

Fehlerdokumentation des Book of Abstracts DHd 2024

Autor:innen

Nicole Majka | Bangor University, Wales | ORCID: 0000-0003-2657-0645

Patrick Helling | Verband Digital Humanities im deutschsprachigen Raum e.V. |
ORCID: 0000-0003-4043-165X

Die DHd Jahreskonferenz 2024 trug den Titel “#Quo vadis DH?”, fand vom 26. Februar 2024 bis zum 1. März 2024 statt und wurde an der Universität Passau ausgerichtet.¹ Die Software ConfTool,² welche zur Konferenzverwaltung genutzt wurde, wurde am 12. Juni 2023 freigeschaltet. Über dieses Tool konnten Autor:innen sich zur Konferenz anmelden und Beiträge einreichen. Autor:innen gaben hierzu Metadaten wie bspw. den *Beitragstitel*, *Autor:innennamen* mit *Affiliation* und *ORCID-ID* und *Keywords* zu einer Einreichung über das ConfTool ein. Die Metadaten wurden schließlich aus dem ConfTool über eine Schnittstelle an den DHConvalidator³ übertragen, in dem ein Template für jede Einreichung inklusive aller zugehörigen Metadaten erzeugt und von den Autor:innen zur Erstellung ihrer Einreichung genutzt wurde. Nachdem die Autor:innen ihren Beitrag im Template geschrieben hatten, wurde dieses über den DHConvalidator prozessiert, sodass eine .dhc-Datei mit einer HTML-, einer Template- und einer TEI-XML-Version der Einreichung sowie eingebetteter Abbildungen erzeugt wurde. Diese wurde schließlich durch die Autor:innen als Beitrag im ConfTool hochgeladen und eingereicht. Mit Hilfe der eingetragenen Metadaten wurden schließlich aus den TEI-XML-Versionen der eingereichten und akzeptierten Beiträge das Book of Abstracts sowie Einzelpublikationen der Beiträge über automatisierte Workflows⁴ generiert (Helling et al. 2022a).

¹ <https://dhd2024.dig-hum.de>, letzter Zugriff 21.05.2024.

² <https://www.conftool.net/de/startseite.html>, letzter Zugriff 24.04.2024.

³ <https://github.com/ADHO/dhconvalidator>, letzter Zugriff 21.05.2024.

⁴ https://github.com/PatrickHelling/DHd2022_BoA_separated, <https://github.com/reborg789/zenodup>, <https://github.com/cceh/zenodup>, letzter Zugriff 24.04.2024.

Für diese automatisierte Erstellung des Book of Abstracts sowie der Einzelpublikationen wurden die eingereichten TEI-XML-Versionen der Beiträge verwendet. In der Regel sind die TEI-XML-Versionen der Beiträge unsauber und können Fehler aufweisen, die händisch korrigiert werden müssen. Diese Fehler sind prinzipiell auf falsche Formatierungen sowie die Übernahme von Formatierungsinformationen beim Kopieren eines Beitrags aus einem anderen Textverarbeitungsprogramm in das Template durch die Autor:innen zurückzuführen (Helling et al. 2022b). Mit diesem Beitrag sollen jene Fehler exemplarisch für die DHd Jahreskonferenz 2024 dokumentiert werden.

Einführung

Zur Jahreskonferenz 2024 wurden insgesamt 144 Beiträge angenommen: 64 Poster, 7 Panels, 19 Workshops, 43 Vorträge und 11 Beiträge des Doctoral Consortiums. Von vorherigen Jahreskonferenzen waren folgende Fehler aus der Bereinigung der Beiträge bereits bekannt:

- Sonderzeichen im Text (oft auch bei Autor:innenamen in der Bibliographie) werden als # dargestellt
- Autor:innennamen sind in der Bibliographie nicht fett gedruckt
- Abbildungen fehlen im Beitrag oder werden fehlerhaft dargestellt
- Tabellen fehlen im Beitrag oder werden fehlerhaft dargestellt
- Bildunterschriften fehlen oder werden fehlerhaft dargestellt

