

### Qualitätsaspekte der Wikipedia

Hammwöhner, Rainer

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Hammwöhner, R. (2007). Qualitätsaspekte der Wikipedia. *kommunikation @ gesellschaft*, 8, 1-24. <https://nbn-resolving.org/urn:nbn:de:0228-200708066>

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# Qualitätsaspekte der Wikipedia

Rainer Hammwöhner (Regensburg)

## Zusammenfassung

Eine kritische Diskussion über ihre Verlässlichkeit begleitet die Entwicklung der *Wikipedia* von Beginn an. Mittlerweile liegen einige Publikationen vor, die sich mit der Qualität dieser neuen Enzyklopädie befassen. In diesem Betrag wird ein kurzer Überblick über mehrere dieser Arbeiten gegeben. Anschließend werden zwei eigene Studien vorgestellt, die sich mit Qualitätsaspekten befassen, die bisher weniger berücksichtigt wurden. In der ersten Studie wird die Abdeckung eines gut umgrenzten Themenkomplexes – Shakespeare's Werk – in mehreren *Wikipedias* untersucht. Die zweite Studie befasst sich mit der Qualität der Wissensorganisation in der *Wikipedia*.

## 1 Einführung

Prinzip und Erfolg der *Wikipedia*<sup>1</sup> sind in mancherlei Hinsicht bemerkenswert. Zunächst fällt vermutlich der rein quantitative Aspekt ins Auge: Die englische *Wikipedia* (Wikipedia 2007) hat mittlerweile (Stand 24.3.2007) mehr als 1.5 Millionen Artikel, weitere 13 in anderen Sprachen verfasste *Wikipedias* verfügen über mehr als einhunderttausend Beiträge. Beeindruckend ist auch das Wachstum der *Wikipedias*, das exponentielle Wachstumsphasen einschließt (Voss 2005, S. 27 ff.). Der eigentliche Grund für die Aufmerksamkeit, welche die *Wikipedia* genießt, liegt jedoch in dem augenscheinlichen Konflikt zwischen dem Anspruch der *Wikipedia*, enzyklopädisches Wissen ohne Einschränkung in Qualität und Reichweite bereitzustellen, auf der einen Seite und dem Prinzip, dieses Wissen von einer offenen, z.T. anonymen Teilnehmergruppe zusammen stellen zu lassen, auf der anderen. Auf diese Herausforderung für den traditionellen Wissens- und Wissenschaftsbetrieb mit seinen Verfahren der Qualitätssicherung erfolgten sowohl vorsichtig affirmative (Kuhlen 2005) als auch polemisch ablehnende Reaktionen (Lanier 2006), wobei letztere auch durch schwerwiegende Qualitätsprobleme in der *Wikipedia* motiviert wurden. Große Beachtung wurde einer Studie über die Qualität der *Wikipedia* zuteil, die im Auftrag von *Nature* erstellt wurde (Giles 2005). Der *Wikipedia* wurde dort eine Qualität zugesprochen, welche jener der *Encyclopedia Britannica* vergleichbar sei. Diese Studie wurde von Seiten der *Britannica* angegriffen, von *Nature* aber verteidigt und mit Materialien untermauert.<sup>2</sup> Weniger als das Ergebnis dieser Debatte ist ihr Verlauf von Interesse. Es wurde deutlich, dass ein Konsens über eine angemessene Methodik zur Evaluierung der *Wikipedia* – wie auch ihrer Konkurrenzprodukte – erst noch zu erzielen ist. Hier ist ein erster Gegenstand für eine Erforschung der *Wikipedia* gegeben. Es überrascht also nicht, dass Fragen zur Qualität einen nennenswerten Umfang der bisher noch nicht sehr umfangreichen, aber durchaus

---

<sup>1</sup> Einen guten Überblick über die Prinzipien der Wikipedia bieten Danowski/Voss (2005).

<sup>2</sup> Verweise auf diese Materialien finden sich in der Online-Fassung von Giles (2005).

qualitätvollen Forschung zur *Wikipedia*<sup>3</sup> einnehmen. Ein erstes methodisches Problem der Qualitätsbewertung zeigt sich in der Auswahl der Qualitätskriterien. Während die Nature Studie ausschließlich auf inhaltliche Korrektheit fokussiert, berücksichtigen andere Studien – etwa Stvilia et al. (2005a) – eine Vielzahl von Qualitätsparametern. Unterschiede bestehen auch im Gegenstand der Bewertung. Die Nature Studie, aber auch die von Stvilia et al. beziehen sich primär auf die Qualität einzelner Artikel, während die Studie von Wiegand (2007) die Gesamtstruktur der Enzyklopädie sowie die Eigenschaften der Software-Umgebung mit einbezieht. In den meisten Studien werden die zu beurteilenden Qualitätskriterien gesetzt bzw. aus einem Grundverständnis der Ziele einer Enzyklopädie abgeleitet, Stvilia et al. (2005b) hingegen leisten eine komplexe Rekonstruktion des Qualitätsmodells der *Wikipedia* auf der Basis einer statistischen Analyse eines Korpus von *Wikipedia*-Artikeln.

In diesem Beitrag wird, nach einer kurzen Einführung in die Prinzipien der *Wikipedia*, ein Überblick über schon publizierte Forschungsergebnisse zur Qualität der *Wikipedia* gegeben. Danach werden zwei eigene Studien vorgestellt, welche die bestehenden Arbeiten ergänzen, aber auch neue Fragen aufwerfen. Diese Studien wurden im Rahmen von studentischen Projekten im Rahmen des Studiengangs Informationswissenschaft an der Universität Regensburg konzipiert und durchgeführt.<sup>4</sup>

## 2 Die *Wikipedia*

Ziel des *Wikipedia*-Projekts ist die Bereitstellung freier enzyklopädischer Information für jedermann. Methodische Grundlage ist die Beteiligung einer offenen Nutzergemeinschaft an diesem Projekt – jeder kann mitwirken. Die Nutzung der von Ward Cunningham konzipierten *Wiki*-Software ist die technische Voraussetzung für die Umsetzung dieses Ansatzes. Im Rahmen dieses Gesamtprojekts ist eine Vielzahl elektronischer Enzyklopädien entstanden, die jeweils in einer Sprache verfasst sind. Jede dieser Enzyklopädien wird von einer eigenen mehr oder minder großen Community getragen, die sich eigene Regeln der Zusammenarbeit setzt. An dieser Stelle kann nicht auf die Eigenheiten spezifischer *Wikipedia*-Projekte eingegangen werden. Es wird deshalb nur das gemeinsame Modell erläutert. Darüber hinaus finden einige innovative Ansätze Berücksichtigung, die in den großen *Wikipedias* eingeführt wurden.

### 2.1 Die Mitarbeiter

Grundsätzlich darf jeder an einer oder mehreren *Wikipedias* mitarbeiten. Dies kann in weitgehender Anonymität – Identifizierung nur durch IP-Nummern – oder in einer durch Zugangsmechanismen kontrollierten Pseudonymität geschehen. In der Mitarbeit an der *Wikipedia* nehmen die Nutzer verschiedene Rollen ein. Als Autoren verfassen sie neue Artikel oder überarbeiten ältere. Im Sinne der Qualitätssicherung übernehmen sie Verantwortung für einzelne oder mehrere Artikel. Grundsätzlich darf jeder Nutzer derartige Tätigkeiten ausführen. In manchen Fällen wird jedoch die Überarbeitung von Artikeln auf authentifizierte

---

<sup>3</sup> Im Zusammenhang mit der deutschen *Wikipedia* wurde für *Wikipedia*-Forschung der Terminus *Wikipedistik* geprägt. Zusammenstellungen von Publikationen über die *Wikipedia* finden sich unter *Wikipedia* (2007)o.

<sup>4</sup> Bei den Teilnehmern der Projektseminare der WS 2005/2006 und 2006/2007 bedanke ich mich für ihr Engagement. Die Ergebnisse, die in diese Veröffentlichung eingeflossen sind, werden jeweils an passender Stelle durch Zitat belegt.

Nutzer beschränkt. Bestimmte Aufgaben, wie etwa das Löschen von Artikeln, bleiben einem engeren Kreis von Administratoren vorbehalten, die aus dem Kreis der Nutzer gewählt werden.

## 2.2 Die Artikel

Die Artikel der *Wikipedia* werden in einer einfachen Auszeichnungssprache verfasst, welche die Definition eines Satzbildes sowie die Integration von Abbildungen und von Hypertext-Verweisen erlaubt. Obschon die Autoren durch keine Richtlinien für bestimmte Artikelformate eingeschränkt werden, haben sich mit der Zeit Konventionen für die Struktur von Biographien, Städtebeschreibungen u.Ä. ausgeprägt.

Durch die verwendete Basissoftware ist festgelegt, dass jede Änderung, die an einem Artikel vorgenommen wird, in einem Versionsraum so protokolliert wird, dass sie dem jeweiligen Autor zurechenbar ist. Zudem sind frühere Fassungen eines Artikels unter Zugriff auf den Versionsraum wieder herstellbar.

