

## Der DHd Data Steward Maßnahmen zur Entwicklung einer nachhaltigen Datenstrategie für die Digital Humanities im deutschsprachigen Raum

### Borges, Rebekka

RebekaBorges@gmx.de  
Rheinische Friedrich-Wilhelms-Universität Bonn, Deutschland

### Debbeler, Anke

adebell1@uni-koeln.de  
Data Center for the Humanities (DCH), Universität zu Köln,  
Deutschland

### Helling, Patrick

patrick.helling@uni-koeln.de  
Data Center for the Humanities (DCH), Universität zu Köln,  
Deutschland

## Der DHd Data Steward

Mit der Ernennung eines Data Stewards hat der Verband Digital Humanities im deutschsprachigen Raum e.V. (DHd) auf der Jahreskonferenz 2020 in Paderborn eine Funktion geschaffen, um eine umfassende Datenstrategie für alle Materialien, Publikationen und Ergebnisse, die im Kontext des DHd-Verbandes entstanden sind/entstehen werden, zu entwickeln.<sup>1</sup> Sie sollen im Sinne der FAIR-Prinzipien (Wilkinson et al. 2016) (1) langfristig gesichert und archiviert als auch (2) nach Möglichkeit nachhaltig publiziert und verfügbar gemacht werden.

Mit diesem Posterbeitrag soll die bisherige Arbeit des DHd Data Stewards präsentiert sowie ein Blick in die Datenzukunft des Verbandes gegeben werden.

## Materialien der DHd-Konferenzen

### Langfristige Verfügbarkeit von DHd-Konferenzwebsites

Einen zentralen Gegenstand und Zugangspunkt einer DHd-Jahreskonferenz stellt die Konferenzwebsite dar. Die Erstellung und das Hosting werden von den lokalen Organisator\*innen übernommen. Nach einer Konferenz fehlt es häufig an notwendigen Mitteln für den langfristigen Betrieb dieser Konferenzwebsites.

Zur Gewährleistung einer dauerhaften Erreichbarkeit wurde damit begonnen, statische HTML-Versionen einzelner Konferenzwebsites zu erzeugen, um mögliche technische Abhängigkeiten aufzulösen und den Kurationsaufwand herunterzufahren. Mit einem Umzug auf die technische Infrastruktur der Alliance of Digital Humanities Organizations (ADHO) sollen die Konfe-

renzwebsites langfristig als Unterseite der Verbandswebsite bereitgestellt.<sup>2</sup>

### Nachhaltige Publikation aller einzelnen DHd-Beiträge

Das Book of Abstracts einer Jahreskonferenz stellt ein wichtiges Publikationsinstrument dar. Seit 2016 wurden die Book of Abstracts via Zenodo publiziert (siehe Burr 2017; Stolz 2017; Vogeler 2018; Sahle 2019; Schöch 2020). Für die erste DHd-Jahreskonferenz 2014 wurde kein Book of Abstracts publiziert. Die Beiträge zur DHd-Jahreskonferenz 2015 sind als PDF-Datei über die Konferenzwebsite verfügbar (siehe Stiegler 2015).

Wenngleich die Books of Abstracts ein wichtiges Schaufenster der deutschsprachigen Digital Humanities (Sahle 2019; Schöch 2020) sind, ermöglichen sie weder eine eindeutige Zitierbarkeit einzelner Beiträge noch ihre Erfassung in digitalen Katalogen. Potentiale und Lösungsansätze zum Umgang mit DHd-Beiträgen wurden bereits in verschiedenen Formaten community-getrieben adressiert (Cremer 2018; Andofer 2019; Andorfer et al. 2019; Lordick 2020; Steyer et al. 2020).

Um die DHd-Beiträge einzeln zitierbar zu veröffentlichen, wurden alle Beiträge gesammelt:

- 2016, 2018-2020: TEI-Dateien via GitHub<sup>3</sup>
- 2017: TEI-Dateien von lokalen Organisator\*innen
- 2015: gesammelte PDF-Datei von lokalen Organisator\*innen
- 2014: gesammelte, unvollständige PDF-Datei von lokalen Organisator\*innen + weitere Beiträge durch Autor\*innen als PDF- und Word-Dateien

Zur Erstellung einzelner PDF-Dateien zu jedem Abstract wurde für die Jahrgänge 2016-2020 auf bestehende Transformationskripte zur Erstellung von Book of Abstracts aus TEI-Dateien zurückgegriffen.<sup>4</sup> Die Skripte wurden angepasst, um für jedes Abstract einzeln eine PDF-Datei zu generieren.<sup>5</sup>

