

Kollaboratives Wissensmanagement im Bildungsbereich und die Zitierfähigkeit von Wiki-Wissen

Martin Haase, Rüdiger Weis*
Universität Bamberg, TFH Berlin

23. Chaos Communication Congress, Berlin 2006

Im Bildungsbereich werden Wikis immer wichtiger – vor allem die Wikipedia und Wikibooks. Dabei ist klar, dass kollaborativ erworbenem Wissen eine gewisse Unsicherheit anhaftet. Der Beitrag gibt einen Überblick über die Probleme und über Maßnahmen, die ergriffen worden sind oder noch ergriffen werden können, um Wikiwissen im Wissenschafts- und Hochschulbereich einzusetzen. Zudem werden Argumente geliefert, warum und wie man Wikiwissen in Schule und Hochschule nutzen kann. Da in Kürze in der deutschen Wikipedia die stabilen (geprüften) Versionen kommen werden, geht die Diskussion über die Zitierbarkeit der Wikipedia in eine neue Runde.

Im Mittelpunkt des Beitrags steht das wissenschaftliche Zitieren der Wikipedia im Kontext der Diskussion um stabile Versionen, Qualitätssicherung und die Sicherung von Expertenwissen:

- Welcher Zusammenhang besteht zwischen wissenschaftlichem Arbeiten und dem Wissenskonsens der Wikipedia?
- Erlaubt die Wikipedia die Bereitstellung von wissenschaftlich verwertbaren Zitaten und wie kann sie präzise zitiert werden?

Die traditionelle Zitation einer Webseite mit einer URL und der Angabe eines Konsultationsdatums wird der schnellen Veränderbarkeit der Inhalte nicht gerecht. Hier müssen andere Formen der Zitation verwendet werden. Der Gebrauch universeller Identifikatoren (zum Beispiel URNs) ist denkbar. Mithilfe einfacher Skripte lassen sich verschiedene Zitationsformate voneinander ableiten, wenn sie einmal (zum Beispiel über RFCs) standardisiert worden sind.

*<mailto:martin.haase@split.uni-bamberg.de> & <mailto:rcw@tfh-berlin.de>; wir bedanken uns bei Leo Sauer mann, Jakob Voss und Patrick Danowski für ihre interessanten Anregungen.

1 Die Wikipedia als Quelle

Wissenschaftliche Arbeiten dienen dazu, neue Erkenntnisse zu gewinnen. Die eigenen Forschungen bauen dabei in der Regel auf frühere Forschungen auf, auf die natürlich Bezug genommen werden muss. Es ist daher unumgänglich, in der eigenen Arbeit andere Arbeiten zu zitieren.

Da neues Wissen immer auf älterem aufbaut, besteht in der Wissenschaft auch ein vitales Interesse an einem Urheberrecht, das das Zitieren möglichst wenig oder im Idealfall gar nicht einschränkt.

Dabei beschränkt sich die wissenschaftliche Zitation nicht auf Text allein, sondern umfasst auch Daten, Abbildungen und in neuerer Zeit auch Ton- und Videodokumente. Leider ist das Urheber- und Verwertungsrecht in vielen Ländern gerade in Bezug auf audiovisuelles Material nicht gerade wissenschaftsfördernd. Umso wichtiger ist es, dass Wissenschaftler ihre Forschungsergebnisse unter einer freien Lizenz publizieren, damit andere auf ihre Forschungen aufbauen können. Vor diesem Hintergrund erklärt sich auch das Interesse deutscher Forschungs- und Wissenschaftsorganisationen an Open Access [1].

1.1 Zitieren aus Lexika

Nicht zitiert wird gewöhnlich aus Lexika (außer natürlich wenn das Lexikon selbst Gegenstand der Arbeit ist): viele Lexika ermöglichen nämlich keinen unmittelbaren Zugang zu Forschungsergebnissen; die Informationen, die sie enthalten, wurden speziell für das Lexikon aufbereitet, und zwar oft nicht nur von Fachleuten, sondern auch von einer Redaktion oder einer Gruppe von Herausgebern, die nicht aus Fachleuten besteht.

Allerdings muss unterschieden werden zwischen *wissenschaftlichen Fachlexika* und so genannten *Konversationslexika*. Manche Fächer verfügen über Fachlexika (für die Linguistik der romanischen Sprachen ist das zum Beispiel das *Lexikon der romanistischen Linguistik* [2]), die sich durchaus um eine Aufarbeitung des Forschungsstands ihrer Disziplin bis zu ihrem Erscheinen bemühen und daher auch gelegentlich in wissenschaftlichen Arbeiten zitiert werden.

