

Jahresbericht 2021

Data Center for the Humanities

Philosophische Fakultät der Universität zu Köln

Impressum

Herausgeber

Data Center for the Humanities (DCH)
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln
info-dch@uni-koeln.de

Redaktion

Jonathan Blumtritt
Anna Fischer
Patrick Helling
Elisabeth Mollenhauer
Claes Neufeind
Felix Rau
Philip Schildkamp

Layout

Julia Sorouri

Druck

Hausdruckerei
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln

Auflage 40 Stück
Stand: Mai 2022

© Data Center for the Humanities 2022

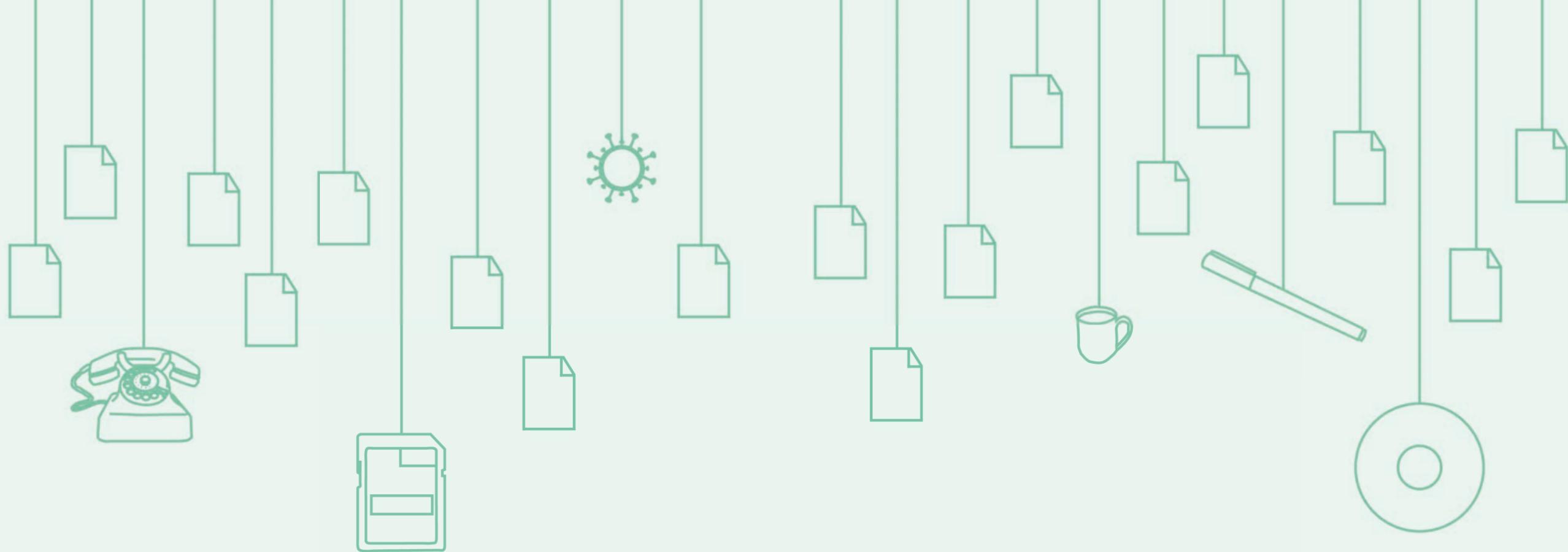
Jahresbericht 2021

Data Center for the Humanities

Philosophische Fakultät der Universität zu Köln

INHALTSVERZEICHNIS

1	Personal & Organisation	8	6	Übernahme von Ressourcen & Archivierung	56
1.1	Beirat	8	6.1	Überführung von Ressourcen in eine dauerhafte Betreuung	56
1.2	Personalentwicklung	8	6.2	Datenpublikation/Daten-Schnittstellen	59
1.3	Team 2021	12	6.3	Archivierung am DCH	59
1.4	Kooperation CCEH	19	6.3.1	Archivierungsprozess	59
2	Einbettung & Kooperationen an der Universität zu Köln	20	6.3.2	Archivierte Forschungsdatenbestände	61
3	Außeruniversitäre Sichtbarkeit, Vernetzung & Kooperationen	32	6.3.3	Datensicherung	62
3.1	Geisteswissenschaftliche Konsortien NFDI	32	7	Betrieb von Diensten & Infrastruktur	64
3.1.1	Text+	34	7.1	Infrastruktur	64
3.1.2	NFDI4Culture	35	7.2	Dienste	65
3.1.3	Geisteswissenschaftliche Konsortien in der Antragsphase	37	8	Drittmittelprojekte	68
3.2	Landesinitiative fdm.nrw	39	8.1	Projektabschluss: DFG-LIS-Projekt SustainLife	68
3.3	GO FAIR Implementation Network Data Stewardship Competence Centers (DSCC-IN), German Chapter GO UNITE!	40	8.2	BMBF-Verbundprojekt QUEST	72
3.4	Research Data Alliance e. V. (RDA DE)	41	9	Veranstaltungen & Publikationen	74
3.5	CLARIN – European Research Infrastructure for Language Resources and Technology	42	9.1	Eigene Veranstaltungen	74
3.6	Digital Endangered Languages and Musics Archives Network (DELAMAN)	42	9.1.1	DHCon 2021	74
3.7	Verband Digital Humanities im deutschsprachigen Raum e. V. (DHd)	43	9.1.2	Virtual Summer School „Deep Learning for Language Analysis“	77
3.7.1	DHd-AG Datenzentren	43	9.1.3	FORGE 2021 – Forschungsdaten in den Geisteswissenschaften	78
3.7.2	Data Steward	43	9.1.4	GO UNITE! Workshop der AG FDM-Beschreibungsmodell	88
3.8	Deutsche Initiative für Netzwerkinformationen e. V. (DINI)	45	9.2	Vorträge, Präsentationen & Teilnahmen	89
3.9	Sonstige Gremien, Mitgliedschaften & Kooperationen	45	9.3	Publikationen	93
4	Kompetenzbereich audiovisuelle Daten	46	10	Awareness, Schulungen & Lehre	94
4.1	Language Archive Cologne (LAC)	47	10.1	Sammelkarten „Daten & Datteln digital“	94
4.2	BMBF-Verbundprojekt QUEST	48	10.2	a.r.t.e.s. Workshop „Forschungsdatenmanagement in den Geisteswissenschaften“	100
4.3	CLARIN Knowledge-Centre for Linguistic Diversity and Language Documentation (CKLD)	48	10.3	Master-Übung „Forschungsdatenmanagement“	100
5	Beratung & Vermittlung	50	Ausblick	103	
5.1	Beratungen 2021	52	Abkürzungsverzeichnis	104	
			Abbildungsverzeichnis	110	



Zusammenfassung

Das Jahr 2021 war vor allem von der Planung der FORGE-Konferenz (Forschungsdaten in den Geisteswissenschaften) geprägt. Unter dem Titel „Mapping the Landscape – Geisteswissenschaftliches Forschungsdatenmanagement zwischen lokalen und globalen, generischen und spezifischen Lösungen“ wurde die Konferenz vom Data Center for the Humanities (DCH), dem Cologne Center for eHumanities (CCEH), dem Institut für Digital Humanities (IDH) und der AG Datenzentren des Verbands Digital Humanities im deutschsprachigen Raum e. V. (DHd) im September 2021 an der Universität zu Köln ausgetragen. Bedauerlicherweise musste die dreitägige Konferenz für Autor*innen und Teilnehmer*innen Covid-19-bedingt digital stattfinden. Allerdings konnte mit der Einrichtung eines „Situation Room“ im Hörsaalgebäude der Wirtschafts- und Sozialwissenschaften für die Mitglieder des Kölner Organisationskomitees der fehlenden Konferenzatmosphäre entgegengewirkt werden. Mit der Ausrichtung der FORGE-Konferenz konnte das Thema Forschungsdatenmanagement (FDM) am Standort Köln und darüber hinaus sichtbar gemacht werden (s. Kap. 9.1.3).

Covid-19-bedingt kam es 2021 zu keiner Wiederaufnahme der Präsenzangebote des DCH: Beratungen fanden wie im Vorjahr digital via E-Mail oder Zoom statt (s. Kap. 5), ebenso wurden jährliche Veranstaltungen, darunter die DHCon und die Summer School Deep Learning for Language Analysis (s. Kap. 9.1 und 9.2) sowie die Lehre weiterhin in den digitalen Raum verlagert (s. Kap. 10.2 und 10.3).

Seit Oktober 2021 wird, neben dem bereits 2020 gestarteten NFDI4Culture, mit Text+ ein weiteres geisteswissenschaftliches Konsortium in der Nationalen Forschungsdateninfrastruktur (NFDI) gefördert. Das DCH und das Language Archive Cologne (LAC) werden mit zentralen Daten und Diensten zu den Datendomänen sprach- und textbasierte Sammlungen sowie lexikalische Ressourcen beitragen (s. Kap. 3.1.1 und 3.1.2).

Mit dem Ende der Juniorprofessur von Dr.'in Brigitte Mathiak und der Berufung von Prof. Dr. Nils Reiter an das IDH gab es 2021 einen Wechsel auf der Leitungsebene: Im Oktober wurden Prof. Dr. Nils Reiter, Prof.'in Dr. Eleftheria Paliou und Prof. Dr. Øyvind Eide durch die Engere Fakultät in die Leitung des DCH gewählt; Nils Reiter

wurde zum Leitungssprecher ernannt (s. Kap. 1.2 und 1.3).

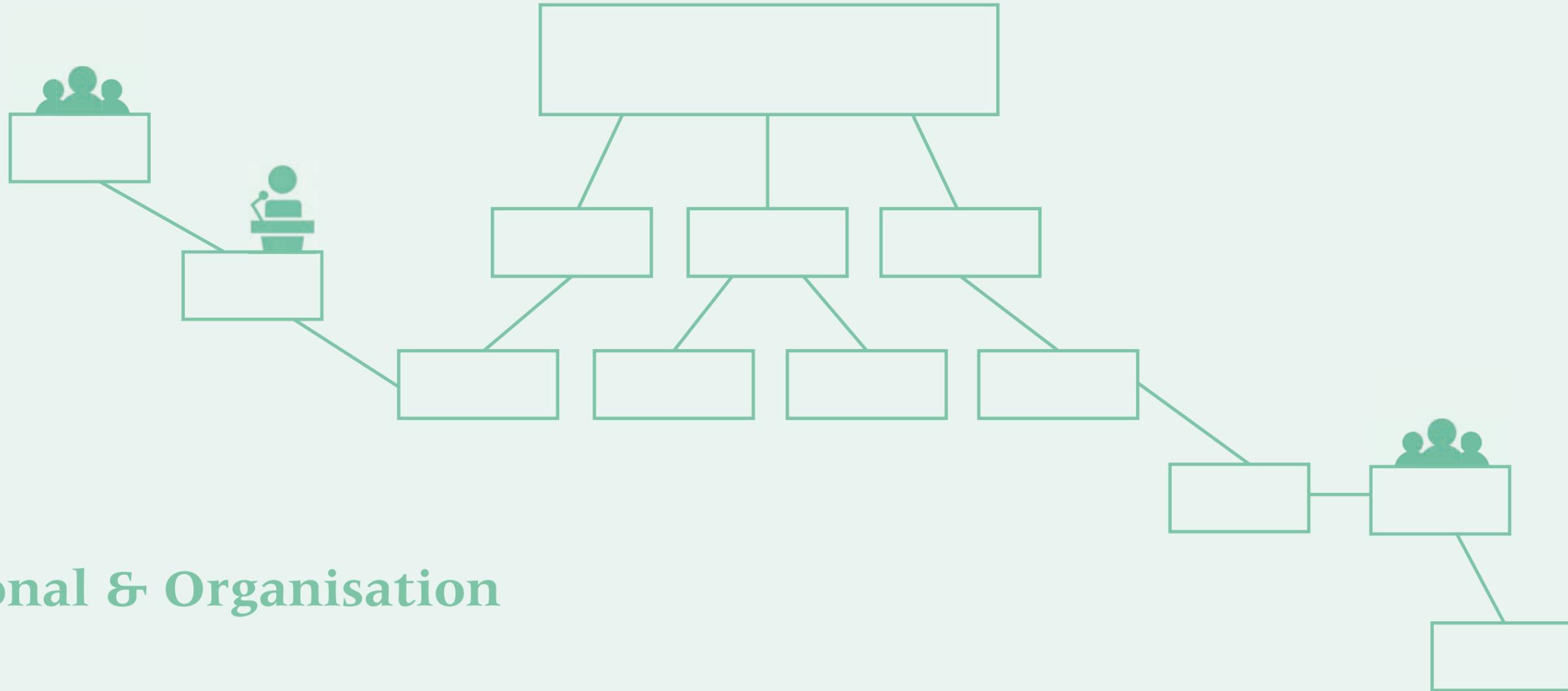
Nach über drei Jahren endete das von der Deutschen Forschungsgemeinschaft (DFG) geförderte Projekt SustainLife im Programm „Wissenschaftliche Literaturverorgungs- und Informationssysteme“ (LIS). Zum Projektabschluss wurden u. a. ein Repositorium der im Projekt bearbeiteten Use Cases zusammen mit einem niedrigschwelligem Zugangportal veröffentlicht (s. Kap. 8.1).

2021 konnte das im Bereich der Archivierung am DCH entwickelte Belegsysteem getestet werden, das in Zukunft der Erstellung von referenzierbaren Einzelnachweisen über archivierte Forschungsdatenbestände, u. a. von Projektpartner*innen, dienen soll (s. Kap. 6.3).

Die 2020 begonnene digitale Sammelkartenreihe zum Forschungsdatenmanagement „Daten & Datteln digital“ ist seit nun vier Semestern fortlaufend und mittlerweile fester Bestandteil der Öffentlichkeitsarbeit des DCH an der Universität zu Köln und darüber hinaus (s. Kap. 10.1).

Mit über 300 neuen Follower*innen 2021 hat sich der Twitter-Account @dch_cologne erneut als eine wichtige Kommunikationsplattform hinsichtlich der Sichtbarkeit von Themen aus dem Bereich des Forschungsdatenmanagements, insbesondere während der andauernden Pandemie, erwiesen. Im Mai 2021 wurde die neugestaltete Webseite des DCH gelauncht und ist nun mit der visuellen Hervorhebung der FDM-Services des DCH noch stärker auf die Bedarfe von Wissenschaftler*innen an der Philosophischen Fakultät ausgerichtet.





1 Personal & Organisation

1.1 Beirat

Aufgrund der personellen Entwicklungen und Veränderungen, insbesondere auf der Leitungsebene des DCH (s. Kap. 1.2), fand 2021 keine Beiratssitzung statt. Für 2022 ist wieder eine Sitzung des Beirats geplant.

1.2 Personalentwicklung

Mit dem Ende der Juniorprofessur am IDH Anfang 2021 hat Brigitte Mathiak auch die Leitungsebene des DCH verlassen. Als Senior Scientist in der Abteilung Knowledge Technologies for the Social Sciences am GESIS – Leibniz-Institut für die Sozialwissenschaften und ab 2022 als Co-Applicant des Teilprojekts „FAIR.rdm – FAIRes Management von afrikanischen archäologischen Forschungsdaten“ im DFG-Schwerpunktprogramm (SPP) 2143 „Entangled Africa: Innerafrikanische Beziehungen zwischen Regenwald und Mittelmeer (ca. 6000 bis 500 Jahre vor heute)“ steht sie auch weiterhin in Kooperation mit dem DCH.

Mitglieder des Beirats 2021:

- **Dr.'in Constanze Curdt**, Regionales Rechenzentrum (RRZK), Universität zu Köln
- **Dr.'in Lisa Dieckmann**, Geschäftsführerin von prometheus – Das verteilte digitale Bildarchiv für Forschung & Lehre, Universität zu Köln
- **Dr. Jens Dierkes**, Universitäts- und Stadtbibliothek Köln (USB)
- **Prof. Dr. Nikolaus P. Himmelmann**, Direktor des Instituts für Linguistik (IfL), Universität zu Köln
- **Dr.'in Ania López** (seit 2020 vertreten von **Dr. Matthias Fingerhuth**), Landesinitiative fdm.nrw der Digitalen Hochschule NRW (DH.NRW)
- **Prof.'in Dr. Eleftheria Paliou**, Archäologisches Institut, Fachgebiet Archäoinformatik/ Direktorin des Cologne Digital Archaeology Lab (CoDArchLab), Universität zu Köln
- **Prof. Dr. Patrick Sahle** (seit 2020), Professur für Digital Humanities, Historisches Seminar, Bergische Universität Wuppertal
- **Prof. Dr. Dr. h.c. Andreas Speer**, Vorstandssprecher des Cologne Center for eHumanities (CCeH), Universität zu Köln

Die Rolle des kommissarischen Leitungssprechers bis zur Wahl der neuen Leitung durch die Engere Fakultät hat Øyvind Eide übernommen. In der Fakultätssitzung am 13. Oktober 2021 wurden schließlich Nils Reiter, seit dem Wintersemester 2021/22 Professor für Digital Humanities – Sprachliche Informationsverarbeitung am IDH, und Eleftheria Paliou, Professorin am Archäologischen Institut und bislang Beiratsmitglied, in die Leitung des DCH gewählt. Sie bilden gemeinsam mit Øyvind Eide die neue Leitungsebene des DCH. Die Sprecherschaft der Leitung wurde von Nils Reiter übernommen.

Im Oktober 2021 wurde das DFG-LIS-Projekt SustainLife nach gut dreieinhalb Jahren Laufzeit abgeschlossen (s. Kap. 8.1). Das DCH musste sich entsprechend von den Kolleg*innen Anna Fischer, zuletzt wissenschaftliche Mitarbeiterin und Projektkoordinatorin, Philip Schildkamp, Projektmitarbeiter seit Projektstart 2018 und Dennis Demmer als selbstständigen Software-Entwickler im Projekt verabschieden. Philip Schildkamp ist als neuer Mitarbeiter am CCEH an der Kooperation zwischen den beiden Einrichtungen beteiligt.

Frank Marian Eilers hat im August 2021 nach zwei Jahren als studentische Hilfskraft im Bereich Archivierung das DCH verlassen. Seit April 2021 unterstützen nun Cosima Granderath-Miegel als wissenschaftliche Hilfskraft (WHB) und Sara Saleh als wissenschaftliche Hilfskraft (WHB) das DCH bei Archivierungsprozessen des DFG-Sonderforschungsbereichs (SFB) 1252 „Prominence in Language“ und komplettieren das LAC-Team. Lukas Lammers ist seit Mai 2021 als wissenschaftliche Hilfskraft (WHB) am DCH und unterstützt das Teilprojekt P11 „Learning through connecting: Datenmanagement als Basis zur nachhaltigen Nutzung von Forschungsdaten“ im SPP 2143 „Entangled Africa“.



1.3 Team 2021

Leitung



Nils Reiter (Sprecher der Leitung)

Prof. Dr. Nils Reiter wurde zum 1. Oktober 2021 als Professor für Digital Humanities – Sprachliche Informationsverarbeitung an das IDH berufen, wo er von 2019 bis 2021 eine Vertretungsprofessur innehatte. Promoviert hat er zu „Discovering Structural Similarities in Narrative Texts using Event Alignment Algorithms“ an der Universität Heidelberg in einem Kooperationsprojekt der Computerlinguistik und der Klassischen Indologie im SFB 619 „Ritualdynamik“. Seine Forschungsinteressen umfassen

Machine Learning und Deep Learning, Computational Literary Studies, quantitative Dramenanalyse, Ontologien und Wissensbasen sowie die Operationalisierung geisteswissenschaftlicher Fragen und Begriffe. Seit 2016 war Nils Reiter wissenschaftlicher Koordinator bzw. Co-Principal Investigator im Center for Reflected Text Analysis (CRETA) an der Universität Stuttgart. Er ist derzeit Projektleiter von „QuaDramA: Tracking Character Knowledge“ (Q:TRACK) im SPP 2207 „Computational Literary Studies“ und im Teilprojekt „Merkmale ästhetischer Reflexionsfiguren: Systematische Annotation und quantitative Analyse“ im SFB 1391 „Andere Ästhetik“ und Vorsitzender des Vereins CRETA e. V. Im Oktober 2021 wurde er in die Leitung des DCH und zum Leitungssprecher gewählt.

 **ORCID: 0000-0003-3193-6170** | <https://nilsreiter.de/>



Eleftheria Paliou

Prof. Dr. Eleftheria Paliou ist seit 2016 Professorin für Archäoinformatik am Archäologischen Institut und Direktorin des Cologne Digital Archaeology Lab (CoDArchLab). Zuvor war sie u. a. Marie-Curie-Stipendiatin an der Universidade de Évora (Portugal) und Alexander-von-Humboldt-Stipendiatin an der Ruprecht-Karls-Universität Heidelberg. An der University of Southampton hat sie zu „The visual consumption of mural painting in Late Bronze Age Akrotiri (Thera, Greece) – A computational

approach to visibility analysis in three-dimensional built spaces“ promoviert. Ihre Forschung umfasst Geoinformationssysteme (GIS), 3D-Modellierung und -Rekonstruk-

tion, räumliche Statistik, Urban Computing, Computersimulation und die Ägäische Bronzezeit. Sie ist u. a. Leiterin der Arbeitsgruppe 4 Archaeological scholarship des COST (European Cooperation in Science & Technology) Action Arkwork im EU-Förderprogramm Horizon 2020 und Co-Leiterin des Projekts „Modellierung prähistorischen Jagdverhaltens: Erweiterung der Jäger-Sammler Archäologie durch indigene Experten“, gefördert vom Bundesministerium für Bildung und Forschung (BMBF). Von Beginn an ist Eleftheria Paliou Mitglied des 2018 gebildeten Beirats des DCH. Im Oktober 2021 wurde sie in die Leitung des DCH gewählt.

<https://eleftheria121.wordpress.com/>



Øyvind Eide

Prof. Dr. Øyvind Eide hat zu „The area told as a story. An inquiry into the relationship between verbal and map-based expressions of geographical information“ als Stipendiat des Research Council Norway am King's College in London promoviert. Bevor er 2017 als Professor für Digital Humanities – Historisch-Kulturwissenschaftliche Informationsverarbeitung an das IDH berufen wurde, war er an der Universität Oslo und an der Universität Passau im Fachbereich Digital Humanities tätig. Er forscht und lehrt u. a. zu

Theorien der Modellbildung in den Geisteswissenschaften und Informationsmodellierung im Bereich digitales Kulturerbe sowie zu digitalen Karten. Øyvind Eide leitet u. a. das internationale ERASMUS+ Projekt ARTEST und ist Co-Applicant des Teilprojekts P11 „Learning through connecting: Datenmanagement als Basis zur nachhaltigen Nutzung von Forschungsdaten“ im SPP 2143 „Entangled Africa: Innerafrikanische Beziehungen zwischen Regenwald und Mittelmeer (ca. 6000–500 Jahre vor heute)“. Seit April 2019 ist er Teil der Leitungsebene des DCH.

 **ORCID: 0000-0002-7766-6287** | <https://oeide.no/>

Fakultätsstellen



Felix Rau (Geschäftsführung)

Felix Rau hat Allgemeine Sprachwissenschaft, Indologie und Sinologie (Älteres China) an der Universität zu Köln studiert. Er war wissenschaftlicher Mitarbeiter am Leiden University Centre for Linguistics (LUCL) und am Institut für Linguistik (IfL) an der Universität zu Köln. Er hat jahrelange praktische Erfahrung mit der Erhebung und Verarbeitung von Forschungsdaten in der linguistischen Feldforschung. Seine Forschungsschwerpunkte umfassen audiovisuelle (AV) Daten in den Geisteswissenschaften,

Metadaten für geisteswissenschaftliche Forschungsdaten, lexikalische Ressourcen und Sprachdokumentation (Schwerpunkt Südasien). Seit 2019 ist er wissenschaftlicher Mitarbeiter und Geschäftsführer am DCH: Er arbeitet in BMBF-Verbundprojekt „QUEST: Quality-Established“ sowie vorher im BMBF-Verbundprojekt „KA³ – Kölner Zentrum Analyse und Archivierung von AV Daten“, am LAC und ist Teil des Beratungsteams. Darüber hinaus widmet er sich der Integration der DCH-Angebote in die Forschungsinfrastrukturen CLARIN-D (Common Language Resources and Technology Infrastructure Deutschland) und CLARIN ERIC (CLARIN European Research Infrastructure Consortium) und dem Arbeitsbereich Zertifizierung.

 **ORCID: 0000-0003-4167-0601**



Patrick Helling (Koordination)

Patrick Helling hat Medienwissenschaften, Medieninformatik und Medienwirtschaft an der Universität zu Köln und der Universität Bayreuth studiert. Seine Forschungsschwerpunkte liegen in der digitalen Transformation geisteswissenschaftlicher Forschung sowie der Erforschung geisteswissenschaftlichen Forschungsdatenmanagements. Als wissenschaftlicher Mitarbeiter und Koordinator des DCH kümmert er sich u. a. um die Vertretung des Datenzentrums in Gremien und Verbänden

und ist Teil des Beratungsteams. Patrick Helling befasst sich insbesondere mit der Formalisierung und Modellierung von Forschungsdatenmanagement in den Geisteswissenschaften und arbeitet im Rahmen seiner Promotion an der Entwicklung eines abstrakten Beschreibungsmodells für Forschungsdatenmanagementprozesse, -strukturen und Bedarfe. Seit März 2020 ist Patrick Helling kooptiertes Vorstandsmitglied

und Data Steward im DHd-Verband. Er ist Mitglied des Research Data Alliance Deutschland e. V. (RDA DE), vertritt das DCH im Verein Deutsche Initiative für Netzwerkinformation e. V. (DINI) und ist im Kontext der Initiative GO FAIR als Koordinator der nationalen Arbeitsgruppe GO UNITE! des Data Stewardship Competence Centers Implementation Networks (DSCC-IN) aktiv.

 **ORCID: 0000-0003-4043-165X**

Projektstellen



Anke Debbeler

Anke Debbeler hat einen Bachelorabschluss in Mathematik sowie Erziehungswissenschaften und Informationsverarbeitung mit den Schwerpunkten Medienpädagogik und linguistische Datenverarbeitung an der Universität zu Köln. Derzeit studiert sie Informationsverarbeitung im Master. Als wissenschaftliche Mitarbeiterin im BMBF-Verbundprojekt KA³ war sie ab Dezember 2019 für die Konvertierung von Metadaten zuständig. Seit Mai 2021 betreut sie als wissenschaftliche Mitarbeiterin das LAC.



Cosima Granderath-Miegel

Cosima Granderath-Miegel hat Literatur- und Sprachwissenschaften mit Schwerpunkt in kognitiver Linguistik an der Rheinisch-Westfälischen Technischen Hochschule (RWTH) Aachen studiert. Derzeit befindet sie sich im Master-Studium der Informationsverarbeitung an der Universität zu Köln. In ihrer Tätigkeit als wissenschaftliche Hilfskraft (WHB) im SFB 1252 „Prominence in Language“ seit April 2021 kuratiert und konvertiert sie in linguistischer Feldforschung erhobene Daten und Metadaten.



Lukas Lammers

Lukas Lammers hat im Bachelor Klassische Archäologie und Katholische Religionslehre an der Westfälischen Wilhelms-Universität Münster studiert. Neben seinem Master-Studium im Fach Archäoinformatik an der Universität zu Köln arbeitet er als wissenschaftliche Hilfskraft (WHB) seit Mai 2021 am DCH und unterstützt das SPP 2143 „Entangled Africa“ beim Datenmanagement sowie der -kuratation und Archivierung.



Elisabeth Mollenhauer

Elisabeth Mollenhauer studiert Kunstgeschichte an der Rheinischen Friedrich-Wilhelms-Universität Bonn. Als wissenschaftliche Hilfskraft (WHB) ist sie für die Redaktion des Jahresberichts des DCH, die Webseite und den Twitter-Account zuständig. Außerdem unterstützt sie das Beratungsteam. In ihren Aufgabenbereich fallen außerdem die Organisation der DHCon und der FORGE-Konferenz 2021.



Sara Saleh

Sarah Saleh hat einen Bachelorabschluss in Geschichte und Informationsverarbeitung an der Universität zu Köln und studiert derzeit dort Informationsverarbeitung im Master. Als wissenschaftliche Hilfskraft (WHB) ist sie seit April 2021 im SFB 1252 „Prominence in Language“ angestellt. Sie arbeitet am Archivierungsworkflow des DCH mit und unterstützt darüber hinaus das LAC.



Nicole Majka

Nicole Majka hat einen Bachelor in Latein und Slavistik und einen Master in Linguistik an der Universität zu Köln absolviert. Sie war als wissenschaftliche Hilfskraft (WHB) in zwei Projekten an der Universität zu Köln tätig: Sie hat das Projekt B01 im SFB 1252 „Prominence in Language“ unterstützt und inventarisiert und analysiert weiterhin für das BMBF-Verbundprojekt QUEST Metadatenformate für AV-Daten als externe Fachkraft. An der Bangor University (Wales) promoviert sie derzeit in der

Linguistik im Bereich des walisischen Spracherwerbs.