Jedem Artikel ist eine Diskussionsseite zugeordnet, auf der die Autoren sich über Defizite des Artikels austauschen können und Meinungsverschiedenheiten beilegen können. Auffällig ist, dass in der deutschen *Wikipedia* nur etwa einem Drittel aller Artikel eine Diskussion zugeordnet ist, während dies in der englischen *Wikipedia* bei etwa zwei Dritteln aller Artikel der Fall ist (Stichprobenumfang jeweils 1000). Dies ist darauf zurückzuführen, dass in der englischen *Wikipedia* ein anderer Gebrauch von den Diskussionsseiten gemacht wird, indem hier etwa Kurzvermerke über die Zugehörigkeit zu Portalen etc. zu finden sind, die in der deutschen *Wikipedia* weniger auftreten. Dies zeigt an einem einfachen Beispiel, dass die Nutzung der gemeinsamen Software in den einzelnen *Wikipedias* auf eine unterschiedliche Weise erfolgen kann. Untersuchungen zu einer *Wikipedia* lassen sich deshalb nur in engen Grenzen auf eine andere übertragen.

## 2.3 Strukturen der Wissensorganisation in der Wikipedia

Jeder Artikel der *Wikipedia* ist über seinen Titel, der den Lexikoneintrag bezeichnet, eindeutig identifizierbar und über eine Zugriffsfunktion erreichbar. Diese Form des Zugriffs erweist sich für das Nachschlagen von Begriffen in den meisten Fällen als zulänglich. Für komplexere inhaltliche Anfragen und für die Verwaltung und Qualitätssicherung der Enzyklopädie sind jedoch elaboriertere Formen der Wissensorganisation erforderlich.

### 2.3.1 Alphabetischer Index

Der alphabetische Index erlaubt einen Zugriff auf die Inhalte der Enzyklopädie, der demjenigen traditioneller Printmedien weitgehend analog ist. Der gegenüber der Titelsuche bestehende informationelle Mehrwert ist hier jedoch gering – etwa zum Aufsuchen von Titeln mit unklarer Schreibung.

### 2.3.2 Themenportale

Themenportale fassen thematisch verwandte Artikel zusammen. Sie leisten damit eine systematische Erschließung der Inhalte der Enzyklopädie. Sie sind nützlich für Nutzer die sich über ein Themengebiet informieren wollen. Themenportale werden häufig von einer eigenen Nutzergruppe im Rahmen eines Projekts betreut. Sie sind damit auch ein Instrument der Qualitätssicherung. Sowohl in der englischen wie auch in der deutschen *Wikipedia* gibt es Portale, die ein differenzierteres Qualitätsmodell nutzen als die Gesamtencyklopädie. Im Portal *Biologie* der deutschen *Wikipedia* etwa werden *validierte Artikel* zusätzlich zu den anderen Qualitätsstufen unterschieden. Derzeit (Stand 04/2007, Wikipedia 2007a) werden in der deutschen *Wikipedia* 379 Portale zur Verfügung gestellt. Portale können hierarchisch durch Unterportale gegliedert sein. Einzelne Lemmata oder auch Unterportale können in mehreren Portalen enthalten sein.

### 2.3.3 Kategoriensystem

Ein weiteres Instrumentarium der systematischen Erschließung der größeren *Wikipedias* sind Kategoriensysteme. Das Kategoriensystem der deutschen *Wikipedia* ist seit 2004 in Entwicklung befindlich (Voss 2006). Das Kategoriensystem stellt ein kontrolliertes, strukturiertes Vokabular zur Verschlagwortung der *Wikipedia*-Artikel zur Verfügung. Der Aufbau dieses Vokabulars erfolgt nach dem gleichen offen kooperativen Verfahren, das auch der Erstellung der Artikel zugrunde liegt. Hier besteht eine erhebliche Diskrepanz zur Sicht der Dokumentationslehre, die den Aufbau von Ordnungssystemen als einen komplexen, aufwendigen Prozess begreift, der planvoll zu erfolgen hat (Gaus 1995, S. 165 ff.).

### 2.3.4 Listen

Vor der Einführung des Kategoriensystems waren Artikellisten das einzige Instrument zur systematischen Erschließung der *Wikipedia*. Auch heute noch werden zahlreiche Listen (Wikipedia 2007b) geführt, die Artikel unter einem wohldefinierten Merkmal zusammenstellen – z.B. Listen der einer Stadt zuzurechnenden Persönlichkeiten oder Publikationsorgane, Kunstwerke einer Epoche etc. Diese Listen werden intellektuell zusammengestellt und intellektuell ausgewertet. Weitere Listen werden von der MediaWiki-Software automatisch generiert. Dies sind z.B. die Listen mit zuletzt geänderten, nicht kategorisierten, besonders langen oder kurzen Artikeln. Es ist offensichtlich, dass derartigen Verzeichnissen für das Qualitätsmanagement eine große Bedeutung zukommt. Dies trifft auch auf die Beobachtungslisten zu. Hier kann jeder authentifizierte Nutzer diejenigen Artikel eintragen, von deren Änderungen er automatisch durch die MediaWiki-Software informiert werden will.

### 2.3.5 Namensräume

Die *Wikipedia* ist in mehrere Namensräume aufgeteilt. Innerhalb eines Namensraums muss ein Artikelname eindeutig sein. Jeweils ein Namensraum ist für Enzyklopädieartikel, für die Diskussionen zu diesen Artikeln, für Bilder, für Diskussionen zu den Bildern, für Kategorien, Diskussionen zu den Kategorien usw. vorgesehen.

### 2.3.6 Hypertext-Verknüpfungen

Als Hypermediasystem verfügt das MediaWiki über die Möglichkeit, interne Dokumente durch Verweise sowohl unter einander als auch mit anderen elektronischen Dokumenten aus dem Internet zu verknüpfen. Interne Verweise dienen – wie in anderen Enzyklopädien auch – zur referentiellen Vernetzung der Lemmata. Externe Verweise führen zumeist zu Quellen, welche die Information aus dem Lexikonartikel absichern. Im Gegensatz zu anderen Hypertexten ist es in der *Wikipedia* nicht grundsätzlich als Fehler anzusehen, wenn eine interne Verknüpfung auf einen nicht existierenden Artikel verweist. Sie deutet vielmehr an, dass hier ein Lemma fehlt, für das noch ein Autor gesucht wird. Im Gegensatz zu konventionellen Enzyklopädien – dies gilt auch für deren digitalisierte Fassungen – sind in der *Wikipedia* die Lemmata auch mit ihren Entsprechungen in anderssprachigen *Wikipedias* verknüpft (*Interlanguage-Links*), so dass der Leser – entsprechende Sprachkompetenz vorausgesetzt – einfachen Zugang zu kulturbedingt abweichenden Sichten auf einen Gegenstandsbereich hat.

### 2.3.7 Vorlagen

Zur Standardisierung der Struktur ihrer Artikel bietet die *Wikipedia* mittlerweile eine große Menge von Formatvorlagen, die Musterstrukturen für Biographien, Städtebeschreibungen und viele andere Dinge mehr definieren. Die deutsche *Wikipedia* stellt derzeit 93 Artikelvorlagen zur Verfügung (Wikipedia 2007c). Diese Formatierungsvorschläge sind eine große Hilfe für Autoren und stellen zusätzlich ein wichtiges Instrument der Qualitätssicherung dar, indem sie eine Vereinheitlichung der Artikelstrukturen befördern.

Weitere Textbausteine (Wikipedia 2007d) dienen der Vereinheitlichung der Kommentierung von Artikeln, wie etwa Qualitätsurteile, Urheberrechtsvermerke usw.

## 2.4 Das Qualitätsmodell der Wikipedia

Die *Wikipedia* verfügt über ein vergleichsweise elaboriertes aber nicht geschlossen ausformuliertes Qualitätsmodell. Es können produkt- und prozessorientierte Qualitätsmerkmale unterschieden werden, die in Dokumentationstexten zur *Wikipedia* definiert werden.

### 2.4.1 Produktorientierte Qualitätsmerkmale

Zunächst wird ein Grundverständnis vorausgesetzt, was die Qualität einer Enzyklopädie auszeichnet. Eine Explizierung dieser Qualitätsmerkmale erfolgt fallorientiert erst dann, wenn ein Dissens über bestimmte Kriterien zu beobachten ist. So entstand eine Kollektion von erläuternden Texten zu einigen zentralen Qualitätskriterien (Wikipedia 2007f), wie Relevanzkriterien, angemessenem Textlayout etc.

Neben dem Ziel, eine Enzyklopädie aufzubauen, werden drei Qualitätsmerkmale in zahlreichen Texten (etwa Wikipedia 2007e) besonders hervorgehoben:

- Neutraler Standpunkt: Die Darstellung von Inhalten soll von einem neutralen Standpunkt aus erfolgen, der konkurrierende Theorien oder Anschauungen berücksichtigt. Eine an der Relevanz des jeweiligen Ansatzes orientierte Gewichtung ist dabei ausdrücklich erlaubt. Diese Forderung soll einerseits einer möglichst

objektiven Darstellung dienen, andererseits ist sie eine wichtige Voraussetzung, Konsens in einer großen Autorenschaft herzustellen, in der selbst extreme Meinungen ihre Vertreter finden.

