

Improving publication processes of the Association for Digital Humanities in the German Speaking Areas (DHd) - The DHd Data Steward and the community-driven Task Force “DHd Abstracts”

Debbeler, Anke

adebell1@uni-koeln.de

Data Center for the Humanities (DCH), University of Cologne, Germany

Helling, Patrick

patrick.helling@uni-koeln.de

Data Center for the Humanities (DCH), University of Cologne, Germany

Borges, Rebekka

borges@uni-bonn.de

Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

Introduction

Since 2012, the Association for Digital Humanities in the German Speaking Areas (DHd) ¹ has represented the field of Digital Humanities (DH) in Germany, Austria, and Switzerland. With over 400 members, it is one of the largest DH associations worldwide.

The DHd association has organized an annual DHd conference together with different local DH institutions since 2014. ² At each conference, an average of around 130 conference papers were presented in the following formats: lecture, panel, workshop, poster, and, since 2020, doctoral consortium.

Usually, after each conference, an individual Book of Abstracts (BoA) is produced script-based ³ and is published by the local organizers via the online storage service Zenodo ⁴ (see Burr 2017; Stolz 2017; Vogeler 2018; Sahle 2019; Schöch 2020; Geierhos 2022). ⁵ These BoAs represent an important showcase for the field of Digital Humanities (see Schöch 2020: V). Nevertheless, the collective publication of conference papers never allowed for individual citation and referencing of individual papers, nor could they be indexed in digital catalogs or used as research data. Potentials of an improved publication strategy as well as of improved scholarly analysis of conference papers have been already addressed by the DHd community (Cremer 2018; Andorfer et al. 2019;

Steyer et al. 2020; Busch et al. 2022) and partially implemented in exemplary fashion (Andorfer 2019; Lordick 2020 ⁶).

For improving the findability, accessibility, interoperability and reusability of the DHd conference papers in the sense of the FAIR-Principles (Wilkinson et al. 2016), the DHd association appointed a Data Steward. ⁷ In this paper we will present publication workflows developed by the DHd Data Steward and illustrate how the collaboration with the DHd community is an essential part of the improvement of the publication strategy within the DHd association. More precisely, we will introduce the DHd Task Force “DHd Abstracts”, which is working on improving the submission and publication processes community-driven.

DHd Data Steward

Central goals of the DHd Data Steward include publishing DHd conference papers as individual publications and making the underlying TEI-XML files, which are generated by the authors during the submission process through the DHConvalidator ⁸, accessible in a standardized way for reuse (Helling et al. 2022; Borges et al. 2022).

To this end, workflows have been developed for pre-processing ⁹ the respective TEI-XML files of the individual DHd conferences and generating an individual PDF file ¹⁰ for each conference paper. In addition, these workflows make it possible to publish the conference papers automatically via Zenodo's REST API with rich metadata, which is extracted script-based directly from the TEI-XML files. ¹¹

With this strategy, 1071 papers of the DHd conferences 2014-2022 have been published as distinct publications in the DHd community on Zenodo ¹² under a CC-BY 4.0 license and with individual Digital Object Identifiers (DOI). The published conference papers were indexed by the Digital Bibliography & Library Project (DBLP) ¹³ and the Index of Digital Humanities Conferences. ¹⁴ Additionally, an strategy for publishing and providing conference papers as reusable research data was developed: for each DHd conference, a GitHub repository was created to make all conference papers available in TEI-XML and PDF format with their metadata. ¹⁵ In order to avoid data loss, each GitHub repository is also published via Zenodo and is archived ¹⁶ through the “cold archiving” processes (Rau, Helling and Barabucci 2022) of the Data Center for the Humanities (DCH) ¹⁷ at the University of Cologne. ¹⁸

While these workflows are a step forward in improving the publication strategy of the DHd association, there are still major challenges to address.

Task Force “DHd Abstracts”

To distribute the responsibilities of the DHd Data Steward the DHd association has introduced the Task Force “DHd Abstracts” in 2022. ¹⁹ It consists of various members of the DHd community and is particularly concerned with the evaluation of the existing submission and publication process and its documentation for submitters.

