

Die Augen der Vielen: Plagiate unter Open Access

„Given enough eyeballs, all bugs are shallow“

Ulrich Herb

Open Access Referent Saarländische Universitäts- und
Landesbibliothek & Wissenschaftsberater

**Plagiate & Co - Wissenschaftliches Fehlverhalten ist (k)ein Kavaliersdelikt
Universität Mainz, 30.05.2012**

Agenda

- Plagiate? Closed Access? Open Access?
- Plagiate unter Open Access
- Genügen die Augen der Vielen?
- Wie offen muss Wissen sein?
- Genügt offener Zugang zu Text-Publikationen?
- Fazit

Plagiate?
Closed Access?
Open Access?

Plagiat

„Ein **Plagiat** (...) ist die Anmaßung fremder geistiger Leistungen. Dies kann sich auf die Übernahme fremder Texte oder anderer Darstellungen (z. B. Zeitungs-, Magazinartikel, Fotos, Filme, Tonaufnahmen), fremder Ideen (z. B. Erfindungen, Design, Wissenschaftliche Erkenntnisse, Melodien) oder beides gleichzeitig (z. B. Wissenschaftliche Veröffentlichungen, Kunstwerke, Romane) beziehen.“

<http://de.wikipedia.org/wiki/Plagiat> [24.05.2012]

Closed Access

Closed Access/ Toll Access

- Nutzung wissenschaftlicher Information gegen Zahlung (Subskription, Pay per View)
- Abtretung meist ausschließlicher Nutzungsrechte der Autoren an Verlage

Plagiate & Closed Access

- Hans-Peter Schwintowski, 2005
- Madonna Constantine, 2008
- Bruno Frey, 2011
- Silvana Koch-Mehrin, 2012
- Pál Schmitt (ungarischer Staatspräsident), 2012
- Karl-Theodor zu Guttenberg: Entlarvt, weil er zahlreiche Online-Quellen nutzte, 2011
- Plagiatsdiskussion um Johannes Hahn, EU-Kommissar für Regionalpolitik und ehemaliger Wissenschaftsminister: Dissertation liegt nur im Print vor, zitierte Quellen ebenfalls.
- Bildautoplugiate, z.B. durch Jan Hendrik Schön in Nature- und Science-Publikationen, die unterschiedliche Effekte belegen sollten
Gerd Fröhlich, G. (2006). Plagiate und unethische Autorenschaft. *Information - Wissenschaft & Praxis*, 57(2), S. 81-89. <http://eprints.rclis.org/6014/1/plagiate.pdf>

Open Access

Open Access = entgeltfreier Zugang zu wissenschaftlichen Informationen.

- Green Road Open Access bezeichnet das Veröffentlichen von wissenschaftlichen Texten, die bereits formal in einem regulären Verlagsangebot erschienen sind (sogenannte Postprints) oder deren Vorabversionen (Preprints) auf Open-Access-Servern (Repositories).
- Golden Road Open Access bezeichnet die Herausgabe wissenschaftlicher Open-Access-Journale bzw. das Publizieren in solchen Zeitschriften oder das Publizieren anderer Dokumentarten wie etwa Monografien in einem Open-Access-Verlag. Der Goldene Weg produziert originäre Publikationen, also *Born Open Access Documents*.

Plagiate unter Open Access

Plagiate unter Open Access

These (von Open Access Gegnern)

- Open Access, der freie Zugang zu Informationen im WWW, fördert Plagiarismus.

Meine Gegenthesen

- Open Access erleichtert die Aufdeckung von Plagiarismus und kann diesen bekämpfen.
- Plagiarismus wird durch Open Access riskanter.
- Im Open Access herrscht transparenterer Umgang mit Plagiaten als im Closed Access.