- **Verifizierbarkeit:** Alle Inhalte eines *Wikipedia*-Artikels sollen durch Quellenangaben abgesichert sein. Da die Reputation der *Wikipedia* nicht durch Verweis auf die fachliche Autorität eines Herausgebergremiums abgestützt werden kann, muss der Nachweis der Korrektheit der *Wikipedia*-Inhalte durch Referenz auf verlässliche Quellen erbracht werden.
- **Keine Originalforschung:** In der *Wikipedia* soll gesichertes Wissen publiziert werden. Sie ist deshalb kein Ort für die Veröffentlichung von Forschungsergebnissen.

Hinweise auf die konkreten Qualitätsmerkmale, die für wichtig gehalten werden, ergeben standardisierte Bearbeitungshinweise (Wikipedia 2007g), die auf Unvollständigkeit oder Unverständlichkeit des Artikels, fehlende Quellenangaben oder Verletzung des Neutralitätsgebots aufmerksam machen. Die englische *Wikipedia* stellt in gedrängter Form die Kriterien zusammen, welche einen exzellenten Artikel auszeichnen (Wikipedia 2007h). Er ist gut geschrieben, umfassend, korrekt, neutral und stabil. Er ist den Richtlinien entsprechend gegliedert, wenn sinnvoll mit Bildern ausgestattet und von angemessenem Umfang. All dies sind keine präzise definierten, exakt im Sinne einer objektiven Evaluation operationalisierbaren Kriterien. Sie dienen jedoch als Argumentationsgrundlage für Entscheidungen im Qualitätssicherungsprozess der *Wikipedia*.

#### 2.4.2 *Der Qualitätssicherungsprozess der Wikipedia*

Der Qualitätsprozess der *Wikipedia* basiert primär auf der Aufmerksamkeit der Nutzergemeinschaft. Jeder ist gehalten, Artikel zu kommentieren, mit Qualitätsvermerken zu versehen oder – günstigstenfalls – zu verbessern. Die Offenheit dieses Prozesses wird von der Sicherheit getragen, dass jede Version eines Artikels wieder herstellbar ist, wenn eine Änderung mit oder ohne Absicht zu einer Verschlechterung seiner Qualität führen sollte.

Um einen Anreiz zur Verbesserung der Artikelqualität zu schaffen und um dem Leser eine Orientierungshilfe zu geben, werden die Artikel der *Wikipedia* in Qualitätsstufen unterteilt. Neben der Normalstufe werden *lesenswerte* und *exzellente* Artikel unterschieden. Manche Portale unterscheiden noch weitere Qualitätsstufen (s.o.). Die Zuweisung einer erhöhten Qualitätsstufe erfolgt auf Vorschlag durch beliebige Nutzer und wird durch einen wohldefinierten Abstimmungsprozess bestätigt oder verworfen. Ein ähnlicher Prozess ermöglicht die Löschung besonders schlechter Artikel, insbesondere solcher, die gegen zentrale Auflagen – etwa die Wahrung der Belange des Urheberrechts – verstoßen. Der Vollzug derartiger Maßnahmen bleibt einer eingeschränkten Nutzergruppe – den Administratoren – vorbehalten, die über besondere Rechte verfügen.

#### 2.5 *Zusammenfassung*

Aus der ursprünglichen Idee, eine frei verfügbare, von einer offenen Nutzergruppe getragene Enzyklopädie zu schaffen, die durch ein Wiki-System zugänglich wird, ist mittlerweile ein nicht nur umfangreiches sondern auch sehr komplex strukturiertes Medienprodukt geworden. Während immer wieder festgestellt und kritisiert wird, wo die *Wikipedia* hinter konventionellen Enzyklopädiën zurückbleibt, wird weniger betont, dass sie in manchen Belangen – etwa der Aktualität oder der Interlingualität – schon über diese hinausgeht. Es ist

daher nicht offensichtlich, dass die an konventionelle Enzyklopädien anzulegenden Qualitätskriterien allein Maßstab für die *Wikipedia* sein können.

### 3 Studien zur Qualität der *Wikipedia*

Zunächst sollen einige Studien vorgestellt werden, die sich mit Aspekten der Qualität der *Wikipedia* befassen. Diese werden jeweils hinsichtlich ihrer Methodik und Reichweite untersucht.

#### 3.1 *Die Nature-Studie (Giles 2005)*

Im Folgenden soll der Ablauf der Studie kurz skizziert werden. Zunächst wurden 50 Artikel ausgewählt, die ein breites naturwissenschaftliches Themenfeld abdecken sollten. Die Darstellungen aus der *Wikipedia* und der Online-Ausgabe der *Encyclopedia Britannica* sollten sich ungefähr im Umfang entsprechen, so dass letztendlich nur 42 Artikel zur Evaluation verblieben. Die Themen waren so ausgewählt, dass in der Redaktion Kompetenz vorhanden war. Zu jedem der Themen wurde ein Experte ausgewählt, der jeweils ein Paar aus einem *Wikipedia*- und einem *Britannica*-Artikel zu bewerten hatte, ohne die Zuordnung zu kennen. Die Gutachter sollten in den Artikeln drei Arten von Fehlern identifizieren: sachliche Fehler, kritische Auslassungen, irreführende Formulierungen. Die Gutachten wurden dann vom Redaktionsteam gesichtet, die Fehlerangaben wurden gewichtet und summiert. Es stellte sich heraus, dass die beiden Enzyklopädien keine Unterschiede hinsichtlich gravierender Fehler aufwiesen (je vier), in Artikeln der *Wikipedia* aber deutlich mehr „kleine“ Fehler zu beobachten waren (162/123). Besagte Studie wurde durch die Redaktion der *Britannica* heftig kritisiert, da Ungereimtheiten hinsichtlich der Textgrundlage und der Bewertung von Fehlern bestünden. Die Kritik wurde von *Nature* zurückgewiesen.

Unabhängig von dieser Kritik ist jedoch zu beachten, dass die Reichweite der Studie sehr begrenzt ist. Folgende Einschränkungen sind zu bemerken:

- Die Untersuchung evaluiert nur die Korrektheit von Lexikoneinträgen. Andere Qualitätsaspekte wie Lesbarkeit und Verständlichkeit, Konsistenz in der Gliederung, thematische Abdeckung etc. blieben unberücksichtigt. Die Vermutung, dass sich die Qualität redaktioneller Arbeit gerade bei diesen Qualitätsaspekten zeigen kann, ist nahe liegend.
- Die ausgewählten Artikel repräsentieren nicht die zu untersuchenden Grundgesamtheiten (Artikel der *Wikipedia* bzw. *Britannica*). Es ist davon auszugehen, dass naturwissenschaftliche Fragestellungen andere *Wikipedia*-Autoren anziehen als etwa Themen der Allgemeinbildung. Die Studie besitzt also Aussagekraft allenfalls für naturwissenschaftliche Themen.

#### 3.2 *Studie von Christian Schlieker (Schlieker 2005)*

In seiner Diplomarbeit befasst sich Christian Schlieker mit elektronischen Enzyklopädien. Neben anderen Fragestellungen, die hier nicht weiter erörtert werden sollen, nimmt eine komparative Studie bezüglich der Qualitätsparameter Abdeckung und informationelle Absicherung breiten Raum ein. Untersucht wurden die deutsche *Wikipedia*, *Encarta* und eine Druckfassung der Enzyklopädie *Der Große Brockhaus* (20. Auflage). Aus jeder der Enzyklopädien wurde eine Zufallsstichprobe von 15 Artikeln gezogen (Schlieker 2005, S. 43



ff). Dabei wurden nur Vollartikel und keine Begriffsklärungen o.Ä. berücksichtigt. Aus der *Wikipedia* wurde die Stichprobe mit Hilfe der Zufallsfunktion gezogen, aus der *Encarta* mangels besserer Alternative durch „zufälliges“ Scrollen im alphabetischen Index bestimmt und aus der *Brockhaus*-Enzyklopädie durch ein mehrstufiges Zufallsverfahren (random walk), das zunächst den Band, dann Spalte, Seite und Lemma auswählt. Das Auswahlverfahren bevorzugt vermutlich für den *Brockhaus* längere Artikel, da dann auf einer Seite geringer Wahlmöglichkeiten bestehen. Bei der *Encarta* ist keine gleichmäßige Berücksichtigung des Alphabets zu erwarten, da „zufälliges Scrollen“ die Endpositionen der Liste benachteiligt. Kritischer für die Aussagekraft der Untersuchung dürfte jedoch der geringe Stichprobenumfang sein.

Die Stichproben benutzte Schlieker (2005, S. 51 ff) zunächst, um die thematische Abdeckung der Enzyklopädien zu untersuchen:

- Er überprüfte in welchem Maße die einzelnen Enzyklopädien die Gesamtmenge von 45 Lemmata abdecken. Es ergab sich die Rangfolge *Brockhaus* - *Wikipedia* – *Encarta* (33, 29, 25 von 45).
- Anschließend nahm er eine Kategorisierung der 45 Lemmata vor, um aus der jeweiligen Abdeckung auf thematische Profile der Enzyklopädien zu schließen. *Brockhaus* wurde ein geringerer Aktualitätsbezug unterstellt als der *Wikipedia*, während bei der *Encarta* eine Bevorzugung amerikanischer Fragestellungen beobachtet wurde.