Due to inconsistencies in the TEI-XML files the local organizers of the DHd conferences need to clean the TEI-XML files extensively to transform these into a comprehensive BoA. Analyzing and evaluating these inconsistencies is currently the main challenge of

the DHd Task Force and it especially requires the interregional exchange over locally generated knowledge during the organization of the DHd conferences.

The intent is to have a definite list of all possible mistakes in the TEI-XML files in order to develop strategies to prevent them in the future. For example, by creating a handout for submitters to reduce the respective inconsistencies in the TEI-XML files the Task Force aims to support the different institutions hosting the DHd conferences.

In addition, the whole submission process is evaluated and possible alternative tools improving the TEI-XML output, e.g., the FidusWriter (Gebhard 2022), are being tested. Furthermore, the Task Force is facing the problem of inconsistent affiliation details and is working on approaches to implement more authority data into the TEI-XML files.

The work of the Task Force is meant to address challenges and problems that affect the whole DHd community. Moreover, it opens up new possibilities to shape workflows that are part of the making of a community event, in this case the DHd conferences, in a way which involves the community of scholars themselves.

Concluding remarks

This contribution outlines how the original publication strategy of the annual DHd conferences was improved by the establishment of the DHd Data Steward. In addition, it describes the role of the community-driven Task Force in the improvement of both the submission processes and the publication strategy in the context of the DHd conferences.

In our talk, we will critically examine the workflows and processes for pre-processing and automated publication. Moreover, we will discuss further potentials for improving the whole submission process and plead for these tasks to be perceived and solved as a DH community task. As a matter of fact, with our paper we intend to share the strategies we developed with a broader DH community, in order to allow for their reuse. Indeed, they were already successfully reused in the contexts of different conferences in the German speaking areas, e.g., by the vDHD conference 2021²⁰ and the FORGE conference 2021.²¹

Notes

1. <https://dig-hum.de/>.
2. With an exception in 2021 due to the SARS-CoV-2 pandemic.
3. Scripts are usually based on <https://github.com/karindalziel/TEI-to-PDF>, e.g., <https://github.com/NinaSeemann/DHd2020-BoA>.
4. <https://zenodo.org/>.
5. Exceptions: In 2014 no BoA was published. In 2015 the BoA was published on the institutional website of the conference hosting institution (see Stiegler 2015).
6. Due to cyberattack this resource is temporarily not available.
7. <https://dig-hum.de/dhd-data-steward>.
8. <https://github.com/ADHO/dhconvalidator>.
9. <https://github.com/reborg789/zenodup>.
10. https://github.com/PatrickHelling/DHd_BoA-separated, based on <https://github.com/karindalziel/TEI-to-PDF>.
11. <https://github.com/cceh/zenodup>.
12. <https://zenodo.org/communities/dhd/?page=1&size=20>.
13. <https://dblp.org/db/conf/dhd/index.html>.
14. https://dh-abstracts.library.cmu.edu/conference_series/30.

15. <https://github.com/orgs/DHd-Verband/repositories>.
16. <https://dch.phil-fak.uni-koeln.de/bestaende/datensicherung>.
17. <https://dch.phil-fak.uni-koeln.de/>.
18. <https://www.uni-koeln.de/>.
19. The DHd Task Forces are a format, that allows to work on common challenges of the DHd community together with the DHd community.
20. <https://vdhd2021.hypotheses.org/>.
21. <https://forge2021.uni-koeln.de/>.