Ehemalige

Brigitte Mathiak

Dr.'in Brigitte Mathiak hat für ihre Dissertation zur Analyse von wissenschaftlichen Texten in der Bioinformatik gearbeitet und ist dann als Teamleiterin an das GESIS – Leibniz-Institut für Sozialwissenschaften in Köln gewechselt. Dort hat sie u. a. zu den Themen Forschungsdatenmanagement und Archivierung gearbeitet und geforscht. Gemeinsam mit Prof. Dr. Dr. h.c. Frank Leymann vom Institut für Architektur von Anwendungssystemen (IAAS) der Universität Stuttgart hat sie von 2018 bis 2021 das DFG-LIS-Projekt SustainLife geleitet. Von 2015 bis 2021 war sie als Juniorprofessorin für Digital Humanities am IDH. 2016 hat sie die Geschäftsführung am DCH übernommen und wurde 2018 in die Leitungsebene des DCH gewählt; von April 2019 bis Januar 2021 war sie Sprecherin der Leitung. Seit Februar 2021 ist sie wieder bei GESIS und ist außerdem Co-Applicant des 2022 startenden Teilprojekts „FAIR.rdm – FAIRes Management von afrikanischen archäologischen Forschungsdaten“ im SPP 2143 „Entangled Africa“, welches auch am DCH angesiedelt ist.

Andreas Witt

Prof. Dr. Andreas Witt ist seit März 2019 Leiter der neuen Abteilung Digitale Sprachwissenschaft und Leiter des Programmbereichs Forschungsinfrastrukturen am Institut für Deutsche Sprache (IDS) in Mannheim. Seine Position am IDS ist verbunden mit einer Professur für Computational Humanities & Text Technologie an der Universität Mannheim. Vor seinem Wechsel beschloss die Philosophische Fakultät, dass Andreas Witt als kooptierter Professor für Sprachliche Informationsverarbeitung / Digital Humanities weiterhin Mitglied des IDH bleibt. Seine Forschungsinteressen umfassen Korpuslinguistik, audiovisuelle Sprachkorpora, Informationsmodellierung und Forschungsdatenmanagement. Er hat intensiv die Vernetzung und Forschungscoopera-

tionen des DCH in der nationalen und internationalen Forschungslandschaft vorangetrieben und war von Juli 2018 bis April 2019 Sprecher der Leitung. Aktuell ist er Principal Investigator im Kölner Teilprojekt des BMBF-Verbundprojekts QUEST und Participant im Oktober 2021 gestarteten NFDI-Konsortium Text+.

Frank Marian Eilers

Frank Marian Eilers studiert Informationsverarbeitung an der Universität zu Köln und ist freiberuflicher Schriftdolmetscher. Seine Aufgabengebiete als studentische Hilfskraft am DCH umfassten die Betreuung der Webseite des DCH, des internen Wiki, die technische und administrative Betreuung der Datenarchivierung sowie die Dokumentation und Optimierung des Archivierungsworkflows.

Anna Fischer

Anna Fischer hat Politikwissenschaft mit der Spezialisierung Spanien-, Portugal- und Lateinamerikastudien am Institut d'Études Politiques in Paris und Informationsverarbeitung in Köln studiert. Im und neben dem Studium hat sie sich in mehreren Projekten der Verarbeitung von Textdaten in romanischen Sprachen gewidmet. Ab Januar 2020 war sie Teil des DFG-LIS-Projekts SustainLife und von September 2020 bis zum Abschluss des Projekts im Oktober 2021 als wissenschaftliche Mitarbeiterin verantwortlich für die Projektkoordination.

Philip Schildkamp

Philip Schildkamp hat Soziologie, Psychologie und Informationsverarbeitung studiert. Als Projektmitarbeiter hat er ab März 2018 am DCH das DFG-LIS-Projekt SustainLife unterstützt und u. a. die Übertragbarkeit von standardisierten Verfahren zum Software-Deployment und -Management aus dem industriellen Umfeld auf das der Digital Humanities (DH) erforscht. Seit März 2021 unterstützt er als wissenschaftlicher Mitarbeiter das CCEH in den Bereichen Systemadministration und -integration sowie in der Entwicklung und Provisionierung.

Dennis Demmer

Dennis Demmer studiert im Bachelor Informationsverarbeitung und Geschichte an der Universität zu Köln. Ab November 2020 bis zum Projektende im Oktober 2021 hat er als selbstständiger Software-Entwickler das DFG-LIS-Projekt SustainLife in den Bereichen Use-Case-Modellierung, Server-Infrastruktur und technische Dokumentation unterstützt.

1.4 Kooperation CCEH



Jonathan Blumtritt

Jonathan Blumtritt hat Neuere Geschichte, Wirtschafts- und Sozialgeschichte und Informatik an der Friedrich-Schiller-Universität Jena und an der Rheinischen Friedrich-Wilhelms-Universität Bonn studiert und war seit 2011 in verschiedenen Digital Humanities Projekten am CCEH beschäftigt. Seit der Gründung des DCH 2012/2013 war er für Koordination, Beratung, Antragsentwicklung und Datenkuration zuständig. Als wissenschaftlicher Mitarbeiter im BMBF-Verbundprojekt KA³ hat er von

2015 bis 2019 die Projektleitung bei der Koordination der technischen Entwicklungen unterstützt. Seit Oktober 2019 ist er Teil der Geschäftsführung des CCEH, nimmt Aufgaben der Koordinierungsstelle Digital Humanities der AWK wahr und engagiert sich intensiv bei den Beteiligungen an den geisteswissenschaftlichen Konsortien in der NFDI in Zusammenarbeit mit dem DCH.

 **ORCID: 0000-0002-1438-379X**



Claes Neuefeind

Dr. Claes Neuefeind hat Informationsverarbeitung, Deutsche Philologie und Philosophie an der Universität zu Köln studiert und war von 2008 bis 2012 Kollegiat an der a.r.t.e.s. Graduate School for the Humanities Cologne; promoviert hat er unter dem Titel „Muster und Bedeutung – Bedeutungskonstitution als kontextuelle Aktivierung im Vektorraum“. Ab 2007 war er als wissenschaftlicher Mitarbeiter am IfL in der Abteilung Sprachliche Informationsverarbeitung tätig. Von März 2018 bis September

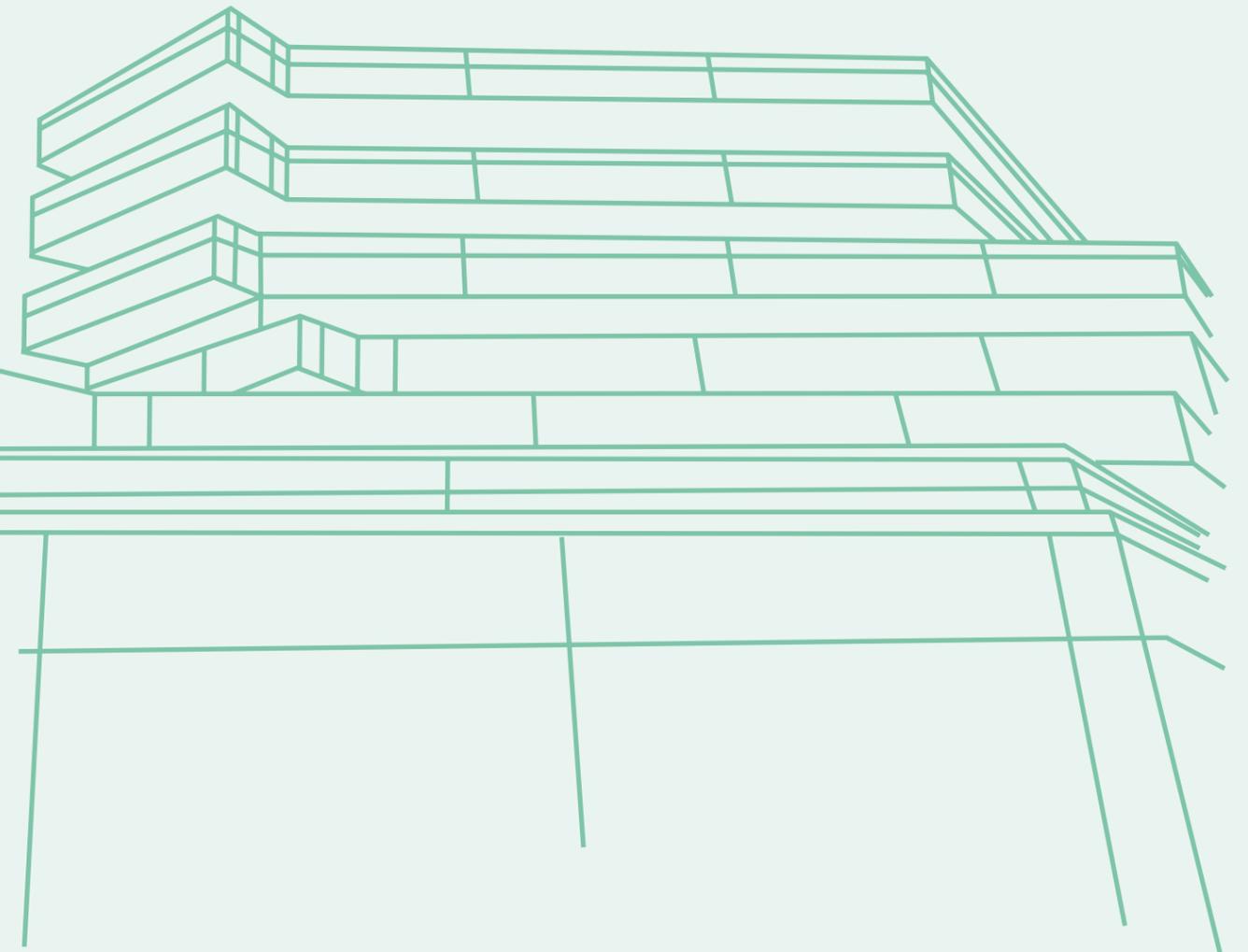
2019 hat er am DCH das DFG-LIS-Projekt SustainLife bearbeitet und war 2019 an der Beantragung der NFDI-Konsortien Text+ und NFDI4Culture beteiligt. Seit Oktober 2019 ist er Teil der Geschäftsführung des CCEH und dort u. a. mitverantwortlich für die Koordinierungsstelle Digital Humanities der AWK.

 **ORCID: 0000-0002-9377-9492**

2 Einbettung & Kooperationen an der Universität zu Köln

Dekanat der Philosophischen Fakultät

Das DCH ist dem Dekanat der Philosophischen Fakultät direkt unterstellt und seit der Verabschiedung einer eigenen Satzung und der Ernennung eines Beirats 2018 noch stärker an der Fakultät verankert. Die Zusammenarbeit wird durch die Abstimmung mit Mitarbeiter*innen am Dekanat stetig verstärkt. Das DCH steht in engem Kontakt mit dem Prodekan für Forschung und wissenschaftlichen Nachwuchs, Prof. Dr. Wolfram Nitsch sowie seiner Mitarbeiterin Henriette Terpe. Die in Zusammenarbeit mit dem Dekanat entstandene HOW-TO-Reihe „Daten und Datteln“ wird weiterhin in Form von digitalen Sammelkarten, die via Twitter publiziert werden, umgesetzt (s. Kap. 10.1). Darüber hinaus wurde im Berichtszeitraum eine „Daten & Datteln“-Sonderserie zum Thema „Digitales Publizieren an der Universität zu Köln“ gemeinsam mit Vertreter*innen der Universitäts- und Stadtbibliothek (USB) aufgesetzt, die aus gemeinsamen Gesprächen mit Wolfram Nitsch bezüglich einer uniweiten Hochschulbibliografie entstanden ist. Auch im Kontext des Aufbaus einer NFDI steht das DCH in regelmäßigem Austausch mit Vertreter*innen des Dekanats. Ziel ist es dabei u. a. ein gemeinsames Austauschforum für lokale Beteiligte an den geisteswissenschaftlichen NFDI-Konsortien mit dem Dekanat zu etablieren sowie die



NFDI selbst stärker an der Fakultät zu verankern. Ein initialer Workshop des geplanten NFDI-Hubs ist für Mai 2022 geplant (s. Ausblick).

Dezernat 7

Das Dezernat 7 „Forschungsmanagement“ der Universität zu Köln und das DCH stehen bei der Vermittlung von Geisteswissenschaftler*innen, die sich in der Antragsphase befinden und Beratungsbedarf bei FDM-Fragen haben, regelmäßig in Kontakt. Darüber hinaus arbeitet das DCH im Kooperationsrahmen mit dem Cologne Competence Center for Research Data Management (C³RDM) mit Vertreter*innen des Dezernats 7 zusammen.

Kommission für Informations- und Kommunikationstechnologie (IuK)

Als stellvertretender Fakultätsvertreter nimmt Øyvind Eide regelmäßig an den Sitzungen der Kommission teil. Mitarbeiter*innen des DCH nehmen zusätzlich an den Sitzungen des IuK-operativ teil. Bei diesen informellen Treffen werden die Themen aus der Kommission auf operativer Ebene vor- und nachbereitet.

Veröffentlichung 2021: Daten & Datteln Sonderedition „Digitales Publizieren an der Universität zu Köln“

Im Sommersemester 2021 hat das DCH in Kooperation mit der USB und dem Dekanat der Philosophischen Fakultät die Sonderedition „Digitales Publizieren an der Universität zu Köln“ bestehend aus drei Karten veröffentlicht. Die Karten wurden über den Fakultätsverteiler an die Mitglieder der Philosophischen Fakultät kommuniziert.

Themen:

- Open Access
- Import von Daten durch den Kölner UniversitätsPublikationsServer (KUPS)
- Import über den Digital Object Identifier (DOI)

Unter <https://dch.phil-fak.uni-koeln.de/daten-und-datteln/sammelkarten> stehen die Karten zum Download zur Verfügung.



Regionales Rechenzentrum (RRZK)

Das DCH nutzt ausschließlich Hardware zur Bereitstellung von Speicher- und Rechenkapazitäten, die das RRZK anbietet, und hat auf operativer und leitender Ebene insbesondere im Rahmen des mittlerweile ausgelaufenen BMBF-Verbundprojekts KA³ eng mit dem RRZK zusammengearbeitet. In Zukunft ist ein intensiverer Austausch bei der Unterstützung des SFB 1252 „Prominence in Language“ geplant. Durch eine umfassende Dokumentation aller genutzten Ressourcen und vorhandener Bestände des DCH soll die Zusammenarbeit mit dem RRZK auf operativer Ebene in Zukunft noch transparenter und effizienter gestaltet werden. Darüber hinaus findet zwischen RRZK und DCH im Rahmen des C³RDM ein regelmäßiger Dialog statt.

Universitäts- und Stadtbibliothek (USB)

Der Austausch zwischen der USB und dem DCH bei FDM-Fragen wird mittlerweile in der Kooperation mit dem C³RDM gebündelt und intensiviert. Zusätzlich stehen das DCH und die Universitäts- und Stadtbibliothek bei den Themen Hochschulbibliografie und Digitales Publizieren im Kontakt. Im Berichtszeitraum fanden mehrere Treffen im Rahmen der Konzeption eines Folgeantrags zum Fachinformationsdienst (FID) Philosophie, der u. a. an der USB angesiedelt ist, statt. Zu betonen ist in diesem Zusammenhang insbesondere der Austausch und die Zusammenarbeit mit Ralf Depping (USB). Darüber hinaus findet im März 2022 ein gemeinsam organisierter Workshop zum Thema „Data Literacy für Geisteswissenschaftler*innen“ statt (s. Ausblick).

Cologne Competence Center for Research Data Management (C³RDM)

Mit dem C³RDM wurde 2018 eine fakultätsübergreifende Anlaufstelle für Fragen zum Thema Forschungsdatenmanagement an der Universität zu Köln eingerichtet. Das C³RDM besteht aus Vertreter*innen des RRZK, der USB und des Dezernats 7. Das DCH steht in engem Kontakt mit den Mitarbeiter*innen des C³RDM. Im Rahmen eines monatlichen Jour Fixe findet ein regelmäßiger Austausch zwischen den beiden Institutionen statt. Im Falle von FDM-Beratungsvorgängen, in denen Schnittmengen zwischen beiden Einrichtungen identifiziert werden können, arbeiten C³RDM und DCH eng zusammen. Entsprechend fanden im Berichtsjahr einige gemeinsame Beratungstermine mit Forschenden statt. Darüber hinaus steht das DCH auch im Kontext der NFDI im Austausch mit den Kolleg*innen des C³RDM.

Cologne Center for eHumanities (CCeH)

Das CCeH ist mit dem DCH auf organisatorischer und strategischer Ebene eng verzahnt. Gemeinsam mit dem prometheus-Bildarchiv, dem DCH und der Theaterwissenschaftlichen Sammlung (TWS) der Universität zu Köln ist das CCeH bereits Participant des im Oktober 2020 gestarteten Konsortium NFDI4Culture. Die Zusam-

menarbeit wurde zuletzt durch den Förderbeginn des NFDI-Konsortiums Text+ im Oktober 2021, an dem beide Institutionen unmittelbar beteiligt sind, noch einmal deutlich verstärkt. Darüber hinaus waren 2021 Vertreter*innen beider Institutionen gemeinsam an einer erneuten Antragstellung für die Konsortien NFDI4Memory und NFDI4Objects beteiligt (s. Kap. 3.1). Beide Institutionen haben im September 2021 gemeinsam mit dem IDH die FORGE-Konferenz veranstaltet (s. Kap. 9.1.3).

Koordinierungsstelle Digital Humanities der Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste (AWK)

Das DCH steht der Koordinierungsstelle Digital Humanities der AWK, die am CCeH angesiedelt ist, seit 2013 als Datenzentrum zur Verfügung und ist am Forschungsdatenmanagement der hier betreuten Projekte beteiligt. Die Koordinierungsstelle fällt personell mit der Geschäftsführung des CCeH zusammen und ist somit eng mit dem DCH verbunden. Insbesondere im Rahmen der Entwicklungen und Konsortialbildungen hinsichtlich einer NFDI ist das DCH über die AWK fest in die Prozesse eingebunden (s. Kap. 3.1).

Institut für Digital Humanities (IDH)

In Zusammenarbeit mit dem IDH hat das DCH im Sommersemester 2020 erneut die Übung „Forschungsdatenmanagement“ für Master-Studierende der Studiengänge Medieninformatik, Informationsverarbeitung und Linguistik veranstaltet (s. Kap. 10.3). Beide Institutionen waren gemeinsam mit dem CCeH zum Ende des Wintersemesters 2020/21 Veranstalter der DHCon 2021, des dritten interaktiven Showcase Events des IDH, der virtuellen Summer School „Deep Learning for Language Analysis“ im August sowie der FORGE-Konferenz im September 2021 (s. Kap. 9.1.1, 9.1.2 und 9.1.3). Gemeinsam mit Øyvind Eide ist das DCH an der Konsortialbildung von NFDI4Objects beteiligt, das in der dritten Ausschreibungsrunde 2021 erneut seinen Antrag eingereicht hat (s. Kap. 3.1.3).

Institut für Linguistik (IfL), Abteilung Allgemeine Sprachwissenschaft (ASW)

Mit dem IfL besteht eine langjährige Zusammenarbeit, die den Ausgangspunkt für den Aufbau des Kompetenzbereichs AV-Daten am DCH bildete. Die laufenden Zertifizierungen als CLARIN B-Centre und dem CoreTrustSeal (CTS) sowie die beiden BMBF-Verbundprojekte KA³ und QUEST sind aus dieser Zusammenarbeit erwachsen. Als neueste Zusammenarbeit ist die Beteiligung des DCH am NFDI-Konsortium Text+ in Kooperation mit dem IfL entstanden (s. Kap. 3.1.1).

Thomas-Institut, Fachinformationsdienst (FID) Philosophie

Mit dem Thomas-Institut bestehen eine langjährige Zusammenarbeit und ein intensiver Austausch auf technischer und strategischer Ebene: Das Projekt Digital

Averroes Research Environment (DARE), das am Thomas-Institut angesiedelt ist, war gemeinsam mit dem DCH Projektpartner im 2020 abgeschlossenen BMBF-Verbundprojekt KA³. Regelmäßige Konsultationen mit Vertreter*innen des Thomas-Instituts sowie der USB fanden im Berichtszeitraum im Rahmen der Konzeption eines Folgeantrags für den FID Philosophie statt.

prometheus – Das verteilte digitale Bildarchiv für Forschung & Lehre

Das DCH hat, gemeinsam mit den Kolleg*innen des CCeH, als geisteswissenschaftliches Datenzentrum das prometheus-Bildarchiv bei der Antragstellung für das Konsortium NFDI4Culture aktiv unterstützt. Mit der bewilligten Förderung des Konsortiums im Oktober 2020 wurde diese Unterstützung weiter ausgebaut. Das DCH übernimmt in diesem Zusammenhang eine aktive Rolle als lokale Partnerinstitution und Participant des Konsortiums (s. Kap. 3.1.2).

NFDI-Hub – Geisteswissenschaftliche Konsortien an der Universität zu Köln

Mit dem Ziel einen runden Tisch zum (regelmäßigen) Austausch über Aktivitäten und Entwicklungen der geisteswissenschaftlichen Konsortien innerhalb der NFDI, dem Dekanat sowie Wissenschaftler*innen an der Philosophischen Fakultät zu etablieren, ist für Mai 2022 eine gemeinsame Vernetzungsveranstaltung geplant. Langfristig sollen mit Hilfe eines lokalen NFDI-Hub die Entwicklungen innerhalb der geisteswissenschaftlichen NFDIs, die an der Philosophischen Fakultät angesiedelt und vertreten sind, nachhaltig in die Fakultät getragen und dort verankert werden (s. Kap. 3.1 und Ausblick).

a.r.t.e.s. Graduate School for the Humanities Cologne

Das DCH hat im Sommersemester 2021 und im Wintersemester 2021/22 erfolgreich mehrtägige Workshops zum Thema „Forschungsdatenmanagement in den Geisteswissenschaften“ für Promovierende an der a.r.t.e.s. Graduate School for the Humanities Cologne veranstaltet (s. Kap. 10.2). Gleichzeitig treibt das DCH aktiv die Platzierung des Themas Forschungsdatenmanagement an der Graduiertenschule voran und präsentiert regelmäßig die Services des Datenzentrums im Promotionscafé.

SFB 1252 „Prominence in Language“

Das DCH unterstützt den SFB 1252 „Prominence in Language“ seit 2017 insbesondere durch eine enge Kooperation mit dem Informationsinfrastrukturprojekt (INF) unter der Leitung von Prof. Dr. Nikolaus P. Himmelmann. In der aktuellen zweiten Phase, die 2021 begonnen hat, wird die Kooperation mit dem INF-Projekt weiter vertieft: Zwei wissenschaftliche Hilfskräfte werden über das INF-Projekt finanziert und sind direkt am DCH angesiedelt. Sie unterstützen die Kuratierung, Archivierung und Publikation der Daten aus den Teilprojekten des SFB 1252 (s. Kap. 1.3 und 6.3.2).

INTERVIEW MIT: **FAHIME SAME & MAXIMILIAN HÖRL** (DOKTORAND*INNEN UND MITARBEITER*INNEN IM INF-PROJEKT DES SFB 1252)

Elisabeth Mollenhauer: Was ist das INF-Projekt des SFB 1252? Mit was für Daten habt ihr zu tun? Was sind eure Aufgaben?

Fahime Same: Um zu verstehen, was ein INF-Projekt ist, muss man zunächst wissen, was ein Sonderforschungsbereich ist. In einem SFB gibt es viele Teilprojekte, die sich alle mit einem übergreifenden Thema unter verschiedenen Gesichtspunkten befassen, in unserem Fall ist das Thema Prominence in Language. Das INF-Projekt ist dabei für ein gutes Datenmanagement verantwortlich: Es bereitet alles für die Datenarchivierung vor, kümmert sich um die Erstellung von Metadaten und organisiert passende Infrastrukturen für die Ablage von Forschungsdaten. Es kann außerdem Strukturen für die gemeinsame Arbeit an Daten schaffen, wie in unserem Fall die Organisation des verwendeten Cloud-Speichersystems. Darüber hinaus übernimmt unser INF-Projekt auch noch einige andere Dinge, die mit Daten zu tun haben. Für die Teilprojekte stellen wir beispielsweise Angebote zu Annotationen zur Verfügung, wie und wo man sie am Besten macht oder wo keine notwendig sind, weil wir die Informationen selbst aus den Daten ziehen können. Wir bieten den Teilprojekten unseres Forschungsbereichs außerdem Unterstützung bei Fragen zur Datenanalyse.

Maximilian Hörl: Wir unterscheiden bei Forschungsdaten aus den verschiedenen Projekten zwischen zwei Gruppen: Wir haben Daten aus Sprachexperimenten und wir haben Projekte, die mit audiovisuellen oder schriftlichen Sprachkorpusdaten arbeiten. In der experimentellen Gruppe gibt es wiederum ganz verschiedene Datentypen, je nachdem mit Hilfe welcher Technik die Experimente durchgeführt werden. Dabei kommen neben Fragebogentechniken auch Eye-Tracking, EEG (Elektroenzephalografie) oder fMRI (funktionale Magnetresonanztomografie) zum Einsatz. Die resultierenden Daten unterscheiden sich sowohl in ihrer Größe, als auch darin, wie man sie analysiert.

Elisabeth Mollenhauer: Was sind eure Erfahrungen in der Zusammenarbeit mit dem DCH? In welchen Bereichen konnte euch das DCH bislang unterstützen? Welche Services des DCH nutzt ihr? Wo wäre darüber hinaus noch Bedarf?

Fahime Same: Von Beginn der ersten Förderphase an standen wir in sehr guter Beziehung zum DCH. Nicht nur auf der technischen Ebene, auch auf der konzeptuellen Ebene hat uns das DCH unterstützen können. Am Anfang ging es beispielsweise darum, wie wir unsere Metadaten sammeln und speichern können. Wir haben uns anfangs etwas verloren gefühlt, weil es einfach so viele Daten gibt und die Projekte so unterschiedlich sind. Das DCH hat uns beispielsweise empfohlen, für den Start einen Fragebogen zu entwickeln, in dem alle datenrelevanten Informationen aus den Projekten erhoben werden und Interviews mit den Beteiligten zu führen, um einen

Überblick über den SFB zu bekommen. Das war auch wichtig für eine gute Kommunikation mit den Projekten. Auf konzeptueller Ebene hat uns das DCH im Hinblick auf die Metadaten aus den einzelnen Projekten geholfen. Anschließend konnte uns das DCH dabei unterstützen, ein passendes Datenspeichersystem für die Archivierung zu finden und uns bei der Archivierung selbst unterstützen.

Maximilian Hörl: Es erleichtert unsere Arbeit, dass wir keinen eigenen Server betreiben müssen. Mittlerweile sind die meisten konzeptionellen Teile unserer Arbeit erledigt, wodurch schon viel geschafft ist. Dafür war die Unterstützung des DCH äußerst wichtig. Was weitere Bedarfe angeht, arbeiten wir noch an ein paar Dingen für die Zukunft: Wir versuchen, die Forschungsdaten aus dem SFB für die Forschungsöffentlichkeit sichtbar zu machen. Ein Überblick über die archivierten Daten auf unserer Webseite wäre sehr nützlich. Deshalb freuen wir uns darauf, das Belegschaftssystem nutzen zu können, das das DCH entwickelt [Anm. d. Redaktion: Kap. 6.3.1], und auf die Möglichkeit, einige Daten im LAC publizieren zu können. Das ist natürlich alles noch in der Entwicklung.

Elisabeth Mollenhauer: Gibt es denn noch weitere Services, die ihr gerne nutzen würdet, die aber aktuell fehlen?

Fahime Same: Wir hoffen, dass die Archivierung im LAC bald reibungslos läuft, damit wir dort mehr Daten ablegen können. Das Belegschaftssystem, das das DCH entwickelt, würde uns darüber hinaus einen guten Überblick über das Archiv geben. Außerdem fördert es die Sichtbarkeit der Daten und es dient Forschenden als Bestätigung darüber, dass ihre Daten archiviert wurden.

Elisabeth Mollenhauer: Es gibt im INF-Projekt in der zweiten Förderphase Hilfskraftstellen, die am DCH angesiedelt sind. Wie bewertet ihr das für die bisherige Zusammenarbeit von SFB und DCH?

Maximilian Hörl: Es wird uns in der Zusammenarbeit mit dem DCH helfen, denn so ist eine Person direkt für uns ansprechbar, die dann mit Priorität unsere Aufgaben bearbeiten kann. Das ist wahrscheinlich ein gutes Modell, aber es ist vielleicht noch zu früh, sich eine abschließende Meinung zu bilden, da die zweite Phase ja gerade erst gestartet ist.

Fahime Same: In der ersten Phase hatten wir gemeinsam mit dem DCH viel Arbeit zu bewältigen. Auf konzeptueller Ebene war es schwer für uns herauszufinden, was eigentlich genau zu tun ist. Wir hatten einige Beratungstermine mit dem DCH, aber es wäre hilfreich gewesen, diese Hilfskraftstellen auch in die Förderanträge der ersten Phase zu schreiben und sie direkt zum Projektstart zu haben. So hätten alle von

Beginn an zusammenwachsen können.

Maximilian Hörl: Die Hilfskräfte sind wichtig für die Arbeit, die aus der Planung resultiert, aber die Unterstützung, die wir hierbei auf konzeptueller Ebene erhalten haben, hätten sie nicht alleine abdecken können.

Elisabeth Mollenhauer: Was würdet ihr anderen geisteswissenschaftlichen Projekten für ihre Datenmanagementplanung raten?