Weiterhin wurde die mediale Aufbereitung der Artikel untersucht. Die Artikel der *Encarta* waren am umfassendsten mit Medienobjekten (Bilder, Audio, Video) ausgestattet, während die *Wikipedia* noch hinter dem *Brockhaus* zurückblieb. Erwartungsgemäß lag die *Wikipedia* hinsichtlich des internen Vernetzungsgrades vorne, gefolgt von *Encarta* und *Brockhaus*. Eine Absicherung der Information durch externe Literaturangaben war am häufigsten beim *Brockhaus* zu finden, gefolgt von *Encarta* und *Wikipedia*.

### 3.3 Die Studie von Hammwöhner, Fuchs, Kattenbeck und Sax (Hammwöhner et al. 2007)

In dieser Studie wurde die deutsche *Wikipedia* mit einer Print-Ausgabe des *Brockhaus* verglichen. Aus jeder Enzyklopädie wurden 50 Lemmata zufällig gezogen, die jeweils in der anderen auch enthalten waren. Um diese Stichprobe zu erhalten, mussten aus der *Wikipedia* 87 und aus dem *Brockhaus* 205 Artikel gezogen werden. Die vergleichbaren Artikel wurden nach folgenden Qualitätsparametern verglichen: Umfang und Vollständigkeit der Artikel, Quantität und Qualität der Quellenangaben sowie formale Korrektheit. Eine inhaltliche Überprüfung der Artikel erfolgte nicht. Zusätzlich wurde überprüft, ob die thematische Zuordnung der insgesamt gezogenen Artikel eine ähnliche Themenverteilung ergibt.

Bezüglich der thematischen Abdeckung und der formalen Absicherung durch Quellenangaben fällt das Urteil dieser Studie zugunsten der *Wikipedia* aus. Deren Artikel waren zudem ausführlicher. Die formale Korrektheit der Artikel – bis hin zur Orthographie – war hingegen im *Brockhaus* deutlich besser. Es lässt sich jedoch feststellen, dass die *Wikipedia* im den 2 Jahren, die zwischen der Studie von Schlieker und dieser Studie verstrichen sind, erhebliche Fortschritte erzielt hat. Zur Themenverteilung lässt sich zusammenfassend anmerken, dass beide Enzyklopädien einen hohen Anteil an biographischer Information enthalten. Während der *Brockhaus* einen höheren Anteil an mathematisch-naturwissenschaftlichen Themen vertrat, hatte die *Wikipedia* einen Bonus bei aktuellen Themen wie Computertechnik, Film

und Musik. Einige der in der *Wikipedia* vertretenen Themen – etwa aus dem Bereich der Esoterik – werden für eine traditionelle Enzyklopädie wohl auch als irrelevant angesehen.

### 3.4 *Die Studien von Stvilia, Twidale, Gasser und Smith (Stvilia et al. 2005a, 2005b)*

Die vorgenannten Studien gingen jeweils von einem impliziten Qualitätsmodell aus, welches die Grundlage für eine vergleichende Bewertung der *Wikipedia* bildete. Die Studien von Stvilia, Twidale, Gasser und Smith entwickeln zunächst ein Qualitätsmodell für die *Wikipedia* (Stvilia et al 2005a). Grundlage dieses Qualitätsmodells ist die in der (englischen) *Wikipedia* vorgenommene Hervorhebung von „Featured Articles“. Grundlage sind zunächst die Qualitätskriterien der *Wikipedia*: „Comprehensive, Accurate, Verifyable, Stable, Well-written, Uncontroversial, Compliance, Appropriate Images with acceptable copyright status, Appropriate Style and Focus“. Diesem Kriterienkatalog stellen sie einen von Crawford vorgeschlagenen gegenüber, der sich auf traditionelle Print-Enzyklopädien bezieht: „Scope, Format, Uniqueness, Authority, Accuracy, Currency, Accessibility“. Auch einen eigenen, umfangreicheren Kriterienkatalog, der hier nicht dargestellt werden soll, setzen sie zu den beiden anderen in Beziehung. Damit ist eine Grundlage für die Systematisierung von Qualitätsstudien sowie den Vergleich verschiedener Qualitätsmodelle gegeben. Im empirischen Teil der Studie stellen die Autoren sodann den Zusammenhang zwischen abstrakten Qualitätsmodellen und der Praxis der *Wikipedia* her. Dies erfolgt aufgrund einer inhaltlichen Analyse der den Lemmata zugeordneten Diskussionsseiten. Dazu wurden aus der *Wikipedia* und der Menge der Featured Articles jeweils 30 Artikel ausgewählt, die über einen hinlänglich umfangreichen Diskussionsteil verfügten. Die in der Diskussion auftretenden Argumente wurden kategorisiert und den Qualitätsdimensionen der *Wikipedia* zugeordnet. Somit wurde eine Einschätzung der Bedeutung der einzelnen Qualitätsaspekte für die Einordnung als Featured Article möglich. Zusätzlich wurde versucht, Textparameter zu isolieren, welche als Symptome für eine hohe Qualität der Artikel anzusehen sind (Umfang und Struktur der Diskussion, Umfang und Überarbeitungsfrequenz der Artikel, Anzahl der Autoren etc.). Dieser quantitative Ansatz wurde in einer weiteren Studie vertieft (Stvilia et al 2005b), so dass ein automatischer Qualitäts-Klassifikator entwickelt werden konnte.

### 3.5 *Weitere Studien mit Qualitätsbezug*

Die Studie von Andrew Lih (2004) befasst sich mit der Verlässlichkeit der *Wikipedia* als journalistische Quelle. Er entwickelt eine einfache Metrik, welche die Qualität von Artikeln erfassen soll. Sie beruht auf der Anzahl von Änderungen, die ein Artikel durchlaufen hat und der Anzahl der beteiligten Autoren. Während eine zunehmende Zahl der Bearbeitungen für eine hohe inhaltliche Konsolidierung eines Beitrags spräche, erhöhe eine große Anzahl von Autoren die Wahrscheinlichkeit von Inkonsistenzen und Argumentationsbrüchen. Die Metrik wurde an einem Testkorpus kalibriert, um dann an Artikeln, die von der Presse zitiert wurden, getestet zu werden. Diese Studie initiierte den Einzug quantitativer Methoden in die Analyse der *Wikipedia*.

Emigh und Herring (2004) vergleichen zwei elektronische Enzyklopädien (*Wikipedia* und *Everything2*) hinsichtlich des verwendeten Sprachregisters. Sie beobachten in *Everything2* eine vergleichsweise freie Sprache, während die *Wikipedia* zunehmend eine eher förmliche Sprache verwendet. Wenngleich diese Studie keinen unmittelbaren Bezug zur Frage der Qualität der *Wikipedia* hat, ist doch von Interesse, dass der Redaktionsprozess eine

Vereinheitlichung der Sprache zur Folge hat. Dieser Effekt ist als relevant für Qualitätskriterien wie Lesbarkeit und sprachliche Konsistenz anzusehen.

Eine sehr aktuelle Studie hat Wiegand (2007) vorgestellt, welche die *Wikipedia* mit drei anderen digitalen Enzyklopädien verglich (Brockhaus, Bertelsmann, Encarta). Anhand von 150 Testbegriffen werden Korrektheit, Vollständigkeit, Verständlichkeit, Aktualität und mediale Aufbereitung der drei Enzyklopädien verglichen. Insgesamt wird nicht ganz deutlich, wie die Auswahl der Artikel motiviert war. Es wurde offensichtlich Wert darauf gelegt, dass einige wichtige Themengebiete repräsentiert waren. Auch die Operationalisierung der Kriterien wird nicht offen gelegt. Zusätzlich zu den inhaltlichen Kriterien wird auch das Bedienkonzept der Zugangssoftware in die Bewertung einbezogen. Auch diese Studie belegt, dass sich die *Wikipedia* auf Augenhöhe mit den anderen Enzyklopädien bewegt, in der Darstellung eher zu ausführlich ist und nicht signifikant mehr Fehler aufweist.

### 3.6 Zusammenfassende Würdigung

Bisher sind noch vergleichsweise wenige Studien zur Qualität der *Wikipedia* erschienen. Diese sind zudem methodisch nicht unangreifbar. Auffällig ist die Diskrepanz zwischen den aus Studien resultierenden Qualitätsurteilen und den zum Teil witzigen Pressemeldungen über aktuelle, zum Teil krasse Fehlleistungen der *Wikipedia* (z.B. Rühle 2006). Dies mag nicht zuletzt an den methodischen Problemen der Studien liegen, hat seinen Hauptgrund aber vermutlich darin, dass diese Studien ihren Blick auf die mittlere Qualität einer Grundgesamtheit richten, die Presse aber immer zu Recht auf besonders problematische Einzelfälle verweist. Darin liegt kein Widerspruch. Eine (traditionelle) Enzyklopädie muss sich der Anforderung stellen, dass auch der schlechteste Artikel gewisse Qualitätsschranken nicht unterschreitet. Insofern sind die oben genannten Studien – ebenso wie die noch folgenden – nur begrenzt geeignet, Werturteile über die *Wikipedia* oder andere Enzyklopädien zu begründen. Sie sind jedoch von Nutzen, die Effekte zu verstehen, die aus den neuen Redaktionsprozessen der *Wikipedia* resultieren. Besonders viel versprechend scheinen dabei solche Ansätze zu sein, die es erlauben, Qualitätsindikatoren automatisch für große Textmengen zu bestimmen. In diese Richtung verweisen die Arbeiten von Stvilia et al. (2005a, 2005b).