Zur persistenten Publikation der DHd-Beiträge wurde das generische Online-Repositorium Zenodo gewählt.<sup>6</sup> Hier wird bereits seit 2019 eine DHd-Community als zentraler Publikationsort kuratiert.<sup>7</sup>

Mit Hilfe weiterer XSL-Transformationskripte wurden für die Jahrgänge 2016-2020 jeweils eine Konferenz-Metadaten-datei generiert, die dem DataCite-Schema und den Anforderungen von Zenodo entspricht (siehe Abb. 1).<sup>8,9</sup>

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata xmlns="http://www.zenodo.org/xml" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.zenodo.org/xml zenodo.xsd">
  <upload_type>publication</upload_type>
  <publication_type>conferencepaper</publication_type>
  <publication_date>2022-03-07</publication_date>
  <title>Der DHd Data Steward - Maßnahmen zur Entwicklung einer nachhaltigen Datenstrategie für die Digital Humanities im deutschsprachigen Raum</title>
  <creators>
    <creator>
      <name>Borges, Rebekka</name>
      <affiliation>Rheinische Friedrich-Wilhelms-Universität Bonn, Deutschland</affiliation>
    </creator>
    <creator>
      <name>Debbeler, Anke</name>
      <affiliation>Data Center for the Humanities, Universität zu Köln, Deutschland</affiliation>
    </creator>
    <creator>
      <name>Helling, Patrick</name>
      <affiliation>Data Center for the Humanities, Universität zu Köln, Deutschland</affiliation>
    </creator>
  </creators>
  <descriptionA single abstract from the DHd-2022 Book of Abstracts.</description>
  <access_right-open>access_right=
  <license>cc-by</license>
  <id>
  <keywords>DHd2022, Forschungsdatenmanagement, FDM, Data Steward, Datenstrategie, Publikation, Archivierung, Veröffentlichung</keywords>
  <contributors>
    <contributor>
      <name>Gierthoes, Michaela</name>
      <affiliation>Universität der Bundeswehr, München</affiliation>
      <type>Editor</type>
    </contributor>
    <contributor>
      <name>Helling, Patrick</name>
      <affiliation>Digital Humanities im deutschsprachigen Raum e.V.</affiliation>
      <type>Editor</type>
    </contributor>
  </contributors>
  <conference_title>DHd 2022 - Kulturen des digitalen Gedächtnisses</conference_title>
  <conference_acronym>DHd 2022</conference_acronym>
  <conference_date>07-03-2022-11-03-2022</conference_date>
  <conference_place>Potsdam</conference_place>
  <conference_url_https>https://www.dhd2022.de/</conference_url_https>
  <committee_doi>
  <references>https://doi.org/45.2353/beispieldoi.3409948</references>
  <notes>Sofern eine editorische Arbeit an dieser Publikation stattgefunden hat, dann bestand diese aus der Einleitung von Bindestrichen in Überschriften, die aufgrund fehlerhafter Silbentrennung entstanden sind, der Vervollständigung von Namen der Autor*innen in das Schema "Nachname, Vorname" und/oder der Trennung von Überschrift und Unterüberschrift durch die Setzung eines Punktes, sofern notwendig.</notes>
</metadata>
```

Abb. 1: Metadatenchema zu den DHd-Beiträgen mit Beispieleinträgen.

Für den Upload der Abstracts auf Zenodo wurde automatisiert für jedes Abstract ein Ordner (Bundle) mit PDF-Datei, TEI-Datei (sofern vorhanden) und Metadatenatz im JSON-Format generiert. Die Publikation dieser Bundles erfolgte automatisiert durch einen eigenen Publikationsworkflow (siehe Abb. 2) über die Zenodo REST API-Schnittstelle in die DHd-Zenodo Community.<sup>10</sup> Die DHd-Beiträge verfügen über Digital Object Identifier (DOI), sind via OpenAIRE auffindbar und wurden durch die computer science bibliography (dblp) katalogisiert.<sup>11, 12</sup>