Zusätzlich zu den schon bekannten Fehlern wurden auch weitere Fehler aufgezeichnet, die seltener vorkommen, aber ebenfalls korrigiert werden mussten. Diese wurden im Folgenden unter der Kategorie "sonstige Fehler" zusammengefasst. Dabei handelte sich unter anderen um folgende Fehler:

- Abstände zwischen Fließtext, Tabellen oder Bildern sind nicht einheitlich
- Überschriften sind nicht als solche hierarchisch formatiert
- fehlende bzw. nicht einheitliche Informationen bei Autor:innen (bspw. Affiliation)
- fehlende Bibliografie
- fehlende Titel

- Zeilenabstand ist aufgrund von Fußnotenreferenzierungen nicht einheitlich
- zu große Abstände im Text, wenn Textteile eingeklammert waren
- Probleme bei URLs (automatische Silbentrennung führt zu ungewollten Bindestrichen in Links und somit zu nicht funktionalen Links)
- fehlende Satzzeichen

Übersicht über alle Beiträge

Die folgende Tabelle 1 enthält eine Zusammenfassung aller Formatierungsfehler der Beiträge zur DHd 2024 vor der Prozessierung zum Book of Abstracts bzw. zu den Einzelpublikationen. Hierbei ist zu beachten, dass nicht alle, aber die Mehrzahl der Beiträge, fehlerhaft waren. Die unterschiedlichen Fehler der Kategorie "sonstige Fehler" wurden aus Gründen der Übersichtlichkeit nicht einzeln aufgeführt. Auch hier gibt es Beiträge, die nicht fehlerhaft waren und solche, die mehr als einen Fehler hatten. Daher wird jeder Beitrag, der mindestens einen Fehler der Kategorie "sonstige Fehler" aufwies, nur einmal in dieser Spalte gezählt.

Tabelle 1: Übersicht der Fehlertypen aller Beiträge.

| | Anzahl Beiträge | Sonderzeichen | Autor:innen nicht fett | Abbildung fehlt | Tabelle fehlt | Fehler bei Bildunterschriften | sonstige Fehler | Fehler gesamt |
|---------------------|-----------------|---------------|------------------------|-----------------|---------------|-------------------------------|-----------------|---------------|
| Poster | 64 | 10 | 36 | 1 | 1 | 29 | 44 | 121 |
| Panels | 7 | 3 | 5 | 0 | 0 | 0 | 7 | 15 |
| Workshops | 19 | 3 | 12 | 1 | 1 | 5 | 12 | 34 |
| Doctoral Consortium | 11 | 2 | 7 | 1 | 0 | 1 | 11 | 22 |
| Vorträge | 43 | 12 | 1 | 0 | 0 | 2 | 6 | 21 |
| Gesamt | 144 | 30 | 61 | 3 | 2 | 37 | 80 | 213 |

Bei den 144 Beiträgen, die angenommen wurden, traten in 30 Beiträgen Sonderzeichenfehler auf, bei denen Sonderzeichen als # angezeigt wurden, bei 61 Beiträgen waren die Autor:innennamen in der Bibliographie nicht fett gedruckt, bei drei Beiträgen fehlten die Abbildungen, bei zwei die Tabellen, 37 Beiträge hatten fehlerhafte Bildunterschriften und 80 hatten Fehler der Kategorie "sonstige Fehler". Insgesamt wurden mindestens 213 Fehler behoben. Diese Zahl ist hier nicht akkurat abgebildet und muss höher angenommen werden, da jeder Beitrag nur einmal in jeder

Kategorie aufgeführt wurde, auch wenn mehr als ein Fehler der jeweiligen Kategorien in einem Beitrag auftrat.

