

Vergütung elektronischer Publikationen in Repositorien – aktueller Stand

Uwe Müller

Seit 2007 können neben gedruckten auch elektronische Publikationen bei der Verwertungsgesellschaft Wort (VG Wort) für die Ausschüttung von Tantiemen gemäß Urheberrechtsgesetz (UrhG) gemeldet werden. Das Verfahren, das dafür zur Anwendung kommt, ist technisch allerdings so unzureichend, dass Betreibern digitaler Repositorien, wie sie für die Veröffentlichung elektronischer wissenschaftlicher Publikationen zum Einsatz kommen, die Teilnahme praktisch unmöglich gemacht wird. Nachdem monatelang nach einer gemeinsamen Lösung gesucht wurde, die auch die Belange von Repository-Betreibern und den dort veröffentlichenden Autoren berücksichtigt, macht die VG Wort nun einen Rückzieher und besteht kompromisslos auf ihrem Standard-Verfahren.

Der Auftrag

Die Hauptaufgabe von Verwertungsgesellschaften leitet sich aus einem Vergütungsanspruch ab, den Urheber schöpferischer Werke gegenüber Herstellern und bestimmten Betreibern von Geräten und Speichermedien haben, die sich potentiell zur Vielfältigkeit dieser Werke eignen¹. Sie organisieren die Verteilung der Gelder, die Hersteller im Rahmen der so genannten Geräteabgabe für Kopierer, Drucker, Plotter, DVD-Brenner und dergleichen abführen müssen. Ähnliche Regelungen und Gesellschaften gibt es auch in anderen Ländern. Die VG Wort ist neben zwölf weiteren in Deutschland zugelassenen Verwertungsgesellschaften (u. a. GEMA für Musikwerke, VG Bild-Kunst für Werke der bildenden Kunst, VFF für Filmproduktionen) diejenige, die für Sprachwerke zuständig ist. Sie ver-

tritt neben Autoren auch Verlage und schützt die erzielten Einnahmen als Tantiemen an die Rechteinhaber aus.

Als Sprachwerke galten in diesem Zusammenhang bis in die jüngere Vergangenheit hinein ausschließlich gedruckte Publikationen, die in der Regel in Verlagen erschienen sind. Da aber Texte zunehmend auch im Internet verbreitet werden, erweiterte die VG Wort 2007 die Anspruchsberechtigung auf online veröffentlichte Dokumente und führte das so genannte Meldesystem für Texte auf Internetseiten (METIS) ein². Damit sollte es Autoren beliebiger urheberrechtlich geschützter Texte ermöglicht werden, am jährlichen Ausschüttungsverfahren teilzunehmen.

Das schließt neben klassischen Verlagsangeboten Blogbeiträge ebenso ein wie Eigen- und Zweitveröffentlichungen im Sinne des grünen Weges zu Open Access³. Und: Die Vergütung für einen Text ist keine einmalige Angelegenheit. Sie erfolgt für jedes Jahr, in dem die vorgegebene Mindestzugriffszahl überschritten wurde und betrug für 2009 je 20 bis 30 EUR⁴. Auf den ersten Blick klingt das nicht nur für Autoren wie eine gute Nachricht. Denn wenn die betreffenden Texte bei einem Verlag oder einem anderen Betreiber – z. B. einem Repository – veröffentlicht werden, wird dieser mit 30 % an den ausgeschütteten Tantiemen beteiligt. Dazu müssen allerdings bestimmte Voraussetzungen erfüllt sein: Zum einen können Texte nur dann gemeldet werden, wenn sie einen gewissen Mindestumfang aufwei-

sen, der derzeit auf 1.800 Zeichen festgelegt ist. Ausgenommen von dieser Regelung ist lediglich die Textgattung Lyrik. Zum anderen dürfen die Texte keinen Kopier- oder Passwortschutz besitzen – relevant vor allem für PDF-Dateien. Schließlich werden nur für solche Texte Tantiemen ausgeschüttet, die im Laufe eines Kalenderjahres über eine Mindestzahl von Zugriffen verfügen⁵. Und hier liegt die größte Herausforderung.