Fahime Same: Wenn man sehr datenaffin ist, kann es passieren, dass man ein bisschen zu perfektionistisch wird. Viele Forschende finden Metadaten nicht sonderlich interessant, aber diejenigen, die in einem INF-Projekt arbeiten, finden sie faszinierend. Wir waren anfangs sehr perfektionistisch hinsichtlich der Metadaten. Aber ab einem bestimmten Punkt ist uns klargeworden, dass wir also erstmal eine Grundlage schaffen sollten, bevor wir anschließend immer noch die Möglichkeit haben die Metadaten weiter zu verbessern. Als Beteiligte an einem INF-Projekt in einem größeren SFB kann ich empfehlen, nicht allzu viel Zeit für die Konzeption einer perfekten Lösung zu verwenden. Man sollte stattdessen früh anfangen Metadaten zu sammeln und sich mit den Daten zu befassen. Dadurch, dass man früh einige Datenbeispiele sieht und sich damit beschäftigt, kann man auch praktische Probleme, die man nicht vorhergesehen hat, schrittweise lösen und wird nicht von einer Flut chaotischer Daten überrollt. Deshalb startet man lieber früher mit einem Ablauf, der noch nicht bis ins Detail perfekt ausgearbeitet ist. Später kann man immer noch verbessern.

Für uns war die Kommunikation außerdem sehr wichtig. Man muss die Projekte früh kennenlernen und den Kontakt pflegen. Wir haben sehr früh Interviews mit den einzelnen Projekten geführt. Das hat uns nicht nur geholfen die Daten kennen zu lernen, sondern auch die Menschen. Gerade in der Zusammenarbeit mit einem INF-Projekt ist gegenseitiges Vertrauen wichtig. Auch hier sollte man so früh wie möglich damit beginnen.

Zum Beispiel wollten wir zu Beginn für alles perfekte Guidelines schreiben, bis wir realisiert haben, dass Hands-On-Workshops viel effektiver sind. Bei 80 oder 90 Forschenden kann man nicht garantieren, dass alle die Guidelines lesen. Guidelines sind sinnvoll, aber man darf darüber nicht vergessen, viele solcher Workshops anzubieten. Sie müssen gar nicht viel Aufwand benötigen, aber sie helfen, Dinge in Gang zu bringen.

Maximilian Hörl: Ich empfehle bei den Lösungen, die man umsetzt, pragmatisch zu sein. Wenn man alle Ausnahmefälle abdecken will, die in einer Situation mit

20 oder mehr Forschungsprojekten zwangsläufig auftreten, dann wird man keine Lösung finden, die für alle funktioniert. Für die Archivierung und die Metadaten empfehle ich, eine Lösung zu wählen, die für die meisten Projekte funktioniert, und sich erst mit den Ausnahmen zu beschäftigen, wenn sie auftreten. Die Lösungsstrategien werden sonst übermäßig kompliziert und womöglich treten Sonderfälle, die zwar möglich wären, in der Praxis dann gar nicht auf. Man verliert viel Zeit und Energie, wenn man jedes Problem im Vorfeld lösen möchte. Also: Praktisch denken, Dinge früh anpacken, und dann wird genug Zeit bleiben, um sich um auftretende Probleme zu kümmern.

Das Interview wurde von Elisabeth Mollenhauer und Felix Rau im Dezember 2021 auf Englisch und via Zoom geführt.



SPP 2143 „Entangled Africa: Innerafrikanische Beziehungen zwischen Regenwald und Mittelmeer (ca. 6000 bis 500 Jahre vor heute)“

Seit 2020 unterstützt das DCH das Projekt P11 „Learning through connecting: Datenmanagement als Basis zur nachhaltigen Nutzung von Forschungsdaten“ des SPP 2143 in seinem Forschungsdatenmanagement und insbesondere bei Datentransformation, -kuratierung und -archivierung. Das Projekt ist eine Kooperation des African Archaeology Archive Cologne (AAArC) mit dem IDH. Seit 2020 wird aus Projektmitteln des AAArC eine wissenschaftliche Hilfskraft (WHB) finanziert, die aktiv die Datentransformation und Datenkuratierung unterstützt und am DCH angesiedelt wird (s. Kap. 1.3 und 6.3.2). In der 2022 startenden zweiten Phase des SPP 2143 wird das DCH zusammen mit dem AAArC das Projekt weiterführen und mit einer Mitarbeiter*innenstelle das archäologische Datenmanagement weiter betreiben (s. Ausblick).

DFG-Graduiertenkolleg (GRK) 2661 „anschießen – ausschließen“ – Kulturelle Praktiken jenseits globaler Vernetzung

Das GRK 2661 wird seit Oktober 2021 durch die DFG gefördert und in einem Verbund der Universität zu Köln, der Kunsthochschule für Medien Köln (KHM) und der Technischen Hochschule Köln (TH) durchgeführt. Das DCH stand bereits während der Antragsphase mit Vertreter*innen des Kollegs in Kontakt und hat in diesem Rahmen eine beratende Rolle eingenommen. Neben einer grundsätzlichen Beratungs- und Unterstützungsvereinbarung über die gesamte Projektlaufzeit hinweg bietet das Datenzentrum auch Workshops und Lehrveranstaltungen zu spezifischen Aspekten des Forschungsdatenmanagements für die Wissenschaftler*innen des GRK 2661 an. Zentral ist dabei die Ausrichtung didaktischer Inhalte an den Forschungsfragen, Methoden und Vorhaben der Forschenden (s. Ausblick)

Center for Data and Simulation Science (CDS)

Bei der Kompetenzvermittlung im Bereich Datenanalyse kooperiert das DCH mit dem CDS, das 2021 auch Mitveranstalter der virtuellen Summer School „Deep Learning for Language Analysis“ war. Das CDS hat die FORGE-Konferenz im September 2021 unterstützt (s. Kap. 9.1.3).

Competence Area (CA) III „Quantitative Modeling of Complex Systems“ (QMCS)

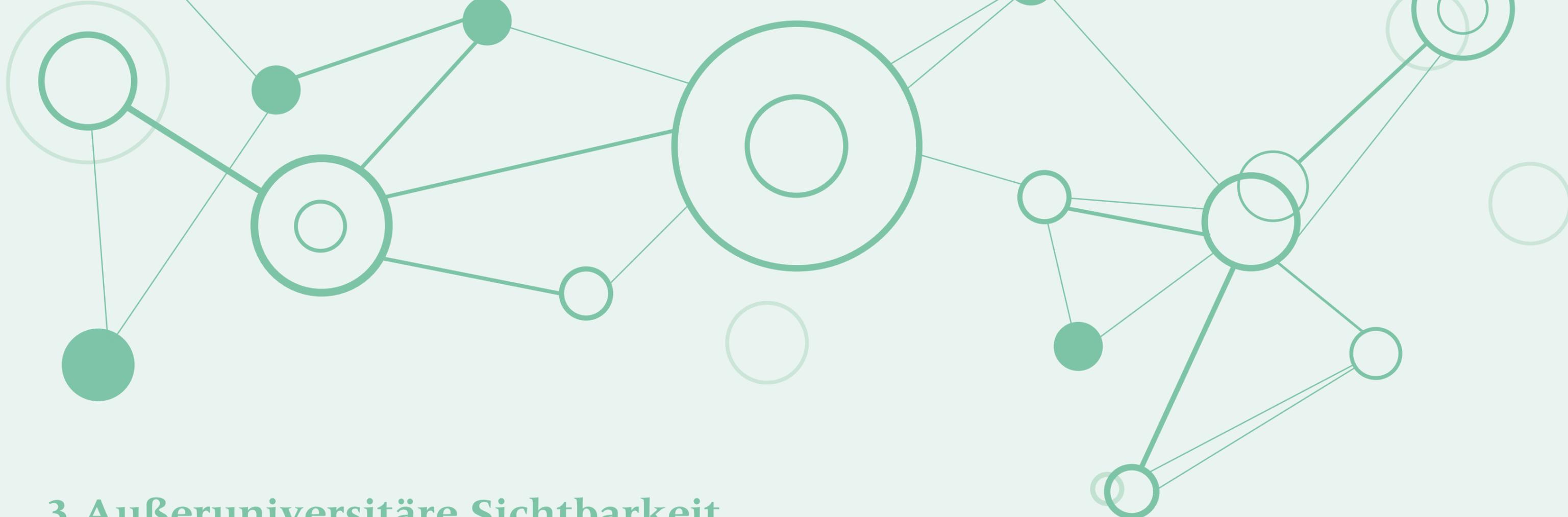
Seit mehreren Jahren kooperiert das DCH bei der Kompetenzvermittlung im Bereich Datenanalyse auch mit der CA III. Die 2021 von IDH, CCeH, CA III, CDS und DCH organisierte virtuelle Summer School „Deep Learning for Language Analysis“ wurde, wie die vorhergehenden Summer Schools, durch die CA III finanziell und organisatorisch gefördert (s. Kap. 9.1.2).

Weitere Akteure

Das DCH hält engen Kontakt zu verschiedenen Projekten und Strukturen an der Philosophischen Fakultät, die datenintensive digitale Forschung betreiben und tauscht sich regelmäßig mit diesen aus:

- [African Archaeology Archive Cologne \(AAArC\)](#)
- [Arachne/Cologne Digital Archaeology Laboratory \(CoDArchLab\)](#)
- [Institut für Linguistik \(IfL\), Abteilung Phonetik](#)
- [Internationales Kolleg „Morphomata: Genese, Dynamik und Medialität kultureller Figurationen“ \(Käte-Hamburger-Kolleg\)](#)





3 Außeruniversitäre Sichtbarkeit, Vernetzung & Kooperationen

Über die Grenzen der Universität hinaus engagieren sich Vertreter*innen des DCH in Gremien, Verbänden und Arbeitsgruppen, kooperieren mit anderen Akteur*innen im geisteswissenschaftlichen Forschungsdatenmanagement und fördern somit die Sichtbarkeit und Vernetzung innerhalb der (Forschungsdaten-)Fachgemeinschaft. Diese außeruniversitäre Zusammenarbeit leistet einen wesentlichen Beitrag zur Etablierung des DCH als Institution in der (deutschsprachigen) Forschungslandschaft und ist Ausgangspunkt für die Entwicklung von Projektanträgen, die gemeinsame Problemstellungen im FDM-Bereich adressieren. Nicht zuletzt bietet die Teilnahme an Fachtagungen und Workshops den Mitarbeiter*innen des DCH die Möglichkeit sich fortzubilden und methodisch auf dem neuesten Stand informationswissenschaftlicher Entwicklungen zu bleiben.

3.1 Geisteswissenschaftliche Konsortien NFDI

Die Universität zu Köln zählt zu den führenden Standorten für digitale Forschung in den Geisteswissenschaften und ist folgerichtig durch verschiedene Wissenschaftler*innen und Institutionen an Konsortialbildungen aller bisherigen geisteswissen-

schaftlichen Konsortien beteiligt. Das DCH hat sich früh an den Formierungsprozessen geisteswissenschaftlicher Konsortien in der NFDI beteiligt, z. B. durch die Teilnahme an der Veranstaltungsreihe „Wissenschaftsgeleitete Forschungsinfrastrukturen für die Geistes- und Kulturwissenschaften in Deutschland“ im Jahr 2018, aus denen die Antragstellungen der Initiativen der Memorandumsgruppe geisteswissenschaftlicher Konsortien (NFDI4Culture, NFDI4Memory, NFDI4Objects, Text+) hervorgegangen sind. Darüber hinaus ist mit Felix Rau ein Vertreter des DCH Mitglied im 2021 gegründeten Verein Geisteswissenschaftliche Forschungsinfrastrukturen e. V. (GWI), in dem sich Mitglieder der verschiedenen geisteswissenschaftlichen NFDIs zusätzlich organisieren und im NFDI e. V. vertreten sind.

Das DCH hat einerseits bereits frühzeitig lokale Vertreter*innen, die an Konsortialbildungen beteiligt waren, unterstützt und andererseits auch selbst eine aktive Rolle bei der jeweiligen Konzeption von NFDI-Anträgen übernommen: Neben dem aktiven Austausch mit Kolleg*innen an der Philosophischen Fakultät, die sich im Rahmen der Antragsstellung der einzelnen NFDI-Konsortien beteiligt haben, stand das Kölner Datenzentrum auch selbst mit Vertreter*innen der NFDI-Konsortien in der Antrags-

phase in Kontakt und hat den Prozess mit individuellen „Letter of Commitment“ unterstützt.

Zu den ab Oktober 2021 geförderten Konsortien gehört neben NFDI4Microbiota, NFDI4Earth, BERD@NFDI, NFDI4DS, DAPHNE4NFDI, FAIRmat, PUNCH4NFDI, NFDI-MatWerk und MaRDI auch das Konsortium für Text- und Sprachdaten Text+, das gemeinsam mit dem 2020 gestarteten Konsortium NFDI4Culture die Geisteswissenschaften innerhalb der NFDI bislang abbildet.

Vertretung: Øyvind Eide, Patrick Helling, Felix Rau, Nils Reiter

3.1.1 Text+

Das Konsortium Text+ (Language- and Text-Based Research Data Infrastructure) wird als eines von zehn Konsortien der zweiten Ausschreibungsrunde zur NFDI auf Empfehlung der Gemeinsamen Wissenschaftskonferenz (GWK) seit dem 1. Oktober 2021 für zunächst fünf Jahre gefördert.

Das Konsortium Text+ legt als Forschungsdateninfrastruktur für Text- und Sprachdaten den Fokus auf drei Datendomänen: Sprach- und textbasierte Sammlungen, lexikalische Ressourcen und Editionen. Mit dieser Ausrichtung adressiert das Konsortium Bedarfe aus verschiedenen geisteswissenschaftlichen Fachdisziplinen, darunter Sprachwissenschaften, Literaturwissenschaften, Philosophie, Klassische Philologie,



Anthropologie, außereuropäische Kulturen und Sprachen sowie sprach- und textbasierte Forschung der Sozial-, Wirtschafts-, Politik- und Geschichtswissenschaften.

Am CCeH an der Universität zu Köln wurde eine zentrale Arbeitsstelle des Konsortiums im Rahmen der langjährigen Kooperation mit der AWK als eine der mitantragstellenden Institutionen des Konsortiums eingerichtet. Die Kölner Arbeitsstelle wird innerhalb des Konsortiums die Datendomäne Editionen koordinieren. Die Universität zu Köln ist darüber hinaus mit dem DCH als Participant beteiligt. Prof. Dr. Dr. h.c. Andreas Speer, Vorstandssprecher des CCeH und Beiratsmitglied des DCH, fungiert als Data Domain Speaker für die Datendomäne Editionen.

Mit den Schwerpunkten Außereuropäische Sprachen und Digitale Lexikografie Außereuropäischer Sprachen sowie dem LAC als dezidiertes Forschungsdatenrepositorium für Sprachdaten tragen das DCH und die Universität zu Köln auch mit zentralen Daten und Diensten in den Datendomänen sprach- und textbasierte Sammlungen und lexikalische Ressourcen zum Konsortium bei. Die Universität zu Köln wird für diesen Bereich von Prof. Dr. Nikolaus P. Himmelmann, Direktor des IfL und Beiratsmitglied des DCH, vertreten.

Applicant Institution:

- Leibniz-Institut für Deutsche Sprache (IDS), Mannheim

Co-Applicant Institutions:

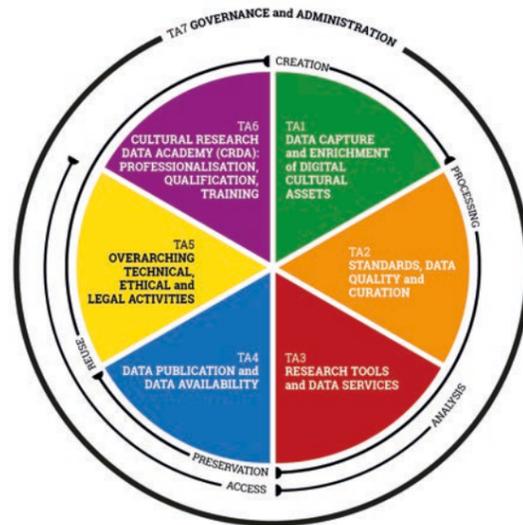
- Berlin-Brandenburgische Akademie der Wissenschaften (BBAW), Berlin
- Deutsche Nationalbibliothek (DNB), Leipzig/Frankfurt a. M.
- Nordrhein-Westfälische Akademie der Wissenschaften und der Künste (AWK), Düsseldorf
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB)

3.1.2 NFDI4Culture



Das Konsortium NFDI4Culture (Consortium for the research data on material and immaterial cultural heritage) wird bereits seit der ersten Ausschreibungsrunde zur NFDI im Oktober 2020 für zunächst fünf Jahre gefördert.

Ziel von NFDI4Culture ist der Aufbau einer bedarfsorientierten Infrastruktur für Forschungsdaten zu materiellen und immateriellen Kulturgütern und die Schaffung einer gemeinsamen Struktur auf nationaler Ebene, die sich um eine nachhaltige Sicherung, Standardisierung und Bereitstellung von Forschungsdaten des kulturellen



Erbes bemüht. Das Konsortium adressiert die Bedarfe eines breiten Spektrums an Fachdisziplinen von der Architektur-, Kunst-, Musik- bis hin zur Theater-, Tanz-, Film- und Medienwissenschaft. Das Konsortium sieht umfangreiche Beteiligungsmöglichkeiten für die Nutzenden der involvierten Fachdisziplinen, aber auch für Kunst- und Kulturschaffende unterschiedlichster Tätigkeitsbereiche und Vertreter*innen der Zivilgesellschaft vor.

Als einer von neun Co-Applicants ist die Universität zu Köln über Dr.'in Lisa Dieckmann, Geschäftsführerin des prometheus-Bildarchivs, vertreten. Daneben sind CCeH und DCH, vertreten von Dr. Claes Neuefeind, und die TWS als Participants am Konsortium beteiligt.

Co-Applicants:

- Akademie der Wissenschaften und der Literatur Mainz
- FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur
- Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB)
- Stiftung Preußischer Kulturbesitz (SPK), Berlin
- Technische Informationsbibliothek (TIB), Hannover
- Ruprecht-Karls-Universität Heidelberg
- Universität zu Köln
- Philipps-Universität Marburg
- Universität Paderborn

Der Schwerpunkt der Konsortialarbeit lag im ersten Jahr vor allem auf dem Aufbau der Community. Neben regelmäßigen Community-Foren der einzelnen Task Areas wurden auch verschiedene Expert*innenforen eingerichtet, u. a. in der Task Area 3 (TA 3) „Research Tools and Data Services“ das Forum „Nachhaltige Softwareentwicklung in NFDI4Culture“, an dem das DCH gemeinsam mit dem CCeH unmittelbar beteiligt ist. Das DCH ist zudem Teil des im Aufbau befindlichen Beratungsnetzwerks des Konsortiums, das durch die in der TA 3 eingerichtete „Consulting agency for sustainable development and operation of research tools and data services in the area of NFDI4Culture“ koordiniert wird.

Seit November 2021 findet ein aktiver Austausch mit Vertreter*innen des NFDI4Culture-Helpdesk über Konzepte zur Entwicklung von Best Practices und Workflows bei FDM-Beratungen statt. Zusätzlich hat das Kölner Datenzentrum mit einem Vortrag über die Integration von Forschungsdatenmanagement in der universitären Lehre zum NFDI4Culture-Workshop „CRDA Forum 2021 – Forschungsdatenmanagement als Teil der Hochschullehre in der Culture Community“ im Juni 2021 beigetragen.

3.1.3 Geisteswissenschaftliche Konsortien in der Antragsphase



NFDI4Memory (Consortium for the Historically-Oriented Humanities)

Das Konsortium NFDI4Memory schließt über die Geschichtswissenschaften hinaus alle historisch arbeitenden Wissenschaften ein.

Die Universität zu Köln ist durch den Lehrstuhl für die Geschichte der Frühen Neuzeit und das Zentrum für LehrerInnenbildung (ZfL) im Konsortium vertreten. Auch das DCH ist, gemeinsam mit dem CCeH, als Participant am Konsortialantrag beteiligt und bringt seine Expertise in den Bereichen Metadatenstandards, Datenkuration, Data Literacy und Beratung in das Konsortium mit ein. NFDI4Memory konnte in der zweiten Antragsrunde zur NFDI nicht berücksichtigt werden. Der Antrag wurde in der dritten Antragsrunde im Oktober 2021 erneut eingereicht und befindet sich aktuell in der Begutachtungsphase.

Applicant Institution:

- IEG – Leibniz-Institut für Europäische Geschichte, Mainz

Co-Applicant Institutions:

- Landesarchiv Baden-Württemberg (LABW), Stuttgart
- Bayerische Staatsbibliothek (BSB), München
- Deutsches Museum von Meisterwerken der Naturwissenschaft und Technik (DM), München
- FIZ Karlsruhe – Leibniz-Institut für Informationsstruktur
- Verband der Historiker und Historikerinnen Deutschlands e. V. (VHD)
- Herder-Institut für historische Ostmitteleuropaforschung (HI), Marburg
- Humboldt-Universität (HU) Berlin
- Martin-Luther-Universität (MLU) Halle-Wittenberg
- Universität Trier (UT)
- Hochschule Darmstadt – University of Applied Sciences (h_da)

Unter den Participants sind für die Universität zu Köln am Antrag beteiligt:

- Prof.'in Dr. Beatrix Busse, ZfL/Prorektorin für Lehre und Studium/Lehrstuhl für Diachrone Englische Sprachwissenschaft, Englisch Seminar I
- Myrle Dziak-Mahler, ZfL
- Prof.'in Dr. Gudrun Gersmann, Lehrstuhl für die Geschichte der Frühen Neuzeit, Historisches Institut
- Prof. Dr. Dr. h.c. Andreas Speer, CCeH
- Prof. Dr. Øyvind Eide, DCH
- Jonathan Blumtritt, CCeH
- Patrick Helling, DCH



NFDI4Objects (Research Data Infrastructure for the Material Remains of Human History)

NFDI4Objects plant ein Konsortium für Wissenschaftler*innen verschiedener Disziplinen, die sich mit materiellem Erbe aus drei Millionen Jahren Menschheits- und Umweltgeschichte befassen.

Dazu gehören Artefakte, Architektur, archäologische Befunde und anthropogen geformte Landschaften sowie biologische oder ökologische Überreste. Über Prof. Dr. Øyvind Eide ist das DCH gemeinsam mit dem IDH, dem CCeH und dem CoDArchLab als Participant am Antrag beteiligt. NFDI4Objects konnte in der zweiten Antragsrunde zur NFDI nicht berücksichtigt werden. Der Antrag wurde in der dritten Antragsrunde im Oktober 2021 erneut eingereicht und befindet sich aktuell in der Begutachtungsphase.

Applicant Institution:

- Deutsches Archäologisches Institut (DAI), Berlin

Co-Applicant Institutions:

- Christian-Albrechts-Universität zu Kiel (CAU)
- Deutsches Archäologisches Institut (DAI), Zentrale wissenschaftliche Dienste, Berlin
- Deutsches Bergbau-Museum Bochum (DBM)
- Generaldirektion Kulturelles Erbe Rheinland-Pfalz (GDKE), Mainz
- Hochschule Mainz – University of Applied Sciences
- Landesamt für Denkmalpflege Baden-Württemberg, Esslingen am Neckar
- Max-Planck-Institut für Wissenschaftsgeschichte (MPIWG), Berlin
- Rheinische Friedrich-Wilhelms-Universität Bonn
- Römisch-Germanisches Zentralmuseum – Leibniz-Forschungsinstitut für Archäologie (RGZM), Mainz
- Römisch-Germanisches Zentralmuseum – Leibniz-Forschungsinstitut für Archäologie (RGZM), Restaurierung und Konservierung, Mainz
- Stiftung Preußischer Kulturbesitz (SPK)
- Gemeinsamer Bibliotheksverbund (GBV), Verbundzentrale (VZG), Göttingen

Unter den Participants sind für die Universität zu Köln am Antrag beteiligt:

- Prof. Dr. Øyvind Eide, IDH/DCH
- Prof.'in Dr. Eleftheria Paliou, Fachgebiet Archäoinformatik, Archäologisches Institut/Direktorin des CoDArchLab
- Prof. Dr. Dr. h.c. Andreas Speer, Vorstandssprecher des CCeH

3.2 Landesinitiative fdm.nrw

Das DCH steht in regelmäßigem Kontakt mit der an der Universität Duisburg-Essen angesiedelten Landesinitiative fdm.nrw der Digitalen Hochschule NRW (DH.NRW). Im Berichtsjahr haben Mitglieder des Datenzentrums regelmäßig an Veranstaltungen der Initiative, u. a. am FDM-Forum, teilgenommen. Mit dem FDM-Forum hat fdm.nrw ein Format aufgesetzt, bei dem Vertreter*innen von FDM-Institutionen und Infrastruktureinrichtungen gemeinsam unterschiedliche Anliegen im Forschungsdatenmanagement sowie Best Practices adressieren und bearbeiten. Im Oktober 2021 war

Patrick Helling mit einem Vortrag zur Entwicklung eines formalen Beschreibungsmodells für FDM-Servicestrukturen und -Bedarfe an dem Format beteiligt (s. Kap. 9.2).

Vertretung: Patrick Helling

3.3 GO FAIR Implementation Network Data Stewardship Competence Centers, German Chapter GO UNITE!

Das DCH ist Gründungsmitglied des im Dezember 2019 an der Universität Leiden offiziell gegründeten GO FAIR DSCC-IN. Ziel des DSCC-IN ist es, vor dem Hintergrund der FAIR-Prinzipien („Findable, Accessible, Interoperable, Reusable“), Kompetenzen und Bedarfe im Bereich Forschungsdatenmanagement an einzelnen FDM-Standorten zu identifizieren und zusammenzutragen. Auf diese Weise soll ein möglichst umfangreiches internationales Expert*innen-Netzwerk aufgebaut werden. Das DSCC-IN bietet dabei Raum für den aktiven Austausch und die direkte Unterstützung zwischen verschiedenen FDM-Standorten und versteht sich insbesondere als Forum zur Bildung und Definition von FDM-Kompetenzstandards.

Das GO FAIR DSCC-IN wird in sogenannten nationalen Chapters organisiert: Diese stellen einerseits die Nähe zu einzelnen FDM-Kompetenzstandorten dar und tragen andererseits nationale Entwicklungen in das gesamte DSCC-IN. Das deutsche DSCC-IN-Chapter GO UNITE! wird von Monika Linne (KonsortSWD) und Patrick Helling koordiniert.

In GO UNITE! wurden insgesamt drei Kernthemen identifiziert, die in Form von Community-getriebenen Arbeitsgruppen adressiert werden:

Vernetzung der Vernetzung

Die deutschsprachige FDM-Landschaft ist insbesondere durch viele unterschiedliche Netzwerke und Initiativen geprägt. Ziel der Arbeitsgruppe ist es, bestehende Netzwerke und Initiativen, die ähnliche FDM-Themenkomplexe bearbeiten, zusammenzubringen und den gemeinsamen Austausch zu befördern.

Entwicklung eines formalen FDM-Beschreibungsmodells

Im Fokus steht die Community-getriebene Evaluation von Möglichkeiten zur Entwicklung eines formalen Beschreibungsmodells für FDM-Servicestrukturen und -Bedarfe. Auf diese Weise sollen eine Konsolidierung von Forschungsdatenmanage-

ment erreicht und Qualitätsstandards definiert werden. Die Basis für diese Entwicklungen stellen bestehende und praktizierte FDM-Servicestrukturen sowie alltägliche FDM-Bedarfe von Forschenden dar.

Internationalisierung

Der Umgang mit Forschungsdaten muss immer auch international gedacht werden. Entsprechend ist es ein grundlegendes Anliegen von GO FAIR und somit auch von GO UNITE!, einen internationalen Austausch über Aspekte des Forschungsdatenmanagements zu befördern.

Im Laufe des Berichtsjahres fanden insgesamt zwei GO UNITE! General Meetings statt. Darüber hinaus wurde im Juli 2021 ein Workshop der Arbeitsgruppe zur Entwicklung eines formalen FDM-Beschreibungsmodells veranstaltet, an dem Patrick Helling als Co-Organisator und mit Impulsvorträgen beteiligt war (s. Kap. 9.1.4 und 9.2).

Vertretung: Patrick Helling

3.4 Research Data Alliance e. V. (RDA DE)

Das DCH ist Gründungsmitglied des Anfang 2018 gegründeten RDA DE. Ziel ist es, einen verstärkten Austausch und eine effizientere Wiederverwendung von Forschungsdaten mittels verbesserter sozialer und technischer Rahmenbedingungen nach einem Bottom-up-Ansatz zu fördern. Das DCH nimmt regelmäßig an der jährlichen RDA DE Tagung und der Mitgliederversammlung teil (s. Kap. 9.2).