Fehler aller Beiträge

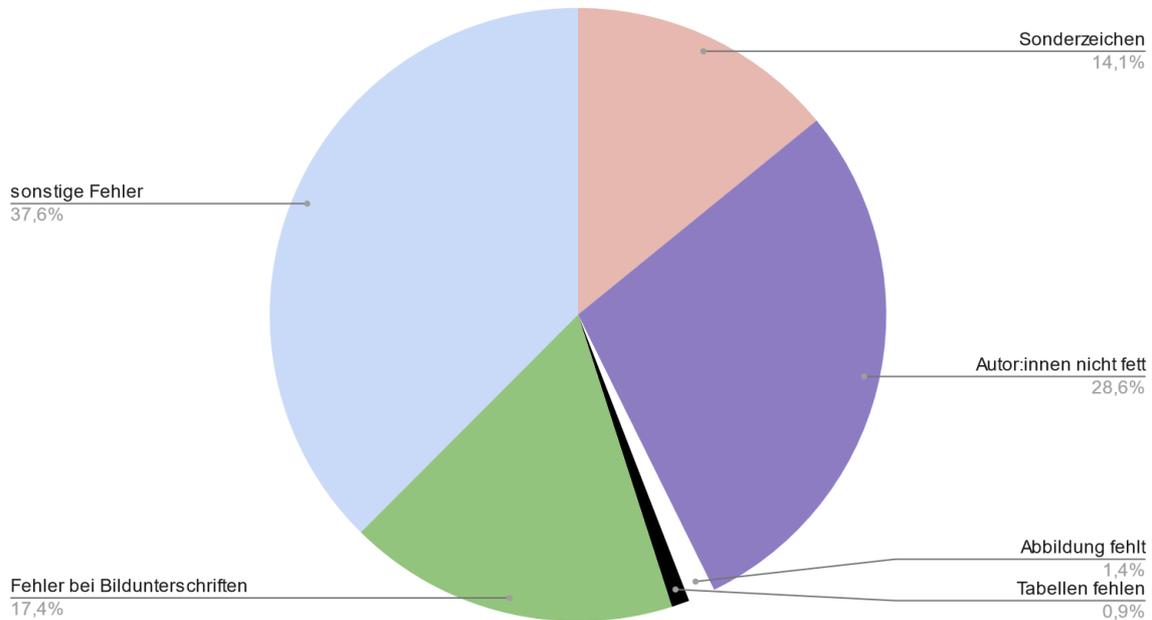


Abbildung 1: Fehleraufteilung aller Beiträge.

In Abbildung 1 wird deutlich, dass die Kategorie “sonstige Fehler” mit 37,6% am häufigsten vorkommt. Dies liegt daran, dass in dieser Kategorie mehr als eine Fehlerart festgehalten wurde. Darauf folgt die Kategorie “Autor:innennamen nicht fett” mit 28,6%. Fehlerhafte Bildunterschriften sind mit 17,4% die dritthäufigste Fehlerkategorie und Sonderzeichenfehler ergeben 14,1% der Fehler. Fehlende Tabellen und Abbildungen liegen gemeinsam bei 2,3%.

Wenn man die Anzahl der Beiträge beachtet, die Abbildungen und Tabellen haben, zeigt sich, dass die Quote der fehlenden Abbildungen und Tabellen vergleichsweise gut ausfällt.

Tabelle 2: Anzahl der Beiträge, die Abbildungen und / oder Tabellen enthalten.

| | Anzahl der Beiträge | Beiträge mit Abbildungen | Beiträge mit Tabellen |
|--------|---------------------|--------------------------|-----------------------|
| Poster | 64 | 37 | 3 |
| Panels | 7 | 0 | 0 |

| | | | |
|---------------------|-----|----|----|
| Workshops | 19 | 8 | 2 |
| Doctoral Consortium | 11 | 2 | 2 |
| Vorträge | 43 | 30 | 9 |
| Gesamt | 144 | 77 | 16 |

Im Vergleich zu Abbildungen kommen Tabellen seltener vor. Von 144 Beiträgen haben nur 16 Beiträge Tabellen, von denen nur zwei nach der automatisierten Generierung fehlten (12,5%). Abbildungen kommen hingegen bei über der Hälfte der Beiträge vor. Von 144 Beiträgen haben 77 Abbildungen, von denen nach der automatisierten Generierung nur drei fehlten (3,8%).

Darüber hinaus sind Fehler in der Formatierung der Bildunterschriften relevant, die aus Gründen der Barrierefreiheit benötigt werden, da sie Lesenden mit eingeschränkter Sehkraft helfen, nachzuvollziehen, was Abbildungen in einem Beitrag zeigen, beispielsweise, wenn ein Screenreader verwendet wird.