Die Zählung

Denn um diese letzte Bedingung überprüfen zu können, müssen die Zugriffe auf die jeweiligen Dokumente gezählt werden. Aus zunächst einmal nahe liegenden Gründen überlässt die VG Wort diese statistische Erhebung nicht den jeweiligen Dokumentanbietern, sondern nimmt die Zählung selbst vor – mithilfe eigener so genannter Zählserver (z.B. <http://vg00.met.vgwort.de>), die von der Firma Spring betrieben werden. Die Idee ist denkbar einfach: Jeder im Zählsystem erfassten Online-Publikation wird ein zentral durch die VG Wort vergebenen eindeutiger Identifier zugeordnet. Jedes Mal wenn dieser Text aufgerufen wird, wird parallel ein HTTP-Request an den Zählserver abgesetzt, der als Parameter den entsprechenden Identifier enthält und somit einen individuell gewerteten Zählimpuls auslöst, etwa so:

<http://vg01.met.vgwort.de/na/7bfbce574c674b57acbed310bff4627d>

Damit die Zählung im Zusammenhang mit dem Aufruf des betreffenden Textdokuments ausgelöst wird, muss sie als so genannter Zählpixel in den HTML-Code eingebettet werden. Dafür wird das für die Einbindung von Grafiken in Webseiten vorgesehene HTML-Element `` genutzt, dessen Attribut `src` eine URL enthält, unter der der Browser per HTTP-Request nach dem Bild sucht:

⁵ Gezählt werden nur Zugriffe aus Deutschland und nur solche, die einem menschlichen Nutzer zugeordnet werden – also keine Suchmaschinenzugriffe usw.

¹ Siehe § 54 UrhG.

² Siehe <http://www.vgwort.de/verguetungen/auszahlungen/texte-im-internet.html>.

³ Mit dem grünen Weg zu Open Access werden Parallelveröffentlichungen beschrieben, die im Sinne des *Self Archiving* durch den Autor selbst auf der eigenen Webseite oder auf einem institutionellen oder fachlich ausgerichteten Repository kostenfrei veröffentlicht werden, siehe http://open-access.net/de/allgemeines/was_bedeutet_open_access/openaccessstrategien/#c399

⁴ Siehe http://www.vgwort.de/fileadmin/quoten/Quoten_2010_f_C3_BC_r_2009.pdf. Im Moment liegt die Untergrenze bei 1.500 Zugriffen (Nutzersessions) pro Kalenderjahr. Bis 2.999 Zugriffe gab's dann 20 EUR, ab 3.000 Zugriffen wurden 25 EUR und ab 10.000 Zugriffen 30 EUR gezahlt.

```
<img src=„http://vg01.met.vgwort.de/na/7bfbce574c674b57acbed310bff4627d“ width=„1“ height=„1“/>
```

Tatsächlich ist das Bild in diesem Falle eher das Abfallprodukt – und es ist auch praktisch unsichtbar, denn der Zählserver liefert ein nur 1x1 Pixel großes transparentes Gif-Image zurück, von dem der Zählpixel auch seinen Namen hat. Für den Nutzer einer Publikation bleibt die Zählung in der Regel vollkommen unbemerkt.

Es leuchtet ein, dass dieses Verfahren für Autoren, die ihre Texte auf der eigenen Webseite veröffentlichen, durchaus praktikabel ist. Erleichternd kommt hinzu, dass die zur Registrierung notwendigen Identifier auf den Registrierungsseiten der VG Wort selbst angefordert werden können und unmittelbar erzeugt werden⁶ und dass dazu – abgesehen von einer Email-Adresse – zunächst keinerlei Angaben benötigt werden. Die vollständige Registrierung mit den Angaben zum Autor und dem Text muss erst im Zuge des Ausschüttungsverfahrens erfolgen – also ohnehin nur, falls die notwendige Zugriffszahl überhaupt erreicht wurde.

Wie die Zählung genau erfolgt, darüber hüllt sich die VG Wort übrigens ebenso in Schweigen wie über die tatsächlich erreichten Zugriffszahlen für die einzelnen Dokumente. Bekannt gegeben wird den Autoren nur, in welche „Nutzungsklasse“ es ihr Text geschafft hat.

