

Offen für alle(s)? Open Identities im Reviewprozess der DHd-Konferenz

Burghardt, Manuel

burghardt@informatik.uni-leipzig.de
Universität Leipzig

Czmiel, Alexander

czmiel@bbaw.de
Berlin-Brandenburgische Akademie der Wissenschaften

Dieckmann, Lisa

lisa.dieckmann@uni-koeln.de
Universität zu Köln

Guhr, Svenja

svenja.guhr@tu-darmstadt.de
Technische Universität Darmstadt

Jacke, Janina

janina.jacke@uni-goettingen.de
Technische Universität Darmstadt

Reiter, Nils

nils.reiter@uni-koeln.de
Universität zu Köln

Scholger, Walter

walter.scholger@uni-graz.at
Universität Graz

Wuttke, Ulrike

ulrike.wuttke@fh-potsdam.de
Fachhochschule Potsdam

Einleitung

Die Begutachtung von Forschungsbeiträgen ist ein zentraler Pfeiler wissenschaftlicher Qualitätssicherung, sei es für Zeitschriften, Konferenzen oder drittmittelfinanzierte Forschungsprojekte. Dafür, *wie* diese Begutachtung konkret abläuft, gibt es unterschiedliche Modelle, Gepflogenheiten, Erfahrungen und Erwartungen.

Das bei DHd-Konferenzen bis inklusive 2020 verwendete Modell sah eine Teilanonymisierung vor, d.h. Autor:innen waren den Gutachter:innen namentlich bekannt, jedoch nicht umgekehrt (sog. *single-blind*-Modell). Die Gutachten selbst (Text und zahlenmäßige Bewertung) wurden nur den Autor:innen, allen Gutachter:innen des Beitrags und dem Programmkomitee mitgeteilt. Dieses Modell war Gegenstand von Diskussionen auf den DHd-

Mitgliederversammlungen 2019 und 2020, bis 2020 beschlossen wurde, für die nächste DHd-Konferenz ein *zero-blind*-Modell zu erproben – mit der Einschränkung, dass die Namen der Gutachter:innen nur den Autor:innen des Beitrags und dem Programmkomitee bekannt sind, aber nicht insgesamt veröffentlicht werden.

Wir möchten mit diesem Panel der DHd-Community die Möglichkeit geben, sich über das Begutachtungsverfahren der DHd-Jahrestagungen zu informieren und auszutauschen. Zum einen ist das Zeitkorsett einer Mitgliederversammlung eng begrenzt, was eine wirkliche Diskussion schwierig macht. Zum anderen betrifft das Begutachtungsverfahren auch zahlreiche Nicht-Mitglieder.

Begriffe und Problemstellung

Für eine zielführende und produktive Diskussion sollen auf dem Panel zunächst einige zentrale Begriffe geklärt werden, die in der Diskussion häufig vorkommen. Unter dem Begriff “open peer review” wird ein Bündel unterschiedlicher konkreter Verfahren zusammengefasst, die unterschiedliche Aspekte des Reviewprozesses öffnen:

Open Reports

Mit “open reports” oder “open reviews” ist gemeint, dass die zu einem Beitrag geschriebenen Gutachten öffentlich einsehbar sind, also über den Kreis der Autor:innen des Beitrages hinaus. Ob diese als eigenständige Publikationen adressier- und zitierbar sind oder dem Beitrag quasi als eine Art Anhang beigelegt werden, ist nicht festgelegt, wirft aber neue Fragen auf: Können die Gutachten als eigenständige Publikation anonym bleiben? Können sie dem Beitrag als Anhang angefügt werden, sodass der Beitrag immer mit den Gutachten gelesen werden wird? Und, im konkreten Verfahren: Wenn die Beiträge nach Kenntnisnahme der Gutachten von den Autor:innen überarbeitet werden, müssen dann auch die Gutachten nochmal überarbeitet werden?

Trotz dieser Fragen liegen Vorteile offener Reviews klar auf der Hand: Spätere Leser:innen der begutachteten Beiträge können kritische oder umstrittene Punkte direkt identifizieren, und nicht zuletzt können offene Reviews auch als Beispiele für Erst-Gutachter:innen dienen, wie Reviews aussehen (können).

