacy Bedarfserhebung für die historisch arbeitenden Disziplinen. www.doi.org/10.5281/zenodo.10417016.
- Grillenberger, Andreas und Ralf Romeike. 2019. *Vorstudie: Hochschulübergreifende Konzepte zum Erwerb von 21st century skills am Beispiel von Data Literacy*. Hg. von Geschäftsstelle Hochschulforum Digitalisierung beim Stifterverband für die Deutsche Wissenschaft e.V. Hochschulforum Digitalisierung. Bd. 43. Hochschulforum Digitalisierung. Berlin / Essen: Edition Stifterverband – Verwaltungsgesellschaft für Wissenschaftspflege mbH. <http://doi.org/10.5281/zenodo.2633091>.
- Gronwald, Marco. 2019. Speichervolumen- und Kostenabschätzung für die Archivierung von Forschungsdaten. Auswertung einer Bedarfsumfrage an der Universität Osnabrück. *B.I.T. online. Zeitschrift für Bibliothek, Information und Technologie mit aktueller Internet-Präsenz*: www.b-i-t-online.de 22, Nr. 2: 135–143. <https://www.b-i-t-online.de/heft/2019-02-fachbeitrag-gronwald.pdf> (zugegriffen: 21. Juni 2024).
- Hartmann, Niklas K., Boris Jacob und Nadin Weiß. 2019. *RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement*. Version 1.0. Potsdam: Zenodo. <https://doi.org/10.5281/zenodo.3585556>.
- Heidrich, Jens, Pascal Bauer und Daniel Krupka. 2018. *Strukturen und Kollaborationsformen zur Vermittlung von Data-Literacy-Kompetenzen. Stand der Forschung*. Arbeitspapier des Hochschulforums Digitalisierung 32. Online. https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr32_Data_Literacy_Kompetenzen_Literatur.pdf (zugegriffen: 21. Juni 2024).
- Iglezakis, Dorothea und Björn Schembera. 2018. Anforderungen der Ingenieurwissenschaften an das Forschungsdatenmanagement der Universität Stuttgart - Ergebnisse der Bedarfsanalyse des Projektes DIPL-ING. *o-bib. Das offene Bibliotheksjournal* 5, Nr. 3 (2018): 46–60. <https://doi.org/10.5282/o-bib/2018H3S46-60>.

- KMK, Kultusministerkonferenz. 2016. *Bildung in der digitalen Welt Strategie der Kultusministerkonferenz*. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf (zugegriffen: 24. Mai 2024).
- Kompetenzverbund lernen:digital. 2023. Wissenschaft und Praxis im Dialog. *Kompetenzverbund lernen:digital*. <https://lernen.digital/> (zugegriffen: 28. Mai 2024).
- Krähwinkel, Esther, Patrick Langner, Robert Lipp und Andre Manuel Pietsch. 2022. HeFDI Data Learning Materials: Forschungsdatenmanagement – eine Online-Einführung. <https://doi.org/10.5281/zenodo.6373595>.
- Krathwohl, David R. 2002. A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice* 41, Nr. 4: 212–218. <https://www.jstor.org/stable/1477405> (zugegriffen: 23. Mai 2024).
- Lachner, Andreas, Katharina Scheiter und Kathleen Stürmer. 2020. Digitalisierung und Lernen mit digitalen Medien als Gegenstand der Lehrerinnen- und Lehrerbildung. In: *Handbuch Lehrerinnen- und Lehrerbildung*, hg. von Colin Cramer, Johannes König, Martin Rothland, und Sigrid Blömeke, 67–75. 1. Aufl. Stuttgart: UTB. <https://doi.org/10.36198/9783838554730> (zugegriffen: 23. Mai 2024).
- Langhanke, Gerald und Wolfgang Stille. 2015. Umgang mit Forschungsdaten an der TUD Erste Schritte zur Bedarfserhebung und Leitlinienentwicklung. <http://docplayer.org/11801705-Umgang-mit-forschungsdaten-an-der-tud-erste-schritte-zur-bedarfserhebung-und-leitlinienentwicklung.html> (zugegriffen: 28. Mai 2024).