Die Probleme

Aus der Sicht von Anbietern größerer elektronischer Publikationsbestände stellt sich das allerdings weniger positiv dar. Zu ihnen zählen beispielsweise Verlage oder eben digitale Repositories, wie sie wissenschaftliche Einrichtungen und Fachcommunities heute allerorten betreiben, um die Publikationen ihrer Mitglieder online bereitzustellen. Sie enthalten nicht selten mehrere zehntausend Dokumente, die – falls sie überhaupt in Form von HTML vorliegen – vielfach auf mehrere Webseiten aufgeteilt sind. Um sie in das Zählverfahren einzubinden, muss jede einzelne HTML-Seite mit einem Zählpixel ausgestattet werden – ein für statische Seiten mühevolleres Unterfangen. Denn es lässt sich nicht ohne weiteres automatisieren. Immerhin bietet die VG Wort Verlagen und anderen Betreibern inzwischen einen Webservice an, mit dem automatisiert Identifier bestellt und heruntergeladen werden können – eine Funktion, die bis vor kurzem ebenfalls fehlte.

Als noch wesentlich problematischer stellt sich der Ansatz bei der Zugriffszählung von

PDF-Dokumenten dar, die bekanntlich den größten Teil vieler Repositories ausmachen. Weil sich darin anders als bei HTML-Seiten keine Bilder einbauen lassen, die von einer externen URL nachgeladen werden, kommt hier auch die eher unauffällige Integration von Zählpixeln nicht in Betracht. Damit auch PDF-Dokumente gezählt werden können, verlangt die VG Wort, dass die HTTP-Anfragen über ihren Zählserver gelenkt werden, bevor die eigentlichen Dateien an den Nutzer ausgeliefert werden können. Soll beispielsweise eine PDF-Datei mit der URL

<http://www.my-repository.de/foo.pdf>

für Zählungen berücksichtigt werden, muss sie fortan via Zählserver aufgerufen werden, also durch eine URL der Art

<http://vg01.met.vgwort.de/na/7bfbce574c674b57acbed310bff4627d?l=http://www.my-repository.de/foo.pdf>

Der erste Teil enthält wiederum den zugeordneten Identifier und ist identisch mit der URL des entsprechenden Zählpixels. Nach der Registrierung des Zugriffs auf dem Zählserver erfolgt von dort der Redirect auf die eigentliche PDF-Datei. Deren URL wird dem Zählserver als GET-Parameter mitgegeben. Aus der Sicht von Repository-Betreibern ist dieses Verfahren vor allem aus zwei Gründen vollkommen inakzeptabel. Zum einen hängt die Erreichbarkeit der eigenen Publikationen unmittelbar von der technischen Verfügbarkeit eines Dritten ab: Wenn der Zählserver der VG Wort nicht verfügbar ist, ist auch das Repository praktisch offline, denn die URLs zu den eigenen Dokumenten zeigen auf den Zählserver.

Zum anderen ist es für viele Betreiber schlicht unvorstellbar, dass die URLs zu den eigenen Dokumenten zu einer fremden Domain gehören – vgwort.de. Zum Branding einer Institution gehört auch, dass sich aus URLs eigener Dokumente und Webseiten deren Identität unmittelbar erkennen lässt. Selbstverständlich müssen diese Zählserver-URLs in dieser Form auch nach außen gegeben werden – also beispielsweise an Suchmaschinen oder per OAI an Nachweis- und Aggregationsdienste. Denn andernfalls werden Zugriffe, die von dort kommen, nicht gezählt. Sollte sich die derzeitige Lösung der VG Wort für die Zählung von Online-Publikationen flächendeckend durchsetzen, wären weite Teile des deutschen Internets quasi zentralisiert, weil ihre Erreichbarkeit von der Erreichbarkeit der Zählserver abhängt.

Schließlich hat die Sache auch noch einen technischen Haken. Denn wenn die PDF-Datei schließlich im Browser des Nutzers erscheint, steht die unveränderte URL im Adressfeld, im Beispiel also

<http://www.my-repository.de/foo.pdf>

Setzt er darauf ein Bookmark und ruft das Dokument beim nächsten Mal auf diesem Wege auf, wird die Datei direkt vom Betreiber geladen – ohne Umweg über den Zählserver, aber eben auch ohne Zählung. Dasselbe kann passieren, wenn die endgültige URL von Suchmaschinen u. ä. verlinkt wird, die bei der Indexierung die Weiterleitung nachvollziehen.