„Given enough eyeballs, all bugs are shallow“

Linus Law, benannt nach dem Namensgeber und Initiator der Open Source Software Linux, Linus Torvalds

Prinzip fördert Qualität

- beim Produzieren von Programmcode
- bei der Sammlung von Geodaten
Roland Ramthun (2012). Offene Geodaten durch OpenStreetMap. In U. Herb (Hrsg.), Open Initiatives: Offenheit in der digitalen Welt und Wissenschaft (S. 159-184). <http://universaar.uni-saarland.de/monographien/volltexte/2012/87/>
- bei Texten
Jim Giles (2005). Internet encyclopaedias go head to head. *Nature* **438**, 900-901 (15 December 2005) [doi:10.1038/438900a](https://doi.org/10.1038/438900a); Published online 14 December 2005
- bei der Aufdeckung von Plagiaten
Guttenplag, http://de.guttenplag.wikia.com/wiki/GuttenPlag_Wiki

...

Junk Journals / Bogus Journals

- *Illegitime wissenschaftliche Selbsthilfe*

Pakistanische Higher Education Commission verlangt von Bewerbern Publikationen in “peer-reviewed, recognized foreign journals”. Folge: Ausgewanderte Pakistanis gründen das *Journal of Applied Pharmacy* in Kanada

Ralf Neumann (2012): Von zweifelhaftem Verhalten der Journals. In: Laborjournal Blog
<http://www.laborjournal.de/wordpress/?p=4223#more-4223>

Jeffrey Beall (2012): The Story of Tahira Mughal. Metadata Blog
<http://metadata.posterous.com/the-story-of-tahira-mughal-assistant-professo>

- Jeffrey Beall's List of predatory Open Access Publishers
<http://metadata.posterous.com/83235355>

Plagiate & Open Access Repositories: arXiv

arXiv, <http://arxiv.org/>:

Open Access Repository für Dokumente aus Physik, Mathematik, Informatik, Statistik und verwandter Fächer, 758.000 Volltexte [24.05.2012].

2007: 65 Plagiate werden vom Server entfernt, <http://arxiv.org/new/withdrawals.aug.07.html>

Statement Mustafa Saltis, der als Mitautor von 40 der 65 Artikel agierte, gegenüber Nature:

“Most of our papers have **been published in the science citation index journals**. Until **now no one has claimed that we plagiarize**.” Geoff Brumfiel (2007). Turkish physicists face accusations of plagiarism. *Nature* **449**, 8 (6. September 2007). [doi:10.1038/449008b](https://doi.org/10.1038/449008b)

Aktuell: Maschineller Test beim Upload von Dateien, intellektuelle Prüfung der Dokumente, Kontaktaufnahme mit Autoren, ggf. Entfernen oder Ablehnung von Dokumenten

Plagiate & Open Access Repositories: arXiv

An exact solution of the moving boundary problem for the relativistic plasma expansion in a dipole magnetic field

H. B. Nersisyan, K. A. Sargsyan, D. A. Osipyan, H. H. Matevosyan

(Submitted on 4 Apr 2012)

An exact analytic solution is obtained for a uniformly expanding, neutral, highly conducting plasma sphere in an external dipole magnetic field with an arbitrary orientation of the dipole moment in the space. Based on this solution the electrodynamic aspects related to the emission and transformation of energy have been considered. The results obtained can be used to treat qualitatively experimental and simulation data, and several phenomena of astrophysical and laboratory significance.

Comments: 11 pages, 2 figures. arXiv admin note: substantial text overlap with [arXiv:physics/0603239](#) and [arXiv:1007.0250](#)

Subjects: **Plasma Physics (physics.plasm-ph)**; Mathematical Physics (math-ph); Space Physics (physics.space-ph)

Cite as: [arXiv:1204.0874v1](#) [physics.plasm-ph]

Submission history

From: Hrachya Nersisyan [[view email](#)]

[v1] Wed, 4 Apr 2012 06:34:47 GMT (103kb)

[Which authors of this paper are endorsers?](#)

- **Entsprechend markierte Dokumente sind *minor cases***
- ***Major cases* werden entfernt**
- **Ca. 1.200 von 758.000 Dokumenten $\approx 0,16\%$ *minor cases***

Plagiate & Open Access Repositories: RePEc

RePEc, <http://repec.org/>

Research Papers in Economics, ca. 1,2 Mio Texte nachgewiesen [24.05.2012]

