

# Viel Licht und noch mehr Schatten: Plagiat in Dissertationen

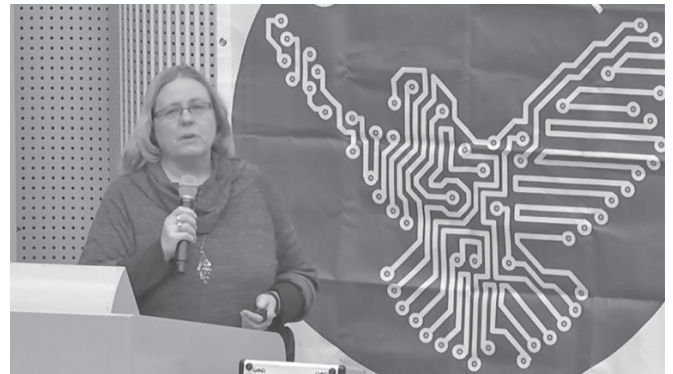
## Zusammenfassung des Vortrags von Debora Weber-Wulff

Seit 2011 werden bei VroniPlag Wiki Plagiate in Dissertationen dokumentiert, aktuell sind 177 Fälle publiziert worden. Obwohl diese Fälle nur die Spitze des Eisbergs sind, reagieren die Hochschulen teilweise sehr träge auf Plagiatsanzeigen, obwohl es durchaus Lichtblicke gibt. Interessanterweise sind die Dissertationen alle veröffentlicht, also eigentlich sichtbar für jeden. Nur, wie macht man Plagiate sichtbar?

FIFFKON 2016

### Kurzer Rückblick auf die Plagiatskandale

Dass es Plagiate gibt, ist nicht neu: Schon aus dem Jahre 1265 ist uns ein Vorfall überliefert; in den letzten Jahren sorgten Plagiate jedoch auch politisch für Aufmerksamkeit. Die Berichterstattung erzeugte dabei den Eindruck, allein die Technik habe ihr Aufdecken ermöglicht. Dies ist allerdings ein Missverständnis, ebenso wie die Annahme, dass es bei der großen Welle der Plagiatsvorwürfe nur um Dissertationen von Politikern gehe. Viel erschreckender ist die lange Liste der aktiven Dozent:innen und Forscher:innen, deren Arbeiten in den vergangenen Jahren als Plagiate enttarnt wurden. Von den insgesamt 177 Vorfällen auf VroniPlag Wiki sind der größte Teil Dissertationen, aber auch Habilitationsschriften und andere wissenschaftliche Arbeiten und Veröffentlichungen sind darunter. Betroffene Universitäten weisen die Verantwortung von sich oder spielen die Fälle herunter mit Hinweisen auf die „Fachkultur“, verfallen jedoch dann in Aktionismus, wenn es um prominente Fälle geht. Auf der anderen Seite stehen die positiven Nachwirkungen: gezielte Plagiatsprävention wie Seminare zu wissenschaftlichem Arbeiten oder das ReFAIRenz-Programm der Uni Konstanz.



Debora Weber-Wulff

gewissermaßen als Quotienten einer Plagiathaftigkeit auswerfen, sind oft eher bedeutungslos und ändern sich auch nicht nachvollziehbar, sobald die Heuristik zufallsbasiert ist. Wenn darauf Abzüge in der Benotung basieren, ist das ein Problem, das alle wissenschaftliche Arbeiten betrifft. Falsch-positive Ergebnisse müssen die Grundannahme sein und gründlich überprüft werden. Dabei bleibt wie in den meisten Fällen computergestützter Entscheidungsfindung die Software nur ein Werkzeug, welches das kritische Lesen der Arbeit nicht ersetzt. Neben Auffälligkeiten wie unterschiedlichen Anführungszeichen oder Links innerhalb der Texte, die durch Kopiervorgänge entstanden sein können, sind Unregelmäßigkeiten in Formulierungen, Grammatik oder Zitierstilen Indizien für Plagiate, die eine genauere inhaltliche Untersuchung nahelegen.