Open Identities

Bei “open identities” handelt es sich um den am kontroversesten diskutierten Aspekt. Grundsätzlich können analog zur Verwendungsweise in den experimentellen Wissenschaften verschiedene Stufen von “blindness” unterschieden werden: Beim *double-blind-Verfahren* (Abb. 1) sind sowohl Autor:innen als auch Gutachter:innen gegenseitig anonym. Aus praktischen Gründen (etwa um Interessenkonflikte bei der Zuordnung zu verhindern) muss eine Instanz existieren, die Autor:innen und Gutachter:innen namentlich kennt (das Programm- oder Organisationskomitee, bei großen Konferenzen auch sog. *area chairs*, die für einen thematischen, abgegrenzten Bereich zuständig sind). Das *double-blind*-Verfahren wird z.B. bei den Konferenzen der *Association for Computational Linguistics* (ACL) verwendet, wobei die Bekanntheit zwischen den Gutachter:innen unterschiedlich gehandhabt wird.

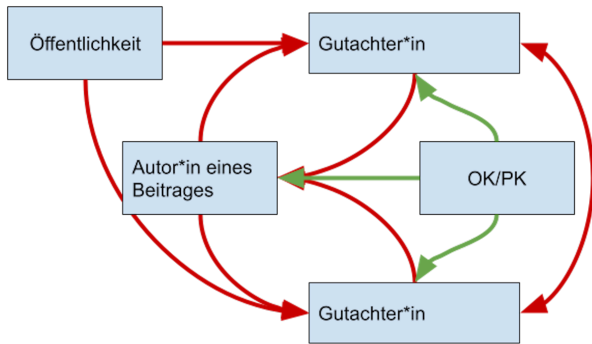


Abb. 1: *Double-blind*-Verfahren. OK/PK: Organisations- bzw. Programmkomitee. Roter Pfeil: Name nicht bekannt, grüner Pfeil: Name bekannt. Die Abbildung bezieht sich auf jeweils einen Beitrag, d.h. der Pfeil zwischen den Gutachter:innen repräsentiert die Bekanntheit zwischen den Gutachter:innen, die für den gleichen Beitrag zur Begutachtung eingeteilt wurden.

Ein sog. **single-blind -Verfahren** wurde seit Anbeginn für die DHd-Konferenzen angewendet. Dabei kennen die Gutachter:innen eines Beitrages dessen Autor:innen, aber nicht umgekehrt (Abb. 2).¹

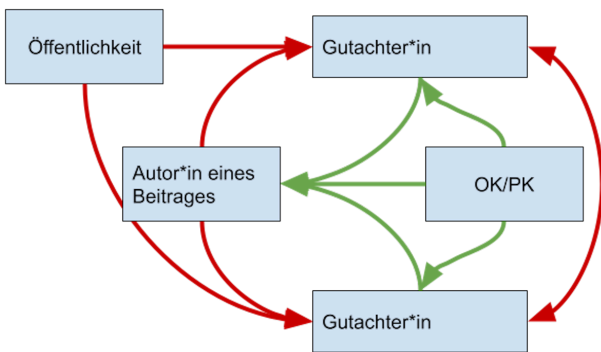


Abb. 2: *Single-blind*-Verfahren, das für die DHd-Konferenzen bis zur DHd 2020 angewendet wurde.

Im **zero-blind -Verfahren** sind sich sowohl Gutachter:innen als auch Autor:innen gegenseitig bekannt (Abb. 3). Dieses Verfahren wird bei der DHd2022 erstmals und testweise angewendet. Unterschiedlich gehandhabt wird, für wen genau die Gutachter:innen namentlich bekannt sind. Bei der DHd2022 werden die Gutachter:innen namentlich nur den jeweiligen Autor:innen (und dem Programmkomitee) bekannt gemacht.

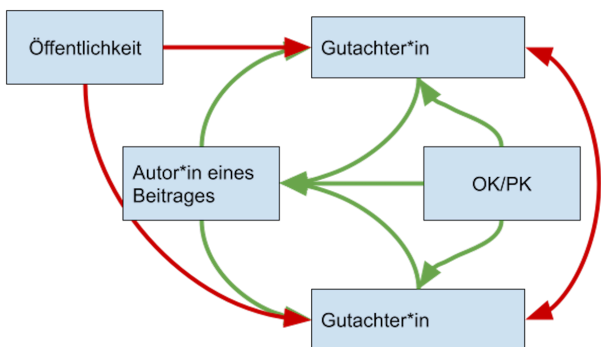


Abb. 3: DHd-Reviewing für die DHd 2022.