Vertretung: Patrick Helling



aus unserem Twitter-Account @dch_cologne

3.5 CLARIN – European Research Infrastructure for Language Resources and Technology

Das DCH ist über das hauseigene Repository LAC in die europäischen CLARIN-Infrastruktur integriert (s. Kap. 4.1): Seit 2018 ist das LAC ein CLARIN-C-Centre und wurde im April 2019 unter Vorbehalt der Zertifizierung mit dem CTS als CLARIN-B-Centre anerkannt. Mit der Ernennung zum B-Centre wird dokumentiert, dass das DCH ein Zentrum in der CLARIN-Infrastruktur ist, das Dienste anbietet und entsprechende technische Standards erfüllt. Der Zertifizierungsprozess des LAC mit dem CTS ist noch nicht abgeschlossen. In einem Verbund mit fünf weiteren europäischen Institutionen betreiben das DCH und das IfL seit 2017 das CLARIN Knowledge-Centre for Linguistic Diversity and Language Documentation (CKLD), das 2018 durch das CLARIN ERIC als CLARIN-K-Centre zertifiziert wurde (s. Kap. 4.3). Das DCH beteiligt sich außerdem an der CLARIN ERIC Component Metadata Infrastructure (CMDI) Task Force.

Vertretung: Felix Rau

3.6 Digital Endangered Languages and Musics Archives Network (DELAMAN)

Seit Juni 2020 ist das DCH mit dem LAC assoziiertes Mitglied des DELAMAN (s. Kap. 4.1). Das 2003 gegründete Netzwerk digitaler Archive für bedrohte Sprachen und Musik ist die wichtigste internationale Organisation für Spracharchive. Das DCH beteiligt sich an der DELAMAN Arbeitsgruppe Open Language Archives Community (OLAC), die sich mit dem Erhalt und der Weiterentwicklung der OLAC-Metadateninfrastruktur beschäftigt.

Vertretung: Felix Rau

3.7 Verband Digital Humanities im deutschsprachigen Raum e. V. (DHd)

3.7.1 DHd-AG Datenzentren

Das DCH ist aktives Mitglied der AG Datenzentren des DHd-Verbands, die insbesondere der Vernetzung und dem Austausch von Informationen und Kompetenzen dient und gleichzeitig einen zentralen Verbund geisteswissenschaftlicher Datenzentren im deutschsprachigen Raum darstellt. Das DCH ist immer wieder an Veröffentlichungen und Stellungnahmen der AG Datenzentren beteiligt und bringt seine Expertise in den Kreis der Arbeitsgruppe ein.

Gemeinsam mit der AG Datenzentren als (Mit-)Veranstalter der FORGE-Konferenz haben das DCH, CCeH und IDH die Konferenz wieder ins Leben gerufen: Neben einer offenen Sitzung für die Mitglieder der AG Datenzentren fand auch ein Workshop „AG Datenzentren: Dokumentation von Forschungsdaten – Erfahrungen und Aufgaben aus der Praxis“ unter der Leitung von Dr.‘in Katrin Moeller, Convenorin der Arbeitsgruppe, statt (s. Kap. 9.1.3).

Vertretung: Patrick Helling, Felix Rau, Jonathan Blumtritt

3.7.2 Data Steward

Mit der Ernennung von Patrick Helling zum Data Steward hat der DHd-Verband auf der Jahreskonferenz 2020 in Paderborn eine zentrale Funktion geschaffen, um eine umfassende Datenstrategie für alle Materialien, Publikationen und Ergebnisse, die im Kontext des DHd-Verbands entstanden sind und entstehen werden, zu entwickeln. Sie sollen im Sinne der FAIR-Prinzipien (1) langfristig gesichert und archiviert als auch (2) nach Möglichkeit nachhaltig publiziert und verfügbar gemacht werden.

Ein zentraler Aspekt der Arbeit des Data Stewards ist die individuelle Publikation aller einzelnen Abstracts der vergangenen und zukünftigen DHd-Jahreskonferenzen, die bisher gesammelt in Book of Abstracts veröffentlicht wurden. Darüber hinaus sollen die Webseiten der vergangenen DHd-Konferenzen in statischer Form auf die nachhaltige Infrastruktur der Alliance of Digital Humanities Organizations (ADHO) migriert werden. Ein weiteres Augenmerk liegt auf der Entwicklung eines nachhaltigen Konzepts zum Umgang mit internen Dokumenten wie Protokollen von Vorstandssitzungen und Materialien aus Mitgliederversammlungen.

Der Data Steward ist kooptiertes Mitglied des DHd-Vorstands und nimmt regelmäßig an Vorstandssitzungen teil.

Vertretung: Patrick Helling

Veröffentlichung 2021: DHd-Konferenz-Abstracts

Mittlerweile konnten insgesamt 917 Abstracts der vergangenen DHd-Jahreskonferenzen einzeln und individuell referenzierbar auf Zenodo publiziert werden. Die Abstracts sind somit nachhaltig verfügbar und konnten u. a. vom Digital Bibliography & Library Project (DBLP) indiziert werden. Die Ausgangsdaten für die Publikation sowie die entwickelten Workflows und Skripte zur automatisierten Publikation via Zenodo wurden über GitHub öffentlich verfügbar und nachnutzbar gemacht. Darüber hinaus konnte mit der Webseite zur DHd-Konferenz 2020 in Paderborn eine erste Konferenz-Webseite in statischer Form in die ADHO-Infrastruktur überführt werden.

GitHub:
<https://github.com/DHd-Verband>

Zenodo:
<https://zenodo.org/communities/dhd/>
<https://zenodo.org/communities/dhd-verbandsaktivitaeten/>

3.8 Deutsche Initiative für Netzwerkinformationen e. V. (DINI)

Das DCH nimmt regelmäßig an DINI/nestor-Veranstaltungen teil und beteiligt sich inhaltlich. Das Kölner Datenzentrum ist auf den Mitgliederversammlungen des Vereins vertreten. Darüber hinaus vertritt Patrick Helling GO UNITE! in regelmäßigen Treffen deutschsprachiger FDM-Initiativen und NFDI-Konsortien, welche von DINI/nestor initiiert und durchgeführt werden.

Vertretung: Patrick Helling

3.9 Sonstige Gremien, Mitgliedschaften & Kooperationen

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Vertretung: Felix Rau, Patrick Helling

Verbund Forschungsdaten Bildung (VerbundFDB)

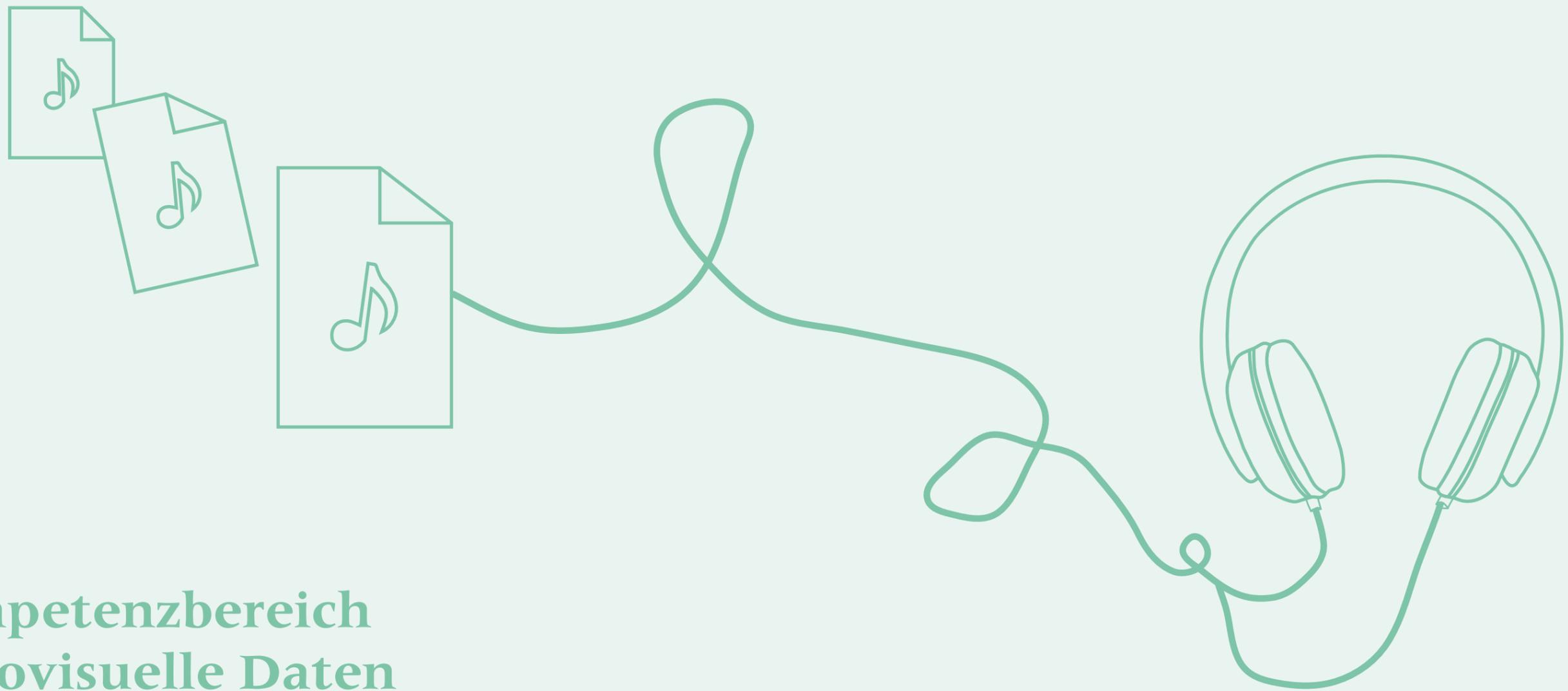
Vertretung: Felix Rau

DINI/nestor-AG Forschungsdaten

Vertretung: Felix Rau, Patrick Helling, Jonathan Blumtritt

RDA Professionalizing Data Stewardship Interest Group (IG)

Mitgliedschaft: Patrick Helling



4 Kompetenzbereich audiovisuelle Daten

Der Kompetenzbereich AV-Daten geht auf die Zusammenarbeit mit dem IfL und dessen Direktor, Prof. Dr. Nikolaus P. Himmelmann, zurück. Im Mittelpunkt stehen der Betrieb und Ausbau einer technischen Infrastruktur zur Archivierung und Publikation von AV-Sprachdaten sowie die Unterstützung von Wissenschaftler*innen durch Beratung, Tools und Standards. In diesem Kontext ist das DCH eng mit CLARIN-D, CLARIN ERIC und DELAMAN vernetzt (s. Kap. 3.5 und 3.6). Im BMBF-Verbundprojekt QUEST erarbeitet das DCH Qualitätskriterien für Metadaten für AV-Daten (s. Kap. 8.2). Im 2021 angelaufenen NFDI-Konsortium Text+ vertritt das DCH den Bereich der AV-Daten in der Datendomäne sprach- und textbasierte Sammlungen (s. Kap. 3.1.1).

Mitarbeiter*innen des Datenzentrums beteiligen sich außerdem regelmäßig an entsprechenden Workshops und Konferenzen, tragen den Themenkomplex der AV-Daten aktiv in die DH-Community und schärfen so stetig das Profil des Kompetenzbereichs (s. Kap. 9.2).

Im Spätsommer 2021 haben IDH, DCH, CCeH, CDS und CA III die virtuelle Summer School „Deep Learning for Language Analysis“ veranstaltet (s. Kap. 9.1.2)

4.1 Language Archive Cologne (LAC)

Das LAC steht als Repository Wissenschaftler*innen aller Disziplinen offen, die mit (annotierten) AV-Sprachdaten arbeiten und ein kompatibles Metadatenmodell nutzen. Ab 2013 wurde das LAC auf der Basis der vom Max-Planck-Institut für Psycholinguistik (MPI-PL) in Nijmegen entwickelten LAT-Software Suite betrieben. 2019 wurde die technische Infrastruktur durch die Entwicklungen aus dem BMBF-Verbundprojekt KA³ abgelöst. Neben Sprachaufnahmen mit linguistischem Fokus werden mittlerweile auch vermehrt Lieder außereuropäischer Kulturen und Aufnahmen oraler Literaturtraditionen in das Repository aufgenommen. Seit Juni 2020 ist das LAC assoziiertes Mitglied im DELAMAN (s. Kap. 3.6). Das Frontend des easy Language Archive Cologne (eLAC) ermöglicht einen an Mobilgeräte und Sprachgemeinschaften angepassten Zugriff auf die Daten des LAC.

4.2 BMBF-Verbundprojekt QUEST

Das 2019 gestartete BMBF-Verbundprojekt QUEST erforscht Qualitätskriterien für AV-Sprachdaten und die begleitenden Metadaten. Das Projekt ist aus den Aktivitäten des CKLD erwachsen, dessen deutsche Mitglieder den Kern des Projekts ausmachen. Das DCH ist im Projekt für die Erarbeitung von Qualitätskriterien für Metadaten verantwortlich und kooperiert hier eng mit dem IDS in Mannheim, dem Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS) in Berlin und den Projektpartnern an der Universität Hamburg (s. Kap. 8.2).

4.3 CLARIN Knowledge-Centre for Linguistic Diversity and Language Documentation (CKLD)

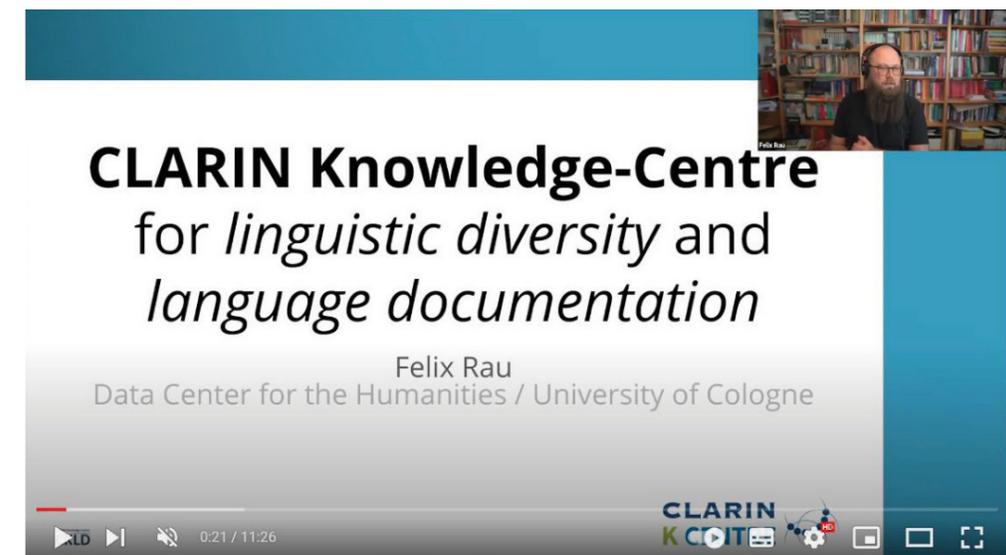
Das CKLD ist ein Zusammenschluss aller europäischen Forschungsdatenzentren, die einen Schwerpunkt in AV-Sprachdaten, Sprachdokumentation und Sprachdiversität haben: Neben den Kölner Institutionen DCH und IfL sind das Hamburger Zentrum für Sprachkorpora (HZSK), das im Akademienprogramm der Union der deutschen Akademien der Wissenschaften angesiedelte Langzeitvorhaben „Grammatiken, Korpora und Sprachtechnologie für indigene nordeurasische Sprachen“ (INEL) in Hamburg, das Endangered Languages Archive (ELAR) und das SOAS World Languages Institute (SWLI) der SOAS University of London sowie das ZAS in Berlin beteiligt. Das CKLD ist als CLARIN-K-Centre zertifiziert (s. Kap. 3.5).

Das CKLD bildet eine zentrale Anlaufstelle für Wissenschaftler*innen, in der die Mitgliedsinstitutionen ihre Expertise gebündelt bereitstellen, Empfehlungen und Informationen zu aktuellen Best Practices, verfügbaren Ressourcen, Werkzeugen und Tutorials rund um die Erhebung, Aufbereitung und Archivierung von Forschungsdaten aus der Sprachdokumentation zur Verfügung stellen und Schulungen und Workshops anbieten, in denen grundlegende und spezialisierte Kompetenzen vermittelt werden. Darüber hinaus unterstützt das Verbundzentrum Wissenschaftler*innen von der frühen Planungsphase bis zur Realisierung von Sprachdokumentationsprojekten oder anderen feldforschungsbasierten Forschungsunternehmungen, berät zu technologischen und datenbezogenen methodologischen Fragen, insbesondere in Bezug auf Ausrüstung, digitale Werkzeuge und Methoden, ist behilflich beim Auffinden von relevanten Datensätzen und vermittelt Zugang zu Expertenwissen hinsichtlich spezifischer Regionen und Sprachfamilien.

Die Kölner Partner betreiben das CKLD-Webportal, das die Angebote des Zentrums an die Öffentlichkeit kommuniziert und über seine Aktivitäten informiert. Über den zentralen Helpdesk können Anfragen an das CKLD gerichtet werden, die intern an die richtige Ansprechperson weitergeleitet werden.

Aus dem CKLD heraus ist das BMBF-Verbundprojekt QUEST entwickelt worden (s. Kap. 8.2).

Im Mai 2021 hat Felix Rau im Rahmen des virtuellen CLARIN Café „One Infrastructure For Many Languages“ das CKLD in einem vorproduzierten Video vorgestellt (s. Kap. 9.2).



Felix Rau im virtuellen CLARIN Café „One Infrastructure For Many Languages“



5 Beratung & Vermittlung

Eine zentrale Säule der Servicestruktur des DCH ist die Beratungstätigkeit. Das Kölner Datenzentrum steht allen Wissenschaftler*innen der Philosophischen Fakultät offen, wenn es um individuelle Fragen rund ums Forschungsdatenmanagement geht.

Mit Hilfe von persönlichen Beratungsgesprächen versuchen die Mitarbeiter*innen des DCH Informationen über das jeweilige Projekt, die entsprechenden Rahmenbedingungen und vorliegenden FDM-Bedarfe zu sammeln, um auf diese Weise passgenau zu beraten, zu vermitteln oder Lösungen zu entwickeln. Zusätzlich führt das DCH auch Gespräche mit Angehörigen anderer Fakultäten und Forschungseinrichtungen, sofern eine große inhaltliche Überschneidung mit den Aktivitäten des DCH und der hier bereits beheimateten Projekte und Bestände besteht, oder ein strategisches Interesse am Ausbau einer Zusammenarbeit vorhanden ist. Grundsätzlich ist es dabei unerheblich, ob es sich im Rahmen von Anfragen um ein großes Verbundprojekt oder eine Promotionsarbeit handelt.



Ein Beratungsvorgang umfasst in der Regel mehrere gemeinsame Sitzungen, eine zeitliche Begrenzung gibt es nicht. Im Allgemeinen unterscheidet das DCH bei den verschiedenen Beratungsvorgängen fünf Kategorien:

1. Allgemeine FDM-Beratungen
2. Antragsberatungen/-beteiligung („ab ovo“)
3. begleitendes Forschungsdatenmanagement („in vitae“), z. B. in laufenden Großprojekten
4. Hilfestellung bei endenden bzw. abgeschlossenen Projekten („post mortem“)
5. Versorgung von „legacy data“, also Beständen, deren Erhebung bereits länger zurück liegt und deren Urheber*innen und Bearbeiter*innen nicht mehr zur Verfügung stehen

Die zeitliche Dimension, d. h. der Projektstatus zum Zeitpunkt der Kontaktaufnahme mit dem DCH, entscheidet dabei wesentlich über die Bandbreite der Empfehlungsoptionen. Eine Kategorisierung und Analyse der Anliegen und Probleme, die im Rahmen der Beratungstätigkeit zu Tage treten, sind Grundlage für ein fundiertes

Verständnis der Gesamtsituation an der Philosophischen Fakultät. Auswertungen, die aus zunehmend schematisierten Abläufen und Protokollen gewonnen werden können, sind der Ausgangspunkt für die Entwicklung von Best Practices. Die Beratungstätigkeit des Datenzentrums ist immer wieder Gegenstand von Vorträgen und Publikationen von Vertreter*innen des DCH und hat auch die Aufmerksamkeit anderer Zentren geweckt, die sich für Erfahrungswerte und Workflows interessieren.

5.1 Beratungen 2021

Auch 2021 wurden Covid-19-bedingt Beratungsanfragen via E-Mail oder individuell vereinbarter Zoom-Sitzungen betreut:

* Projektstatus zum Zeitpunkt der Kontaktaufnahme

** Promotionsanfragen

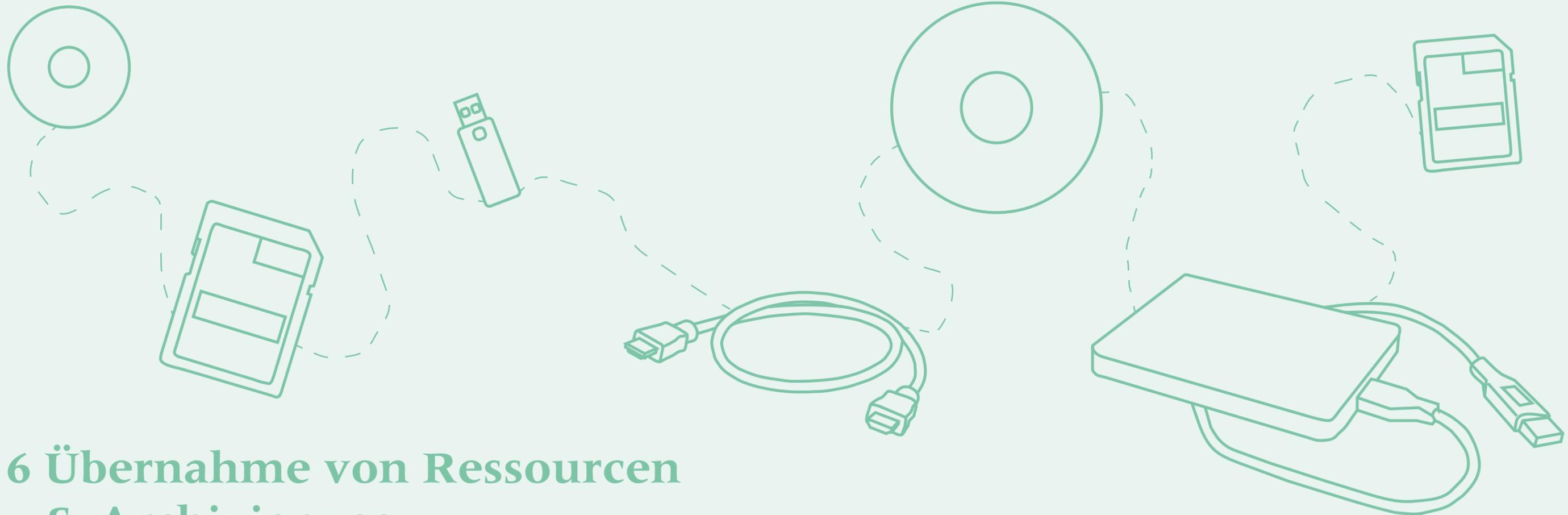
*** Studierendenanfragen

Fachrichtung	Status*	Anfrage
Externe Forschungseinrichtung: Digital Humanities	In vitae	Forschungsmanagement/ Metadaten
Medienkulturwissenschaften	Post mortem	Nachhaltige Verfügbarmachung, Nachnutzung von Datenbanken
Interdisziplinäre Forschungs- einrichtung	Ab ovo	Antragsbeteiligung, Kooperation Korpusaufbau
Geschichtswissenschaften	Ab ovo	Datenaustausch, -speicherung und Archivierung/Zugänglichkeit, Vermittlung

Geschichtswissenschaften	In vitae	Nachhaltige Publikation von wissenschaftlichen Ergebnissen als Webseiten
Ethnologie	Ab ovo	Management und Sicherheit von Forschungsdaten
Medienkulturwissenschaften/ Digital Humanities	Ab ovo	Qualitative Expert*innen- interviews
Linguistik	Ab ovo	Langzeitarchivierung von Forschungsdaten
Geschichtswissenschaften	Ab ovo	Antragsunterstützung, Umgang mit Forschungsdaten
Philosophie	Ab ovo	Antragsunterstützung und Antragsbeteiligung
Altertumswissenschaften	In vitae	Unterstützung bei der Entwicklung/Findung eines Metadatenschemas/-formats zur Beschreibung von digitalisierten Grabungs- und Felsbilddokumen- tationen
Linguistik/Digital Humanities**	Ab ovo	Umgang mit personenbezogenen Daten
Ethnologie	In vitae	Datenmanagement und Datensicherheit

Ethnologie	Ab ovo	Antragsunterstützung
Digital Humanities***	In vitae	Anonymisierung und Datenschutz bei qualitativen Interviewdaten
Anglistik**	In vitae	Datenschutz Interviewtranskripte
Romanistik	In vitae	Archivierung, Publikation und Zugänglichkeit von Sprachdaten
Graduiertenkolleg	In vitae	Fachspezifische Workshops FDM
Ruhr-Universität Bochum: Slavistik	Post mortem	Dauerbetrieb Lebendes System





6 Übernahme von Ressourcen & Archivierung

Digitale Ressourcen können nach Projektende im Rahmen gegebener finanzieller Möglichkeiten durch das DCH weiter betreut werden. Abhängig von der Art der Ressourcen und den gegebenen Bedarfen wurden am Datenzentrum unterschiedliche Modelle der Betreuung entwickelt: Das DCH kann zunächst ein basales Monitoring sicherstellen. Häufig ist jedoch eine grundlegende Überarbeitung oder Kuratierung von Beständen notwendig, was nicht in jedem einzelnen Fall geleistet werden kann. Insgesamt verwaltet das DCH gemeinsam mit dem CCEH und dem RRZK 8 500 GB aktive Daten und 40 000 GB langzeitarchivierte Daten aus geisteswissenschaftlichen Forschungsprojekten. AV-Daten werden zusätzlich durch das LAC kuratiert, publiziert und archiviert (s. Kap. 4.1).

6.1 Überführung von Ressourcen in eine dauerhafte Betreuung

Das DCH übernimmt regelmäßig Ressourcen, die im Rahmen unterschiedlicher Forschungsprojekte entstanden sind und überführt diese in eine dauerhafte

Betreuung. Nicht mehr aktiv gepflegte Webseiten, wie z. B. Konferenz-Webseiten, werden in Form von Programmcode und statischen HTML-Seiten (Hypertext Markup Language) gesichert.

eXist-Applikation aus dem Promotionsprojekt „Eine elektronische Referenzgrammatik von Àbèsàbèsì“

Im Rahmen seines Promotionsprojekts am IfL hat Jonas Lau für eine elektronische Referenzgrammatik von Àbèsàbèsì eine auf eXist basierende Applikation entwickelt. Diese Applikation wurde nach Abschluss des Projekts auf eine vom CCEH betriebene virtuelle Maschine migriert und seitdem gemeinsam von DCH und CCEH gehostet.

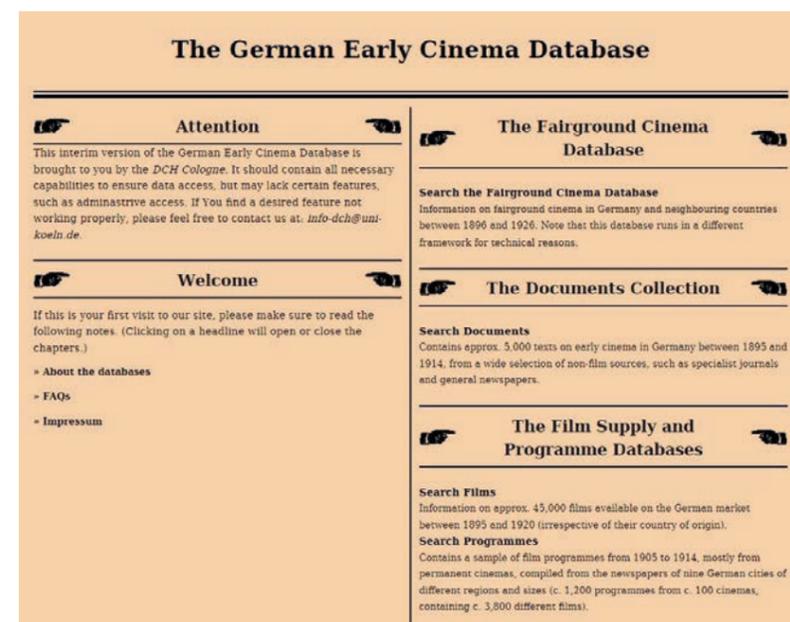
Webseite des Internationalen Kollegs „Morphomata: Genese, Dynamik und Medialität kultureller Figurationen“

Das vom BMBF als Käte-Hamburger-Kolleg im Zeitraum von 2009 bis 2021 geförderte Internationale Kolleg „Morphomata: Genese, Dynamik und Medialität kultureller Figurationen“ hat über die gesamte Laufzeit eine zentrale Webseite auf der Infrastruktur des RRZK als Zugangspunkt zu Entwicklungen und Aktivitäten des Kollegs

gepflegt. Zentral war dabei die Einbindung und Verfügbarmachung von Video- und Audioaufzeichnungen von Veranstaltungen, Vorträgen und Lectures. Zum Ende der Förderlaufzeit wurde eine statische Version der Webseite erzeugt, welche nun in die Betreuung des DCH übergegangen ist.

The German Early Cinema Database

Die durch das DCH betreute Datenbank „The German Early Cinema“ wurde im Berichtszeitraum auf eine neue virtuelle Maschine migriert. Zusätzlich haben mehrere Gespräche mit Kolleg*innen des Open Access Repositoriums media/rep/ und NFDI4Culture bezüglich der Überführung der Datenbank in die Infrastrukturen des NFDI-Konsortiums stattgefunden.