There has always been plagiarism

Übersicht über historische Plagiatsvorwürfe „Historioplak Wiki“

Jahr	Namen	Ort	Fach	Art	Seite in Historioplak Wiki
1265	Arnold Andreas Bull	Kopenhagen	Medizin	D	Arnold Andreas Bull (Kopenhagen 1265)
1836	MacCadden, James Jasper	Berlin	Medizin	D	Jasper MacCadden (Berlin 1836)
1851	Schweitzer, Carl F.	München	Maschinenbau	M	Carl F. Schweitzer (München 1851)
1856	D'Allemand, David	Marburg	Philosophie	DD	David D'Allemand (Marburg 1856)
1857	Schweitzer, C. / Andree, I.	Leipzig	Maschinenbau	M	Carl Schweitzer / I. Andree (Leipzig 1857)
1861	Reigler, Julius	Strasbourg	Medizin	D	Julius Reigler (Strasbourg 1861)
1865	Hilcken, Wenzel	Marburg	Pharmazie	DD	Wenzel Hilcken / Aloys Wilhelm Josten (Marburg 1865)
1865	Josten, Aloys Wilhelm	Marburg	Pharmazie	DD	Wenzel Hilcken / Aloys Wilhelm Josten (Marburg 1865)
1868	Gericke, Curt	Göttingen	Geometrie	D	Curt Gericke (Göttingen 1868)
1871	von Höries, Ludwig II.	Berlin	Rechtswiss.	M	Ludwig von Höries (Berlin 1871)
1872	Heidrich-Land, Michael	Berlin	Medizin	D	Michael Heidrich-Land (Berlin 1872)
1873	Dabbs, Wilhelm	Rostock	Geschichtswiss.	DD	Wilhelm Dabbs (Rostock 1873)
1880	Lütkenmeyer, Adolf Otto	Leipzig	Chemie	DD	Adolf Otto Lütkenmeyer (Leipzig 1880)
1880	Wienberg, Art.	Leipzig	Chemie	DD	Art. Wienberg (Leipzig 1880)

Übersicht über historische Plagiatsvorwürfe

### Methoden zur Aufdeckung

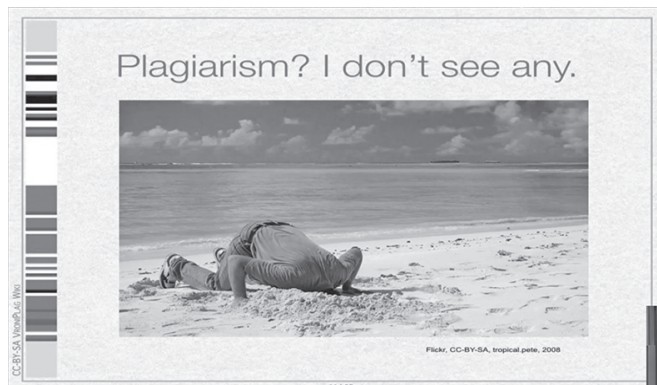
Wie aber nun können Plagiate sichtbar gemacht werden? Hochschullehrer:innen wünschen sich am liebsten eine einfache und zuverlässige Software, die Plagiate „errechnet“. Dadurch ist ein Markt entstanden, auf dem Softwarelösungen mit großen Versprechungen angepriesen werden; am Ende eines so automatisierten Prüfdurchlaufs steht dann eine Zahl, die in Einzelfällen bei Journals sogar schon als Ausschlusskriterium für Textzulassungen verwendet wird. Allein: Ob eine Originalität vorliegt oder nicht, darüber lässt sich nicht so einfach eine zweifelsfreie Aussage treffen, auch ist nicht jede mögliche Quelle der Texte zugänglich – es können höchstens Hinweise auf eventuelle Plagiate geliefert werden. Die Zahlen, die diese Programme