- Plagiatspolicy: <http://plagiarism.repec.org/procedure.html>
- Liste der *Offender*: <http://plagiarism.repec.org/offenders.html>

Überschneidung mit Empfehlungen des Nature Specials Science publishing: How to stop plagiarism. (2012). *Nature*, 481(7379), 21-23. [doi:10.1038/481021a](https://doi.org/10.1038/481021a)

- Harold Garner: Flag plagiarized studies
- Yuehong Zhang & Ian McIntosh: Blacklist repeat offenders

Plagiate & Open Access: OA Plagiatssuche

Open Access Plagiarism Search OAPS, <https://oaps.eu/>

- Nach Registrierung können Dokumente ins System geladen und auf Plagiatsverdacht geprüft werden.
- Zusätzlich: Abfrage über Schnittstelle möglich
- Korpus: 12 Mio. Dokumente

Zwischenfazit

- Der Annahme, ubiquitäre Verfügbarkeit von Texten im WWW begünstige Plagiarismus, steht entgegen, dass Plagiarismus im Open Access riskanter (weil leicht zu entlarven) wird.
- Plagiatsfälle werden durch Geheimhaltung genährt, Plagiateure durch Closed Access geschützt
- Open Access Communities pflegen transparenten Umgang mit Open Access

Genügen die Augen der Vielen? Genügt entgeltfreie Nutzbarkeit?

Augen genügen nicht immer ...

- Automatische Plagiatssuche (arXiv, OAPS, CrossCheck, etc.) benötigt Zugriff auf DRM-freie Volltexte
- Dateien müssen maschinell verarbeitbar sein
- Geschäftsmodell vieler Subskriptionsverlage setzt auf restringierten Zugang zu Dokumenten und Verknappung von Information

Restriktion & Verknappung

- *Article Rental Scheme*, das nur das reine Lesen des Artikels erlaubt: Download, Ausdruck oder Kopieren von Inhalten sind nicht möglich, die Zugriffsdauer ist auf 24 Stunden beschränkt.
- Pixel-PDFs, DRM-versiegelte PDFs, ...
- Verlage (Elsevier, Wiley, Macmillan, ...) untersagen via Lizenzbedingung *Text Mining* und verhindern Forschung und Innovation

Alok Jha (2012). Text mining: what do publishers have against this hi-tech research tool? *The Guardian*, (23. Mai 2012).

<http://www.guardian.co.uk/science/2012/may/23/text-mining-research-tool-forbidden>

Vom Burger zur Kuh

„Extracting information from general PDFs is impossible, and likened to ‘converting a hamburger back to a cow’ (I am sometimes credited with this aphorism but I didn't create it). A generic PDF may be a bitmap, contain only vector strokes, and may have ‘order backwards in words’. However for scientific publications which are largely mechanised there is quite a lot that can be done.”

Peter Murray-Rust (2012). Extracting and indexing information from scientific literature ("the PDF Cow").

Open Science Mailinglist, 18. April 2012,

<http://web.archiveorange.com/archive/v/6YDTDOvDHdJQMqr4OhYo>

Open Access & Open Knowledge: Wie offen muss das Wissen sein?

Open Access: frei oder offen?

Diskussion um *Gratis Open Access*, *Libre* oder *True Open Access*:

- *Gratis*: Reine entgeltfrei Nutzungsmöglichkeit
- *True*: Texte (oder Daten) können analog der Open Source Prinzipien genutzt werden

Urheberrechte bleiben in allen Szenarien gewahrt!

Open Definition: Versuch, Open Source Prinzipien auf alle Informationsinhalte zu übertragen.

Open Knowledge

Offenes Wissen nach Open Definition ist Wissen, das (entgelt- und restriktions)frei

- benutzt (z.B. gelesen, analysiert)
- weiterverwendet (z.B. neu ausgewertet, modifiziert und mit anderen Daten kombiniert)
- weiterverteilt und kopiert, also zur Nutzung durch andere angeboten werden kann.