Ganz anders sieht es mit Konversationslexika aus: Diese sind nicht für Wissenschaftler verfasst, sondern für interessierte Laien. Schon die Bezeichnung *Konversationslexikon* drückt dies aus: Solche Lexika stellen Wissen zur Verfügung, das es ermöglicht, sich an gebildeter Konversation zu beteiligen. Da sie keinen Anspruch auf Wissenschaftlichkeit erheben, enthalten sie auch keine oder nur wenige Quellenangaben, und vor allen in deutschen Lexika dieses Typs wird nicht angegeben, von wem ein Artikel verfasst wurde. Zudem sorgt eine Redaktion für die Einheitlichkeit und allgemeine Verständlichkeit von Artikeln, wodurch fachliche Informationen noch einmal modifiziert und mitunter verkürzt und vereinfacht dargestellt werden. In diesem Zusammenhang von „gesichertem Wissen“ zu sprechen, erscheint wenig angebracht. Es ist auch nicht Aufgabe eines Konversationslexikons, sich um Wissenschaftlichkeit zu bemühen. Didaktische und oft auch kommerzielle Erwägungen stehen im Vordergrund.

Auch von „geprüftem Wissen“ kann eigentlich keine Rede sein: Die Redaktion wird von Fachleuten eingereichte Beiträge sicherlich sprachlich und stilistisch Korrektur le-

sen, vielleicht auch Datumsangaben oder Namensschreibungen auf offensichtliche Fehler überprüfen; sofern sie jedoch nicht selbst aus Fachleuten besteht, kann sie das in den Produktionsprozess eingespeiste Fachwissen gar nicht ihrerseits einer genauen fachlichen Überprüfung unterziehen. Sie muss sich letztlich auf die Kompetenz der vom Verlag eingeworbenen Experten verlassen. Eine Überprüfung wäre vielleicht durch ein *Peer Review* möglich, also durch eine Überprüfung der zu veröffentlichenden Artikel durch andere Fachleute, aber solche Maßnahmen liegen jenseits der Ansprüche eines Konversationslexikons.

Es ist einleuchtend, dass Konversationslexika nicht als Quelle für wissenschaftliche Arbeiten in Frage kommen. Das sollte Studierenden als angehenden Wissenschaftlern schon am Anfang ihres Studiums klar sein, aber auch wissenschaftlichen Laien (Journalisten, Schülern usw.) sollte bewusst sein, dass Konversationslexika nicht vollkommen verlässliche Informationen enthalten und dass weitere Recherchen nötig sind, um sich in ein Thema auch jenseits gebildeter Konversation einzuarbeiten.

1.2 Wissenskonsens

Die Wikipedia ist in ihrer Zielsetzung nur zum Teil mit einem Konversationslexikon vergleichbar. Artikel, die von der Gemeinschaft der Wikipedianer als gut oder lesenswert bezeichnet werden, oder Artikel, die in Schreibwettbewerben prämiert wurden, sind in aller Regel deutlich ausgebauter als es in einem Konversationslexikon möglich oder nötig wäre. Gemeinsam ist allerdings, dass sich Konversationslexika und die Wikipedia an ein Publikum von interessierten Laien wenden. Daher müssen auch Wikipedia-Artikel für Nicht-Fachleute verständlich sein. Damit verbunden ist sicher ein Verlust an fachlicher Präzision. Dieser wiegt aber weniger schwer, wo Fachleute und Laien gemeinsam an Artikeln arbeiten: ein interessierter Laie kann in den Produktionsprozess des Artikels eingreifen, wenn dieser unverständlich wird, gleichzeitig kann ein Experte fachliche Ungenauigkeiten sofort richtig stellen.

Sicher ist die Zusammenarbeit zwischen Fachleuten und Laien nicht immer reibungslos, aber vielversprechender als die Trennung von Fachautoren und Redaktion, wie sie bei gedruckten Lexika üblich ist, zumal ein *Peer Review*, also die Kontrolle der Experten- und Laienarbeit durch andere Fachvertreter in jedem Fall möglich ist.