Insgesamt hatten 37 Beiträge Fehler bei den Bildunterschriften, wovon die Posterbeiträge den größten Anteil mit 29 fehlerhaften Beiträgen stellen. Das heißt, dass fast die Hälfte (48,1%) der Beiträge mit Bildern falsch formatierte Bildunterschriften hatten.

Die nächsten Kapitel gehen auf die Verteilung der Fehler innerhalb der einzelnen Beitragskategorien ein.

Poster

Es wurden insgesamt 64 Posterbeiträge angenommen. Nach der automatisierten Generierung der Beiträge im PDF-Format hatten zehn der Beiträge Sonderzeichenfehler, bei 36 Beiträgen waren die Autor:innennamen in der Bibliographie nicht fettgedruckt, in einem Beitrag fehlten Abbildungen und einem weiteren fehlten Tabellen, bei 29 Beiträgen gab es Fehler bezüglich der Formatierung der Bildunterschriften und 44 der Beiträge hatten Fehler der Kategorie "sonstige Fehler" (siehe Abbildung 2). Insgesamt wurden mindestens 121 Fehler behoben.

Fehler der Posterbeiträge

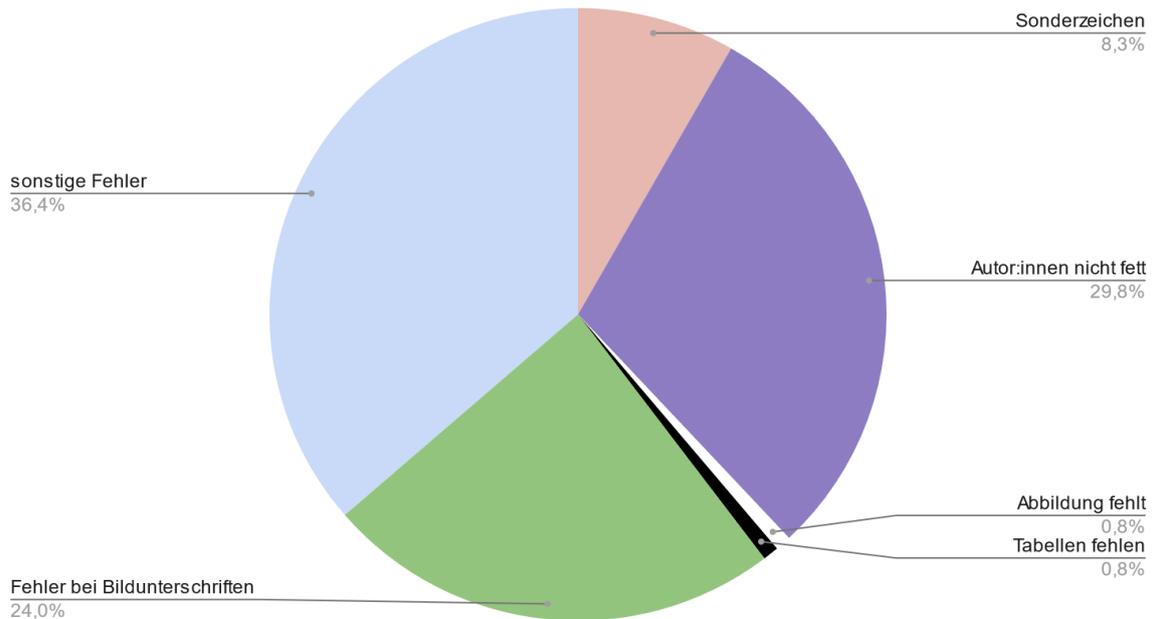


Abbildung 2: Fehleraufteilung der Posterbeiträge.

Panels

Bei den 7 Panelbeiträgen, die angenommen wurden, wurden nur drei Fehlerkategorien identifiziert. Drei Beiträge hatten Sonderzeichenfehler, bei fünf Beiträgen waren die Autor:innennamen in der Bibliographie nicht fett gedruckt und alle Beiträge hatten einen Fehler der Kategorie "sonstige Fehler" (siehe Abbildung 3). Insgesamt wurden hier mindestens 15 Fehler behoben.