Weitere Bedingungen:

- Zugänglichmachung in einer technisch leicht zu handhabenden und veränderbaren Form
- Verwendung offener Dateiformate

<http://opendefinition.org/okd/deutsch/>

Ulrich Herb (2012). Offenheit und wissenschaftliche Werke: Open Access, Open Review, Open Metrics, Open Science & Open Knowledge. In U. Herb (Hrsg.), *Open Initiatives: Offenheit in der digitalen Welt und Wissenschaft*.

<http://universaar.uni-saarland.de/monographien/volltexte/2012/87>

**Genügt offener Zugang zu Text-
Publikationen?**

**Oder muss deren Kontext
ebenfalls offen gestaltet sein?**

Open Review: Plagiatsvermeidung & Qualitätssicherung

Offene Begutachtung und die Verhinderung von Plagiaten

Unterschiedliche Ausformungen

- Einreichungen werden Open Access gestellt
- Gutachten werden Open Access gestellt
- Gutachterkreis variabel (bestellte Peers, registrierte Experten, ...)
- Anonyme oder nicht-anonyme Gutachter

Atmospheric Chemistry and Physics (ACP)

<http://www.atmospheric-chemistry-and-physics.net>

Ulrich Pöschl (2006). Open Access & Collaborative Peer Review: Öffentliche Begutachtung und interaktive Diskussion zur Verbesserung von Kommunikation und Qualitätssicherung in Wissenschaft und Gesellschaft. In S. Hornbostel & D. Simon (Hrsg.), *Wie viel (In-) Transparenz ist notwendig? : Peer Review revisited* (iFQ Working Papers, S. 43-46). Institut für Forschungsinformation und Qualitätssicherung. http://www.forschungsinform.de/publikationen/download/working_paper_1_2006.pdf

Retraction: Transparenter Umgang mit entdeckten Plagiaten?

Retractions

- erfolgen in aller Regel zögerlich
 - sind schlecht dokumentiert (v.a. im Print, aber auch online)
 - im Closed Access schwer nachzuvollziehen
 - in Datenbanken selten nachgepflegt
- Ralf Neumann (2011): Schlecht gepflegt und dann noch frech. *Laborjournal Blog* (21.07.2011), <http://www.laborjournal.de/wordpress/?p=3224>

Open Access zu Informationen über Retractions:

- Retraction Watch, <http://retractionwatch.wordpress.com/>

Fazit

Fazit I

„Der Widerstand gegen das Neue hat noch jede wissenschaftliche Innovation behindert; warum sollte das gerade bei „Open Access“ anders sein? (...)

Open Access ist für Forschungskommunikation (Preprint-Server) wie Wissenschaftskommunikation (peer reviewed Journale, Archivierung referierter Postprints) problemlos geeignet, ermöglicht aber wesentlich besser als konventionelle Papierpublikationen die Aufdeckung von Plagiat und Täuschung und die Kennzeichnung als gefälscht überführter Publikationen.“

Fröhlich, G. (2009). Die Wissenschaftstheorie fordert Open Access. *Information - Wissenschaft & Praxis*, 60(5), 253-258. <http://eprints.rclis.org/16840/>

Fazit II

Open Access und Open Science erleichtern das Aufdecken von Plagiaten (und die Qualitätssicherung allgemein), nicht nur im Falle wissenschaftlicher Text-Publikationen, sondern auch bei:

- Bildplagiaten
- Forschungsdaten
- Artikeleinreichungen
- Projekteinreichungen
- ... allen erdenklichen Items, die im Forschungsprozess anfallen

Allerdings:

- Informationen müssen nach Prinzipien der Open Definition verfügbar gemacht werden, rein kostenloser Zugang ist weder *offen* noch *transparent*.
- Maschinelle Auswertung muss *rechtlich* (Lizenz) und *technisch* (DRM-Freiheit) möglich sein.

Vielen Dank für Ihre Aufmerksamkeit.

u.herb@sulb.uni-saarland.de
<http://www.sulb.uni-saarland.de/de/service/publikationsangebote/>

SAARLÄNDISCHE
UNIVERSITÄTS-UND
LANDESBIBLIOTHEK



u.herb@scinoptica.com
<http://www.scinoptica.com>

[scinoptica]