Konferenz-Webseite zur FORGE 2021

Die Konferenz-Webseite wird durch das DCH als Wordpress-Instanz auf dem Web-space des RRZK betrieben (s. Kap. 9.1.3). Für das Jahr 2022 ist es geplant eine statische Version der Webseite zu erzeugen und in dieser Form in eine dauerhafte Betreuung durch das DCH zu überführen.

6.2 Datenpublikation/Daten-Schnittstellen

Die Datenpublikation über gut dokumentierte Programmierschnittstellen, Application Programming Interfaces (API), ist zentraler Bestandteil unserer Datenpublikations- und Nachhaltigkeitsstrategie. Wir stellen hierbei Daten über REST- und GraphQL-Schnittstellen zur Verfügung, die so durch lokale und externe Projekte nachgenutzt werden können. Diese Publikationsform ermöglicht eine flexible Nachnutzung in vielfältigen Kontexten und Strukturen. Einen Schwerpunkt bildet momentan die Publikation von lexikalischen Ressourcen, vor allem von außereuropäischen Sprachen.

C-SALT APIs for Dictionaries

Die Schnittstellen des Cologne South Asian Languages and Texts (C-SALT) für Lexika ermöglichen einen flexiblen und nachnutzungsfreundlichen Zugriff auf einige der wichtigsten lexikalischen Ressourcen, die an der Universität zu Köln gehostet werden.

<https://api.c-salt.uni-koeln.de>

Kosh – APIs for Dictionaries

Kosh ist ein generisches Open-Source-Framework für die Erstellung von API für XML-Datenbanken (Extensible Markup Language). Kosh bietet Zugriff auf lexikalische Ressourcen über REST- und GraphQL-Schnittstellen. Das Angebot wird in den Projekten „VedaWeb“ und „Das zoroastrische Mittelpersische: Digitales Corpus und Wörterbuch“ (MPCD) genutzt. Kosh-basierte Services werden vom DCH auch in die NFDI Text+ eingebracht (s. Kap. 3.1.1).

<https://cch.github.io/kosh>

6.3 Archivierung am DCH

Das DCH sorgt für die Archivierung von größeren Datensammlungen, die in unterschiedlichen Forschungsprojekten an der Universität zu Köln entstanden sind. Für die längerfristige Speicherung von großen Datenmengen stellt das RRZK die seit Jahrzehnten bewährte Technologie der Magnetbandarchivierung zur Verfügung, die sich durch hohe Beständigkeit und vergleichsweise geringe Kosten pro Speichereinheit auszeichnet und daher weltweit häufig als Backup- und Archivsystem eingesetzt wird. Bandarchivierung eignet sich zur Sicherung großer Mengen an Rohdaten (z. B. aus Digitalisierungsprojekten) und zur Gewährleistung der von Projektförderern erwarteten Aufbewahrungsfristen. Die Technologie ist auch Mittel der Wahl bei der Speicherung sensibler Daten, die nicht in öffentlich zugänglichen Systemen exponiert werden und deren Weitergabe nur unter streng kontrollierten Bedingungen

erfolgen soll. Am RRZK wird das Produkt Tivoli Storage Manager (TSM) von IBM betrieben. Die Speicherung erfolgt redundant an zwei verschiedenen Standorten, um die Ausfall- und Katastrophensicherheit zu erhöhen. Für die kurzfristige Speicherung und Verwaltung von Daten, die sich noch in Bearbeitung befinden, werden zusätzlich hochverfügbare Speichermedien des RRZK genutzt, z. B. der Scale-OutFile Service (SoFS).

Das DCH organisiert und operationalisiert die systematische Archivierung von Datenbeständen mit zeitlicher Nähe zur Digitalisierung und Erschließung in laufenden Projekten. Es führt aber auch die Archivierung als Notfallmaßnahme durch, wenn Datenbestände an der Philosophischen Fakultät bekannt werden, für die keine Ansprechpartner*innen an der Universität zu Köln ausfindig gemacht werden können.

Das vom RRZK betriebene und vom mittlerweile abgeschlossenen BMBF-Verbundprojekt KA³ weiterentwickelte LAC steht zusätzlich für AV-Daten zur Verfügung. Das LAC sichert und publiziert AV-Daten sowie dazugehörige Annotationen und detaillierte Metadaten. Diese werden online bereitgestellt und können zur weiteren Nutzung heruntergeladen oder gestreamt werden (s. Kap. 4.1).

6.3.1 Archivierungsprozess

Seit einer 2018 durchgeführten Optimierung und Dokumentation der einzelnen Arbeitsschritte wird der Archivierungsprozess am DCH durch eine Pipeline unterstützt, die sicherstellt, dass die Dateisysteme in das in der Archivierung gängige BagIt-Format überführt, Prüfsummen erstellt und basale Metadaten angelegt werden. Dies ermöglicht es, Mitarbeiter*innen und Hilfskräfte systematischer anzulernen, um dem gestiegenen Archivierungsbedarf zu begegnen. Seit 2019 wurde der Prozess weiter standardisiert und den stetig steigenden Anforderungen entsprechend angepasst. Zentral war dabei die Erstellung von Dokumentationen im internen Wiki des DCH, die die Arbeitsabläufe im Sinne der fortlaufenden Wissensvermittlung und -weitergabe umfassend beschreiben. Sie dienen als Nachschlagewerk und gewährleisten den fortlaufenden Betrieb auch bei wechselndem Personal. Diese Dokumentationen umfassen neben Tipps zu Best Practices ein umfassendes Troubleshooting, eine Schritt-für-Schritt-Anleitung und eine Sammlung von Linux-Bash-Skripten, die Prozesse beschleunigen und vereinfachen.

Veröffentlichung 2021: Belegzettel in der Archivierung

Ab 2020 wurde schrittweise die Erstellung eines automatisierten Dokumentations- und Belegsystems getestet. Ziel ist es, sowohl die hausinterne Dokumentation über die archivierten

Projekte und ihre Eckdaten zu verbessern als auch Kooperationspartner*innen einen Nachweis über den Verbleib ihrer Datensätze zu ermöglichen. Dafür wird im Workflow des Archivierungsprozesses ein vollautomatisch erstellter Beleg, sowohl für die interne als auch externe Dokumentation, generiert. Aufbauend auf dem Belegzettel-Workflow wurde 2021 auch ein Workflow aufgebaut, der es ermöglicht, Datensätze mit einem Digital Object Identifier (DOI) zu verknüpfen. Für die Registrierung einer DOI werden Informationen zum Datensatz auf der DCH-Webseite veröffentlicht und gleichzeitig grundlegende Metadaten über den DOI-Service der USB beim DOI-Registrar DataCite registriert. Dieser angepasste Workflow ermöglicht es, Daten öffentlich nachzuweisen ohne die Daten selbst zu veröffentlichen.

6.3.2 Archivierte Forschungsdatenbestände

Itinera Nova – Redaktions- und Präsentationssystem für die mittelalterlichen Schöffenakten der Stadt Leuven

Leitung: Hadewijch Masure (Stadtarchiv Leuven), Dr. Claes Neufeind (CCeH, Universität zu Köln)

Partner: Stadtarchiv Leuven; Historisch-Kulturwissenschaftliche Informationsverarbeitung (HKI), Universität zu Köln (2009–2012), CCeH/DCH (seit 2016)

Das Projekt des Stadtarchivs Leuven umfasst die Digitalisierung und Verfügbarmachung von ca. 950 000 Seiten aus Schöffenakten aus der Zeit zwischen 1362 und 1795 als Scans und indizierte Volltexte bzw. Transkriptionen. Das DCH hat in Kooperation mit dem CCeH die Datensicherung von Manuskript-Digitalisaten, die von 2009 bis zu seiner Emeritierung 2016 von Prof. Dr. Manfred Thaller betrieben wurde, übernommen und führt die Bandarchivierung des weiter andauernden Digitalisierungsvorhabens fort.

Cologne Online Tamil Lexicon

Leitung: Dr. Thomas Malten (Institut für Südasiens- und Südostasien-Studien (SASOA), Universität zu Köln)

Partner: SASOA; CCeH/DCH

Das Cologne Online Tamil Lexicon ist eine 1999 publizierte digitale Edition des Tamil Lexicon, das zwischen 1924 und 1936 an der University of Madras (Chennai, Tamil Nadu, Indien) entstanden ist. Die Textdatei und der PDF-Scan des Wörterbuchs sind Teil des C-SALT-Webportals der Universität zu Köln.

SFB 1252 „Prominence in Language“

Leitungssprecher: Prof. Dr. Klaus von Heusinger

Partner: Informationsinfrastruktur (INF) des SFB 1252, Universität zu Köln; DCH

Der an der Universität zu Köln angesiedelte interdisziplinäre Sonderforschungsbereich mit 20 Teilprojekten arbeitet an einer Beschreibung und Modellierung von sprachlichen Prominenzrelationen (s. Interview, Kap. 2). Das DCH hat 2020 die Forschungsdaten von 15 Projekten des SFB 1252 archiviert und zusätzlich die Daten des Teilprojekts B05 „Prominenzbedingte Strukturen in austronesischen Sprachen mit symmetrischen Diathesen“ ins LAC übernommen.

SPP 2143 „Entangled Africa: Innerafrikanische Beziehungen zwischen Regenwald und Mittelmeer (ca. 6000 bis 500 Jahre vor heute)“

Koordination: PD Dr. Jörg Linstädter, Deutsches Archäologisches Institut (DAI),

Kommission für Archäologie Außer-europäischer Kulturen (KKAK)

Partner: DAI; Institut für Ur- und Frühgeschichte, Universität zu Köln; IDH; DCH

Das SPP 2143 der DFG ist mit 12 Teilprojekten am DAI in Berlin und an Instituten und weiteren wissenschaftlichen Einrichtung in Berlin, Frankfurt, Hamburg, Kapstadt, Leipzig, Münster, N'Djamena (Chad), Potsdam und Sana'a (Jemen) angesiedelt. Das Teilprojekt P11 „Learning through connecting“ ist eines von zwei an der Universität zu Köln angesiedelten Projekten des SPP 2143 und soll Datenmanagement und Langzeitarchivierung für das gesamte Schwerpunktprogramm sicherstellen. Im Frühjahr 2021 ist das DCH in das Kurations- und Datenmanagementprojekt P11 eingestiegen, um mit den Wissenschaftler*innen des AAARc und des IDH die Projekte des SPP 2143 bei der Veröffentlichung, Dokumentation und Archivierung ihrer Daten in der iDAI-Welt, den Datenportalen und -services des DAI, und anderen archäologischen Repositorien zu unterstützen. In der zweiten Phase des SPP 2143 wird das DCH vom IDH die Rolle des technischen Partners übernehmen und gemeinsam mit dem AAARc im Rahmen des Teilprojekts „FAIR.rdm – FAIRes Management von afrikanischen archäologischen Forschungsdaten“ das Datenmanagement innerhalb des SPP 2143 betreuen.

6.3.3 Datensicherung

Dora Telugu

Die nicht-transkribierten, nicht-annotierten Aufnahmen von Felix Rau dokumentieren die unbekannte dravidische Sprachvariation Dora Telugu. Sie wird im Großraum Koraput im nördlichen Andhra Pradesh und im südlichen Odisha im (zentral-)

östlichen Indien gesprochen. Die Aufnahmen stammen aus einer Forschungsreise von 2009 im Munchinput Mandal in Andhra Pradesh im Rahmen einer ungeplanten explorativen Datenerhebung. Sie enthalten eine Übersetzung in Desia Odia, einer lokalen indoarischen Sprache.

doi:10.18716/dch/a.00000001

Kuvi

Die Datensammlung enthält zwei Aufnahmen von Felix Rau und sind aus einer ungeplanten explorativen Datenerhebung von 2009 mit Kuvi-Sprecher*innen in Albir, einem Dorf in Andhra Pradesh, südöstliches Indien und in Mokaput in Odisha, östliches Indien entstanden. Kuvi gehört zu den dravidische Sprachen. Die Aufnahmen sind nicht transkribiert und nicht annotiert und enthalten eine Übersetzung in Desia Odia, einer indoarischen Sprache. In der Datensammlung liegen außerdem vier Bilddateien mit Notizen von Felix Rau von einer Aufnahme mit einem Kuvi-Sprecher in Karlapodor in Odisha, östliches Indien.

doi:10.18716/dch/a.00000002

Dora Telugu Recordings

Description

This recording documents an unidentified Dravidian variety called "Dora Telugu". This variety is considered a separate language by the speakers and their neighbours. It is spoken in the Greater Koraput region in Northern Andhra Pradesh and Southern Odisha. The recording was made during a field trip in 2009 in the Munchinput mandal of the Visakhapatnam district. The variety is distinct from Konda-Dora (ISO 639-3 kfc, Glottolog kond1295) and is locally considered a variety of Telugu, but distinct from Telugu (ISO 639-3 tel, Glottolog telu1262) and its dialects. The variety documented could be identical or related to Mukha-Dora (ISO 639-3 mmk, Glottolog mukh1238) or Manna-Dora (ISO 639-3 mju, Glottolog mann1245), but its linguistic identity and affiliation is so far undetermined. The recording documents an unplanned and explorative elicitation session. The contact language used in the recording is Desia Odia (ISO 639-3 dso/ort Glottolog adiv1239). The recording is not transcribed or annotated, translations are given in the audio in Desia Odia.

Core team

Felix Rau ([ORCID](#))

Cite as

Rau, Felix 2021. *Dora Telugu Recordings*. Data Center for the Humanities. <https://doi.org/10.18716/dch/a.00000001>

Archive Information

DOI: <https://doi.org/10.18716/dch/a.00000001>
Date of the archival: 2021-04-15 15:35
Size of the package: 462M
Contact Name: Frank Marian Eilers (DCH) - info@dch.uni-koeln.de

Screenshot des Belegzettels über die archivierte Datensammlung „Dora Telugu“ auf der Webseite des DCH.

7 Betrieb von Diensten & Infrastruktur

7.1 Infrastruktur

Das DCH betreibt keine eigene IT-Hardware zur Bereitstellung von Speicher- oder Rechenkapazitäten. Alle Dienste des DCH sowie die Archivierung, der Betrieb und die Pflege der gehosteten Ressourcen werden auf der Infrastruktur des RRZK ausgeführt. Mittelbar oder unmittelbar greift das DCH damit auf eine Vielzahl von Dienstschichten zurück, die vom RRZK zur Verfügung gestellt werden. Dies umfasst die Netzwerkinfrastruktur, Speicherverwaltung, Bandarchivierung, Server-Virtualisierung, Systemadministration und das Webspaces-Hosting. Darüber hinaus wird vom DCH direkt oder indirekt zentral durch das RRZK betriebene Software genutzt. Hierzu zählen u. a. relationale Datenbanken, Build Automation Software und Versionsmanagementsysteme. Bei der Verwaltung der Ressourcen kommen zusätzlich ca. 8 500 GB hochverfügbarer Speicher, 40 000 GB Archivspeicher, 13 virtuelle Maschinen und ca. 20 gehostete Webspaces zum Einsatz.

7.2 Dienste

Das DCH betreibt in eingeschränktem Umfang Software zur Unterstützung der Arbeitsorganisation bzw. um zentrale Anforderungen aus Projekten zu erfüllen:

IIIF- und Bildserver

Seit 2016 betreibt das DCH einen einfachen Bildserver mit Implementierung durch das International Image Interoperability Framework (IIIF). Der Dienst hostet Bildmaterial für die Webdarstellung in verschiedenen betreuten Ressourcen.

XML-Datenbanken

An der Philosophischen Fakultät werden in größerem Maßstab XML-Datenbanken in Projekten aus den Digital Humanities eingesetzt. Gemessen an populären Software-Produkten besetzen XML-Datenbanken eine Nischenfunktion. Hosting-Dienste für XML-Datenbanken sind daher nicht weit verbreitet. Das DCH unterhält Instanzen von eXist-db und BaseX, um diesen Bedarf lokal zu decken.

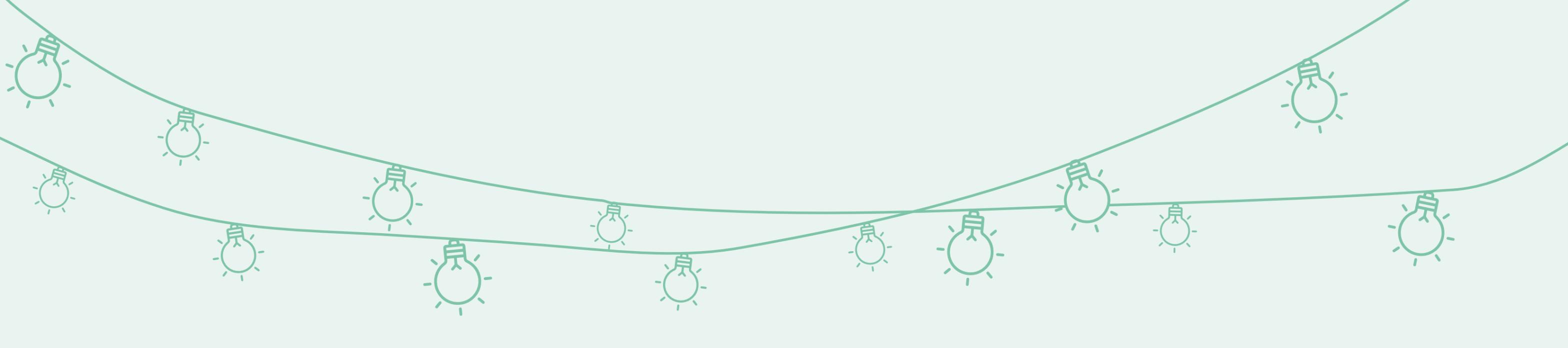
Nextcloud

Seit 2015 betreibt das DCH eine Owncloud-Instanz. 2017 wurde die Migration auf den Nextcloud-Fork vollzogen. Owncloud bzw. Nextcloud lässt sich als selbstverwaltete Alternative zu populären kommerziellen Cloud-Storage-Diensten wie Dropbox, OneDrive oder Google Drive einsetzen. Die Software ist durch Plugins erweiterbar, um zusätzliche Funktionalitäten der täglichen Arbeitsorganisation zu integrieren. Die Nextcloud kommt überwiegend als internes Arbeitsmittel am DCH, CCEH und deren Kooperationsprojekten zum Einsatz. Unter anderem werden hierdurch die Zusammenarbeit und der Datenaustausch mit externen Kooperationspartnern unterstützt. Gleichzeitig beobachtet und evaluiert das DCH laufend andere Dienste, die den Betrieb einer eigenen Instanz funktional gleichwertig ablösen können.

Matomo

Das DCH betreibt seit 2013 eine Matomo-Analytics-Instanz (vormals Piwik). Die Anwendung erfasst Webseitenzugriffe und generiert Benutzungsstatistiken. Matomo ist beliebt als quelloffene und transparentere Alternative zu Google Analytics. Der Dienst wird auch zentral vom RRZK angeboten. Zum Zeitpunkt der Einrichtung am DCH war dies noch nicht der Fall. Das DCH evaluiert, ob eine funktional äquivalente Migration der eigenen Webseitenstatistiken in die zentrale Instanz des RRZK möglich ist.





8 Drittmittelprojekte

Entsprechend der eigenen Satzung führt das DCH Projekte zur Erforschung von Konzepten und Methoden im Forschungsdatenmanagement durch. Ziel ist es dabei u. a., Lösungsstrategien für identifizierte Lücken in der Versorgungslandschaft geisteswissenschaftlichen Forschungsdatenmanagements zu entwickeln und der wissenschaftlichen Community zur Verfügung zu stellen.

8.1 Projektabschluss: DFG-LIS-Projekt SustainLife

 **Partner:** Data Center for the Humanities (DCH), Universität zu Köln; Institut für Architektur von Anwendungssystemen (IAAS), Universität Stuttgart

 **Leitung:** Prof. 'in Dr. Brigitte Mathiak (DCH), Prof. Dr. Dr. h.c. Frank Leymann (IAAS)

 **Team:** Anna Fischer (DCH; ab 2020), Philip Schildkamp (DCH), Dennis Demmer (DCH; ab 2020), Claes Neuefeind (DCH; bis 2019),

Dr. Uwe Breitenbücher (IAAS), Dr.'in Johanna Barzen (IAAS), Lukas Harzenetter (IAAS), Björn Müller (IAAS; ab 2019)

 **Förderung:** DFG (LIS-Programm; Ausschreibung „e-Research-Technologien“); Projektnummer: 379522012

 **Laufzeit:** März 2018 – Februar 2021 (verlängert bis Oktober 2021)

Im Oktober 2021 endete die verlängerte Projektlaufzeit des DFG-LIS-Projekts „SustainLife – Erhalt lebender, digitaler Systeme für die Geisteswissenschaften“. Das vom IAAS an der Universität Stuttgart und vom DCH an der Universität zu Köln kollaborativ durchgeführte Forschungsprojekt war dem Ziel gewidmet, Verfahren und Technologien aus dem Cloud-Deployment auf das Feld der Digital Humanities zu übertragen, um damit Management und Provisionierung von Forschungsanwendungen zu optimieren und deren Sicherung und nachhaltigen Betrieb zu realisieren. Hierfür wurden die standardisierte Modellierungssprache TOSCA (Topology and Orchestration Specification for Cloud Applications) und die Open-Source-Laufzeitumgebung OpenTOSCA eingesetzt und um benötigte Funktionalitäten erweitert (z. B. zur Automatisierung von Software-Aktualisierungen).

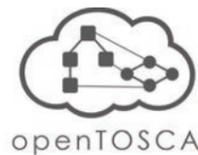
Im Gesamtverlauf des Projekts konzentrierte sich die Stuttgarter Projektgruppe schwerpunktmäßig auf den Ausbau und die Weiterentwicklung von OpenTOSCA, während der Kölner Partner sich der Anbindung des Projekts an die DH-Community sowie der Auswahl und Umsetzung weiterer Use Cases widmete. Indem Modelle und Strategien für Betrieb und Wartung von DH-Software erforscht wurden, leisteten das DCH und das IAAS einen wichtigen Beitrag zur Etablierung eines Nachhaltigkeitsstandards für Forschungsanwendungen in den Geisteswissenschaften.

Auf der FORGE-Konferenz im September 2021 präsentierte die Projektgruppe einen Beitrag zum Thema „Modellierung als Nachhaltigkeitsstrategie für Forschungsanwendungen“. Anschließend war die Projektgruppe auf der zweiten internationalen Konferenz der European Association for Digital Humanities (EADH) mit einer Präsentation zum Thema „Modeling as a Sustainability Strategy for DH Software Applications“ vertreten.

SustainLife-Output: Repositorium aller Use Cases und TOSCA Artefakte

Alle vom DCH modellierten Use Cases sind in einem speziell dafür angelegten GitLab-Repositorium gesammelt worden. Diese Sammlung von TOSCA-kompatiblen Vorlagen für Programmbausteine sowie gesamte Applikationen können ohne Weiteres in das OpenTOSCA-Ökosystem eingebunden und nachgenutzt werden.

<https://gitlab.cceh.uni-koeln.de/sustainlife/tosca-definitions-dch>



SustainLife-Output: Use-Case-Modellierung und Dokumentation

Zusätzlich zu den in TOSCA modellierten Anwendungen und Komponententypen haben wir unsere Erfahrungen während der Modellierung der Use Cases im Use-Case-Repositorium dokumentiert. Dabei wurde nicht nur der technische Aufbau der einzelnen Anwendungen erläutert, sondern auch festgehalten, welche Probleme auftraten und wie diese zu lösen sind. Hierbei zeichnete sich ab, dass sich der Modellierungsaufwand einer einzelnen Anwendung durch die Wiederverwendbarkeit von Komponenten erheblich reduziert werden kann, andererseits die technischen Anforderungen an Modellierende relativ hoch sind.

SustainLife-Output: Eclipse Winery als Selbstbedienungsportal

Um einen niedrighschwelligem und langfristigen Zugang zu den modellierten Use Cases und TOSCA-kompatiblen Vorlagen zu gewährleisten, wird vom DCH eine verstetigte Instanz von Eclipse Winery betrieben. Innerhalb dieser webbasierten Applikation lassen sich alle im Projekt modellierten Komponenten visuell und mit Metadaten versehen betrachten und herunterladen.

<https://sustainery.dch.phil-fak.uni-koeln.de>

8.2 BMBF-Verbundprojekt QUEST

 **Partner:** Hamburger Zentrum für Sprachkorpora (HZSK); Akademieprojekt „Grammatiken, Korpora und Sprachtechnologie für indigene nordeurasische Sprachen“ (INEL), Hamburg; Institut für Deutsche Gebärdensprache und Kommunikation Gehörloser (IDGS), Universität Hamburg; Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Berlin; Archiv für Gesprochenes Deutsch (AGD), Institut für Deutsche Sprache (IDS), Mannheim; Allgemeine Sprachwissenschaft (AWS), Institut für Linguistik (IfL), Universität zu Köln; Data Center for the Humanities (DCH), Universität zu Köln

 **Leitung:** Prof. Dr. Manfred Krifka (ZAS), Prof.'in Dr. Beáta Wagner-Nagy (Institut für Finnougristik/Uralistik, Universität Hamburg), Dr. Thomas Schmidt (IDS), Prof. Dr. Andreas Witt (DCH/IDS)

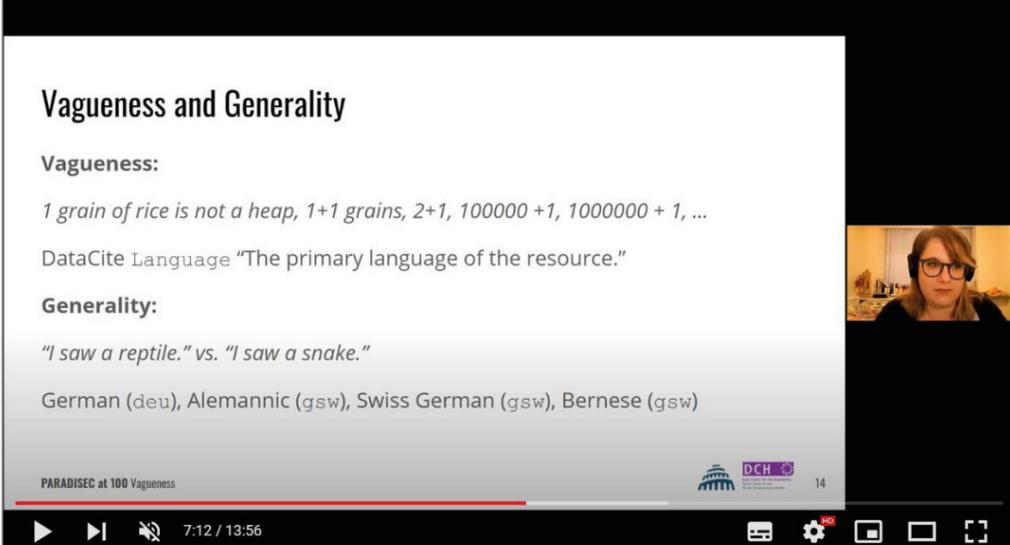
 **Team:** (nur Universität zu Köln): Felix Rau (DCH), Dr.'in Gabriele Schwiertz (IfL), Nicole Majka (DCH)

 **Förderung:** BMBF (Ausschreibung „Förderung von Forschungsvorhaben zur Entwicklung und Erprobung von Kurationskriterien und Qualitätsstandards von Forschungsdaten“); Förderkennzeichen: 16QK09B

 **Laufzeit:** 1. Juni 2019 – 31. Mai 2022

Das BMBF-Verbundprojekt „QUEST: Quality-Established“ erarbeitet seit Juni 2019 gemeinsam mit den fünf in Deutschland ansässigen CKLD-Partnern und dem IDS in Mannheim Qualitätsstandards und Kurationskriterien für annotierte AV-Sprachdaten. Darauf aufbauend sollen mit QUEST Verfahren der Qualitätssicherung für die Erstellung und Kuration solcher Ressourcen und deren Nachnutzungspotentiale entwickelt und erprobt werden. Das Projekt fokussiert dabei auf digitale Forschungsdaten, die in der Sprachdokumentation, Gebärdensprach- und Mehrsprachigkeitsforschung erhoben werden, deren Nachnutzungspotenzial aber weit über diese Disziplinen hinausweist. Das Kölner Team ist mit der Entwicklung von Qualitätsstandards und Kurationskriterien für Metadaten betraut.

Die aktuellsten Ergebnisse präsentierten Nicole Majka und Felix Rau (DCH) sowie Gabriele Schwiertz (IfL) im Februar 2021 virtuell auf der Konferenz des Pacific and Regional Archive for Digital Sources in Endangered Cultures, PARADISEC @100 (s. Kap. 9.2).



Der Vortrag „Vagueness in metadata“ wurde für die PARADISEC @100 als Video vorproduziert (hier mit Nicole Majka im Bild).