Bibliography

- Andorfer, Peter** (2019): dhd-boas-app. <https://dhd-boas-app.acdh-dev.oeaw.ac.at/> (last accessed: 19/04/2023).
- Andorfer, Peter / Cremer, Fabian / Steyer, Timo** (2019): “DHd 2019 Book of Abstracts Hackathon”, in: Book of Abstracts DHd 2019 Conference. Frankfurt am Main und Mainz. DOI: 10.5281/zenodo.4622102.
- Andorfer, Peter / Busch, Anna / Cremer, Fabian / Helling, Patrick / Henrich, Andreas / Henrich / Lordick, Harald / Mischke, Dennis / Steyer, Timo** (2021): Bericht zur vDHD2021-Veranstaltung: Zukunftslabor DHd-Abstracts. DHd-Blog (blog). <https://dhd-blog.org/?p=15980> (last accessed: 19/04/2023).
- Borges, Rebekka / Debbeler, Anke / Helling, Patrick** (2022): “Der DHd Data Steward - Maßnahmen zur Entwicklung einer nachhaltigen Datenstrategie für die Digital Humanities im deutschsprachigen Raum“, in: Book of Abstracts DHd 2022 Conference. Potsdam. DOI: 10.5281/zenodo.6328017.
- Burr, Elisabeth** (ed.) (2017): DHd 2016 Modellierung - Vernetzung - Visualisierung. Die Digital Humanities als Fächerübergreifendes Forschungsparadigma. Konferenzabstracts. Leipzig. DOI:10.5281/zenodo.3679331.
- Busch, Anna / Cremer, Fabian / Lordick, Harald / Mischke, Dennis / Steyer, Timo** (2022): “Strukturen und Impulse zur Weiterentwicklung der DHd-Abstracts“, in: Book of Abstracts DHd 2022 Conference. Potsdam. DOI: 10.5281/zenodo.6328089.
- Cremer, Fabian** (2018): “Nun sag, wie hältst Du es mit dem Digitalen Publizieren, Digital Humanities?“. Digitale Redaktion (blog). <https://editorial.hypotheses.org/113> (last accessed: 19/04/2023).
- Gebhard, Henning** (2022): Fidus Writer als Alternative zum DH ConValidator? Ein Prototyp“, in: Book of Abstracts DHd Conference 2022. Potsdam. DOI: 10.5281/zenodo.6327991.
- Geierhos, Michaela** (ed.) (2022): DH 2022 Kulturen des digitalen Gedächtnisses. Konferenzabstracts. Potsdam. DOI: 10.5281/zenodo.6304590.
- Helling, Patrick / Debbeler, Anke / Borges, Rebekka** (2022): “Konferenzbeiträge strategisch publizieren: Automatisierte Workflows zur individuellen Veröffentlichung von Konferenzbeiträgen am Beispiel des Verbands Digital Humanities im deutschsprachigen Raum e.V.“, in: O-Bib. Das Offene Bibliotheksjournal / Herausgeber VDB, 9(3): 1–17. DOI: 10.5282/o-bib/5835.
- Lordick, Harald** (2020): DH(d) Konferenzbeiträge, Online: <http://www.steinheim-institut.de/dhd/> (last accessed: 27/10/2022).
- Rau, Felix / Helling, Patrick / Barabucci, Gioele** (2022): “Pragmatic Research Data Management in the Humanities: Dark and Cold Archiving at the Data Center for the Humanities“, in: Book of Abstracts DH Conference 2022. Tokio. DOI: 10.5281/zenodo.6967064.

Sahle, Patrick (ed.) (2019): DHd 2019 Digital Humanities: multimedial & multimodal. Konferenzabstracts. Frankfurt am Main. DOI: 10.5281/zenodo.2596095.

Schöch, Christof (ed.) (2020): DHd 2020 Spielräume: Digital Humanities zwischen Modellierung und Interpretation. Konferenzabstracts. Paderborn. DOI: 10.5281/zenodo.3666690.

Steyer, Timo / Andorfer, Peter / Cremer, Fabian (2020): “Abstract Enhancement. Potentiale der DHd-Konferenzabstracts als Daten/Publication“, in: Book of Abstracts DHd 2020 Conference. Paderborn. DOI: 10.5281/zenodo.4621706.

Stigler, 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/> (last accessed: 19/04/2023).

Stolz, Michael (ed.) (2017): DHd 2017 Digitale Nachhaltigkeit. Konferenzabstracts. Bern. DOI: 10.5281/zenodo.3684825.

Vogeler, Georg (ed.) (2018): DHd 2018 Kritik der digitalen Vernunft. Konferenzabstracts. Köln. DOI: 10.5281/zenodo.3684897.

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: 10.1038/sdata.2016.18.