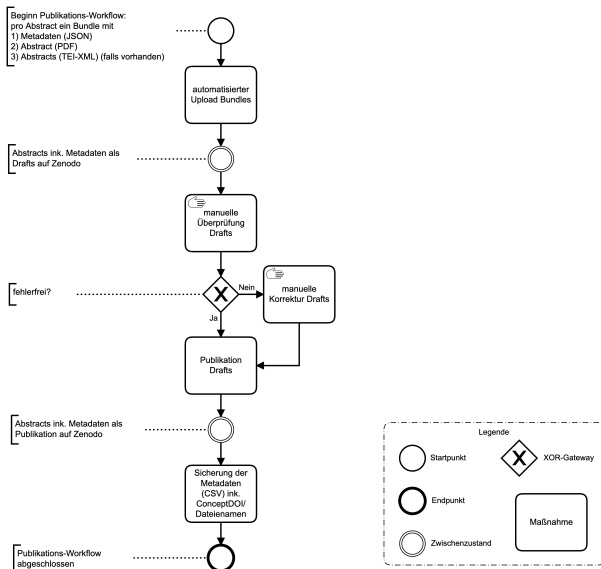


Abb. 2: Der Publikationsworkflow.

## Ausblick und nächste Schritte

Der Umzug von alten Konferenzwebsites auf die ADHO-Infrastruktur soll vorangetrieben werden. Empfehlungen für die technische Umsetzung zukünftiger Konferenzwebsites soll deren Umzug auf die ADHO-Infrastruktur vereinfachen. Zusätzlich bedarf es einer Policy zur Speicherung, Archivierung und ggf. Publikation von DHd-Materialien die außerhalb von der Jahreskonferenzen entstanden sind. Zuletzt sollen die entwickelten Workflows und die technische Realisation zur Sicherung und Publikation der einzelnen DHd-Abstracts im Rahmen einer durch die Community getragenen TaskForce zur kontinuierlichen Unterstützung der wechselnden Konferenz-Organisator\*innen weiterentwickelt und optimiert werden. Die Einrichtung einer solchen TaskForce ist im Kontext der DHd 2022 Konferenz in Potsdam geplant. Ein Aufruf zur Beteiligung der Community wird im Vorfeld kommuniziert.<sup>13</sup> Ein wichtiger Ansatz ist hier die Aufbereitung der TEI-Dateien, damit diese auch im Index of Digital Humanities Conferences (Weingart, Eichmann-Kalwara und Lincoln 2020) katalogisiert werden. Eine entsprechende Initiativgruppe hat sich bereits in einem Workshop auf der vDHd-Konferenz 2021 (Andorfer, Busch, Cremer et al. 2021) formiert.<sup>14</sup>

## Fußnoten

- <https://dig-hum.de/> [letzter Zugriff 10. November 2021].
- Erste Website, die auf die ADHO-Infrastruktur umgezogen wurde: DHd-Jahreskonferenzwebsite 2020, Online: <https://dhd2020.dig-hum.de/> [letzter Zugriff 10. November 2021].
- <https://github.com/peertrilcke/dhd2016-boa> ; <https://github.com/GVogeler/DHd2018> [letzter Zugriff 10. November 2021]; Das GitHub-Repositorium mit den TEI-Dateien zur DHd-Jahreskonferenz 2019 ist nicht mehr verfügbar; <https://github.com/NinaSeemann/DHd2020-BoA> [letzter Zugriff 10. November 2021].
- <https://github.com/karindalziel/TEI-to-PDF> [letzter Zugriff 10. November 2021].
- <https://github.com/araborn/DHd2018> [letzter Zugriff 10. November 2021].
- <https://zenodo.org/> [letzter Zugriff 10. November 2021].
- <https://zenodo.org/communities/dhd> [letzter Zugriff 10. November 2021].
- <https://github.com/reborg789/zenodup> [letzter Zugriff 10. November 2021].
- <https://schema.datacite.org/> [letzter Zugriff 10. November 2021].
- <https://github.com/cceh/zenodup> [letzter Zugriff 10. November 2021].
- <https://www.openaire.eu/> [letzter Zugriff 10. November 2021].
- <https://dblp.org/db/conf/dhd/index.html> [letzter Zugriff 10. November 2021].
- Weitere Kontaktmöglichkeit zum Data Steward: <https://dig-hum.de/dhd-data-steward> [letzter Zugriff 17. November 2021].
- Dieses Poster ist komplementär zum Posterbeitrag „Strukturen und Impulse zur Weiterentwicklung der DHd-Abstracts“ der Initiativgruppe auf der DHd 2022.

## Bibliographie

- Andorfer, Peter** (2019): *dhd-boas-app*, Online: <https://dhd-boas-app.acdh-dev.oeaw.ac.at/> (letzter Zugriff: 14. Juli 2021).
- Andorfer, Peter / Busch, Anna / Cremer, Fabian / Henrich, Andreas / Helling, Patrick / Lordick, Harald / Mischke, Dennis / Steyer, Timo** (2021): "Bericht zur vDHd2021-Veranstaltung: Zukunftslabor DHd-Abstracts". DHd-Blog, Online: <https://dhd-blog.org/?p=15980> (letzter Zugriff: 14. Juli 2021).