Wegen des andauernden Entwicklungsprozesses, in dem sich die *Wikipedia* befindet, ist es sinnvoll, auch weiterhin Vergleiche zu traditionellen Enzyklopädien zu ziehen. Die oben erwähnten Studien von Schlieker und Hammwöhner et al. verdeutlichen, dass in diesem Bereich nur temporäre Aussagen möglich sind. Es ist aber zu verzeichnen, dass in der *Wikipedia*-Forschung noch große Lücken existieren. So gibt es bisher kaum vergleichende Forschung, die Korpora aus verschiedenen *Wikipedias* gegenüberstellen. Zu wenig untersucht ist auch der Paratext der *Wikipedia*. Dies gilt vor allem für die Texte, die der Wissensorganisation dienen. Vernachlässigt hat die *Wikipedia*-Forschung bisher auch weitgehend die Perspektive der Nutzer. Bisherige Studien beruhen überwiegend auf Korpusauswertungen, ohne dass Nutzertests durchgeführt wurden. Im folgenden sollen einige Studien vorgestellt werden, die sich mit diesen bisher vernachlässigten Themen befassen. Wegen der Begrenzungen, denen diese Arbeiten unterlagen, werfen sie vermutlich mehr Fragen auf als sie beantworten, zeigen jedoch hoffentlich, welche interessanten Forschungsfelder der Wikipedistik noch harren.

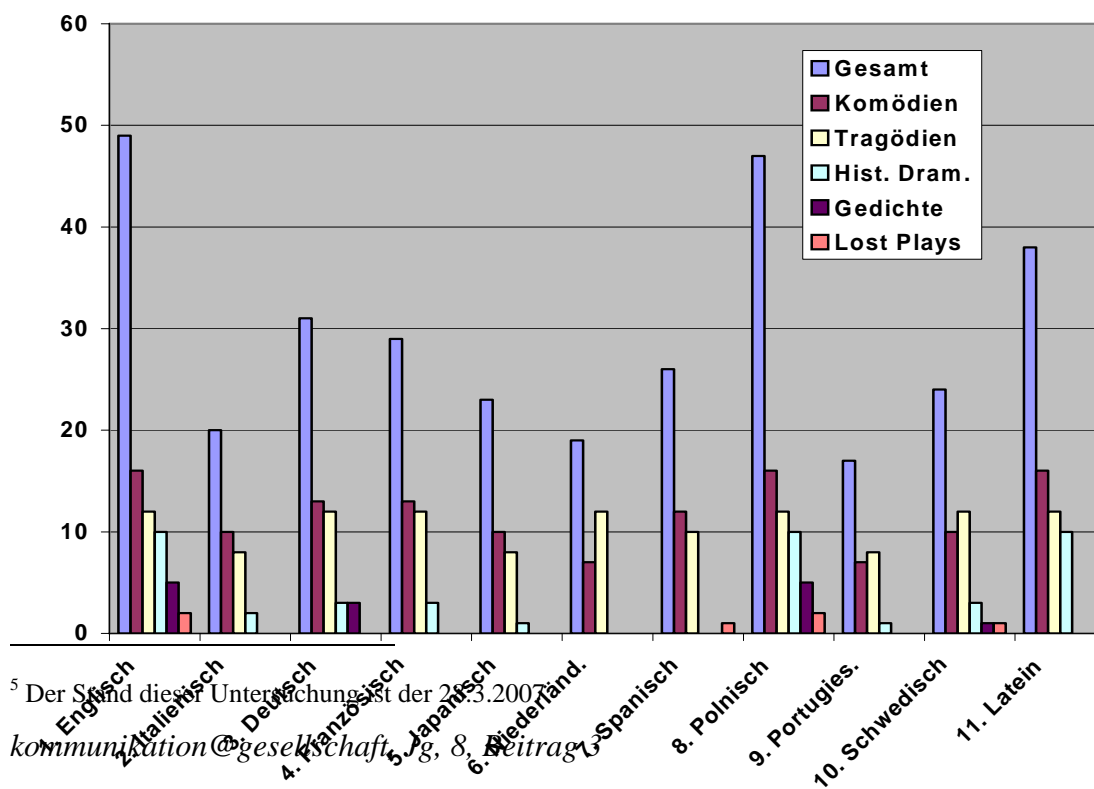
## 4 Shakespeare's Werk in der Wikipedia

Diese erste Studie sollte an einem thematisch sehr eingeschränkten Gegenstandsbereich durchgeführt werden, in dem das Projektteam – Studenten der Anglistik – Beurteilungskompetenz aufwiesen. Damit war die Möglichkeit gegeben, die *Wikipedia*-Artikel auf ihre fachliche Qualität hin zu überprüfen. Zusätzlich wurde hiermit die Möglichkeit eröffnet, die Repräsentation dieses Gegenstandsbereichs in mehreren *Wikipedias* zu überprüfen und zu vergleichenden Urteilen zu kommen. Die Wahl fiel auf die Darstellung von Shakespeare's Werk in der *Wikipedia* – biographische und historische Angaben wurden nicht berücksichtigt.

### 4.1 Der weitere Rahmen

Zunächst wurde untersucht, wie Shakespeare's Werk – das sind 16 Komödien, 10 historische Dramen, 12 Tragödien, 5 Gedichte, 2 verlorene Dramen und 4 apokryphe Texte – in der Gesamt-*Wikipedia* repräsentiert ist<sup>5</sup>. Dabei wurde zunächst nur festgestellt, zu welchen dieser 48 Werke in welchen der *Wikipedias* ein Lemma vorhanden ist. Es ließen sich 44 *Wikipedias* finden, in denen mindestens eines der Werke vertreten war. Die einzige *Wikipedia*, die eine vollständige Abdeckung erzielte war – nicht ganz überraschend – die englische. Für die anderen *Wikipedias* ergab sich folgendes Bild: Tabelle 1 zeigt die Abdeckung von Shakespeare's Werk für die 11 vollständigsten *Wikipedias* auf diesem Gebiet. Auffällig ist, dass die Rangfolge im Grad der Abdeckung weitgehend mit der Rangfolge der Gesamtgröße der *Wikipedias* übereinstimmt. Ausreißer stellen hier die polnische und die lateinische *Wikipedia* dar. Eine etwas andere Reihung – dargestellt durch die Numerierung in der Spaltenbeschriftung des Diagramms – ergibt sich, wenn man den Umfang der Artikel zu den Werken berücksichtigt. Die Artikel zu den 7 bestrepräsentierten Werken wurden nach Umfang geordnet. Die Reihung der *Wikipedias* ergibt sich dann aus der Rangsumme ihrer Artikel. Hier fällt die ausgezeichnete Positionierung der italienischen *Wikipedia* auf, die auch den einzigen exzellenten Artikel zu Shakespeare's Werk aufweist (*Romeo and Juliet*).

Tabelle 1: Repräsentation von Shakespeare's Werk verschiedenen Wikipedias

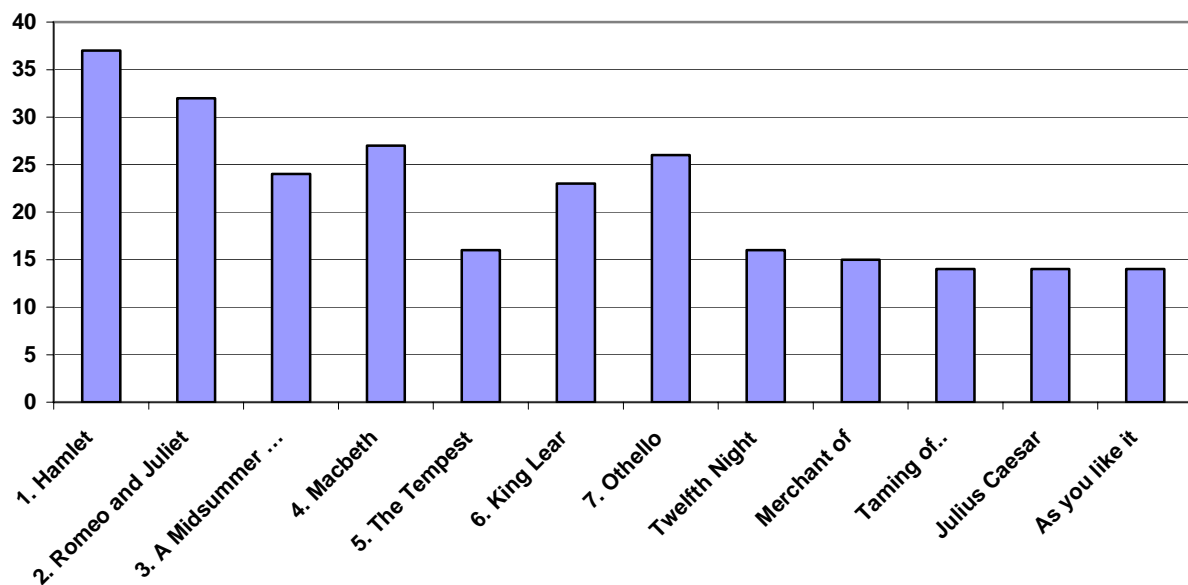


<sup>5</sup> Der Stand dieser Untersuchung ist der 23.2.2007.  
 kommunikation@gesellschaft.org, 8. Beitrag  
<http://nbn-resolving.de/urn:nbn:de:0228-200708066>

Ordnet man die Dramen nach der Anzahl der *Wikipedias*, in denen sie repräsentiert sind, so ergibt sich Tabelle 2 (Einen Artikel zum Hamlet etwa findet man in 37 *Wikipedias*). Auffällig ist, dass kein Drama in allen der 44 untersuchten *Wikipedias* enthalten ist, dass sich aber eine klare Stufung im Grad der Repräsentation ergibt. Bemerkenswert ist der hohe Grad an Abdeckung selbst für die weniger prominenten Werke. Zum Vergleich: Goethes *Faust I* erzielt 12 Nennungen, ebenso wie die *Blechtrommel* und die *Dreigroschenoper*, gefolgt vom *Steppenwolf* (11) und den *Buddenbrooks* (9). Spitzenreiter dieser unsystematischen Auswahl aus deutscher Literatur ist das Nibelungenlied mit 17 Nennungen. Es wird deutlich, das Shakespeare's Werk in den größeren *Wikipedias* in einem Maße durch eigene Lemmata vertreten ist, das auch von größeren Print-Enzyklopädien nicht zu erwarten ist.