VroniPlag Wiki (<http://de.vroniplag.wikia.com/>) nun verbindet beide Ansätze: Hinter dem Namen verbirgt sich eine Gruppe, die zwar auch Software nutzt, aber hauptsächlich selber liest, Gefundenes dokumentiert und darüber diskutiert. Ausgangspunkt kann dabei der Hinweis eines Whistleblowers mit manchmal nützlichen Zeugenberichten aus den unterschiedlichsten Motiven sein. Aber auch reine Neugier, Spielerei beim Durchforsten und Bauchgefühl können in Zufallsfunden von Plagiaten enden. Darüber hinaus können die algorithmenbasierten Analysen des Textmining als Methode hilfreich sein, um dementsprechende Quellen aufzuspüren. Wenn die Quelle jedoch nicht zu finden ist, kann auch kein Plagiat nachgewiesen werden.

Ein großer Vorteil in Deutschland ist, dass alle Dissertationen publiziert werden müssen. Als ein leichter Einstieg in die Plagiatsaufdeckung stellte sich die Medizin heraus, weil dort verhältnismäßig kurze Texte schnell und oft frei zugänglich publiziert werden. Während die Fakultäten selbst, in denen die betroffenen Texte erschienen sind, leicht auszumachen sind, ist das automatisierte Durchsuchen ihrer Websites (sog. Crawlten) kompli-

zierter. Besonders aufwändig ist die Vorarbeit zum eigentlichen Plagiats-Scan: die Daten aus etwa via Texterkennung nicht ganz fehlerfrei ausgelesenen PDF-Dokumenten zu bereinigen. Erst danach kann z. B. der Sim-Algorithmus von Dick Grune den Abgleich verschiedener Publikationen innerhalb einer Uni, zwischen verschiedenen Hochschulen und auch mit Wikipediaeinträgen gewinnbringend durchführen. Die große Anzahl an Kombinationen von Vergleichstexten stellt sich für die Prüfung dabei immer wieder als erhebliche Herausforderung heraus. Insbesondere gilt es abzuwägen zwischen einer mitunter langen Suchlaufzeit und der Qualität der Ergebnisse bzw. der Anzahl der in die Suche einbezogenen Texte. Plagiatsprüfungssoftware kann dabei freilich nicht nur exakte Übernahmen ganzer Sätze oder Abschnitte erkennen, sondern auch in der Arbeit umsortierte Teilsätze oder -abschnitte. Hilfreich für den Prüfer ist dabei auch die Darstellung: Beim Vergleich zweier Texte markieren farbige Unterlegungen jeweils gleiche Passagen visuell und machen so auf den ersten Blick Plagiate und auch die Plagiatsdichte sichtbar.

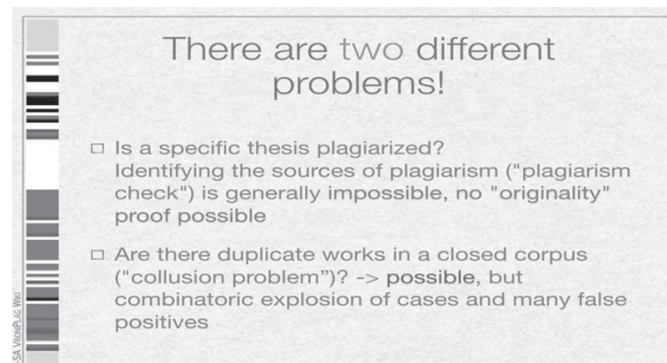
Ein hoher Prozentsatz an Übereinstimmung allein sagt jedoch wenig aus, so ist z. B. Zusammenarbeit kein Plagiat oder natürlich ist es möglich, die gleiche Arbeit selbst auch mehrfach an verschiedenen Stellen zu veröffentlichen. Auf der anderen Seite können auch Übersetzungen Plagiate sein, was weitaus schwieriger automatisiert zu erkennen ist. Der aufwändigste Schritt bei der Plagiatsprüfung ist jedoch die sich den Ergebnissen der computergestützten Prüfung anschließende manuelle Auswertung und Dokumentation, die durch mindestens zwei Personen gemeinsam geschieht und in einem Bericht mit Quellencluster und ggf. Plagiatsketten mündet.