Open Participation

Bei den meisten Begutachtungsverfahren ist eine Instanz vorgesehen, die Gutachter:innen Beiträge zuweist, basierend auf Kompetenz, Interesse und/oder der Vermeidung von Interessenkonflikten. Bei "open participation" entscheiden Gutachter:innen selbst, ob sie einen Beitrag begutachten oder nicht, wobei die Kenntnis über mögliche Beiträge vorausgesetzt wird (d.h. die Beiträge müssen den potentiellen Gutachter:innen zugänglich sein). In der Praxis zieht dieses Verfahren weitere Fragen nach sich: Wie wird sichergestellt, dass alle Beiträge begutachtet werden und auch gleich viele Gutachten erhalten? Was ist die Motivation von Gutachter:innen, Beiträge zu begutachten, und wird dadurch eine (neue) Schieflage eingeführt?

Wir werden uns im Folgenden auf den Aspekt der *identities* und hier auf die Optionen *single*- und *zero-blind* konzentrieren, da sie für die DHd-Konferenzen im Mittelpunkt der Diskussion stehen, und typischerweise die größten Kontroversen auslösen (vgl. Ross-Hellauer/Görögh, 2019). Im Panel allerdings können auch andere Aspekte im Verlauf der Diskussion aufgegriffen werden.

Erfahrungen

In verschiedenen Disziplinen wurde und wird mit unterschiedlichen Begutachtungsverfahren experimentiert. Die Erfahrungen dabei sind unterschiedlich, und oft wird darüber vor allem anekdotisch oder in semi-öffentlichen Kreisen berichtet (wie z.B. von der DH2020). Ein Grund dafür ist sicher, dass es schwer ist, belastbare, vergleichende Studien zum Thema durchzuführen, da der Review-Prozess als sensibel gilt und eine große Zahl an Faktoren die Ergebnisse beeinflussen. Ein solches Experiment – das sich nicht mit open vs. blind beschäftigt hat, sondern mit der Konsistenz von Begutachtung im Allgemeinen – ist das sog. NeurIPS-Experiment (Cortes/Lawrence 2021): Dabei wurden 10% oder 170 der Einreichungen bei der *machine-learning*-Konferenz NeurIPS zwei unabhängigen Programmkomitees zugewiesen, die selbstständig ihre Entscheidung getroffen haben (angenommen wurden Beiträge, wenn sie von mind. einem Komitee angenommen wurden). Bei ca. einem Viertel der Beiträge kamen die Komitees zu unterschiedlichen Entscheidungen. Auch wenn dieses Experiment sich nicht mit open vs. blind beschäftigt, zeigt es vielleicht einen Weg auf, wie generell eine Qualitätssicherung des Reviewprozesses aussehen könnte. Zu Bedenken ist allerdings, dass eine Übertragung von Verfahren aus anderen Disziplinen schon deswegen schwierig ist, weil diese sich in Größe, Diversität, Kompetitivität und Publikationspraxis massiv unterscheiden.

Argumente für die Bekanntmachung von Gutachter:innen (pro Open Identities)

Eine einseitige Anonymität der Gutachter:innen, wie sie bislang bei den DHd-Jahrestagungen praktiziert wurde, bringt Nachteile für die Autor:innen mit sich: Ein in den Mitgliederversammlungen 2019 und 2020 benannter Kritikpunkt ist die Abgabe von "Einsatz-Gutachten" ohne ernsthafte Auseinandersetzung mit den eingereichten Beiträgen, geschweige denn konstruktiven Verbesserungen.

rungsvorschlägen für die Autor:innen. Im Gegensatz dazu steigen Umfang und Konstruktivität der Gutachten in offenen Begutachtungsprozessen nachweislich, da die Verbindlichkeit und Rechenschaftspflicht der Bewertungen von Gutachter:innen durch die Offenlegung ihrer Identität deutlich zunimmt (Besançon/Rönning/Löwgren 2020:7 und Pucker/Schilbert/Schumacher 2019:3). Eine mildere Bewertung der Beiträge ging damit bei der DH2020 zumindest nicht einher (Guiliano/Estill 2020:13, interner Bericht). Des Weiteren konnte auch kein Rückgang der Bereitschaft zur Begutachtung bei der DH2020 festgestellt werden: Nur fünf der vorjährigen Gutachter:innen verweigerten die Begutachtung mit Hinweis auf das offene Verfahren, während die Gesamtzahl der Gutachter:innen von 779 auf 974 anstieg.