Ein Indikator für die Bedeutung, die einem Stück in einer nationalen Wikipedia zugemessen wird, ist der Umfang der Darstellung (Anzahl der Worte). So ergibt sich für jede Wikipedia eine Rangfolge der Dramen. Die Summe dieser Ränge kann dann als ein Hinweis auf eine interkulturelle Rezeptionsintensität angesehen werden. Diese Rangbildung wurde, bezogen auf die elf in Tabelle 1 erwähnten *Wikipedias*, zunächst nur für die Dramen 1-7 durchgeführt und in Tabelle 2 durch die Numerierung der Bezeichner angedeutet. *The Tempest* wurde also in weniger *Wikipedias* besprochen als *King Lear*, dort aber in größerer Ausführlichkeit.

**Tabelle 2: Repräsentation einzelner Dramen in den Wikipedias**



Seinen ersten Platz verdankt der *Hamlet* vor allem seiner Wertschätzung in der deutschen, polnischen und niederländischen *Wikipedia*, die wiederum dem zweitplatzierten Drama *Romeo and Juliet* vergleichsweise wenig Platz einräumten. Auffällig ist auch die eindeutige Erstplatzierung von *A Midsummer Night's Dream* in der französischen *Wikipedia*.

#### 4.2 Ein Vergleich der deutschen mit der englischen Wikipedia

Das Hauptgewicht dieses Teils der Studie<sup>6</sup> lag auf dem Vergleich der Qualität, in der Shakespeare's Werk in der englischen und deutschen *Wikipedia* repräsentiert ist. Grundlage dieser Studie war ein Korpus aller *Wikipedia*-Artikel der oben genannten Werke aus der englischen und der deutschen *Wikipedia*<sup>7</sup>. Nicht berücksichtigt wurden dabei die verlorenen Dramen und die apokryphen Texte. Es war nunmehr eine Kollektion von Qualitätskriterien zu definieren und so zu operationalisieren, dass sie anhand des Korpus überprüft und zu einem Vergleich herangezogen werden konnten. Es wurde von vornherein darauf verzichtet, eine zusammenfassende Globalbewertung für jeden Artikel zu erstellen.

##### 4.2.1 Vollständigkeit

Vollständigkeit einer Enzyklopädie kann auf zwei Ebenen untersucht werden. Zum einen ist die angemessene Repräsentation des Gegenstandsbereichs durch Lemmata zu gewährleisten – dieses Teilkriterium wurde schon in der Voruntersuchung beleuchtet. Es geht aber auch um die Vollständigkeit der einzelnen Artikel. Einen ersten, vergleichsweise leicht zu erhebenden Hinweis auf Vollständigkeit liefert der Umfang der Artikel. Mit wenigen Ausnahmen – der deutschsprachige Artikel zu *Hamlet* und *Twelfth Night* sind geringfügig länger als der englische – sind die Artikel der englischen *Wikipedia* deutlich länger. Im Durchschnitt ist Verhältnis der Artikellängen 2:1, für einzelne Artikel – etwa zu Richard III. – ist es deutlich ungünstiger (etwa 8:1).

<sup>6</sup> Die hier aufgeworfenen Fragen wurden überwiegend von Johannes Molz und Manuel Burghardt (2006) bearbeitet.

<sup>7</sup> Das Korpus wurde am 1.2.2006 extrahiert.

Ein weitaus aussagekräftigerer Indikator für die Vollständigkeit eines Artikels ist das Vorhandensein einer kanonischen Menge von Angaben. Dies ist möglich, da eine adäquate Beschreibung eines Bühnenstücks einige zentrale Gesichtspunkte zu berücksichtigen hat. Dies sind Informationen über a) den Inhalt, b) die handelnden Figuren, c) Adaptionen des Stücks, d) Entstehungszeitraum, e) Veröffentlichungsgeschichte, f) Rezeptions- und Aufführungsgeschichte. Für die fünf Gedichte sind andere Angaben zu erwarten. Über handelnde Personen oder Aufführungsgeschichte ist hier nichts zu sagen. Statt dessen ist eine Strukturanalyse des Gedichts zu wünschen.

Anhand dieses Katalogs wurde die Artikel der beiden *Wikipedias* so bewertet, dass für das Vorhandensein von Information zu den einzelnen Rubriken jeweils auf eine standardisierte Weise Punkte vergeben wurden. Hier erreichte die englische *Wikipedia* im Durchschnitt 40,7% der erreichbaren Punktzahl, die deutsche nur 29,3%. Beide Werte erwecken zunächst einen nachteiligen Eindruck. Es ist deshalb wichtig zu betonen, dass konventionelle Print-Enzyklopädien hier vermutlich erheblich schlechtere Werte erzielen, da sie Einzeldarstellungen künstlerischer Werke allein aus Platzgründen nicht in vergleichbarem Umfang vornehmen können. Jedenfalls fällt das Urteil zugunsten der englischen *Wikipedia* aus.

#### 4.2.2 Formale Korrektheit

Ein Enzyklopädieartikel erfordert eine angemessene Sprache. Dies bedingt die Auswahl eines angemessenen Sprachregisters, setzt aber auch orthographische Korrektheit voraus. Fehler in Orthographie und Interpunktion lassen sich vergleichsweise leicht und objektiv feststellen – für die deutsche *Wikipedia* wurde die neue Rechtschreibung vorausgesetzt. Die Identifikation unangemessener Sprache – etwa saloppe oder umgangssprachliche Formulierungen – beruhte, vermutlich nicht unanfechtbar, allein auf dem Sprachgefühl der beiden Juroren. In beiden Fällen steht wieder die englische *Wikipedia* besser da. In der englischen *Wikipedia* fallen 0,5 orthographische Fehler auf 1000 Worte, in der deutschen sind es 4,4. Erstere weist nur zwei Artikel auf, deren Stil zu bemängeln war (*Measure for Measure*, *Richard III*), während bei letzterer nur 7 Artikel völlig ohne Kritik blieben. Anzumerken ist, dass kein Zusammenhang zwischen der formalen Korrektheit eines Artikels und seiner inhaltlichen Qualität festzustellen war. Ein Schluss von formalen Fehlern auf allgemeine Mängel an Sorgfalt ist also, zumindest in diesem Korpus, nicht möglich.

#### 4.2.3 Neutraler Standpunkt

Die Wahrung eines neutralen Standpunkts gehört zu den Grundprinzipien der Arbeit an der *Wikipedia*. Dieses Kriterium ist also von zentraler Bedeutung. Die Inhalte der Artikel dieses Korpus scheint aber zunächst nur wenig Risiken zu tendenziöser Aussage zu bieten. Verschiedentlich schleichen sich jedoch wertende Aussagen über das zu beschreibende Stück in die Beschreibung ein. Häufigste Ursache der Kritik ist jedoch die Verwechslung von Fakten und Theorien. Im Fall von William Shakespeare und seinen Werken sind jegliche Datumsangaben – etwa zu Uraufführungen – als Teil wissenschaftlicher Theorien anzusehen und deshalb in diesen Zusammenhang zu stellen. Jegliche absolute Datumsangabe – selbst wenn sie vom aktuellen Stand der Shakespeare-Forschung gestützt werden sollte – stellt eine Verletzung der Forderung nach einem neutralen Standpunkt dar. Derartige Fehler sind jedoch allein in der deutschen *Wikipedia* (4 Fälle) anzutreffen. In der englischen *Wikipedia* ist in einem Fall eine wertende Aussage zu einem Drama (*Cymbeline*) zu verzeichnen.

#### 4.2.4 Fachliche Korrektheit

Zunächst wurde die fachliche Korrektheit der in den Beiträgen getroffenen Aussagen überprüft. Basis der Beurteilung war dabei die Fachliteratur.<sup>8</sup> Bewertet wurden alle Angaben zu Inhalt und Einordnung der Werke. Während in der englischen Wikipedia 4 Fehler in drei Artikeln beobachtet wurden, waren es in der deutschen 8 Fehler in 6 Artikeln.