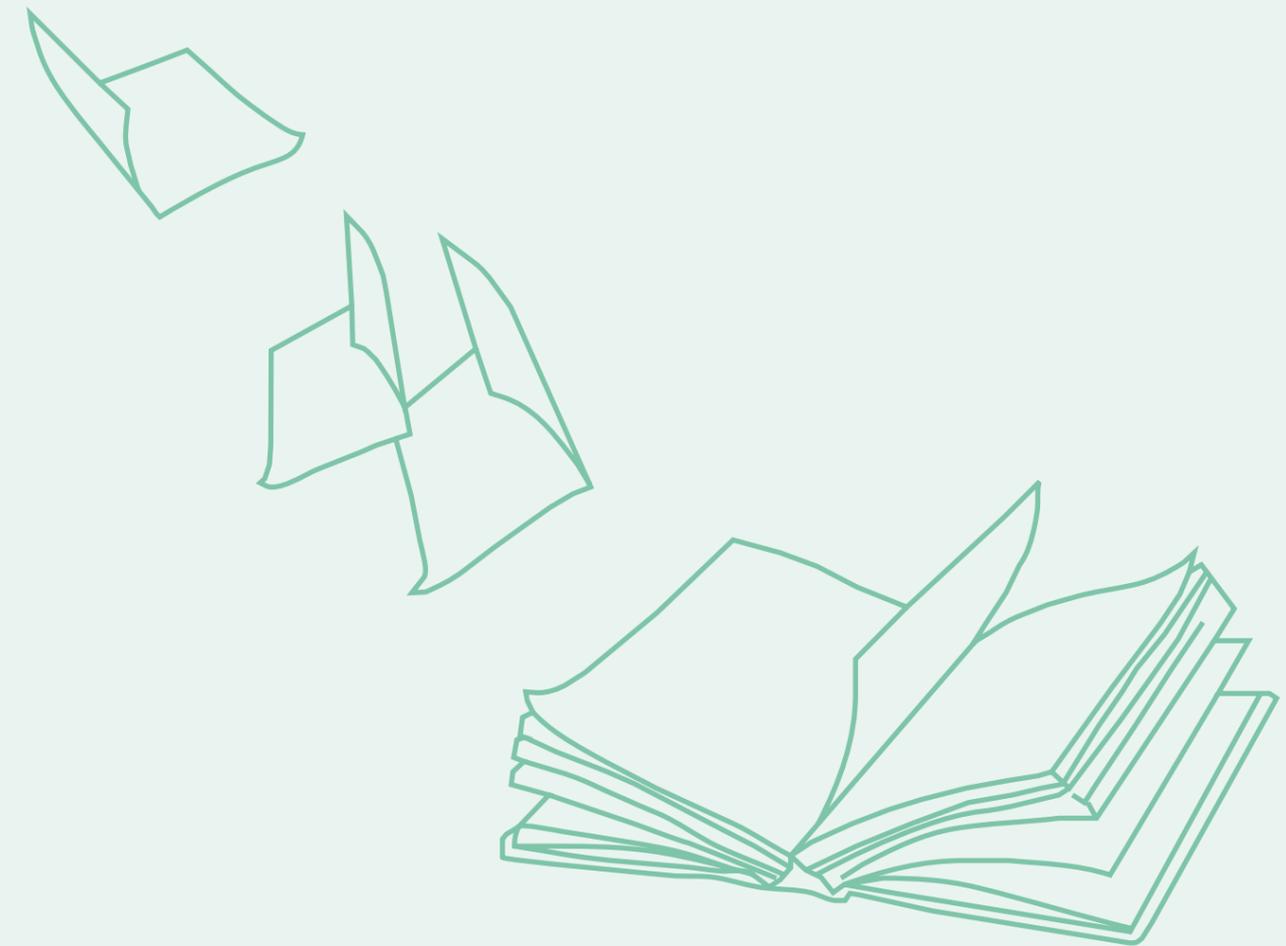
9 Veranstaltungen & Publikationen

9.1 Eigene Veranstaltungen

9.1.1 DHCon 2021

Im Februar 2021 fand das dritte Showcase Event des IDH statt, das Covid-19-bedingt auf den virtuellen Raum ausweichen musste. Organisiert wurde die DHCon 2021, wie im vergangenen Jahr, von Jun.-Prof.'in Dr. Brigitte Mathiak und Elisabeth Mollenhauer. Valentina Hiseni (IDH) war für die Entwicklung der DHCon-Welt auf Gather mit verschiedenen Themenbereichen und Räumen zuständig.

Nach der Begrüßung von Prof. Dr. Øyvind Eide und Jun.-Prof.'in Dr. Brigitte Mathiak in der virtuellen Aula konnten sich die Besucher*innen mit einem Avatar durch die Räume bewegen und die darin platzierten Beiträge anschauen. Wie in jedem Jahr war das Seminar „Computerspiele“ von Brigitte Mathiak aus dem Wintersemester 2020/2021 wieder tragender Bestandteil der DHCon. Die Kursteilnehmer*innen präsentierten ihre entwickelten Computerspiele in diesem Jahr als Let's-Play-Videos.



Vorab hatten die Seminarteilnehmer*innen die Möglichkeit, ihren Raum eigens und auf ihr Spiel angepasst zu gestalten.

Auch aus weiteren Veranstaltungen aus dem Sommersemester 2020 und dem Wintersemester 2020/2021 präsentierten Studierende ihre Kursergebnisse als Poster- und Videobeiträge:

- Seminar „Analyse sozialer Medien mit NLP-Methoden“ (Leitung: Prof. Dr. Nils Reiter)
- Übung „Auditory Spaces: Sound als Gestaltungsmittel in VR und AR“ (Leitung: Jun.-Prof. Dr. Marcus Erbe und Dr. Jan G. Wieners)
- Seminar „Basissysteme Informationsverarbeitung“ (Leitung: Susanne Kurz M.A.)

Ergänzt wurden die studentischen Leistungen durch Präsentationen von Konferenzreichtungen und Projekten der Mitarbeiter*innen an IDH, CCEH und DCH, die die Bereiche Digitale Editionen und Karten, kulturelle Objekte, Virtual Reality (VR) und Augmented Reality (AR), Archäologie, AV-Sprachdaten, Forschungsdatenma-

nagement und DH-Podcasts umfassten.

Die vierte DHCon wird voraussichtlich im März 2022 stattfinden.



aus unserem Twitter-Account @dch_cologne

9.1.2 Virtual Summer School „Deep Learning for Language Analysis“

Vom 30. August bis 3. September 2021 wurde für Studierende und Promovierende der Fachbereiche Linguistik und Digital Humanities vom CDS, IDH, DCH und CCeH erneut eine internationale Summer School veranstaltet. Gefördert wurde sie von der CA III „Quantitative Modeling of Complex Systems“ (QMCS) der Universität zu Köln. Administrative Unterstützung erhielten die Organisatoren Dr. Jürgen Hermes (IDH), Prof. Dr. Nils Reiter (IDH), Dr. Claes Neufeind (CCeH) und Felix Rau (DCH) von Dr. Christian Fischer (CA III) und Dr. Alexander Heinlein (CDS). In diesem Jahr konnte die Summer School auf zwei NVIDIA DGX-1 Server zugreifen, die das CDS für die Universität zu Köln erworben hatte.

Nach einer Begrüßung von Prof. Dr. Øyvind Eide konnten die ca. 50 Teilnehmer*innen wieder aus dem parallel laufenden Programm zwischen den Tracks „Deep Learning with Audio & Speech Data“ unter der Leitung von Paul Wallbott vom Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme in Sankt Augustin (IAIS) und „Text Analysis with Deep Learning“ unter der Leitung von Nils Reiter wählen. Die Übungen wurden von Moritz Eßer (CCeH) und Thomas Oehlke (CCeH) begleitet, die für Rückfragen virtuell zur Verfügung standen. An den letzten beiden Tagen hatten die Teilnehmer*innen, wie in den Vorjahren, die Möglichkeit den Workshop „Building Transformer-Based Natural Language Processing Applications“ zu absolvieren, den das NVIDIA Deep Learning Institute (NVIDIA DLI) zur Verfügung gestellt hatte.

Die Summer School schloss am Freitagnachmittag mit einem Panel mit ehemaligen Absolvent*innen des Master-Studiengangs Informationsverarbeitung: Ekaterina Kruchinina (Cerence), Timothée Schmude (CCeH/molytix GmbH & Co. KG) und Nasrin Saef (frogs42 Gesellschaft für künstliche Intelligenz mbH).

Vgl. hierzu den von Jürgen Hermes verfassten Blog-Beitrag auf <https://experimentales.hypotheses.org/4601>.

9.1.3 FORGE 2021 – Forschungsdaten in den Geisteswissenschaften

Unter dem Konferenztitel „MAPPING THE LANDSCAPE – Geisteswissenschaftliches Forschungsdatenmanagement zwischen lokalen und globalen, generischen und spezifischen Lösungen“ fand vom 8. bis 10. September 2021 die Wiederauflage der FORGE-Konferenz (Forschungsdaten in den Geisteswissenschaften) an der Universität zu Köln statt.

Ausgerichtet wurde die FORGE durch das Data Center for the Humanities (DCH), das Cologne Center for eHumanities (CCeH) und das Institut für Digital Humanities (IDH) der Universität zu Köln mit Prof. Dr. Dr. h.c. Andreas Speer (Vorstandssprecher des CCeH) und Prof. Dr. Øyvind Eide (IDH und DCH) als Vorsitzende des Kölner Organisationskomitees. Die FORGE ist eine Veranstaltungsreihe der AG Datenzentren des Verbands Digital Humanities im deutschsprachigen Raum e. V. (DHD). Zum lokalen Kölner Organisationsteam gehörten Jonathan Blumtritt, Patrick Helling, Elisabeth Mollenhauer, Dr. Claes Neuefeind, Felix Rau, Julia Sorouri, Rebekka Borges, Moritz Eßer, Nicole Majka, Thomas Oehlke und Lukas Sept.

Die virtuelle Konferenz wurde durch die Deutsche Forschungsgemeinschaft (DFG), die AG Datenzentren des DHD-Verbands, das Center for Data and Simulation Science (CDS) an der Universität zu Köln und die Nordrhein-Westfälische Akademie der Wissenschaften und der Künste (AWK) gefördert und unterstützt.

Geschichte der FORGE

Die erste FORGE wurde 2015 vom Projekt „Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit“ (gwin) an der Universität Hamburg veranstaltet und von den Kolleg*innen Hagen Peukert, Stefan Thiemann, Iris Vogel, Kai Wörner und dem gwin organisiert. 2016 fand die FORGE unter dem Konferenztitel „Jenseits der Daten – Nachhaltigkeit für Forschungsanwendungen und Software“ erneut in Hamburg statt.

Die Konferenz übernahm seit Beginn die Funktion eines offenen Forums für den aktiven Austausch über aktuelle Themen und Problemstellungen im geisteswissenschaftlichen Forschungsdatenmanagement (FDM). Durch diese fachliche Ausrichtung ist es der FORGE gelungen, Perspektiven auf das Management von Forschungsdaten um domänenspezifische Bedingungen, Herausforderungen und Lösungsansätze zu komplettieren.

Seit der letzten Ausrichtung der FORGE im Jahr 2016 hat sich die Forschungsdatenlandschaft spürbar weiterentwickelt. Die Anzahl der zugänglichen Angebote auf



Data Center for the Humanities @dch_cologne · 9. Sep. ...
Wir eröffnen die #FORGE2021 mit Grußworten von Prof. Dr. Dr. h.c. Andreas Speer @thomasinstitut @CCeHum, Prof. Dr. @oeide @IdhTweets, Prof.'in Dr. @EvelynGius @DHDInfo und Dr.'in Katrin Moeller und heißen alle herzlich willkommen!

#FDM @UniCologne @Philfak_koeln



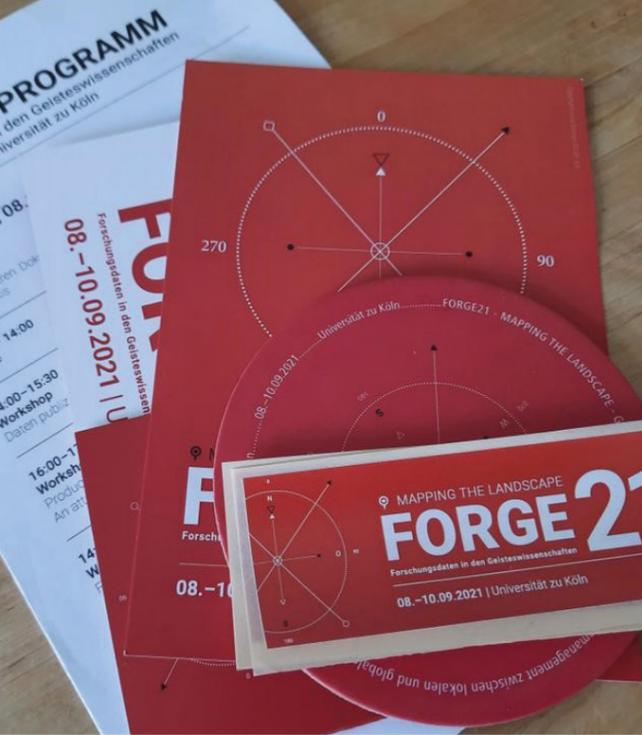
13 24

aus unserem Twitter-Account @dch_cologne

unterschiedlichen Ebenen, lokal und global, nimmt zu. Gleichzeitig ist ein Nebeneinander von Ausdifferenzierung und Kohäsion feststellbar. Entsprechend gab es sowohl in der AG Datenzentren als auch an der Universität zu Köln immer wieder Überlegungen dazu, die Veranstaltungsreihe wieder aufzunehmen. Schlussendlich fiel die Entscheidung hierzu im Rahmen des Workshops „Interoperabilität“ der AG Datenzentren, der im Januar 2020 an der Goethe-Universität Frankfurt stattfand.

Mit der Neuauflage der FORGE sollte schließlich erneut ein Ort der aktiven Verhandlung von generischen und spezifischen Perspektiven auf das Forschungsdatenmanagement angeboten werden, mit dem zur Schärfung des geisteswissenschaftlichen Profils im Umgang mit Forschungsdaten beigetragen werden kann.

Ziel war dabei ausdrücklich die Beförderung einer kritischen Auseinandersetzung mit individuellen Praktiken, Strukturen und Lösungsansätzen. Durch Bezugnahme auf die Gestaltung und Umsetzung von Angeboten, Workflows, Implementierungen, eine strukturbedingte, fachliche Kompetenzverteilung oder eine standortabhängige und/oder wissenschaftspolitische, strategische Ausrichtung einer Einrichtung sollte die eigene Positionierung zu alternativen Strukturideologien provoziert und ausgehandelt werden.



Die FORGE 2021 in Köln diente somit sowohl der Kartierung der geisteswissenschaftlichen FDM-Landschaft als auch der eigenen Verortung in einem lebendigen FDM-Ökosystem.

Virtuelle Durchführung

COVID-19-bedingt wurde die FORGE 2021 für Teilnehmende als rein digitale Veranstaltung via Zoom, Gather und weiteren kollaborativen Tools, die in den Workshops genutzt wurden, umgesetzt.

Zur Gewährleistung eines reibungslosen Ablaufs, einer optimalen Organisation der Konferenz (insbesondere während der Sessions), einer besseren Kommunikation im Organisationsteam und um die Universität zu Köln als Austragungsort sichtbar werden zu lassen, wurde in einem Hörsaal der Universität ein coronakonformer „Situation Room“ nach dem Vorbild der 42. Kölner Mediaevistentagung 2020 (KMT) eingerichtet.

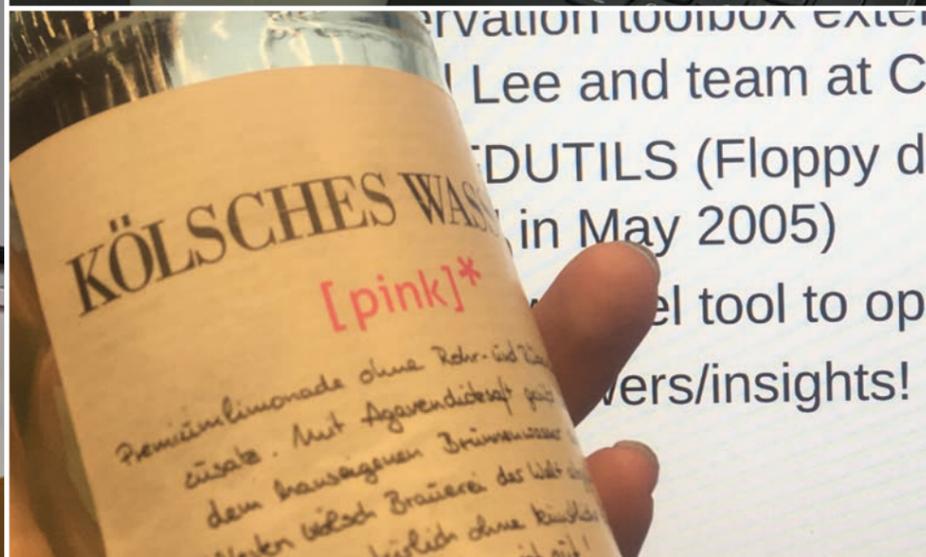
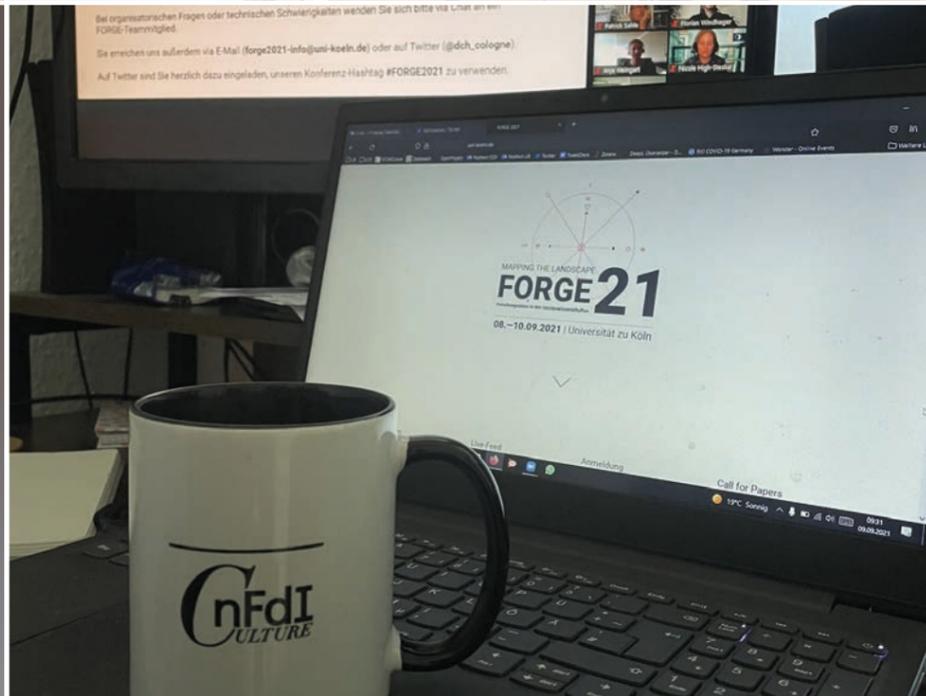
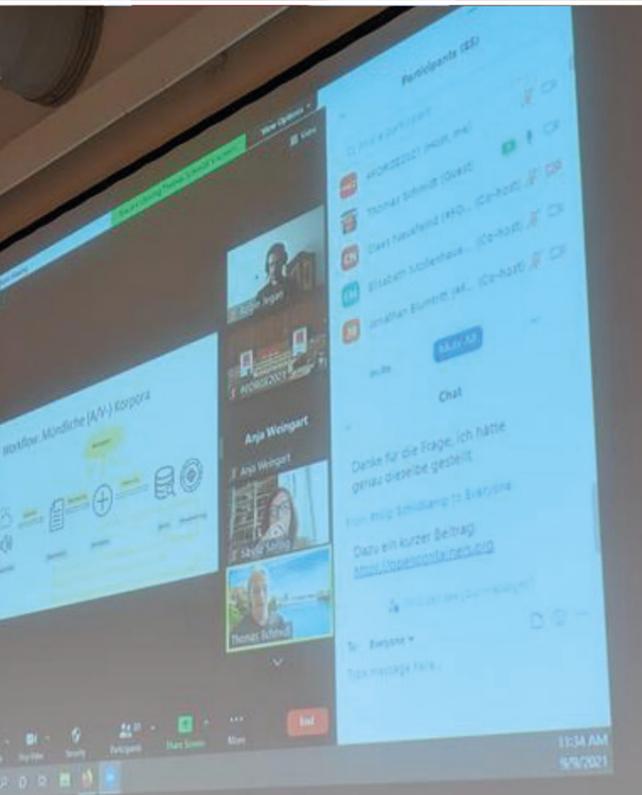
Dazu wurde für den Hörsaal ein Raumkonzept entwickelt, um die einzelnen Funktionen innerhalb des Organisationsteams räumlich zu trennen und damit einen sicheren und sinnvollen Rahmen zu schaffen. So gab es neben einfachen Arbeitsplätzen eine Technikstation und einen „Speaker“-Bereich. Letzterer wurde durch eine qualitativ hochwertige Kamera abgefilmt und direkt in den virtuellen Veranstaltungsraum übertragen. Durch Konferenzposter und das im Raum arbeitende Organisationsteam konnte die Universität zu Köln nach außen wahrgenommen und ein lokaler Bezug zum Veranstaltungsort geschaffen werden.

Während der Konferenz wurden mithilfe der am Institut für Digital Humanities (IDH) entwickelten Web-App autoChirp über den Twitter-Account des DCH @dch_cologne und dem Hashtag #FORGE2021 automatisierte Tweets über die verschiedenen Konferenzbeiträge generiert und so über den Beginn einzelner Sektionen und Veranstaltungen informiert.

Programm der FORGE 2021

Zum Call for Papers wurden insgesamt 40 Beiträge zu unterschiedlichsten Aspekten des geisteswissenschaftlichen Forschungsdatenmanagements eingereicht und in einem „double blind“-Peer-Review-Verfahren durch ein fachlich besetztes Programmkomitee begutachtet. Insgesamt wurden 7 Workshops, 18 Vorträge und 11 Poster auf der Konferenz präsentiert bzw. durchgeführt.

Die Workshops fanden, der Hauptkonferenz vorgelagert, am Mittwoch, dem 8. September 2021 in teilweise parallelen Sessions statt. Nach der offiziellen Eröffnung der Konferenz mit Grußworten der Vorsitzenden des Kölner Organisationskomitees,



Prof. Dr. Dr. h.c. Andreas Speer und Prof. Dr. Øyvind Eide, sowie der zweiten Vorsitzenden des DHD-Verbands, Prof. 'in Dr. Evelyn Gius, und der Convenorin der AG Datenzentren, Dr. 'in Katrin Moeller, wurden am Donnerstag, dem 9. September 2021 neben den ersten drei Vortragssektionen auch eine virtuelle Postersession, eine Sitzung der AG Datenzentren sowie eine Pub-Lecture als Abendveranstaltung durchgeführt. Die Postersession fand auf der digitalen Plattform Gather statt. In der eigens für die Konferenz entwickelten virtuellen Inselwelt mit separaten Räumen für die verschiedenen Poster konnten sich die Teilnehmenden als Avatare fortbewegen und miteinander in Interaktion treten.

Die Pub-Lecture als Abendveranstaltung der Konferenz wurde – mit Bezug zum Konferenzthema – vorab mit der Ankündigung „Bring (and show) your own local drink!“ auf Twitter beworben. Als Vortragender konnte Dr. Dirk von Suchodoletz von der Universität Freiburg gewonnen werden, der unter dem Titel „Unwiederbringliche Daten: Von kontextfreien Magnetströmen bis zu verschlüsselten Flash-Speichern“ über alte Datenträger und Datenrettungsversuche berichtete. Währenddessen twitterten die Teilnehmenden der Pub-Lecture unter dem Hashtag #FDMSocial Fotos ihrer lokalen Biere und Limonaden. Nach dem Vortrag bestand die Möglichkeit, den Abend in einer der virtuellen Bars auf Gather gemeinsam ausklingen zu lassen.

Am Freitag, dem 10. September 2021 fanden schließlich die letzten drei Vortragssektionen statt. Zusätzlich gab es eine Podiumsdiskussion zum Konferenzthema mit Beiträgen von

- **Arianna Ciula**, PhD (stellvertretende Direktorin des King's Digital Lab am King's College London),
- **Alexis Michaud**, PhD (Centre National de la Recherche Scientifique, Direktor des Langues et Civilisations à Tradition Orale),
- **Alexander Czmiel** (Berlin-Brandenburgische Akademie der Wissenschaften, TELOTA Digital Humanities),
- **Valentina Pasquale**, PhD (Istituto Italiano di Tecnologia Genova, CO-Chair des GO FAIR Data Stewardship Competence Centers Implementation Network), und
- **Prof. Dr. Georg Vogeler** (Leiter des Zentrums für Informationsmodellierung – Austrian Centre for Digital Humanities).

Das Programm sowie die Zusammenfassungen zu den Beiträgen sind online verfügbar. Alle Abstracts wurden in einem gemeinsamen Book of Abstracts sowie als einzeln referenzierbare Publikationen in der Zenodo-Community zur FORGE veröffentlicht. Die Pub-Lecture und die Podiumsdiskussion wurden aufgezeichnet und im Anschluss an die Konferenz auf YouTube veröffentlicht.

Auswertungen der Konferenz

Insgesamt wurden 40 Abstracts im Rahmen des Call for Papers eingereicht, von denen letztlich 6 Workshops, 18 Vorträge und 11 Poster als Beiträge auf der Konferenz präsentiert bzw. durchgeführt wurden (Abb. 1). Ein Workshop der AG Datenzentren wurde außerhalb des Review-Verfahrens angeboten. Gegenüber den Workshops (19,4%) und Postern (30,6%) machten die Vorträge die Hälfte des Konferenzprogramms (50,0%) aus (Abb. 2).

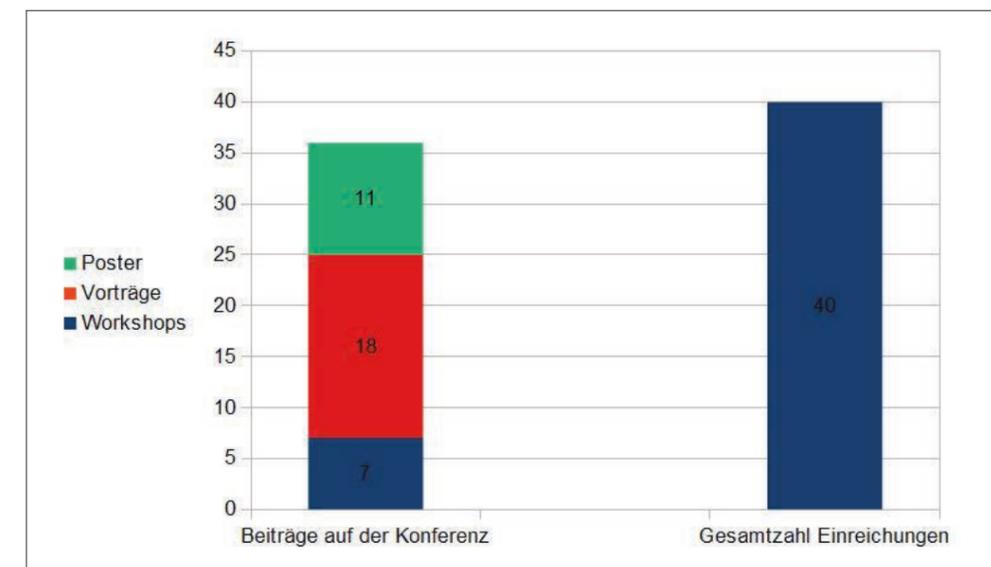


Abb. 1: Einreichungen und Beiträge auf der Konferenz.

Die überwiegende Mehrheit der präsentierten Beiträge kam aus Deutschland (32), außerdem kamen drei Beiträge aus Österreich und einer aus der Schweiz und Frankreich (Abb. 3). Die Mehrheit der Beitragenden waren Teil eines Teams, lediglich 8 reichten ihre Beiträge alleine ein (Abb. 4).

Insgesamt haben sich 305 Personen zur Konferenz angemeldet. Die überwiegende Mehrheit der Teilnehmenden kam aus Deutschland (282). Aus der Schweiz gab es 10 Anmeldungen, aus Österreich 6, aus Italien und Großbritannien je 2 und aus Frankreich, Peru und den USA je eine Anmeldung.

Der Anteil der promovierten und nicht-promovierten Teilnehmenden lag bei 43% (131) zu 57% (174). Unter den promovierten Teilnehmenden belief sich der Frau-

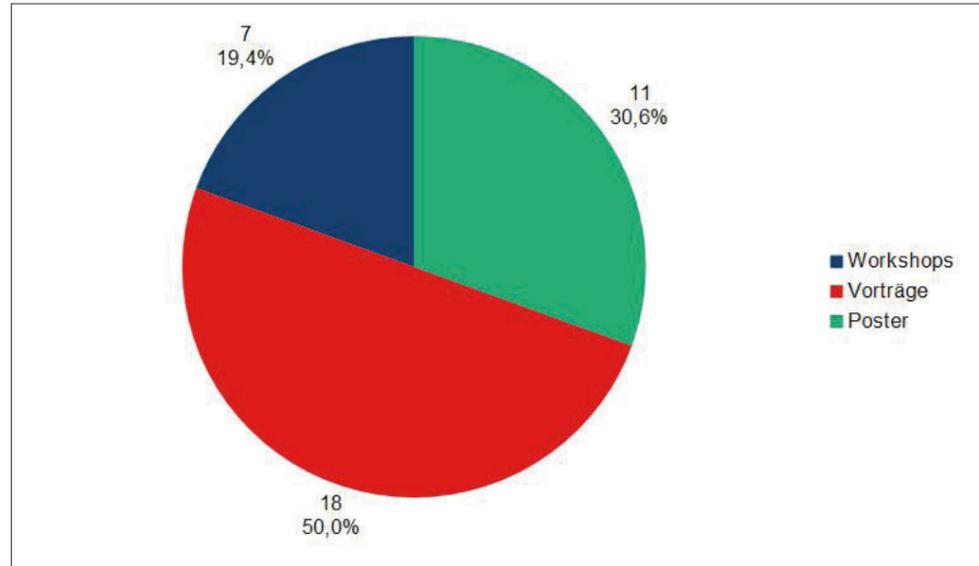


Abb. 2: Anteile der Beitragsformate auf der Konferenz.