- Andorfer, Peter / Cremer, Fabian / Steyer, Timo** (2019): "DHd 2019 Book of Abstracts Hackathon", Beitrag auf der *DHd 2019 Digital Humanities multimedial and multimodal*. 6. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum" (DHd 2019). Frankfurt am Main und Mainz, Online: <http://doi.org/10.5281/zenodo.4622102>.
- Burr, Elisabeth** (ed.) (2017). *DHd 2016 Modellierung - Vernetzung - Visualisierung. Die Digital Humanities als Fächerübergreifendes Forschungsparadigma. Konferenzabstracts*. Leipzig, Online: <http://doi.org/10.5281/zenodo.3679331>.
- Cremer, Fabian** (2018): „Nun sag, wie hältst Du es mit dem Digitalen Publizieren, Digital Humanities?“. *Digitale Redaktion Blog*, Online: <https://editorial.hypotheses.org/113> (letzter Zugriff: 14. Juli 2021).
- Lordick, Harald** (2020): *DH(d) Konferenzbeiträge*, Online: <http://www.steinheim-institut.de/dhd/> (letzter Zugriff: 14. Juli 2021).

**Sahle, Patrick** (ed.) (2019): *DHd 2019 Digital Humanities: multimedial & multimodal. Konferenzabstracts*. Frankfurt am Main, Online: <https://doi.org/10.5281/zenodo.2596095>.

**Schöch, Christof** (ed.) (2020): *DHd 2020 Spielräume: Digital Humanities zwischen Modellierung und Interpretation. Konferenzabstracts*. Paderborn, Online: <https://doi.org/10.5281/zenodo.3666690>.

**Steyer, Timo / Andorfer, Peter / Cremer, Fabian** (2020): „Abstract Enhancement. Potentiale der DHd-Konferenzabstracts als Daten/Publication“, Beitrag auf der *DHd 2020 Spielräume: Digital Humanities zwischen Modellierung und Interpretation*. 7. Tagung des Verbands "Digital Humanities im deutschsprachigen Raum" (DHd 2020). Paderborn, Online: <http://doi.org/10.5281/zenodo.4621706>.

**Stiegler, Johannes** (ed.) (2015): *DHd 2015 Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Information und Interpretation. Konferenzabstracts*. Graz, Online: <https://dhd2015.uni-graz.at/de/nachlese/book-of-abstracts/> (letzter Zugriff: 14. Juli 2021).

**Stolz, Michael** (ed.) (2017): *DHd 2017 Digitale Nachhaltigkeit. Konferenzabstracts*. Bern, Online: <http://doi.org/10.5281/zenodo.3684825>.

**Vogeler, Georg** (ed.) (2018): *DHd 2018 Kritik der digitalen Vernunft. Konferenzabstracts*. Köln, Online: <http://doi.org/10.5281/zenodo.3684897>.

**Weingart, Scott B. / Eichmann-Kalwara, Nickoal / Lincoln, Matthew** (2020): *The Index of Digital Humanities Conferences*. Carnegie Mellon University, Online: <https://dh-abstracts.library.cmu.edu/> (letzter Zugriff: 14. Juli 2021).

**Wilkinson, Mark D. / Dumontier, Michel / Aalbersberg, IJsbrand Jan / Appleton, Gabrielle / Axton, Myles / Baak, Arie / Blomberg, Niklas / Boiten, Jan-Willem / da Silva Santos, Luiz Bonino / Bourne, Philip E. / Bouwman, Jildau / Brookes, Antony J. / Clark, Tim / Crosas, Mercè / Dillo, Ingrid / Dumon, Oliver / Edmunds, Scott / Evelo, Chris T. / Finkers, Richard / Gonzalez-Beltran, Alejandra / Gray, Alasdair J.G. / Groth, Paul, Goble, Carole / Grethe, Jeffrey S. / Heringa, Jaap / A.C't Hoen, Peter / Hooft, Rob / Kuhn, Tobias / Kok, Ruben / Kok, Joost / Lusher, Scott J. / Martone, Maryann E. / Mons, Albert / Packer, Abel L. / Persson, Bengt / Rocca-Serra, Philippe / Roos, Marco / van Schaik, Rene / Sansone, Susanna-Assunta / Schultes, Erik / Sengstag, Thierry / Slater, Ted / Strawn, George / Swertz, Morris A. / Thompson, Mark / van der Lei, Johan / van Mulligen, Erik / Velterop, Jan / Waagmeester, Andrea / Wittenburg, Peter / Wolstencroft, Katherine / Zhao, Jun / Mons Barend** (2016): "The FAIR Guiding Principles for scientific data management and stewardship" in: *Scientific Data* 3, Article number: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>.