#### 4.2.5 Informationelle Absicherung

Zu den Grundmaximen der *Wikipedia* gehört auch, dass jegliche Angaben angemessen durch Referenzen zu belegen sind. Sie ist folglich kein Ort, um Originalforschung zu publizieren. Die Artikel sowohl aus dem deutsch- als auch dem englischsprachigen Teil des Korpus sind überwiegend in einem Maße mit Literaturangaben ausgestattet, das für die deutsche *Wikipedia* als ausreichend angesehen werden kann. Hier wird der englischen *Wikipedia* ihre Ausführlichkeit jedoch zum Problem. Weitergehende Ansprüche in der Darstellung bewirken einen erhöhten Bedarf an Quellenangaben und –zitationen. Dies ist in den seltensten Fällen erfolgt. Im Verlauf einiger Monate nach Ziehung des hier genutzten Korpus wurden die Texte der englischen *Wikipedia* im Rahmen des *Wikipedia*-Projekts *Englisches Theater der elisabethanischen Zeit* (Wikipedia 2007i) gesichtet und mit kritischen Vermerken versehen. Während Kritik an Form und Sprache nur in einem Fall ausgesprochen wurde (*Rape of Lucrece*), erhielten 6 von 10 Tragödien, 9 von 16 Komödien sowie 2 von 5 Gedichten Hinweise zu fehlenden Referenzen und nicht belegter Originalforschung. Unter den Historiendramen wurde nur eines in diesem Sinne kritisiert (*Henry IV*, I). Dies ist aus dem Umstand erklärlich, dass die Historiendramen auch in der englischen *Wikipedia* weniger ambitioniert und ausführlich dargestellt werden. Die gründliche und durchgängige kritische Bewertung aller Artikel des Korpus im Rahmen eines *Wikipedia*-Projekts zeigt die Kritikfähigkeit der *Wikipedia*-Community. Dieser Vorgang verdeutlicht jedoch zusätzlich, dass ein erkanntes Problem nicht in allen Fällen ein gebanntes ist. Seit Monaten sind die Kritikhinweise in die Artikel integriert. Trotzdem sind die beobachteten Mängel bisher (Stand 30.3.2007) nicht behoben. Dies ist um so auffälliger, als manche der Artikel zahlreiche Änderungen durchliefen – im Fall von *Romeo and Juliet* sind es Hunderte.

#### 4.2.6 Diskussion

Offensichtlich ist, dass die englische *Wikipedia* nach fast allen Qualitätskriterien besser zu beurteilen war. Dies ist nicht erstaunlich, ist William Shakespeare doch der englische Nationaldichter. Zu denken geben die Ergebnisse der Vorstudie. Shakespeare's Werk ist in allen größeren *Wikipedias* in hohem Maße repräsentiert, man könnte sagen, dass es zu einem Welt-Kulturgut geworden ist. Ähnliche Präsenz in den *Wikipedias* erreichen wohl nur noch Werke der Weltliteratur wie die *Ilias* (42), die *Odysee* (47), *Don Quichote* (42) oder – *Harry Potter und der Stein der Weisen* (40). Bei weitem übertroffen werden diese Werte jedoch von der Präsenz der Bibel in der Wikipedia (134).

Der Grad, in dem Shakespeare repräsentiert ist, korreliert sehr auffällig mit der Größe der *Wikipedia*. Vermutlich hängt es auch mit der Größe der Community zusammen. Darauf deutet hin, dass 3,8 mal so viele Autoren an 4,4 mal so viel Versionen der englischen *Wikipedia*

---

<sup>8</sup> Etwa Boyce 1996; Crystal 2004; Kermode 2001.



gearbeitet haben wie an der deutschen. Die englischen Autoren sind auch die aktiveren. Dem aktivsten englische Autor sind 225, dem aktivsten deutschen nur 34 Artikelversionen zuzurechnen.

Insgesamt bieten beide *Wikipedias* demjenigen, der an Einstiegsinformation über Shakespeare's Werk interessiert ist, verlässliche und brauchbare Information. Die Texte sind, bei gewissen Abstrichen für die deutsche *Wikipedia*, auch gut lesbar. Die Artikel sind jedoch, auch nach den Maßstäben der *Wikipedia*, nicht über dem Mittelmaß anzusiedeln. Keiner der Texte des Korpus ist bisher als exzellent eingestuft worden. Erinnerung sei an die exzellente Darstellung von *Romeo and Juliet* in der italienischen *Wikipedia*.

Kommentierungsbedürftig ist auch der sehr unterschiedliche Umfang, in dem die Werke in den einzelnen *Wikipedias* besprochen wurden. Sind die in Tabelle 2 deutlich gewordenen Unterschiede in der Rangfolge der Dramen nur Effekte einer zufälligen Bearbeitungsabfolge oder geben sie – dafür spricht der klare Abstand zwischen den Werken – nationale Präferenzen wider? Dies wäre noch zu untersuchen. Was im einen Fall nur als mangelnde Qualität einzelner Beiträge gewertet werden könnte, stellte im anderen Fall eine Anpassung an den Informationsbedarf der Leser dar. Offensichtlich ist jedoch, dass bei diesen deutlichen Unterschieden von einer Qualität *der Wikipedia* nicht gesprochen werden kann.

## **5 Eine Evaluation der Wissensorganisation der *Wikipedia*<sup>9</sup>**

In Abschnitt 2.3 dieses Beitrags wurden bereits die Strukturen der Wissensorganisation beschrieben, die in der *Wikipedia* aus einer Menge von Lemmata einen rezipierbaren Text entstehen lassen. Im folgenden soll untersucht werden, inwieweit diese Strukturen den Prozess der Informationssuche geeignet unterstützen. Die zusätzlichen Fragen, ob sie auch dem Autoren- oder dem Qualitätssicherungsprozess dienlich sind, sollen allenfalls am Rande angesprochen werden. Methodisch ist die Studie zweigeteilt. Zunächst werden die Wissensorganisationsstrukturen im Sinne einer Expertenevaluation untersucht. Besonderes Gewicht wird dabei auf das Kategoriensystem gelegt. Im weiteren wird anhand eines Nutzertest festgestellt, inwieweit Nutzer diese Strukturen zur effizienten und effektiven Befriedigung ihres Informationsbedarfs ausnutzen können.

### *5.1 Expertenevaluation*

Bevor ausführlicher auf das Kategoriensystem eingegangen wird, werden die anderen Formen der Wissensorganisation der *Wikipedia* kurz angesprochen.

---

<sup>9</sup> Die folgenden Ausführungen beruhen weitgehend auf Lehnerer (2006).

### 5.1.1 *Alphabetischer Index*

Aufgrund des Umfangs der *Wikipedia* bedarf der alphabetische Index einer hierarchischen Gliederung. In der deutschen *Wikipedia* ist dies analog zu den Print-Enzyklopädien gelöst, indem jeweils Teilbereiche aus dem Alphabet herausgegriffen werden, die etwa eine gleiche Anzahl an Lemmata enthalten. Aufgrund des Umfangs der *Wikipedia* und der Notwendigkeit auch die Listen der nächsten Zugriffsebene nicht zu stark anwachsen zu lassen, enthält die oberste Ebene 70 Unterteilungen schon für den Anfangsbuchstaben A. Dadurch wird die Darstellung sehr unübersichtlich, die entstehenden Gruppen wirken arbiträr (was sie ja auch sind). Da die *Wikipedia* ständig wächst, wird es erforderlich, die Aufteilung des Index ständig dem Wachstum anzupassen. Als Folge wird es dem Leser unmöglich, die entstehende Struktur zu konzeptualisieren. Dieses Problem wird in zahlreichen anderen Wikipedias, etwa der englischen, französischen, polnischen und niederländischen, günstiger gelöst, indem die Einstiegsebene des alphabetischen Index durch eine fest dimensionierte Buchstabenmatrix gebildet wird.

Probleme entstehen auch aus dem Sortieralgorithmus. Personennamen tauchen unter dem Vornamen, der nicht immer bekannt sein dürfte, im Index auf. Großbuchstaben stehen vor Kleinbuchstaben – *TGV* z.B. also vor *Tansania*. Schwierig ist auch die Einordnung von Sonderzeichen, die durchgängig zuletzt angeordnet werden. *Tyrus* steht also vor *Tötensen*. Dies ist zwar auf jeder Seite des Index dokumentiert, entspricht aber nicht den Gewohnheiten und ist deshalb schwer nachvollziehbar und führt zu Fehlern.

### 5.1.2 *Themenportale*

Themenportale bieten einen systematischen Zugang zu den Lexikoninhalten. Sie sind hervorragend geeignet, einen Überblick über ein Themengebiet zu erzielen. Richtet sich die Suche auf einen konkreten Gegenstand, kann ein Problem daraus erwachsen, dass die Themenportale keine vollständige Abdeckung der Enzyklopädie garantieren. Es kann also vorkommen, dass ein vorhandener Artikel bei thematischer Suche nicht gefunden wird.

Verwirren kann die uneinheitliche Gestaltung der Portalseiten, die für jedes Portal einen neuen Orientierungsprozess erfordert. Positiv, nicht nur für den Autor sondern auch für den Informationssuchenden, ist, dass viele Portale explizite Defizitlisten führen, d.h. Listen von Lemmata, die noch in der Enzyklopädie fehlen.