Fehler der Panelbeiträge

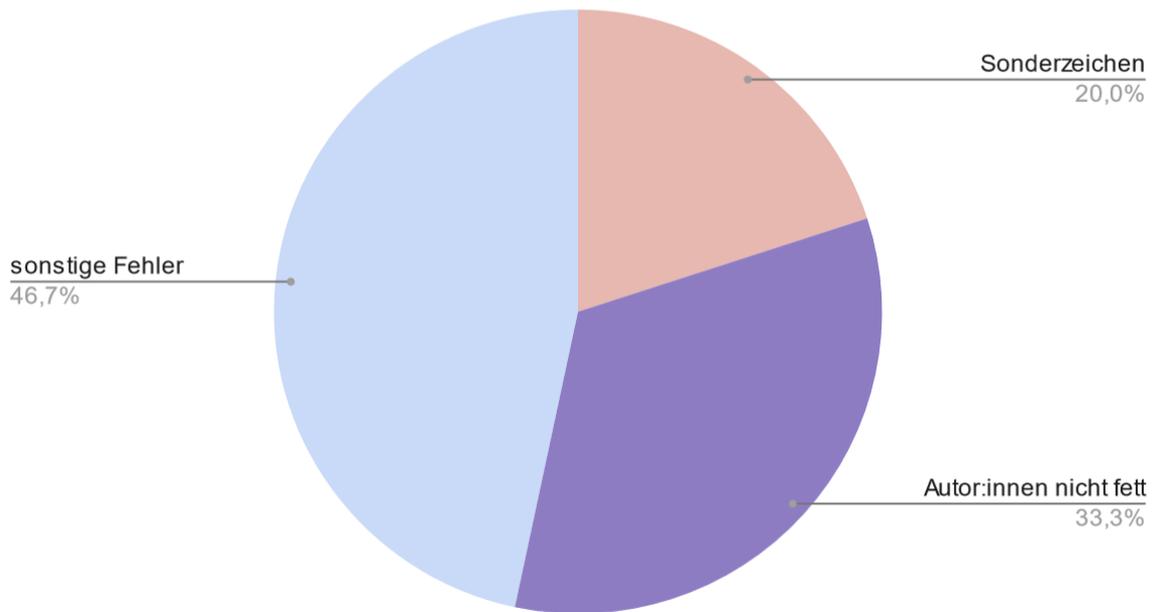


Abbildung 3: Fehleraufteilung der Panelbeiträge.

Workshops

Es wurden 19 Workshopbeiträge angenommen. Von diesen hatten drei Beiträge Sonderzeichenfehler, bei zwölf Beiträgen waren die Autor:innennamen nicht fett gedruckt, in einem Beitrag fehlten Abbildungen, in einem anderen Tabellen, fünf Beiträge hatten fehlerhafte Bildunterschriften und zwölf Beiträge hatten Fehler der Kategorie "sonstige Fehler" (siehe Abbildung 4). Hier wurden insgesamt mindestens 34 Fehler behoben.