### Ergebnisse der Plagiatsprüfungsverfahren

Eine Erkenntnis der bisherigen Prüfungen ergibt, dass in vielen Fällen tatsächlich seitenlang Wikipediainhalte übernommen werden: Mindestens 58 Doktorarbeiten sind überführt worden. Es zeigt sich zudem, dass Plagiate ein universelles Problem sind, das in jedem Fach und an jeder Hochschule auftritt. Entsteht am Ende der Prüfungen aber nun mehr Licht oder mehr Dunkelheit?

Die sich aus den Plagiatsfeststellungen ableitenden und zu erwartenden Entscheidungen der Unis über die betroffenen Texte bzw. Autor:innen lassen oft unbestimmt lange auf sich warten. Sie bedeuten dann entweder den Entzug des Dokortitels oder enden in dem Beschluss, dass der Titel nicht (ausreichend) beeinträchtigt ist. Bedauerlicherweise versuchen Universitäten selbst nach Entzug des Dokortitels oft, die Einstufung der betroffenen Texte als Plagiate möglichst geheim zu halten. Da unter Beachtung der Persönlichkeitsrechte etwas wie ein entsprechender Google-Eintrag über eine Aberkennung durchaus diffizil ist, bietet sich als diskreter Mindestvorgang jedoch etwa an, bei den Bibliotheken, die die Texte führen, einen Zusatzvermerk zur Arbeit im Katalog anzulegen, der z. B. auf den Entzug des Doktorgrades für die Dissertation hinweist. Einen fragwürdigen Umgang mit den Texten pflegen darüber hinaus manche Universitäten, die das Einreichen einer korrigierten Version erlauben, in denen die im VroniPlag-Bericht nachgewiesenen Mängel behoben sind.



### Was zu tun bleibt

Natürlich ist jedoch nicht das klassische Plagiat – die exakte Übernahme von Textpassagen – ein Problem, sondern auch das Verwenden der immer gleichen Ideen stellt keine gute wissenschaftliche Praxis dar. Hinsichtlich der Aufklärung und auch zukünftiger Plagiatsvorbeugung bleibt also noch viel zu tun. Ein weiterer Punkt ist dabei von großer Bedeutung: Maschinell können immer nur Textübereinstimmungen gefunden werden, letztendlich kann aber nur ein Mensch entscheiden, ob tatsächlich ein Plagiat vorliegt, am besten sogar vor Veröffentlichung des Textes. Dafür brauchen Doktorarbeitsbetreuer jedoch auch genügend Zeit, damit ihre ggf. aus der (Zeit-)Not entstandene Fahrlässigkeit nicht auch für sie ein Karriereende bedeutet. Plagiate müssen nicht nur als isoliertes Phänomen, sondern als Symptom eines kranken Systems gesehen werden, das auf schnelles Studium und einen hohen jährlichen Paper-Output statt auf tatsächliche Textqualität setzt. Statt nur die Auswirkungen zu behandeln, müssen auch die tatsächlichen Gründe für die Praxis des Plagierens angegangen werden – auch mit soziologischer Forschung und mehr Aufklärung zum Thema wissenschaftliches Arbeiten.



### Debora Weber-Wulff

**Debora Weber-Wulff** hat eine Professur für Medieninformatik an der Hochschule für Technik und Wirtschaft Berlin und arbeitet zu den Themen Plagiatsforschung, Medieninformatik, E-Learning, Softwaretechnik, Usability, Gender & Informatik und Ethik. Sie schreibt unter dem Namen *WiseWoman* an der Wikipedia und dem *VroniPlag Wiki* mit.