Befürworter:innen eines anonymen Begutachtungsverfahrens argumentieren häufig damit, dass sich Autor:innen durch die Offenlegung ihrer Forschung sowie die Reviewer:innen durch die Formulierung ihrer Reviews angreifbar machen und eine Anonymisierung des Reviewverfahrens als Schutzmechanismus fungiere. Gerade in einer verhältnismäßig überschaubaren Fachgemeinschaft wie den Digital Humanities kann diese Anonymität jedoch nur geringfügig sichergestellt werden, wodurch eine tatsächliche Anonymität der Autor:innen fragwürdig ist oder nur durch umständlichen Anonymisierungsaufwand in den Beiträgen gewährleistet werden.

Gerade bei einer fachlich weit gestreuten Konferenz wie der DHd-Reihe – und einem ebenso breit gefächerten Pool an Gutachtenden – dürfte die Offenlegung der Identitäten und damit der fachlichen Expertisen dazu beitragen, dass Gutachter:innen nur Beiträge begutachten, zu denen eine fachliche Nähe gegeben ist (Pucker/Schilbert/Schumacher 2019:3), was nicht nur konstruktive Rückmeldungen begünstigt, sondern auch negative Gutachten aufgrund fachlicher Unkenntnis verhindert.

Nicht zuletzt kann die offene Kommunikation zwischen Autor:innen und Gutachter:innen nicht nur zu einer Qualitätssteigerung des eingereichten Beitrages und zu weiterführenden Diskussionen führen, sondern auch Impulse für zukünftige Forschungs- und Publikationsvorhaben sowie Kollaborationen geben (Besançon/Rönning/Löwgren 2020: 5). Offene Identitäten im Reviewprozess ermöglichen somit generell eine transparentere Gestaltung, die eine bessere Nachvollziehbarkeit der Annahmen und Ablehnungen von Beiträgen zur DHd-Jahrestagung gewährleistet.

Argumente für die Anonymität von Gutachter:innen (Contra Open Identities)

Mit der Ablehnung von Beiträgen müssen naturgemäß auch schlechte Nachrichten überbracht werden. Autor:innen wiederum reagieren unterschiedlich auf ablehnende Bewertungen, was von den Gutachter:innen schlecht oder gar nicht vorgesehen werden kann. 'Blindness' erlaubt es den Gutachter:innen, ihre Gutachten unabhängig von solchen Erwägungen zu schreiben. Prekär beschäftigte Nachwuchswissenschaftler:innen müssen keine Angst haben, Beiträge etablierter Wissenschaftler:innen negativ zu beurteilen, selbst wenn diese ggf. für das berufliche Fortkommen (z.B. für Forschungsanträge oder Empfehlungsschreiben) relevant sind. Es ist, damit dieser Effekt eintritt, auch nicht entscheidend, ob die etablierten Wissenschaftler:innen in solchen Positionen sind, oder ihre Macht für diese Art von 'Racheaktionen' ausnutzen würden: Die 'Beißhemmung' tritt ja ein, wenn Gutachter:innen negative Reaktionen nur befürchten, unabhängig davon, ob sie eintreten.

Ein möglicher langfristiger Effekt ist die Verkleinerung des Pools an Gutachter:innen, sowie dessen fortschreitende Ent-Diversifikation, da z.B. Professuren nach wie vor überdurchschnittlich häufig von Männern besetzt werden.

Selbst wenn Indizien dafür sprechen, dass offene Gutachter:innen-Namen zu besseren Reviews führen, besteht die Gefahr, dass Gutachten zu oberflächlich bleiben, weil sich Gutachter:innen nicht mehr trauen, subjektive und ggf. spekulative Aspekte anzumerken, welche die Qualität von Beiträgen erheblich steigern können. Insgesamt ist vermutlich eine starke Tendenz zu positiven Reviews zu erwarten, was die Unterscheidbarkeit der Qualität der Beiträge deutlich erschweren würde. Dieser Effekt wurde bereits anhand eines kontrollierten Experiments des *British Journal of Psychiatry* empirisch untermauert (Walsh et al., 2018): "Reviewers who signed were more likely to recommend publication." Wenn mehr Beiträge positiv (und damit ähnlich) bewertet werden, müssen mehr Entscheidungen durch das Programmkomitee getroffen werden. Dies steigert nicht nur den Arbeitsaufwand für wenige Personen, sondern ist auch für Autor:innen weniger transparent.