Fehler der Workshopbeiträge

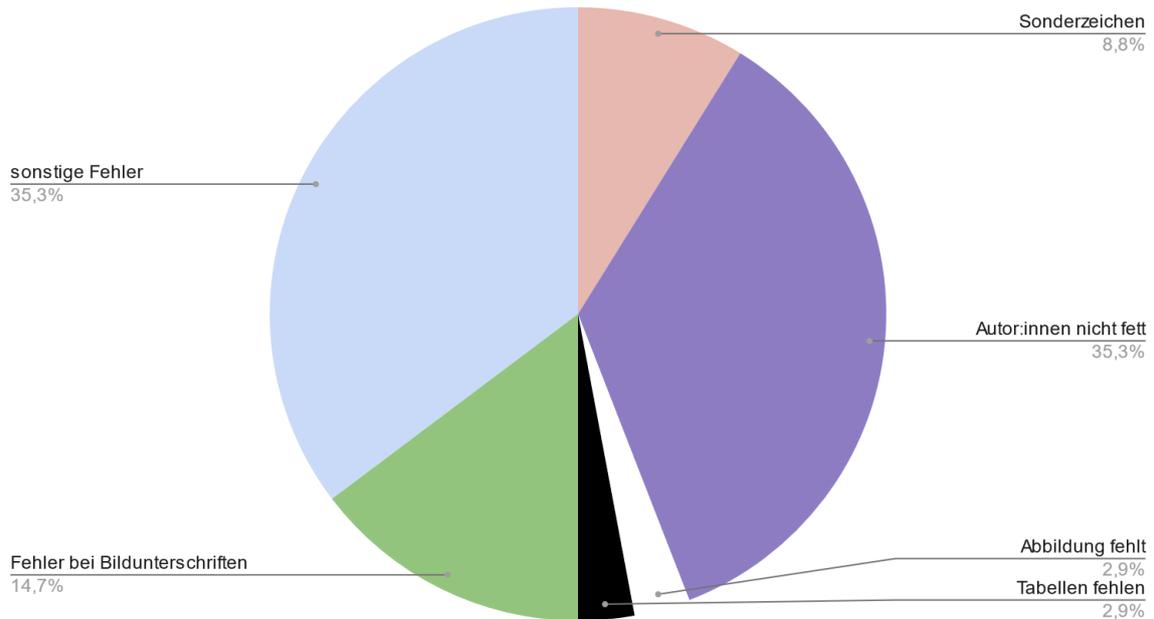


Abbildung 4: Fehleraufteilung der Workshopbeiträge.

Vorträge

Insgesamt wurden 43 Vorträge angenommen. Von diesen hatten zwölf Beiträge Sonderzeichenfehler, sechs Beiträge hatten Fehler der Kategorie "sonstige Fehler", zwei Beiträge hatten fehlerhafte Bildunterschriften und in einem Beitrag waren die Namen der Autor:innen in der Bibliographie nicht fettgedruckt. In keinem der Beiträge fehlten nach der automatischen Generierung der PDFs die Abbildungen oder Tabellen (siehe Abbildung 5).

Fehler der Vorträge

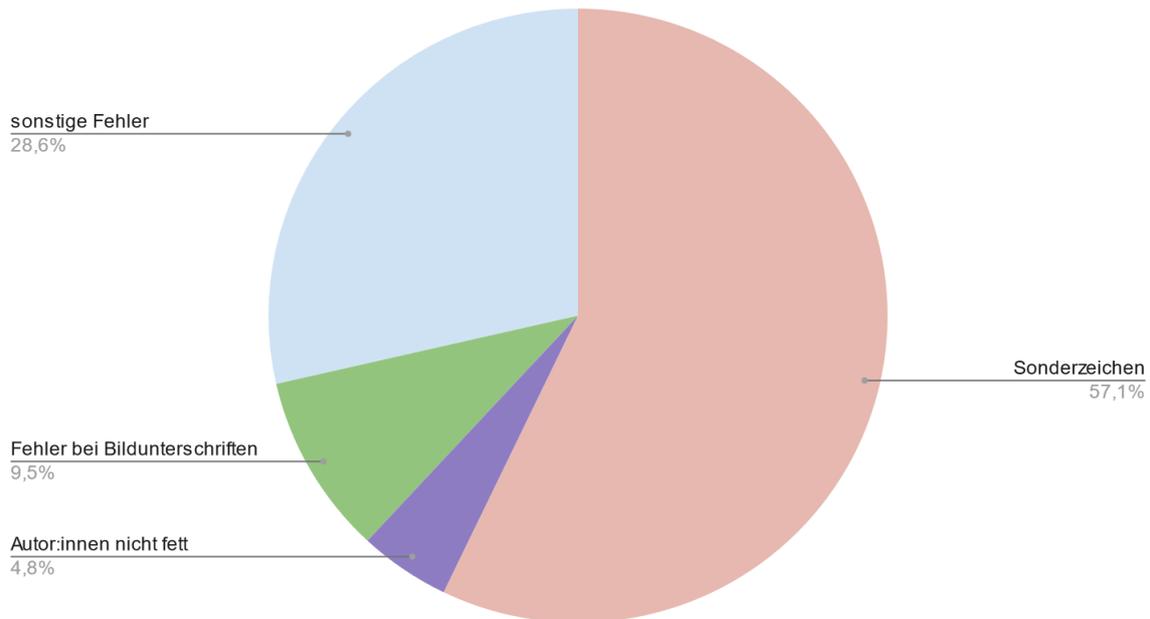


Abbildung 5: Fehleraufteilung der Vorträge.