Der Redirect-Mechanismus ist im Internet übrigens nicht unbekannt: Er kommt in ähnlicher Weise für so genannte Adserver zur Anwendung, die den Erfolg von Online-Werbung messen. Sie sind elektronischen Werbeanzeigen und der Webseite des beworbenen Produkts zwischengeschaltet und bilden u. a. die Grundlage für die Abrechnung der Werbekosten. Allerdings weisen die beiden Anwendungsfälle deutliche Unterschiede auf. Beispielsweise schlägt das Bookmarking-Problem im Falle der Online-Werbung nicht zu Buche, weil ja nur solche Zugriffe auf eine Webseite gezählt werden sollen, die von einer bestimmten Online-Anzeige kommen. Die Nutzung elektronischer Publikationen soll dagegen unabhängig davon gezählt werden, von welcher Seite der Leser kommt oder ob er sie per Lesezeichen ansteuert.

Die Alternative

Die beschriebenen Probleme waren seit der Vorstellung des Zählverfahrens offensichtlich, und so hat es bereits ab 2007 intensive Bemühungen gegeben, zumindest für den Bereich der wissenschaftlichen Repositories zu einer anderen Lösung zu kommen. Vor allem die DINI-Arbeitsgruppe „Elektronisches Publizieren“⁷ ist hier aktiv geworden. Im April 2007 fand an der Humboldt-Universität ein erster gemeinsamer Workshop mit dem bei der VG Wort für die Realisierung betrauten technischen Leiter statt. Als Ergebnis waren die zentralen Schwierigkeiten benannt und auch schon alternative Ansätze skizziert worden. Im Hinblick auf die Zugriffszählung liefern sie auf einen seitens des Repositories ausgelösten Zählimpuls auf dem Zählserver hinaus, der jedoch von der Auslieferung des angefragten Dokuments entkoppelt wäre.

Weil es aufseiten des technischen Dienstleisters der VG Wort kurz darauf zu personellen Veränderungen kam, wurde daraus jedoch zunächst nichts. Immerhin entwickelte die VG Wort ein Konzept zur maschinellen Beantragung und Übertragung der Identifier per Webservice an Verlage und andere

⁶ Siehe <https://tom.vgwort.de/portal/>.

⁷ Siehe <http://www.dini.de/ag/e-publ/>.

Dienstbetreiber. In Bezug auf das Zählverfahren herrschte jedoch zunächst Stillstand. Ende 2008 nahm dann eine gemeinsame Gruppe mit Vertretern von DINI und der VG Wort die Arbeit zur Entwicklung eines Konzeptes für ein alternatives Zählverfahren auf. Wie bereits die ersten Überlegungen lief die darin beschriebene so genannte „Proxy-Lösung“ darauf hinaus, dass der für die Zählung verantwortliche HTTP-Request parallel zur Auslieferung des Dokuments direkt durch den Anbieter erfolgt. Die Vorteile dieser Lösung liegen auf der Hand: Der Browser des betreffenden Nutzers bleibt von dem Vorgang unbehelligt, die URLs der zu zählenden Dokumente bleiben unverändert, und die Verfügbarkeit bleibt vom Zählserver der VG Wort unabhängig.

Weil sich durch die Entkoppelung des Zählimpulses von der Dokumentauslieferung die Möglichkeit zu Manipulationen seitens des Betreibers ergibt – schließlich können Zählimpulse auch ohne tatsächliche Anforderung der zugeordneten Publikation erfolgen – wurde an entsprechende Vereinbarungen zwischen der VG Wort und den Repositories gedacht, in der das gegenseitige Wohlverhalten zugesichert wird.

Obwohl das Konzept in beiderseitigem Einvernehmen erstellt und die technischen Unzulänglichkeiten des METIS-Verfahrens auch mehrfach bestätigt wurden, kam seitens der VG Wort kurz vor der prototypischen Realisierung mit einigen Repositories Ende September nun die Absage an die tatsächliche Umsetzung dieses Alternativverfahrens. Wegen der Gefahr der missbräuchlichen Nutzung und dem hohen Aufwand für den Abschluss von Vereinbarungen habe man sich dafür entschieden, ausschließlich das Standardverfahren einzusetzen.