Im Gegensatz zu landläufigen Konversationslexika lässt sich aus der Versionsgeschichte eines Artikels auch immer genau ersehen, wer jeweils mitgearbeitet hat. Natürlich werden bei der Wikipedia überwiegend Pseudonyme verwendet und es ist auch möglich, als Autor anonym zu bleiben, aber für jeden Mitarbeiter (auch für anonyme IP-Adressen) lassen sich die Beiträge auflisten, die der jeweilige Autor bearbeitet hat (über den Link *Benutzerbeiträge*). Damit lässt sich auch bei Autoren, die sich nicht mit einer eigenen Benutzerseite ausführlich vorstellen, leicht nachvollziehen, woran sie sonst gearbeitet haben. Es ist somit möglich, ein Kompetenzprofil zu ermitteln. Da eine solche Recherche aufwendig ist, soll eine entsprechende Kennzeichnung von Artikeln erfolgen, für die in einer so genannten stabilen (also nicht-vandalisierten) Version ausgewiesene Experten verantwortlich zeichnen.

Auch wenn die Wikipedia in Bezug auf ihre Zuverlässigkeit den herkömmlichen Lexika zumindest in Hinblick auf nicht-vandalisierte Versionen überlegen ist (und Van-

dalismus wird in der Regel ohnehin schnell entdeckt), bleibt die Frage nach ihrer Zitierbarkeit im wissenschaftlichen Kontext, denn auch in der Wikipedia haben wir es mit Informationen zu tun, die zum Zweck der Verbreitung aufbereitet wurden, also nur mittelbar den wissenschaftlichen Forschungsstand widerspiegeln.

Allerdings bieten die in der Wikipedia enthaltenen Informationen einen Vorteil, dem bisher in der Diskussion wenig Beachtung geschenkt wurde, der aber für Wissenschaftler durchaus von Interesse ist: Ausgebaute Wikipedia-Artikel enthalten Informationen über ihren Gegenstand, auf die sich eine Gruppe von Autoren, Ko-Autoren und Lesern geeinigt haben: einen *Wissenskonsens*. Anhand der Versionsgeschichte lässt sich zudem ermitteln, wie dieser Wissenskonsens zustande gekommen ist. Insofern ist die Wikipedia eigentlich eine Primärquelle, die den Zugang zu dem vermittelt, was eine breitere Öffentlichkeit für richtig hält als zum Beispiel nur ein einzelner Experte, eine Experten-Gruppe oder eine fachwissenschaftliche Schule. Der Wissenskonsens ist natürlich auch interessant, wenn es um die Gewinnung neuer Erkenntnisse geht, denn auch dieser wissenschaftliche Fortschritt muss ja wieder in den Wissenskonsens einfließen. Das gilt besonders für die an den Hochschulen betriebene Forschung, die ja in einen Lehrkontext eingebunden ist. Die Wikipedia ist in diesem Sinn durchaus zitierbar – nicht, um sich weitere Recherchen zu ersparen, sondern als Quelle des Wissenskonsens. Mit der Wiedergabe von Wikipedia-Inhalten ist es in einer wissenschaftlichen Arbeit natürlich nicht getan, denn Erkenntnisgewinn muss immer jenseits des Wissenskonsens' liegen. Das gilt natürlich nicht nur für die Wissenschaft. Auch von einem Journalisten ist zu erwarten, dass er *neue* Informationen vermittelt. Die Recherche muss also weitergehen, denn was in der Wikipedia steht, ist ohnehin bekannt. Aufgabe der Schule ist es, zu vermitteln, wie Schüler sich den Wissenskonsens erschließen, hinterfragen und erweitern können. Schon deshalb hat die Wikipedia ihren Platz im Schulunterricht.

Es ist manchen Fachwissenschaftlern Recht zu geben, wenn sie sagen, dass Wikipedia-Inhalte für sie nicht relevant seien; umgekehrt ist allerdings auch nach der Relevanz von wissenschaftlichen Inhalten zu fragen, die sich im breiteren Wissenskonsens der Wikipedia nicht wiederfinden. Sicher kann nicht jedes Detail wissenschaftlicher Forschung sofort Aufnahme in die Wikipedia finden, da eine Online-Enzyklopädie sich ja an interessierte Laien wendet, aber kein Fach kann darauf verzichten, seine Inhalte in angemessener Form einer breiteren Öffentlichkeit zu vermitteln. Ob das gelingt, zeigt sich nicht zuletzt an den Inhalten der Wikipedia.