Doctoral Consortium

Für das Doctoral Consortium wurden elf Beiträge angenommen. Zwei von diesen Beiträgen hatten fehlerhafte Sonderzeichen, bei sieben Beiträgen waren die Autor:innennamen in der Bibliographie nicht fettgedruckt, bei einem Beitrag fehlten Abbildungen und bei einem Beitrag gab es einen Fehler in der Bildunterschrift. Fehlende Tabellen traten bei dieser Beitragskategorie nicht auf, wohingegen jeder Beitrag einen Fehler der Kategorie "sonstige Fehler" hatte (siehe Abbildung 6).

Fehler der Beiträge für das Doctoral Consortium

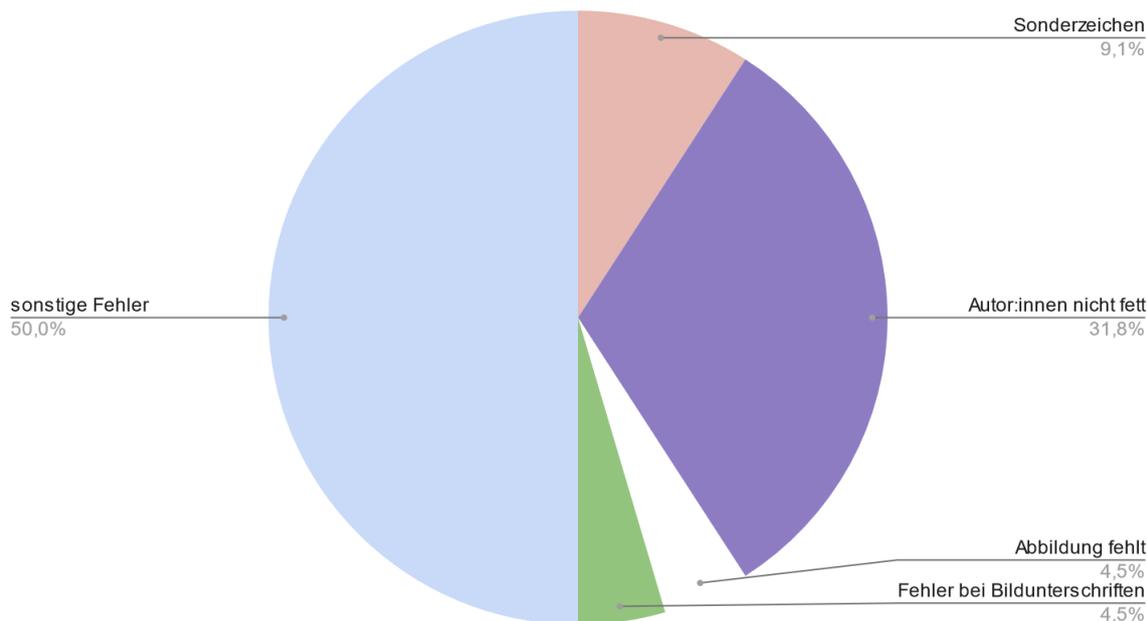


Abbildung 6: Fehlerverteilung bei den Beiträgen des Doctoral Consortiums.

Korrektur der Fehler

Bei der Korrektur der Fehler waren sowohl Mitarbeiter:innen und Hilfskräfte der Universität Passau als auch des DHd Data Stewards involviert. Nach einer Unterweisung, worauf bei der Korrektur zu achten ist, die auf den Erfahrungen der Vorjahre basierte, wurden die Beitragskategorien auf verschiedene Personen aufgeteilt und die Beiträge wurden mehrmals auf Fehler untersucht und korrigiert.

Bereits bekannte Fehler, wie die der Sonderzeichen, falsch formatierten Autor:innennamen, fehlenden Abbildungen und Tabellen sowie Fehler in der Formatierung der Bildunterschriften konnten schnell erkannt und korrigiert werden.