Analyse der DHd2022-Erfahrungen

Nach Abschluss des Review-Prozesses werden die Reviews vergleichend analysiert, um auf die Argumente pro und contra *open identities* unterstützend/entkräftend reagieren zu können. Die Reviews sowie Bewertungen der vergangenen DHds in Köln (2018) und Paderborn (2019) stehen als Vergleichswerte zur Verfügung. Die Bewertungen können direkt quantitativ verglichen werden. Bei den Reviews kommen Faktoren wie die Länge, genannte Referenzen, oder auch *sentiment*-Scores in Betracht (die aber mit Vorsicht zu interpretieren sind). Darüber hinaus soll ein Interview mit der Ombudsstelle in die Analyse einfließen.

Besetzung und Ablauf des Panels

Die Panelist:innen werden mit einleitenden Impulsvorträgen aus ihrer Expertise und Erfahrungen die Ausgangslage und Problemfelder skizzieren und nach einem Austausch innerhalb des Panels in eine offene Diskussion mit dem Plenum übergehen.

Der Ablauf des Panels ist wie folgt geplant (Gesamtdauer 90 Minuten):

- Vorbereitete Beiträge (zusammen ca. 30 Minuten)
 - Einleitung und Begriffsklärung (Alexander Czmiel)
 - Pro Bekanntgabe der Gutachter:innen (Svenja Guhr, Walter Scholger)
 - Pro Anonymität der Gutachter:innen (Janina Jacke, Nils Reiter)
 - Analyse der DHd2022-Erfahrungen (Lisa Dieckmann)
- Diskussion innerhalb des Panels (30 Minuten)
- Plenardiskussion (30 Minuten)

Fußnoten

1. Theoretisch denkbar wäre auch eine Anonymität in umgekehrte Richtung, also bekannte Gutachter:innen und anonyme Autor:innen. Dieses Verfahren wird von Pontille/Torny (2014) als "Blind review" bezeichnet.

Bibliographie

Besaçon, Lonni / Rönnerberg, Niklas / Löwgrén, Jonas et al. (2020): “Open up: a survey on open and non-anonymized peer reviewing”, in: *Res Integr Peer Rev* 5(8) 10.1186/s41073-020-00094-z

Burghardt, Manuel / Dieckmann, Lisa / Reiter, Nils / Steyer, Timo / Scholger, Walter / Trilcke, Peer / Wuttke, Ulrike (2021): *Besseres Reviewing für die DHd (Version 1.0)*. Zenodo 10.5281/zenodo.4633633

Cortes, Corinna / Lawrence, Neil D. (2021): Inconsistency in Conference Peer Review: Revisiting the 2014 NeurIPS Experiment. arXiv:2109.09774v1. <https://arxiv.org/abs/2109.09774>

Guiliano, J. / Estill, L. (2020): DH2020 Report to ADHO CCC
Nature America Inc. (1999): “Pros and cons of open peer review”, in: *Nat Neurosci* 2: 197–198 10.1038/6295

Pontille, David / Tornay, Didier (2014): “The blind shall see! The question of anonymity in journal peer review”, in: *Ada: A Journal of Gender, New Media, and Technology* 4 10.7264/N3542KVV

Pucker, Boas / Schilbert, Hanna Marie / Schumacher, Sina Franziska (2019): “Integrating Molecular Biology and Bioinformatics Education”, in: *Journal of Integrative Bioinformatics* 16(3) 10.1515/jib-2019-0005

Ross-Hellauer, Tony / Görögh, Edit (2019): “Guidelines for open peer review implementation”, in: *Res Integr Peer Rev* 4(4) 10.1186/s41073-019-0063-9

Ross-Hellauer, Tony (2017): “What is open peer review? A systematic review”, in: *F1000Research* 6: 588 10.12688/f1000research.11369.2

Walsh, E. / Rooney, M. / Appleby, L. / Wilkinson, G. (2000): “Open peer review: A randomised controlled trial”, in: *British Journal of Psychiatry*, 176(1): 47-51 10.1192/bjp.176.1.47