Status quo

Obwohl die Meldung und Zählung elektronischer Publikationen bei der VG Wort nun schon seit vier Jahren möglich ist, wird das Verfahren von Verlagen und größeren Betreibern bisher kaum eingesetzt. In einigen Repositories sind vereinzelt Zählmarken zu finden, die in der Regel auf Drängen der jeweiligen Autoren und zu experimentellen

Zwecken eingebaut wurden – vor allem als Zählpixel in HTML-basierte Publikationen. Lediglich der Publikationsserver der Fraunhofer-Gesellschaft⁸ hat das Verfahren im produktiven Einsatz. Über OAI und an Suchmaschinen werden jeweils nur die URLs gegeben, die zur Startseite der einzelnen Dokumente führen. Von dort aus gelangt man über den Umweg des VG-Wort-Zählservers zur PDF-Datei. Abgesehen von diesem Link sind die ungeliebten auf vgwort.de lautenden URLs zwar nicht zu sehen und nirgendwo sonst zu finden. Der Preis allerdings ist, dass die Suchmaschinen die PDF-Dateien nicht indexieren (dürfen) und Recherchen nach Textteilen aus dem Volltext etwa bei Google erfolglos bleiben. Und: PDF-Dateien, die bei Nutzern über ein gespeichertes Lesezeichen aufgerufen werden, werden nicht gezählt.

Für Autoren, deren Publikationen auf Plattformen veröffentlicht sind, die das Zählverfahren (noch) nicht unterstützen – also beispielsweise auf wissenschaftlichen Repositories – bietet die VG Wort eine so genannte Sonderausschüttung an. Sie liegt deutlich unter den Werten der regulären Ausschüttung, erfolgt jedoch unabhängig vom tatsächlichen Umfang der Nutzung⁹. Allerdings lehnt die VG Wort derartige Anträge ab, sobald sich auf dem entsprechenden Repository auch nur eine Zählmarke befindet.

Nach dem vorläufigen Scheitern der Verhandlungen über ein praktikables Zählverfahren wird – beispielsweise im Rahmen der DINI-AG Elektronisches Publizieren – unter Betreibern von Repositories diskutiert, wie damit weiter umzugehen ist. Einerseits fordern Autoren die Möglichkeit der Teilnahme am Melde- und Ausschüttungsverfahren für ihre Publikationen ein. Auch wenn die eigenen Statistiken nahe legen, dass die allerwenigsten wissenschaftlichen Texte die Mindestzahl an Zugriffen überhaupt erreichen, könnten potentielle Vergütungen ggf. als zusätzlicher Anreiz gelten, eigene Veröffentlichungen auf Repositories abzulegen.

⁸ Siehe <http://publica.fraunhofer.de/>.

⁹ Für 2009 wurden für die erste Veröffentlichung eines Autors 12 EUR gezahlt, für jede weitere 3 EUR.

Auf der anderen Seite lässt sich das mangelhafte Zählverfahren nicht ohne entsprechende Beeinträchtigungen einsetzen. Einige Betreiber und Softwarehersteller haben mit unterschiedlichen technischen Workarounds experimentiert, die aber alle ihre Nachteile haben – wie unter anderem das Beispiel der Fraunhofer-Gesellschaft zeigt.

Fazit

Derzeit beharren die Verantwortlichen bei der VG Wort auf ihrer technisch unzureichenden Lösung zur Zählung elektronischer Dokumente. Es bleibt zu hoffen, dass dort ein Umdenken einsetzt und man mit sachgerechteren Lösungen auch dem eigenen Auftrag besser gerecht wird. Angesichts der Probleme, die das derzeitige Verfahren mit sich bringt, und der andererseits als eher gering einzuschätzenden finanziellen Erträge für Autoren und Anbieter kann Betreibern von Repositories derzeit nicht empfohlen werden, daran teilzunehmen. Auch die Implementierung des Systems mit technischen Winkelzügen zur Umgehung einiger der beschriebenen Schwierigkeiten kann die grundsätzlichen Probleme nicht lösen und bedeutet in jedem Falle Zugeständnis seitens des Betreibers bzw. der Autoren.

■ AUTOR

DR. UWE MÜLLER

ist wissenschaftlicher Mitarbeiter am Institut für Bibliotheks- und Informationswissenschaft an der Humboldt-Universität zu Berlin und seit Anfang 2010 einer der beiden Sprecher der DINI-AG „Elektronisches Publizieren“. Humboldt-Universität zu Berlin
Institut für Bibliotheks- und Informationswissenschaft
Lehrstuhl für Informationsmanagement
Dorotheenstraße 26
Postanschrift:
Unter den Linden 6, 10099 Berlin
u.mueller@cms.hu-berlin.de