Weil die Kategorie "sonstige Fehler" eine offene Kategorie ist, in der bisher noch nicht bekannte Fehler verzeichnet wurden, wurde dort alles aufgenommen, was ungewöhnlich erschien. Da unterschiedliche Personen unterschiedliche Beiträge bearbeiteten, wurde individuell entschieden, was als Fehler aufgenommen wurde und was nicht. Diese Fehler wurden gemeinsam besprochen und manche dieser Fehler wurden als korrekturwürdig eingeordnet, andere wiederum nicht. Dies führte bei manchen Fällen zu doppelter Arbeit, da schon korrigierte Beiträge nochmals

bearbeitet werden mussten, je nachdem ob ein Fehler bereits als solcher gewertet wurde, bevor er offiziell im Plenum als solcher deklariert wurde. Als Beispiel hierfür kann der vergrößerte Zeilenabstand genannt werden, der bei manchen Beiträgen entstand, wenn Fußnoten in der Zeile vorhanden waren. Dieser Fehler wurde bspw. automatisiert korrigiert, weshalb hierzu keine Statistik existiert.

Alle Beiträge wurden dreimal auf Fehler überprüft, bevor sie auf Zenodo hochgeladen und veröffentlicht wurden.

Fazit

Bei der DHd Jahreskonferenz 2024, die von der Universität Passau ausgerichtet wurde, wurden 144 Beiträge in den Kategorien Poster, Panels, Workshops, Vorträge und Doctoral Consortium angenommen. Diese Beiträge wurden mehrmals auf Fehler überprüft, bevor sie auf Zenodo hochgeladen und publiziert wurden. Insgesamt wurden mindestens 213 Fehler korrigiert, wobei diese Zahl höher angenommen werden muss, da jeder Beitrag, der einen oder mehrere Fehler einer Kategorie beinhaltet, nur einmal in der jeweiligen Kategorie gezählt wurde.

Bereits bekannte Fehler wurden effizient erkannt und behoben. Fehler der Gruppe "sonstige Fehler" wurden nach Auffinden gemeinsam evaluiert und je nach Häufigkeit und Auswirkungen auf die PDF-Datei individuell als zu korrigierender Fehler gewertet. Grundsätzlich verlief der Prozess gut, kann aber dennoch verbessert werden. Eine regelmäßige Auswertung der Kategorie "sonstige Fehler" und die Erweiterung der Liste schon bekannter Fehler kann zu einem gezielteren Korrekturansatz führen, wenn die Korrigierenden wissen, worauf geachtet werden muss.

Zudem wäre eine standardisierte und detailliertere Art der Fehlerdokumentation während der Korrektur hilfreich, um genauere Auswertungen zu erstellen.

Eine mögliche Option, um Fehler in den TEI-XML-Versionen der Beiträge zu vermeiden, ist die Einschränkung von Formatierungsmöglichkeiten für die Autor:innen. Dies könnte beispielsweise durch die Etablierung eines alternativen Tools zur Erstellung von Beiträgen zu DHd Jahreskonferenzen erreicht werden. Ein

potenzieller Kandidat hierfür stellt der FidusWriter⁵ da, der sich diesbezüglich aktuell in der Evaluierungs- und Anpassungsphase befindet.

Referenzen

Helling, Patrick, Anke Debeller und Rebekka Borges: "Konferenzbeiträge strategisch publizieren. Automatisierte Workflows zur individuellen Veröffentlichung von Konferenzbeiträgen am Beispiel des Verbands Digital Humanities im deutschsprachigen Raum e.V." in O-Bib. 2022a. Das Offene Bibliotheksjournal / Herausgeber VDB 9(3), 1-17. DOI: <https://doi.org/10.5282/o-bib/5835>.

Helling, Patrick, Rebekka Borges, Ingo Börner, Anna Busch, Fabian Cremer, Anke Debeller, Henning Gebhard, Harald Lordick und Timo Steyer: "Der DHd-Verband und seine Abstracts - Betrachtungen des Einreichungsprozesses zu DHd-Jahrestagungen", DHd-Blog. 2022b. Online: <https://dhd-blog.org/?p=18599> (letzter Zugriff 21.05.2024).

⁵ <https://www.fiduswriter.org/>, letzter Zugriff 21.05.2024.