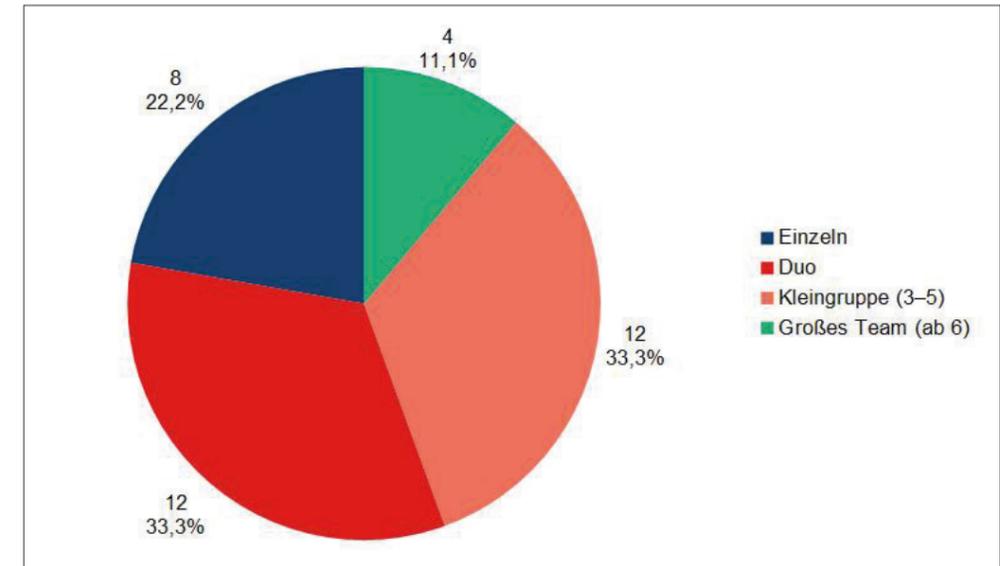


Abb. 4: Anteile der Beitragenden-Konstellationen.

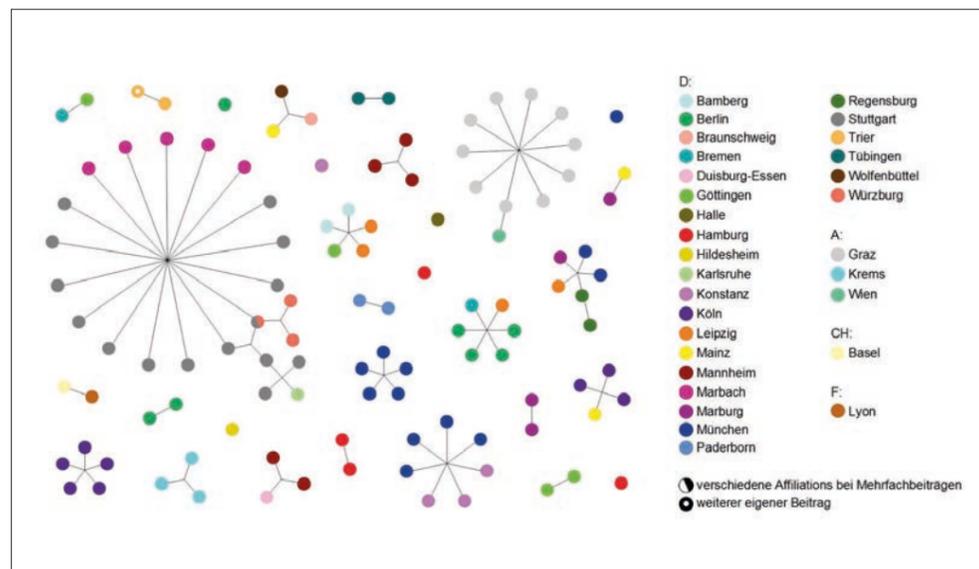


Abb. 3: Netzwerke der Beitragenden (Design: Elisabeth Mollenhauer).

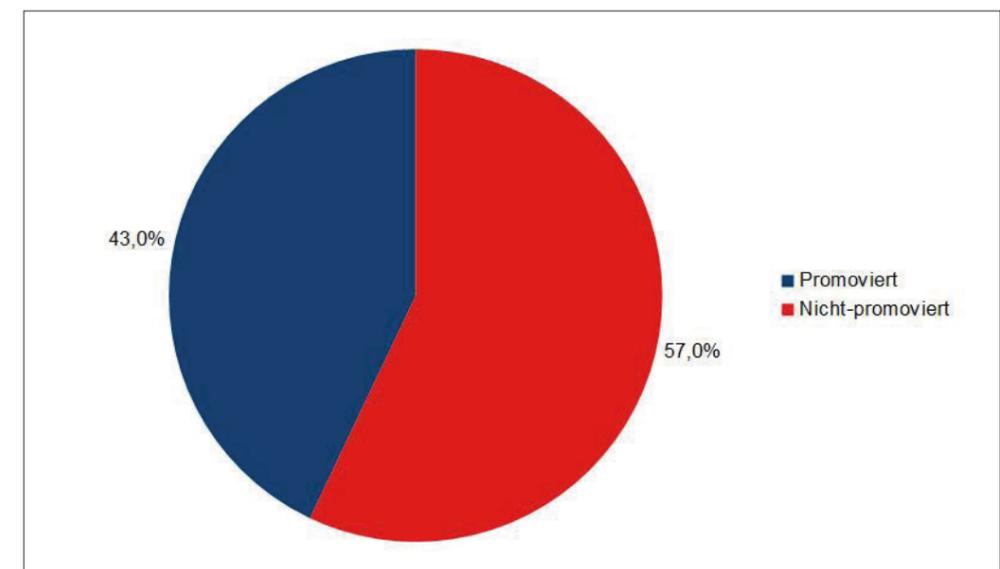


Abb. 5: Anteile der promovierten und nicht-promovierten Teilnehmenden.

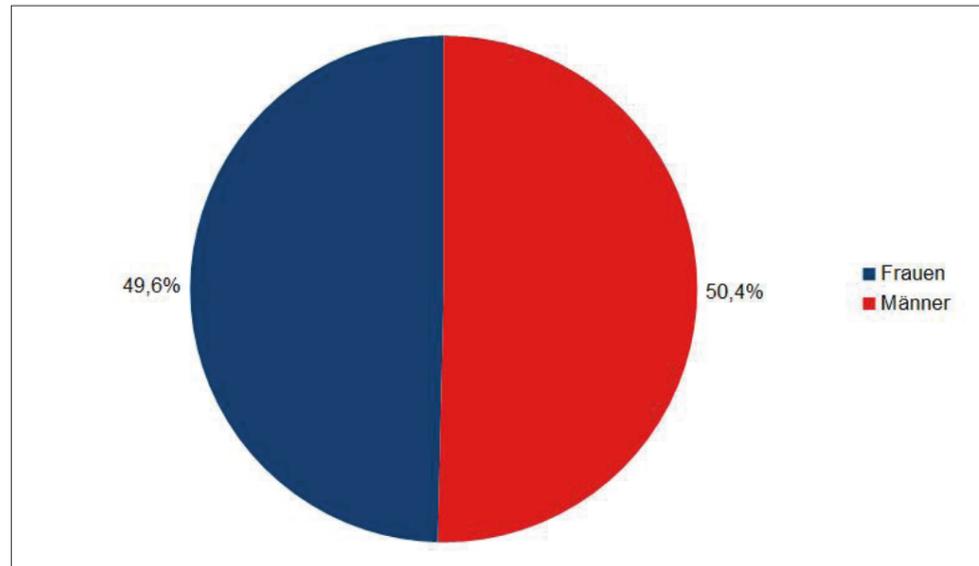


Abb. 6: Anteile der promovierten Teilnehmenden.

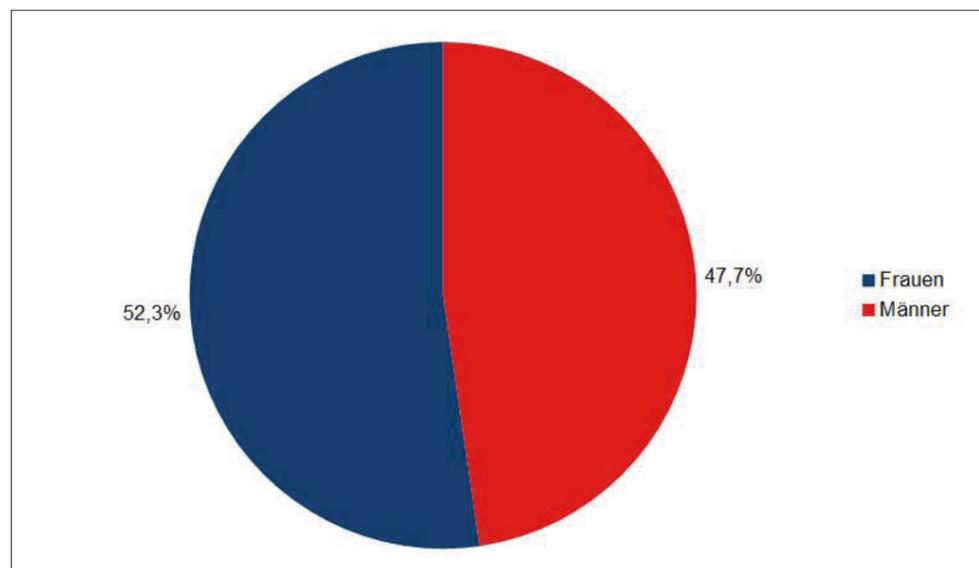


Abb. 7: Anteile der nicht-promovierten Teilnehmenden.

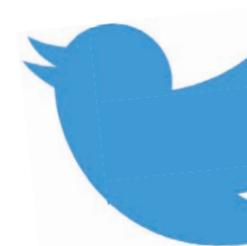
enanteil auf 49,6% (65) und der Männeranteil auf 50,4% (66). Unter den nicht-promovierten Teilnehmenden betrug der Frauenanteil 52,3% (91) zu 47,7% (83) Männeranteil (Abb. 5–7).

Ausblick

Die FORGE ist als regelmäßige Veranstaltung im Kontext der AG Datenzentren des DHD-Verbands geplant und soll alle zwei Jahre an wechselnden Standorten in Deutschland, Österreich und der Schweiz stattfinden.

Auf dem DHD-Blog erschienen am 25. November 2021 unter dem Titel „Wiederauflage der Konferenz FORGE – Forschungsdaten in den Geisteswissenschaften. Nachbericht zur FORGE 2021 an der Universität zu Köln (8. bis 10. September 2021)“.

*Autor*innen: Jonathan Blumtritt, Rebekka Borges, Moritz Eber, Patrick Helling, Nicole Majka, Elisabeth Mollenhauer, Claes Neufeind, Thomas Oehlke, Felix Rau, Lukas Sept, Julia Sorouri.*



aus unserem Twitter-Account @dch_cologne

9.1.4 GO UNITE! Workshop der AG FDM-Beschreibungsmodell

Unter dem Titel „Was ist was? – Entwicklung eines formalen Beschreibungsmodells für Service- und Bedarfsstrukturen im Forschungsdatenmanagement“ wurde im Juli 2021 durch die GO UNITE! AG FDM-Beschreibungsmodell ein interaktiver Workshop mit über 30 Vertreter*innen unterschiedlichster Infrastruktureinrichtungen durchgeführt (s. Kap. 3.3). Ziel des Workshops waren die Diskussion und der Austausch über Möglichkeiten, Chancen und Hürden bei der Entwicklung eines formalen Beschreibungsmodells für FDM-Servicestrukturen. Zu diesem Zweck wurden im Vorfeld der Veranstaltung individuelle Beschreibungen von Servicestrukturen der Institutionen der Teilnehmenden gesammelt. Im Rahmen von mehreren Breakout-Sessions wurde schließlich versucht gemeinsam zu eruieren, welche Möglichkeiten es gibt, diese Beschreibungen abstrakt zu formalisieren, um sich auf diese Weise Best Practices und Qualitätsstandards für FDM-Servicestrukturen zu nähern.

Im Rahmen des Workshops konnte ein grundlegender Bedarf für die standardisierte Beschreibung von FDM-Services im Rahmen des Workshops identifiziert werden. Insgesamt wurden drei Perspektiven auf das Vorhaben definiert, welchen sich in drei unterschiedlichen Folgeworkshops fokussiert gewidmet werden soll:

1. Analyse und Modellierung von Bedarfskategorien/-strukturen auf Seiten der Forschenden
2. Mapping vorhandener FDM-Beschreibungsmodelle
3. Entwicklung von Best Practices zur Durchführung und Dokumentation von FDM-Beratungsgesprächen und -vorgängen

Die Folgeworkshops sind für die erste Jahreshälfte 2022 geplant.

9.2 Vorträge, Präsentationen & Teilnahmen

*Vortragende*r bei mehreren Co-Autor*innen

*kursiv: externe Co-Autor*innen*

virtuell: 

DAI-Workshop „Rekontextualisierung und Aufbereitung archäologischer Forschung mit iDAI.welt Systemen“ (1. Arbeitswerkstatt): Teil 2

13.01.2021. Teilnahme: Felix Rau.

IT-Zertifikat der Philosophischen Fakultät „Digital Objects“ (Leitung: Susanne Kurz M.A.)

13.01.2021. „Forschungsdatenmanagement in den Geisteswissenschaften – Digitales Publizieren“. Gastvortrag: Patrick Helling.

Daten & Datteln...Live! Vorträge für Mitglieder der Philosophischen Fakultät 14.01.2021.

„Digitales Publizieren mit persistenten Identifiern“. Vortrag: Felix Rau.

NFDI4Culture „Culture Community Plenary (CCP)“

04.–05.02.2021. Teilnahme: Felix Rau, Patrick Helling.

DAI-Workshop „Rekontextualisierung und Aufbereitung archäologischer Forschung mit iDAI.welt Systemen“ (1. Arbeitswerkstatt): Teil 3

10.02.2021. Teilnahme: Felix Rau.

GO UNI Strategy Meeting

11.02.2021. Organisation: *Monika Linne*, Patrick Helling.

PARADISEC @100, Sydney Conservatorium of Music

17.–19.02.2021. „Vagueness in Metadata“. Vortrag: *Nicole Majka, Gabriele Schwiertz*, Felix Rau. doi:10.5281/zenodo.4506936.

RDA DE e. V. Tagung und Mitgliederversammlung

23.–26.02.2021. Teilnahme: Patrick Helling, Felix Rau (nur Tagung).

Winterschool „Digitale Editionen – Masterclass“ der Bergischen Universität Wuppertal, des GRK 2196 „Dokument – Text – Edition“ und des Instituts für Dokumentologie und Editorik e. V. (IDE)

26.–28.02.2021. Teilnahme: Elisabeth Mollenhauer.

Ada Lovelace Institute/AI Council Launch „Exploring legal mechanisms for data stewardship“

04.03.2021. Teilnahme: Patrick Helling .

QURATOR 2021 (Conference on Digital Curation Technologies) „Exploring the Future of Digital Communication“, Scientific Workshop – Part 2

08.–12.03.2021. Teilnahme: Felix Rau.

TextGrid e. V. Mitgliederversammlung

09.03.2021. Teilnahme: Felix Rau.

CLARIAH-DE Vollversammlung

09.03.2021. Teilnahme: Felix Rau.

FDM-Workshop der Akademien

11.03.2021. Teilnahme: Felix Rau, Patrick Helling.

EU Open Data Explained Training „Data Catalog Vocabulary (DCAT) and DCAT-AP“, Webinar 2 – Basic User

22.03.2021. Teilnahme: Felix Rau.

vDHd 2021 „Experimente“

26.03.2021. „Die DHd-Abstracts im Zukunftslabor“. Workshop: *Peter Andorfer, Anna Busch, Fabian Cremer, Andreas Henrich, Patrick Helling, Harald Lordick, Dennis Mischke, Timo Steyer, Scott B. Weinhart.*

FAIRsFAIR Workshop „National policy and support actions for research data skills – impact and experiences“

28.04.2021. „GO UNITE German Chapter of the Data Stewardship Competence Centers Implementation Network (DSCC-IN)“. Vortrag: Patrick Helling.

Launch des Anonymisierungstool „QualiAnon“ – Qualiservice

18.05.2021. Teilnahme: Patrick Helling.

Fakultätssitzung der Philosophischen Fakultät

19.05.2021. Vorstellung des DCH und C³RDM. Vortrag: Patrick Helling, *Jasmin Schenk, Florian Willems.* Teilnahme: Felix Rau.

NFDI4Memory Konsortialtreffen

27.05.2021. Teilnahme: Patrick Helling.

Treffen der DHd-AG Datenzentren

28.05.2021. Teilnahme: Patrick Helling, Felix Rau.

CLARIN Café „One Infrastructure For Many Languages“

31.05.2021. Teilnahme und Vorstellung des CKLD: Felix Rau.

forschungsdaten.info live: Datenschutz

02.06.2021. Teilnahme: Patrick Helling.

NFDI4Culture-Workshop „Forschungsdatenmanagement als Teil der Hochschullehre in der Culture Community“

17.–18.06.2021. „Forschungsdatenmanagement in der universitären Lehre – Der Kölner Ansatz“. Vortrag: Patrick Helling.

FAIR Festival 2021

21.–23.06.2021. Teilnahme: Patrick Helling.

IT-Zertifikat der Philosophischen Fakultät „Digital Objects“ (Leitung: Susanne Kurz M.A.)

01.07.2021. „Forschungsdatenmanagement in den Geisteswissenschaften – Digitales Publizieren“. Gastvortrag: Patrick Helling.

GO FAIR Data Stewardship Competence Centers Implementation Network (DSCC-IN) Meeting

06.07.2021. Teilnahme: *Monika Linne,* Patrick Helling.

GO UNITE! Workshop „Was ist was? – Entwicklung eines formalen Beschreibungsmodells für Service- und Bedarfsstrukturen im Forschungsdatenmanagement“

07.07.2021. Organisation: Patrick Helling, *Marina Lemaire, Stefan Kellendonk, Katrin Moeller, Monika Linne.* Impulsvorträge: *Jens Dierkes, Andreas Mühlichen,* Patrick Helling, *Anna Lehmann, Marina Lemaire.*

ORCID DE Online-Seminar „Research Organization Registry“ (ROR)

18.08.2021. Teilnahme: Patrick Helling.

FORGE 2021 – Forschungsdaten in den Geisteswissenschaften „MAPPING THE LANDSCAPE – Geisteswissenschaftliches Forschungsdatenmanagement zwischen lokalen und globalen, generischen und spezifischen Lösungen“ 

08.–10.09.2021.

Organisationskomitee: Øyvind Eide, Andreas Speer, Jonathan Blumtritt, Patrick Helling, Elisabeth Mollenhauer, Claes Neufeind, Felix Rau, Julia Sorouri, Rebekka Borges, Moritz Eßer, Nicole Majka, Thomas Oehlke, Lukas Sept.

„Modellierung als Nachhaltigkeitsstrategie für Forschungsanwendungen“.

Vortrag: Anna Fischer*, Philip Schildkamp*, Dennis Demmer*, Claes Neufeind, Brigitte Mathiak. doi:10.5281/zenodo.5379654.

DINI/nestor Jour Fixe der FDM-Initiativen 

15.09.2021. Teilnahme: Patrick Helling.

DH Showreel des IDH im Rahmen der vDHd 2021 „Experimente“ 

16.09.2021. Vorstellung des DCH und der „Daten & Datteln“-Sammelkarten: Patrick Helling, Felix Rau.

Second International Conference of the European Association for Digital Humanities – Interdisciplinary Perspectives on Data (EADH2021) 

21.–25.09.2021. „Modeling as a Sustainability Strategy for DH Software Applications“. Vortrag: Anna Fischer, Lukas Harzenetter, Philip Schildkamp, Uwe Breitenbücher, Claes Neufeind, Frank Leymann, Brigitte Mathiak (als Videobeitrag umgesetzt von Dennis Demmer, Anna Fischer und Philip Schildkamp).

DSCC-IN National Chapters' Meeting 

04.–05.10.2021. „GO UNITE! German Chapter of the Data Stewardship Competence Centers Implementation Network (DSCC-IN)“. Vortrag: Patrick Helling.

GO UNITE! Herbst-Workshop 

21.10.2021. Organisation: *Monika Linne*, Patrick Helling.

FDM-Forum der Landesinitiative fdm.nrw 

28.10.2021. „Hat der Bedarf schon Struktur? Überlegungen zu Bedarfs- und Servicekonzepten im (geisteswissenschaftlichen Forschungsdatenmanagement)“. Vortrag: Patrick Helling.

DINI/nestor Jour Fixe der FDM-Initiativen 

17.11.2021. Teilnahme: Patrick Helling.

NFDI4Culture Community Plenary 

17.–19.11.2021. Teilnahme: Patrick Helling.

IT-Zertifikat der Philosophischen Fakultät der Universität zu Köln „Digital Objects“ (Leitung: Susanne Kurz M.A.)

09.12.2021. „Forschungsdatenmanagement in den Geisteswissenschaften – digitales Publizieren“. Gastvortrag: Patrick Helling.

9.3 Publikationen

*kursiv: externe Co-Autor*innen*

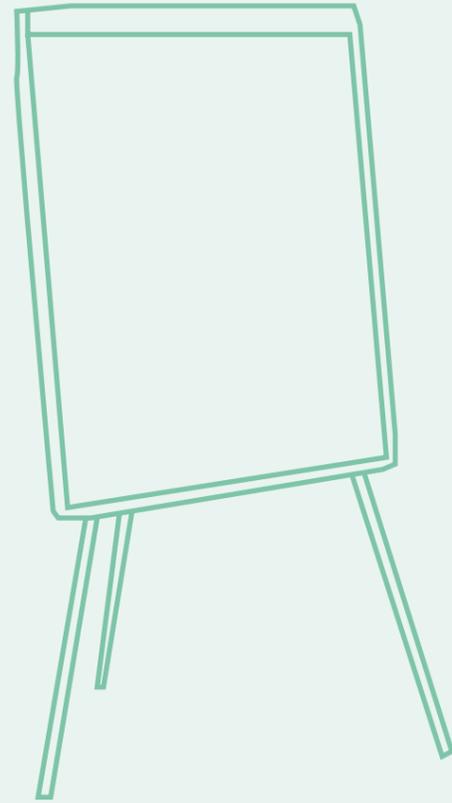
Monika Linne, Patrick Helling, Anja Busch: „Pragmatic – Concrete – Community-driven: The Credo of GO UNI“, in: GO FAIR News, 23. Februar 2021.

Peter Andorfer, Anna Busch, Fabian Cremer, Andreas Henrich, Patrick Helling, Harald Lordick, Dennis Mischke, Timo Steyer: „Bericht zur vDHd2021-Veranstaltung: Zukunftslabor DHd-Abstracts“, in: DHd-Blog, 28. Mai 2021.

Patrick Helling, Andreas Speer, Øyvind Eide (Hrg.): **„FORGE 2021 – Forschungsdaten in den Geisteswissenschaften: MAPPING THE LANDSCAPE – Geisteswissenschaftliches Forschungsdatenmanagement zwischen lokalen und globalen, generischen und spezifischen Lösungen. Konferenzabstracts“**, Köln 2021. doi:10.5281/zenodo.5289380.

Patrick Helling, *Monika Linne, Anja Busch: „GO Unite! Community Workshop – A Recap“, in: GO FAIR News, 2. November 2021.*

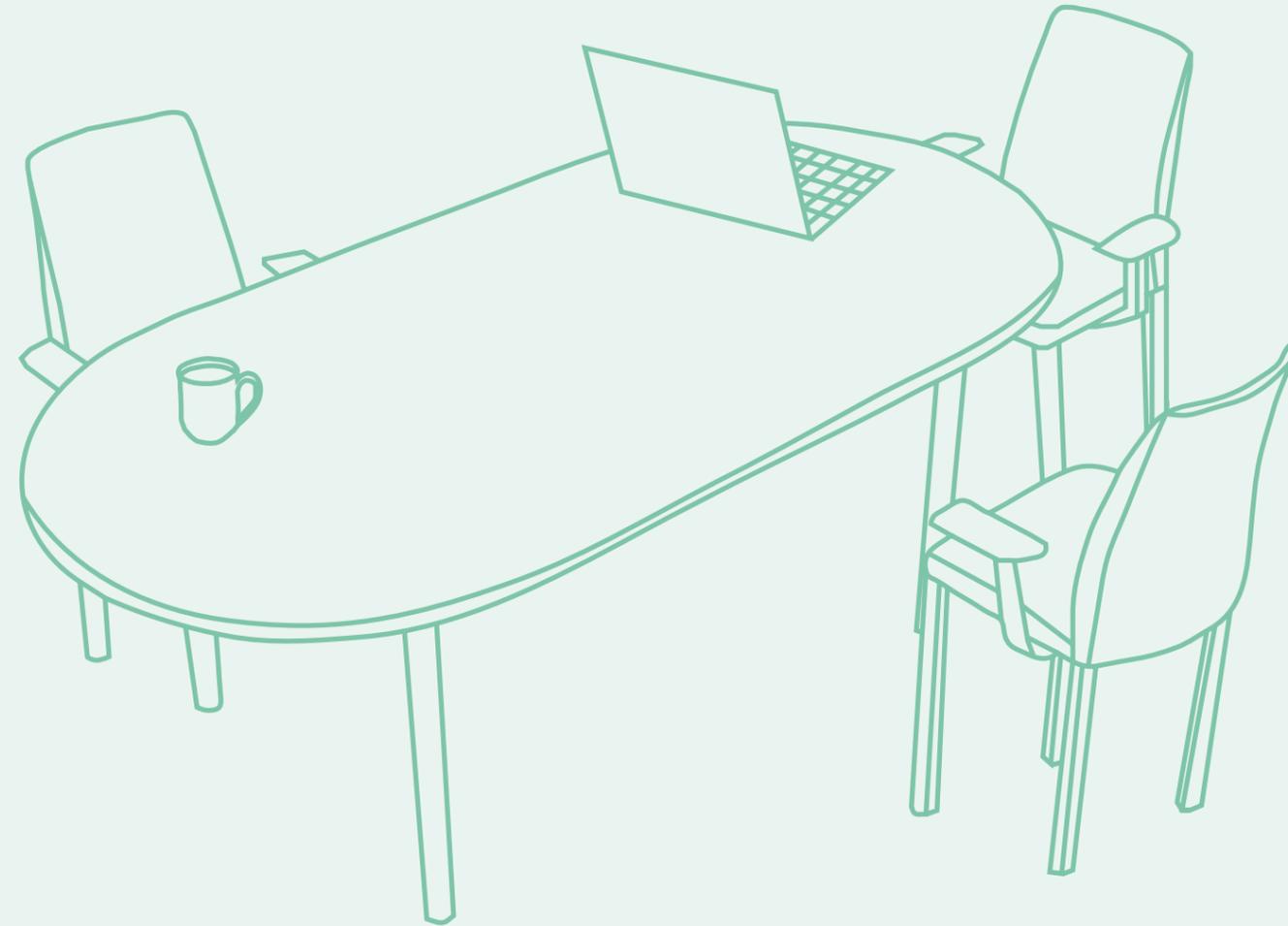
Jonathan Blumtritt, Rebekka Borges, Moritz Eßer, Patrick Helling, Nicole Majka, Elisabeth Mollenhauer, Claes Neufeind, Thomas Oehlke, Felix Rau, Lukas Sept, Julia Sorouri: **„Wiederauflage der Konferenz FORGE – Forschungsdaten in den Geisteswissenschaften. Nachbericht zur FORGE 2021 an der Universität zu Köln (8. bis 10. September 2021)“**, in: DHd-Blog, 25. November 2021.



10 Awareness, Schulungen & Lehre

10.1 Sammelkarten und Vorträge „Daten & Datteln digital“

Ein Schwerpunkt des DCH liegt auf der Vermittlung von FDM-Kompetenzen sowie der Sensibilisierung von Geisteswissenschaftler*innen für die Relevanz von Forschungsdatenmanagement: Weiterhin wird das Thema Forschungsdatenmanagement in die Ausbildung von Promovierenden an der Philosophischen Fakultät und in die Master-Studiengänge Informationsverarbeitung, Medieninformatik und Linguistik integriert (s. Kap. 10.2 und 10.3). Um das Thema noch stärker an der Philosophischen Fakultät und darüber hinaus sichtbar zu machen und in der wissenschaftlichen Praxis der Fakultätsmitglieder zu verankern, hat das DCH gemeinsam mit dem Dekanat die FDM-Sammelkartenreihe „Daten & Datteln“ ins Leben gerufen.



Unsere Daten & Datteln-Sammelkarten

Wintersemester 2020/2021:

HOW TO KEEP IT SIMPLE – Langlebigkeit und Pflege von Forschungsdaten

HOW TO KEEP IT SAFE – Backup und Datensicherheit

HOW TO KEEP IT RUNNING – Nachhaltige Softwarekuratierung

HOW TO KEEP IT CITABLE – Persistente Identifier: DOI, ORCID und Co.