2 Zitation von URLs

Betrachten wir als Beispiel für die Zitation eines Wikipedia-Artikels mit einem eindeutigen, unveränderlichen Link (*Permalink*) die Zitation der beim Schreiben dieses Beitrages aktuellen Version des Artikels zur Internetzitation:

Artikel: Zitieren von Internetquellen. In: Wikipedia, Die freie Enzyklopädie. Bearbeitungsstand: 19. November 2006, 15:23 UTC. URL: http://de.wikipedia.org/w/index.php?title=Zitieren_von_Internetquellen&oldid=24000614 (Abgerufen: 19. November 2006, 16:27 UTC)

Die bisherigen Zitationsvorschläge von permanenten Wiki-URLs erscheinen stark verbesserungsbedürftig. Unter anderem widersprechen folgende Punkte einer durchdachten URL-Wahl (siehe auch: [3]):

- Verwendung der physikalischen Verzeichnisstruktur,
- Bindung an PHP und
- widersinnige Verwendung von `oldid` für die aktuelle Artikelversion.

Angeregt sei eine einfachere URL, für die nur der Artikeltitel und die Artikel-*id* Verwendung finden:

http://de.wikipedia.org/Zitieren_von_Internetquellen&id=24000614

Erfolgt die Zitation über die eindeutige *id* ist die zusätzliche Angabe des Abrufzeitpunktes redundant; er kann angegeben werden, wenn dies als Information für den Leser nützlich ist. Natürlich ist auch die Nennung des Artikeltitels redundant, da der Artikel durch die *id* hinreichend identifiziert wird, aber das Weglassen des Artikeltitels erschwert dem Leser das Verständnis.

3 Uniform Resource Name

Ein *Uniform Resource Name* (URN, engl.: ‚einheitlicher Ressourcen-Name‘) ist ein dauerhafter, ortsunabhängiger Bezeichner für eine Ressource. Aus den angeführten Gründen erscheint die Definition einer eigenen `urn:wikipedia` erstrebenswert. Zur Spezifizierung einer Artikelversion scheinen folgende Daten auszureichen:

- Sprachversion,
- Artikeltitel,
- Versions-ID.

Die Erstellungszeit des Artikel sollte optional anfügbar sein. Eine Angabe der genauen Abrufzeit erscheint weniger notwendig.

urn: NID: wikipedia		
	language:	•
	title:	•
	id:	•

Aus den Datenfeldern der URN kann einfach die URL generiert werden.

```
PermanentURL=
  language+"."
+ "wikipedia"
+ ".org/w/index.php?title="
+ title
+ "&oldid="
+ id
```

Beispiel

Die Zitation des Wikipedia-Artikels *Uniform Resource Name* erfolgt als URN:

`urn:wikipedia:de:Uniform_Resource_Name:24062643`

für:

`http://de.wikipedia.org/w/index.php?title=Uniform_Resource_Name&oldid=24062643`

Stabile Versionen

Für die Einführung von Artikeln mit geprüften stabilen Versionen ist die Ergänzung eines `authors`-Feld erwägenswert, in das die für die stabile Version verantwortlichen Autoren (Experten) eingetragen werden.

4 Weitere Anregungen

Wenn auch viele Einschätzungen bezüglich der „kollektiven Intelligenz“ der Wikipedia-projekte Themen der weit über den Wissenschaftsbereich hinausgehende Diskussionen sind, werden selbst erbitterte Kritiker einräumen, dass Wikikpedia für die Erforschung von Wissensdiskussionen besonders geeignet ist. Die große Menge von Artikeln, Mitarbeitern und Lesern bietet neue Möglichkeiten für statistische Analysen. Hilfreich wären hierfür unter anderem weitere statistische Angaben zu den einzelnen Artikeln; vor allem die folgenden Werte sind aufschlussreich:

- Besucherzahl insgesamt,
- Besucherzahl seit der letzten Änderung.

Diese Angaben könnten in interessanter Weise mit anderen Eigenschaften, wie Verteilung und Struktur der Links, zu statistischen Bewertungen herangezogen werden. Inwieweit ein erweitertes Tagging von Seiten ermöglicht werden sollte, sei als Diskussionspunkt ebenfalls angeregt.

Literatur

- [1] *Berliner Erklärung zu Open Access*.
http://www.mpg.de/pdf/openaccess/BerlinDeclaration_dt.pdf, Stand: 15.11.2006
- [2] Holtus, Günter *et al.* (eds.): *Lexikon der Romanistischen Linguistik*. Tübingen: Niemeyer 1988–.
- [3] Tim Berners-Lee, “Cool URIs don’t change”, <http://www.w3.org/Provider/Style/URI>
- [4] rfc3986, “Uniform Resource Identifier (URI): Generic Syntax”, <http://www.ietf.org/rfc/rfc3986.txt>