- Lemaire, Marina, Lea Gerhards, Stefan Kellendonk, Katarina Blask und André Förster. 2020. *Das DIAMANT-Modell 2.0. Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft*. eSciences Working Papers 05. Trier. <https://doi.org/10.25353/ubtr-xxxx-f5d2-fffb>.
- Lemaire, Marina, Yvonne Rommelfanger, Jan Ludwig, Alexander Lürken-Uhl, Benjamin Merkler und Peter Sturm. 2016. *Umgang mit Forschungsdaten und deren Archivierung. Bericht zur Online-Bedarfserhebung an der Universität Trier*. eSciences Working Papers 02. Trier: Universität Trier. <http://nbn-resolving.de/urn:nbn:de:hbz:385-10156>.
- Nelson, Megan R. Sapp. 2017. A Pilot Competency Matrix for Data Management Skills: A Step toward the Development of Systematic Data Information Literacy Programs. *Journal of eScience Librarianship* 6, Nr. 1: 1–11. <https://doi.org/10.7191/jeslib.2017.1096>.
- Paulmann, Johannes, John Carter Wood, Klaus Ceynowa, Fabian Cremer, Silvia Daniel, Daniel Fähle, Barbara Fichtl, u. a. 2022. *NFDI4Memory. Consortium for the historically oriented humanities. Proposal for the National Research Data Infrastructure (NFDI)*. Mainz. <https://doi.org/10.5281/zenodo.7428489>.
- Petersen, Britta, Claudia Engelhardt, Tanja Hörner, Juliane Jacob, Tatiana Kvetnaya, Andreas Mühlichen, Hermann Schranzhofer, u. a. 2023. *Lernzielmatrix zum Themenbereich Forschungsdatenmanagement (FDM) für die Zielgruppen Studierende, PhDs und Data Stewards*. Version 2. <https://doi.org/10.5281/zenodo.8010617>.
- Schmidt, Birgit und Kathleen Shearer. 2016. *Librarians' competencies profile for research data management*. Hg. von Joint Task Force on Librarians' Competencies in Support of E-Research and Scholarly Communication. https://www.coar-repositories.org/files/Competencies-for-RDM_June-2016.pdf (zugegriffen: 22. Juli 2019).
- Schöch, Christof, Jody Perkins und Luise Borek. 2015. TaDiRAH: Release version 0.5.3. <https://doi.org/10.5281/zenodo.32492>.
- Schweibenz, Werner. 2004. Zielgruppenorientiertes Interaktionsdesign mit Personas. *Information - Wissenschaft und Praxis* 55, Nr. 3: 151–158. https://www.researchgate.net/profile/Werner-Schweibenz/publication/326609895_Zielgruppenorientiertes_Informationdesign_mit_Personas/links/5b58ca12458515c4b244c1f3/Zielgruppenorientiertes-Informationdesign-mit-Personas.pdf (zugegriffen: 28. Mai 2024).
- Slowig, Benjamin, Mirjam Blümm, Konrad U. Förstner, Marvin Lanczek, Birte Lindstädt, Rabea Müller, Ulrike Nickenig, Stephanie Rehwald und Lioba Schreyer. 2022. Zertifikatskurs Forschungsdatenmanagement in NRW. *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB*: 1–10. <https://doi.org/10.5282/O-BIB/5833>.
- Stürz, Roland A., Christian Stumpf, Antonia Schlude, Ulrike Mendel und Danilo Harles. 2023. *Das bidt-Digitalbarometer.international*. Hg. von Bayerisches Forschungsinstitut für Di-

- gitale Transformation bidt. bidt Analysen und Studien. München. <https://www.doi.org/10.35067/xypq-kn68>.
- Swiatek, Cécile, Ciara McCaffrey, Thorsten Meyer, Anna Svenbro, Helene Brinken, Frank Egerton, Anna Wojciechowska und Karin Clavel. 2020. *LIBER Open Science Training Methods and Practices Across European Research Libraries - Survey Analysis*. <https://doi.org/10.5281/zenodo.3903142>.
- Verheul, Ingeborg, Melanie Imming, Jacquelyn Ringerma, Annemie Mordant, Jan-Lucas van der Ploeg und Martine Pronk. 2019. Data Stewardship on the map: A study of tasks and roles in Dutch research institutes. <https://zenodo.org/doi/10.5281/zenodo.2669149>.