Sommersemester 2021: FDM-Ressourcen & -Services

Language Archive Cologne (LAC)

Audiosammlung „Oral Tales of Mongol Bards“

KA³-Audioservices

VedaWeb

Alle Karten stehen als Download verfügbar unter: <https://dch.phil-fak.uni-koeln.de/daten-und-datteln/sammelkarten>.



5

Forschungsdaten sollten langfristig auffindbar, zugänglich, interoperabel und wiederverwendbar sein.



Dies zu gewährleisten kann aufwendig sein, wenn Daten nicht unabhängig von externen Systemen in offenen Formaten zugänglich sind. Ohne aktive Pflege halten

- Daten in gut dokumentierten Formaten *Jahrzehnte*,
- Daten, die spezifische Software benötigen, *Jahre*,
- Daten, die eine Web-App benötigen, *Monate*.

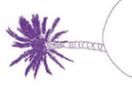
Sorgen Sie für eine bessere Lesbarkeit und Verständlichkeit Ihrer Daten. Einfach und nachhaltig!

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung zu Aspekten der Langlebigkeit und Pflege Ihrer Forschungsdaten suchen, gerne via E-Mail.

info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de



Daten & Dateien



HOW TO KEEP IT SIMPLE

Langlebigkeit und Pflege von Forschungsdaten

6

Datenverlust kann katastrophale Konsequenzen für die wissenschaftliche Arbeit haben. Ein durchdachter Backup-Workflow kann das verhindern und ist daher elementarer Bestandteil guten Forschungsdatenmanagements.



Grundlage ist die 3-2-1-Regel: Daten sollten immer

- dreimal vorliegen (Original und 2 Kopien),
- auf zwei Medientypen gespeichert und
- einmal als Offsite-Kopie gesichert werden.

Die Sicherung sollte dabei regelmäßig, automatisiert und möglichst über institutionelle Infrastruktur geschehen.

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung zu Backup und Sicherheit Ihrer Forschungsdaten suchen, gerne via E-Mail.

info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de



Daten & Dateien



HOW TO KEEP IT SAFE

Backup und Datensicherheit



8

Digitale Forschungsdaten sollten dauerhaft zitierbar sein. Allerdings können sich digitale Speicherorte im Laufe der Zeit verändern und Links können kaputt gehen.



Persistente Identifier, z. B. Digital Object Identifier (DOI), helfen, digitale Publikationen und Forschungsdaten

- unabhängig von ihrem Speicherort,
- eindeutig und dauerhaft identifizierbar und
- verlässlich zitierbar zu machen.

Zusätzlich ermöglichen Systeme wie ORCID eine persistente und unabhängige Identifizierung von Autor*innen.

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung bei der Nutzung eines persistenten Identifier suchen, gerne via E-Mail.



info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de

Daten & Dateien

HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

HOW TO KEEP IT CITABLE

Persistente Identifier: DOI, ORCID und Co.



info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de

Daten & Dateien

HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Digitale Tools und Web-Anwendungen können sowohl Ergebnisse von Forschungsprozessen als auch zentrale Zugangsschichten zu Forschungsdaten sein.



Die Pflege komplexer Anwendungen ist zeit- und kostenintensiv. Um den Betrieb auch nach Projektende noch sicherstellen zu können, sollte Software möglichst

- unabhängig von proprietären Systemen funktionieren,
- auf etablierte und verbreitete Technologien aufbauen,
- auf dynamische Komponenten verzichten und
- umfangreich dokumentiert werden.

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung zu Aspekten nachhaltiger Softwarekuratierung suchen, gerne via E-Mail.

HOW TO KEEP IT RUNNING

Nachhaltige Softwarekuratierung





9

Mit dem Language Archive Cologne (LAC) unterhält das DCH gemeinsam mit dem Institut für Linguistik (IfL) der Universität zu Köln ein Forschungsdatenrepositorium für alle geisteswissenschaftlichen Fachdisziplinen, die mit audiovisuellen (AV) Daten arbeiten. Das LAC bietet

- die fachspezifische Behandlung und langfristige Archivierung sowie
- eine webbasierte, offene Verfügbarkeit von

qualitativ hochwertigen, digitalen Sprachressourcen. Es unterstützt so aktiv einen nachhaltigen Umgang mit Forschungsdaten im Sinne guter wissenschaftlicher Praxis.



Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung bei der Archivierung und Verfügbarmachung von AV-Daten suchen, gerne via E-Mail.

info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de



FDM-Ressourcen & -Services

Language Archive Cologne (LAC)



HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Daten & Dateien



10

Die Audiosammlung „Oral Tales of Mongol Bards“ bietet einen Einblick in die Welt der oralen Literaturtradition der Inneren Mongolei. Sie wurde durch das DCH gemeinsam mit dem Cologne Center for eHumanities (CCEH) kuratiert, im Language Archive Cologne (LAC) archiviert und publiziert. Die Aufnahmen und annotierten Transkriptionen sind

- über eine ansprechende Präsentationswebseite sowie
- über das LAC für weitere wissenschaftliche Forschung langfristig verfügbar und nachnutzbar.



Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung bei der Kuratation, Archivierung und Verfügbarmachung von audiovisuellen (AV) Daten suchen, gerne via E-Mail.

info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de



FDM-Ressourcen & -Services

Audiosammlung „Oral Tales of Mongol Bards“



HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Daten & Dateien



FDM-Ressourcen & -Services



KA3-Audioservices



info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de

Daten & Dateien

HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung bei der Transkription und Analyse von audiovisuellen (AV) Daten suchen, gerne via E-Mail.

Die Spracherkennung wurde durch das IAS Fraunhofer-Institut im Rahmen des BMBF-Verbundprojekts KA3 entwickelt. In Zusammenarbeit mit dem Kölner Rechenzentrum (RRZK) werden diese Services bereitgestellt.



- automatische orthografische Transkription,
- automatische Kapitielerkennung mit Schlagwörtern,
- freie akademische Nutzung sowie
- eine einfache Anwendung.

Mit den KA3-Audioservices stellt das DCH webbasierte Dienste zur Analyse deutschsprachiger Audioaufnahmen bereit. Die Services ermöglichen

11

HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Daten & Dateien



FDM-Ressourcen & -Services



VedaWeb



info-dch@uni-koeln.de
@dch_cologne
https://dch.phil-fak.uni-koeln.de

Daten & Dateien

HOW TO-Reihe zum Forschungsdatenmanagement in den Geisteswissenschaften

Kontaktieren Sie uns, wenn Sie Beratung oder Unterstützung bei der linguistischen Erforschung altindischer Texte suchen, gerne via E-Mail.

Die Funktionalitäten wurden beispielhaft an der Rigveda, dem ältesten Teil der vier Vedas des Hinduismus, umgesetzt. VedaWeb ist Teil des Verbundportals „Cologne South Asian Languages and Texts“ (C-SALT).



- metrisch und morphologisch annotiert verfügbar,
- mit mehreren Übersetzungen und
- lexikografisch sowie korpuslinguistisch durchsuchbar gemacht werden können.

Zur Unterstützung bei der Erforschung altindischer Texte stellt das DCH mit VedaWeb ein webbasiertes, freizugängliches Portal zur Verfügung, mit dessen Hilfe jene

12



10.2 a.r.t.e.s. Workshop „Forschungsdatenmanagement in den Geisteswissenschaften“

Um Promovierende frühzeitig mit dem Thema Forschungsdatenmanagement vertraut zu machen und die Sensibilität für das Thema bereits zu Beginn einer akademischen Karriere zu fördern, hat das DCH im Sommersemester 2021 und im Wintersemester 2021/22 erneut Workshops unter dem Titel „Forschungsdatenmanagement in den Geisteswissenschaften“ an der a.r.t.e.s. Graduate School for the Humanities Cologne veranstaltet. Die Sitzungen fanden virtuell am 23. und am 30. Juni, bzw. am 1. und 8. Dezember 2021 statt.

Ziel des Workshops ist es, die Promovierenden mit FDM-Grundlagen vertraut zu machen und ausgewählte FDM-Konzepte am jeweils eigenen Dissertationsprojekt zu erlernen. Kernbereiche des Lehrkonzepts sind der Forschungsdatenlebenszyklus, die Kuratierung von Forschungsdaten, insbesondere die Verwendung von Metadaten und Repositorien, Open Access, die FAIR-Prinzipien sowie Persönlichkeitsrechts- und Urheberrechtsfragen. Darüber hinaus wird ein besonderer Fokus auf Hard- und Software-Services, die an der Universität zu Köln angeboten werden, gelegt. Die Veranstaltungen enden mit dem Verfassen eines Datenmanagementplans für das eigene Dissertationsprojekt.

Die Workshops gehören zum Wahlpflicht-Bereich des a.r.t.e.s.-Praxisangebots und werden auch 2022 wieder stattfinden.

Kursleitung: Patrick Helling, Felix Rau

10.3 Master-Übung „Forschungsdatenmanagement“

Im Sommersemester 2021 hat das DCH zum dritten Mal die Übung „Forschungsdatenmanagement“ für Studierende der Master-Studiengänge Informationsverarbeitung, Medieninformatik und Linguistik angeboten. In der Übung werden grundlegende Konzepte, Richtlinien und Standards des geisteswissenschaftlichen Forschungsdatenmanagements eingeführt und gleichberechtigt die Perspektive der Forschungspraxis sowie der Datenkuratierung und des Datenmanagements illustriert. Ein besonderes Augenmerk liegt dabei auf der Vielfalt der Methoden und den spezifischen Anforderungen des Forschungsdatenmanagements in den Geisteswissen-

schaften. Zur praktischen Aneignung der behandelten FDM-Aspekte erhalten alle Studierenden der Übung zu Beginn individuelle Datensätze. An ihnen gilt es im Verlauf der Veranstaltung unterschiedlichste praktische Übungen durchzuspielen, um diese schließlich in einem umfassenden Datenmanagementplan zu dokumentieren.

Im Gegensatz zum Vorjahr wurden die Kursinhalte in diesem Jahr nicht in Form von vorproduzierten Videos zur Verfügung gestellt, sondern wieder in wöchentlichen Lehrveranstaltungsterminen besprochen. Dabei waren die Sessions jeweils in einen inhaltlichen und einen praktischen Teil unterteilt.

Die Übung wird im Sommersemester 2022 erneut angeboten.

Kursleitung: Jonathan Blumtritt, Patrick Helling, Felix Rau

Kursprogramm:

Einführung in das Forschungsdatenmanagement

Allgemeine Aspekte des Forschungsdatenmanagements

Forschungsdaten in den Geisteswissenschaften

Konzeptuelle sowie funktionale Beschreibung und Abgrenzung von Daten

Daten, Metadaten und Paradata

Umfassende Diskussion der Funktionsweisen und Logik von Metadaten(-formaten)

Workflows, Versionierung und Nachvollziehbarkeit

Reproduzierbarkeit, Dokumentation und Strategien im Umgang mit Forschungsdaten

Rechtliche Fragestellungen, Datensicherheit

Diskussion urheberrechtlicher und datenschutztechnischer Aspekte beim Umgang mit Forschungsdaten

Ethische Fragestellungen

Diskussion ethischer Herausforderungen im Forschungsdatenmanagement

Repositorien

Eigenschaften und Funktionsweisen

Archivierung, Publikation und Nachnutzung

Behandlung von Publikations- und Nachnutzungsstrategien

Datenkuration

Zentrale Aspekte der Datenkuration in den Geisteswissenschaften

Normdaten

Eigenschaften und Funktionsweisen von Normdaten

Datenmanagementpläne

Umfassende Einführung in die Funktionsweise von Datenmanagementplänen

Projektmanagement

Backupstrategien und Grundfunktionen der Versionierung, sinnvolle Daten- und Ordnerorganisation

Ausblick

Die Entwicklungen im Rahmen der NFDI werden auch im Jahr 2022 eine zentrale Rolle am DCH spielen: Die Entscheidung über eine mögliche Bewilligung der Konsortien NFDI4Memory und NFDI4Objects, die als zwei weitere geisteswissenschaftliche Initiativen in der dritten Förderrunde eingereicht wurden, wird im Laufe des Jahres erwartet. Das DCH ist als geisteswissenschaftliches Datenzentrum an beiden Anträgen beteiligt. Darüber hinaus werden die Arbeiten im Rahmen des 2021 gestarteten Konsortiums Text+ im kommenden Jahr deutlich intensiviert werden.

Zusätzlich sollen mit der Bildung eines NFDI-Hubs an der Philosophischen Fakultät die verschiedenen Institutionen und Personen, die an geisteswissenschaftlichen NFDI-Konsortien beteiligt sind, lokal zusammengebracht werden. Ziel ist es, sowohl einen Austausch vor Ort zu ermöglichen als auch die NFDI grundsätzlich an der Fakultät nachhaltig zu verankern und für Forschende an der Universität zu Köln zu öffnen.

Mit dem Start der zweiten Förderphase des SPP 2143 „Entangled Africa“ und dem Start des Teilprojekts „FAIR.rdm – FAIRes Management von afrikanischen archäologischen Forschungsdaten“ wird das DCH gemeinsam mit dem AAARc das Datenmanagement innerhalb des SPP 2143 betreuen.

Die Aktivitäten des DCH im Rahmen des German Chapter GO UNITE! des GO FAIR DSCC-IN sollen 2022 noch einmal gesteigert werden: So sind z. B. insgesamt drei Workshops zu verschiedenen Perspektiven auf die standardisierte Beschreibung von FDM-Bedarfsstrukturen und auf die Entwicklung von Beratungs- und Betreuungsworkflows in Zusammenarbeit mit Kolleg*innen des Servicezentrums eSciences der Universität Trier geplant.

Für den Jahresbeginn 2022 ist ein Workshop zum Forschungsdatenmanagement in den Geisteswissenschaften für Promovierende am GRK 2661 „anschließen – ausschließen“ geplant. Außerdem plant das DCH in Kooperation mit der USB für März 2022 die Durchführung eines ersten Workshops zu Data Literacy für Master-Studierende an der Philosophischen Fakultät. Der Workshop dient insbesondere der Erprobung von Konzepten zur Platzierung des Themas in den Lehr- und Lernangebotsstrukturen an der Fakultät und soll die Zusammenarbeit mit der USB weiter fördern.



aus unserem Twitter-Account @dch_cologne

Abkürzungsverzeichnis

AAArC	African Archaeology Archive Cologne, Universität zu Köln
ADHO	Alliance of Digital Humanities Organizations
AGD	Archiv für Gesprochenes Deutsch, IDS – Leibniz-Institut für Deutsche Sprache, Mannheim
API	Application Programming Interface
AR	Augmented Reality
a.r.t.e.s.	a.r.t.e.s. Graduate School for the Humanities Cologne, Universität zu Köln
ASW	Allgemeine Sprachwissenschaft, Institut für Linguistik, Universität zu Köln
AV	Audiovisuell
AWK	Nordrhein-Westfälische Akademie der Wissenschaften und der Künste, Düsseldorf
BBAW	Berlin-Brandenburgische Akademie der Wissenschaften, Berlin
BERD@NFDI	NFDI Consortium for Business, Economic and Related Data
BMBF	Bundesministerium für Bildung und Forschung
BSB	Bayerische Staatsbibliothek, München
C ³ RDM	Cologne Competence Center for Research Data Management, Universität zu Köln
CA	Competence Areas, Universität zu Köln
CAU	Christian-Albrechts-Universität zu Kiel
CCeH	Cologne Center for eHumanities, Universität zu Köln
CCP	Culture Community Plenary, NFDI4Culture
CDS	Center for Data and Simulation Science, Universität zu Köln
CKLD	CLARIN Knowledge-Centre for Linguistic Diversity and Language Documentation
CLARIAH-DE	Zusammenschluss der Common Language Resources and Technology Infrastructure Deutschland und der Digital Research Infrastructure for the Arts and Humanities Deutschland (DARIAH-DE)
CLARIN	Common Language Resources and Technology Infrastructure
CLARIN ERIC	CLARIN European Research Infrastructure Consortium
CLARIN-D	Common Language Resources and Technology Infrastructure Deutschland
CMDI	Component Metadata Infrastructure
CoDArchLab	Cologne Digital Archaeology Laboratory, Archäologisches Institut, Universität zu Köln

CRDA	Cultural Research Data Academy, NFDI4Culture
CRETA	Center for Reflected Text Analysis e. V.
C-SALT	Cologne South Asian Languages and Texts, Universität zu Köln
CTS	CoreTrustSeal
DAI	Deutsches Archäologisches Institut, Berlin
DAPHNE4NFDI	DAta for PHoton and Neutron Experiments
DARE	Digital Averroes Research Environment, Thomas-Institut, Universität zu Köln
DBLP	Digital Bibliography & Library Project
DBM	Deutsches Bergbau-Museum Bochum – Leibniz-Forschungsmuseum für Georessourcen
DCAT	Data Catalog Vocabulary
DCAT-AP	Data Catalog Vocabulary Application Profile
DELAMAN	Digital Endangered Languages and Musics Archives Network
DFG	Deutsche Forschungsgemeinschaft
DH	Digital Humanities; ADHO (Alliance of Digital Humanities Organizations) Digital Humanities Conference
DHd	Verband Digital Humanities im deutschsprachigen Raum e. V.
DH.NRW	Digitale Hochschule Nordrhein-Westfalen, Hagen
DINI	Deutsche Initiative für Netzwerkinformation e. V.
DIPF	DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation (ehem. Deutsches Institut für Internationale Pädagogische Forschung), Frankfurt a. M.
DM	Deutsches Museum von Meisterwerken der Naturwissenschaft und Technik, München
DOI	Digital Object Identifier
DSCC-IN	Data Stewardship Competence Centers (GO FAIR Implementation Network)
EADH	European Association for Digital Humanities
eLAC	easy Language Archive Cologne
ELAR	Endangered Languages Archive, SOAS (School of Oriental and African Studies) University of London
FAIR	FAIR-Prinzipien („Findable, Accessible, Interoperable, Reusable“)
FAIRmat	FAIR Data Infrastructure for Condensed-Matter and Chemical Physics of Solids (NFDI)
FDM	Forschungsdatenmanagement
fdm.nrw	Landesinitiative für Forschungsdatenmanagement Nordrhein-Westfalen der Digitalen Hochschule NRW (bis 2019 LNFDI)

FID Fachinformationsdienst
 FIZ FIZ Karlsruhe – Leibniz-Institut für Informationsstruktur (ehem. Fachinformationszentrum Karlsruhe)
 FORGE Konferenz „Forschungsdaten in den Geisteswissenschaften“
 GBV Gemeinsamer Bibliotheksverbund
 GDKE Generaldirektion Kulturelles Erbe Rheinland-Pfalz, Mainz
 GESIS GESIS – Leibniz-Institut für Sozialwissenschaften (ehem. Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen), Köln
 GIS Geoinformationssysteme
 GRK Graduiertenkolleg, Deutsche Forschungsgemeinschaft
 GWI Geisteswissenschaftliche Forschungsinfrastrukturen e. V.
 gwin Projekt „Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit“, Zentrum für nachhaltiges Forschungsdatenmanagement, Universität Hamburg
 GWK Gemeinsame Wissenschaftskonferenz, Bonn
 h_da Hochschule Darmstadt – University of Applied Sciences
 HI Herder-Institut für historische Ostmitteleuropaforschung – Institut der Leibniz-Gemeinschaft, Marburg
 HKI Lehrstuhl Historisch-Kulturwissenschaftliche Informationsverarbeitung, Universität zu Köln
 HTML Hypertext Markup Language
 HU Humboldt-Universität Berlin
 IAAS Institut für Architektur von Anwendungssystemen, Universität Stuttgart
 IAIS Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme, Sankt Augustin
 iDAI Datenportale und -services des Deutschen Archäologischen Instituts
 IDE Institut für Dokumentologie und Editorik e. V.
 IDGS Institut für Deutsche Gebärdensprache und Kommunikation Gehörloser, Universität Hamburg
 IDH Institut für Digital Humanities, Universität zu Köln
 IDS IDS – Leibniz-Institut für Deutsche Sprache (ehem. Institut für Deutsche Sprache), Mannheim
 IEG IEG – Leibniz-Institut für Europäische Geschichte (ehem. Institut für Europäische Geschichte), Mainz
 IfL Institut für Linguistik, Universität zu Köln
 IIF International Image Interoperability Framework
 INEL Akademieprojekt „Indigenous Northern Eurasian Languages“ („Grammatiken, Korpora und Sprachtechnologie für indigene

nordeurasische Sprachen“), Akademie der Wissenschaften in Hamburg, Universität Hamburg, Hamburger Zentrum für Sprachkorpora
 INF Informationsstruktur des Sonderforschungsbereichs 1252, Universität zu Köln
 IuK Kommission für Informations- und Kommunikationstechnologie, Universität zu Köln
 KA³ BMBF-Verbundprojekt „KA³ – Kölner Zentrum Analyse und Archivierung von AV-Daten“
 KKAK Kommission für Archäologie Außereuropäischer Kulturen, Bonn
 KonsortSWD Consortium for the Social, Behavioural, Educational and Economic Sciences (NFDI)
 KUPS Kölner UniversitätsPublikationsServer
 LABW Landesarchiv Baden-Württemberg, Stuttgart
 LAC Language Archive Cologne
 LUCL Leiden University Centre for Linguistics
 LIS Förderprogramm „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ der Deutschen Forschungsgemeinschaft
 MaRDI Mathematical Research Data Initiative (NFDI)
 MLU Martin-Luther-Universität Halle-Wittenberg
 MPCD DFG-Projekt „Das zoroastrische Mittelpersische: Digitales Corpus und Wörterbuch“
 MPI-PL Max-Planck-Institut für Psycholinguistik, Nijmegen
 nestor Kompetenznetzwerk digitale Langzeitarchivierung (Network of Expertise in Long-Term Storage and Availability of Digital Resources in Germany)
 NFDI Nationale Forschungsdateninfrastruktur e. V.
 NFDI4Culture Consortium for the research data on material and immaterial cultural heritage
 NFDI4Memory Consortium for the Historically Oriented Humanities
 NFDI4Objects Research Data Infrastructure for the Material Remains of Human History
 NFDI4DS NFDI4DataScience
 NFDI4Earth NFDI Consortium Earth System Sciences
 NFDI-MatWerk National Research Data Infrastructure for Material Sciences & Engineering
 NFDI4Microbiota National Research Data Infrastructure for the Research of Microbiota
 NVIDIA DLI NVIDIA Deep Learning Institute

OLAC	Open Language Archives Community
ORCID	Open Researcher and Contributor ID
PARADISEC	Pacific and Regional Archive for Digital Sources in Endangered Cultures, University of Sydney, University of Melbourne, Australian National University
PUNCH4NFDI	Particles, Universe, NuClei and Hadrons for the NFDI
QuaDramA	Projekt „Quantitative Drama Analytics“, VolkswagenStiftung/ Deutsche Forschungsgemeinschaft
QUEST	BMBF-Verbundprojekt „QUEST: Quality-Established“
QMCS	Quantitative Modeling of Complex Systems (Competence Area III, Universität zu Köln)
Q:TRACK	Projekt „QuaDramA: Tracking Character Knowledge“, Schwerpunktprogramm 2207 „Computational Literary Studies“
RDA	Research Data Alliance
RDA DE	Research Data Alliance Deutschland e. V.
RDM	Research Data Management
RGZM	Römisch-Germanisches Zentralmuseum – Leibniz-Forschungsinstitut für Archäologie, Mainz
ROR	Research Organization Registry Community
RRZK	Regionales Rechenzentrum, Universität zu Köln
SASOA	Institut für Südasiens- und Südostasien-Studien, Universität zu Köln
SFB	Sonderforschungsbereich, Deutsche Forschungsgemeinschaft
SLUB	Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
SoFS	Scale-Out File Service
SPK	Stiftung Preußischer Kulturbesitz, Berlin
SPP	Schwerpunktprogramm, Deutsche Forschungsgemeinschaft
SUB	Niedersächsische Staats- und Universitätsbibliothek Göttingen
SWLI	SOAS World Languages Institute, SOAS (School of Oriental and African Studies) University of London
Text+	Language- and Text-Based Research Data Infrastructure (NFDI)
TIB	TIB – Leibniz-Informationszentrum Technik und Naturwissenschaften und Universitätsbibliothek, Hannover (ehem. Technische Informationsbibliothek)
TOSCA	Topology and Orchestration Specification for Cloud Applications
TSM	Tivoli Storage Manager

TWS	Theaterwissenschaftliche Sammlung, Universität zu Köln
UT	Universität Trier
USB	Universitäts- und Stadtbibliothek Köln
vDHD	Virtuelle Jahrestagung des Verbands Digital Humanities im deutschsprachigen Raum e. V. 2021
VerbundFDB	Verbund Forschungsdaten Bildung
VHD	Verband der Historikerinnen und Historiker Deutschlands e. V.
VR	Virtual Reality
VZG	Verbundzentrale des Gemeinsamen Bibliotheksverbunds
XML	Extensible Markup Language
ZAS	Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin
ZfL	Zentrum für LehrerInnenbildung, Universität zu Köln

Abbildungsverzeichnis

Titelblatt:	Illustration von Julia Sorouri	Seite 35:	Logo von NFDI4Culture
Seite 4–5:	Illustration von Julia Sorouri	Seite 36:	Grafik von NFDI4Culture (https://nfdi4culture.de/task-areas.html)
Seite 7:	Foto © Fabian Stuertz	Seite 37:	Logo von NFDI4Memory
Seite 8 und 9:	Illustration von Julia Sorouri	Seite 38:	Logo von NFDI4Objects
Seite 11:	Foto von Oliver Broll	Seite 41:	Screenshot eines Tweets (https://twitter.com/dch_cologne)
Seite 12–14:	Fotos (Nils Reiter, Eleftheria Paliou, yvind Eide, Felix Rau) von Marcello Perathoner	Seite 44:	Illustration von Julia Sorouri
Seite 14–15:	Foto (Patrick Helling) von Foto Kuhweide, Foto (Anke Debbeler) von Lukas Sept, Foto (Cosima Granderath-Miegel) von FOTOKISTE.de	Seite 46–47:	Illustration von Julia Sorouri
Seite 16:	Foto (Lukas Lammers) von Lukas Lammers, Foto (Elisabeth Mollenhauer) von Fotoagentur Wolf, Foto (Sara Saleh): privat	Seite 49:	Screenshot eines YouTube-Videostill
Seite 17:	Foto (Nicole Majka) von Felix Rau	Seite 50-51:	Illustration von Julia Sorouri
Seite 19:	Fotos (Jonathan Blumtritt, Claes Neuefeind) von Marcello Perathoner	Seite 55:	Foto © Fabian Stuertz
Seite 20–21:	Illustration von Julia Sorouri	Seite 56-57:	Illustration von Julia Sorouri
Seite 22:	Sammelkarten von Patrick Helling, Felix Rau (Text), Elisabeth Mollenhauer (Layout), Julia Sorouri (Palme)	Seite 58:	Screenshot von https://earlycinema.dch.phil-fak.uni-koeln.de/
Seite 29:	Fotos (Fahime Same, Maximilian Hörl) von Frida Berg	Seite 63:	Screenshot von https://dch.phil-fak.uni-koeln.de/bestaende/datensicherung/dora-telugu
Seite 31:	Foto © format2d	Seite 64–65:	Illustration von Julia Sorouri
Seite 32–33:	Illustration von Julia Sorouri	Seite 67:	Foto © Raimond Spekking / CC BY-SA 4.0 (via Wikimedia Commons)
Seite 34:	Grafik von Text+ (https://twitter.com/Textplus_NFDI), erstellt mit venngage.com	Seite 68–69:	Illustration von Julia Sorouri; Icons von The Noun Project (https://thenounproject.com)
		Seite 70:	Logo von OpenTOSCA
		Seite 71:	Screenshot von https://sustainery.dch.phil-fak.uni-koeln.de/
		Seite 72:	Icons von The Noun Project (https://thenounproject.com)

- Seite 73: Screenshot eines YouTube-Videostill
- Seite 74–75: Illustration von Julia Sorouri
- Seite 76: Screenshot eines Tweets (https://twitter.com/dch_cologne) mit Banner von Julia Sorouri
- Seite 79: Screenshot eines Tweets (https://twitter.com/dch_cologne) mit Banner von Julia Sorouri
- Seite 80: von links oben nach rechts unten: Fotos von Ulrike Wuttke (CC-BY 4.0), Thomas Oehlke, Felix Rau, Franziska Fritzsche (CC-BY 4.0), Elisabeth Mollenhauer
- Seite 83–86: Excel-Grafiken erstellt von Elisabeth Mollenhauer, Abb. 3: Design von Elisabeth Mollenhauer
- Seite 87: Screenshot eines Tweets (https://twitter.com/dch_cologne) mit Zoom-Foto
- Seite 89–93: Icons von shashank singh von The Noun Project (<https://thenounproject.com>)
- Seite 94–95: Illustration von Julia Sorouri
- Seite 96–99: Icons von AFY Studio von The Noun Project (<https://thenounproject.com>), Sammelkarten von Patrick Helling, Felix Rau (Text), Elisabeth Mollenhauer (Layout), Julia Sorouri (Palme)
- Seite 102: Screenshot eines Tweets (https://twitter.com/dch_cologne) mit Zoom-Foto