### 5.1.3 *Listen, Namensräume*

Schon der Zugang zu den vielen Listen, die in der *Wikipedia* verborgen sind, wird dem Leser erschwert. Die Startseite der *Wikipedia* liefert keine Hinweise auf ihre Existenz. Sie sind offensichtlich nur für den internen Gebrauch gedacht, werden z.T. auch durch das Kategoriensystem überflüssig. Ihre Bedeutung für den Prozess der Qualitätssicherung ist unbestritten, auf den hier aber nicht weiter eingegangen werden soll.

Auch von der Existenz von Namensräumen muss der Information suchende Nutzer nichts erfahren. Die Namensräume, aus denen er Information bezieht – Artikel, Diskussion, Benutzer – werden durch die Benutzungsoberfläche so gut verständlich zugänglich gemacht, dass die zugrunde liegende Struktur nicht konzeptualisiert werden muss.

### 5.1.4 *Das Kategoriensystem der Wikipedia*

*kommunikation@gesellschaft, Jg, 8, Beitrag 3*

<http://nbn-resolving.de/urn:nbn:de:0228-200708066>

Die Einführung eines umfassenden Ordnungssystems in die *Wikipedia(s)* stellt eine enorme Herausforderung für die wiki-spezifische Form der Kollaboration dar. Aus der traditionell dokumentarischen Perspektive betrachtet, würde eine derartige Aufgabe langwierige Vorarbeiten und einen stringenten Entwurf erfordern (Gaus 1995, S. 165 ff). Dieses Problem war den Initiatoren des *Wikipedia*-Projekts bewusst (Wikipedia 2007j). Man sah sich nicht in der Lage, diesen Aufwand zu betreiben, vermutete in der Community auch weder die Fachkompetenz ein solches Ordnungssystem zu erstellen, noch diejenige ein eingeführtes Ordnungssystem (etwa die Universelle Dezimalklassifikation (UDK) oder die Regensburger Verbundklassifikation) kompetent zu benutzen. Ohne dass eine Entscheidung für ein konkretes Ordnungsprinzip explizit getroffen worden wäre – Debatten um Klassifikationen, Facettenklassifikationen, strukturierte Schlagwortverzeichnisse wurden geführt – begann durch das Vorpreschen aktiver Nutzer die Entwicklung des Ordnungssystems, das in der *Wikipedia* als Kategoriensystem bezeichnet wird.

Es ist leicht zu sehen, dass es sich um ein strukturiertes Vokabular handelt, dass zur Verschlagwortung von Artikeln durch Begriffskombinationen verwendet wird. Um einen vollgültigen Thesaurus handelt es sich nicht, da keine Äquivalenzrelationen unterstützt werden. Eine textuelle Beschreibung des Ordnungssystems gibt es nicht, die Verschlagwortungsregeln sind sehr allgemein und unvollständig (Wikipedia 2007k).

Beeindruckend ist der Umfang, den die kooperativ erstellten Kategoriensysteme mittlerweile erreicht haben. Im Oktober 2006 waren standen 33.800 Kategorien zur Auszeichnung von 485.000 Artikel der deutschen *Wikipedia*, im Juni 2006 141.000 Kategorien für 1.300.000 Artikel der englischen *Wikipedia* zur Verfügung (Lehnerer 2006). Dies ist allerdings kein Qualitätskriterium. Würde das Kategoriensystem wie ein Thesaurus konsequent und fachgerecht zur Verschlagwortung per Begriffskombination gebraucht, könnte das Schlagwortinventar erheblich geringer sein. Eine große Fachdatenbank wie Inspec etwa kommt mit einem Thesaurus von ca. 9.000 Deskriptoren (IEEE 2004) aus, um 9 Millionen Referenzen bei einer Zuwachsrate von 500.000 Referenzen pro Jahr zu erfassen. Eine spezifische Fachdatenbank ist nur begrenzt mit einer Enzyklopädie zu vergleichen. Während im einen Fall die Themen viel breiter gestreut sind, besteht im anderen Fall ein größerer Bedarf an Erschließungstiefe. Dieses Missverhältnis (1:5 hinsichtlich der Dokumente und 15:1 bezüglich der Deskriptoren, bezogen auf die englische *Wikipedia*) ist dennoch interpretationsbedürftig. Folgende Erklärungen bieten sich an:

- Das Kategoriensystem der *Wikipedia* wird in vielen Bereichen eher wie eine Klassifikation benutzt (Voss 2006), das Potential der Erschließung durch Mehrwortkombination wird deshalb nicht ausgeschöpft.
- Die Kategorisierung dient nicht nur der inhaltlichen Erschließung für die Informationssuche, sondern verfolgt auch noch weitere Ziele, etwa verschiedene Formen statistischer Auswertung (Wikipedia 2007j). Über den genauen Umfang der Kategorisierungsziele besteht Unklarheit in der Community (ebenda).
- Die Community verfügt über wenig dokumentarische Fachkenntnisse. Während man bei jedem Nutzer eine Grundintuition über Aufbau und Ziel einer Enzyklopädie voraussetzen kann – aber auch diese wird immer wieder sehr kontrovers diskutiert – besteht sie bezüglich eines Ordnungssystems wohl kaum. Auch die Unterschiede zwischen den verfügbaren Ordnungsprinzipien dürften nicht klar sein. Auch diese

Problematik scheint in den Diskussionen um das Kategoriensystem auf (Wikipedia 2007j).

Die oben genannten Probleme können durch weitere Fakten und Beispiele verdeutlicht werden. Dass *Johann Wolfgang von Goethe* aufgrund seiner vielseitigen Natur nur durch 35 Kategorien angemessen erfasst werden kann, ist noch plausibel. Warum aber findet er sich gleichzeitig in der Kategorie *Reiseschriftsteller* und *Reiseliteratur*? Warum ist *Reiseschriftsteller* eine Unterkategorie von *Reiseliteratur*? Und ist – letzte Frage – Johann Wolfgang von Goethe, nur weil er Reiseliteratur verfasst hat, automatisch auch ein Reiseschriftsteller? Auf diese Fragen kann man dogmatische oder pragmatische Antworten geben. Auch diese sind schon in der internen Diskussion der *Wikipedia* vertreten (Wikipedia 2007j). Die Pragmatiker scheinen sich durchgesetzt zu haben.

Die Kategorie *Mann* und *Frau* enthalten 125.814 bzw. 21.477 Einträge (Wikipedia 2007m). Für eine Statistik über die Repräsentation der Geschlechter in der *Wikipedia* mag diese Information von hohem Interesse sein. Für die Informationssuche sind Kategorien mit derartig wenig diskriminierender Funktion weitgehend wertlos. Im Gegenzug scheinen Kategorien wie *Verfassung der Vereinigten Staaten*, die nicht mehr als 8 Artikel enthalten als überspezialisiert. Eine Kategorie *Grundgesetz*, die aus Gründen der Konsistenz zu erwarten wäre, fehlt – sinnvollerweise. Schwach besetzte Kategorien sind im Kategoriensystem der *Wikipedia* sehr zahlreich vertreten (Wikipedia 2007n).

Eine Spezialliste (Wikipedia 2007l) verzeichnet nur 138 undefinierte Kategorien, die als Schlagworte vergeben wurden. Dies sind überwiegend Kategorien zur Auszeichnung von Benutzerseiten. Insofern scheint der Prozess von Definition und Nutzung von Kategorien mittlerweile akzeptiert zu sein.

Die oben zusammengestellten, z.T. eher anekdotischen Beobachtungen wurden durch eine Zufallsstichprobe<sup>10</sup> untermauert. Zur Beschreibung von 100 Artikeln wurden 171 Kategorien eingesetzt. Jeder Artikel war durchschnittlich 3 Kategorien zugewiesen. 2% der Artikel waren nicht kategorisiert. 30% der Artikel verfügten über eine Diskussionsseite. Nur in einem Fall war die Kategorisierung Gegenstand der Diskussion. Im Durchschnitt 2,5 Bearbeitungen befassten sich pro Artikel mit der Zuweisung von Kategorien, wobei insgesamt 262 Kategorien zum Einsatz kamen. Während frühere Versionen der Artikel einen freieren Umgang mit Kategorien zeigten, wurde spätere Änderungen vorwiegend der Vereinheitlichung gewidmet, in deren Verlauf auch einige Kategorien ganz gelöscht wurden.

## 5.2 Nutzung der Wissensorganisation der Wikipedia

Der vorangehende Abschnitt hat gezeigt, dass die *Wikipedia* zahlreiche und differenzierte Strukturen der Wissensorganisation anbietet. Gleichwohl wurden einige Ansatzpunkte für Kritik deutlich, die aus Sicht der Dokumentations- und Informationswissenschaft vergleichsweise deutlich ausfallen könnte. Eine verlässliche Basis für ein Urteil kann hier aber nur durch Nutzertests geschaffen werden. Dies gilt schon für konventionelle Informationssysteme, viel mehr jedoch für solche, die durch neue Formen der Partizipation die Informationsarbeit überhaupt auf eine neue Grundlage stellen. Der im folgenden vorgestellte Nutzertest soll dieser Notwendigkeit Rechnung tragen. Er kann hinsichtlich der

---

<sup>10</sup> Dazu wurden am 7. 7.2006 100 Artikel extrahiert.

Anzahl als auch der Auswahl der Probanden nicht als Grundlage für eine empirische Studie über die *Wikipedia* dienen. Er zeigt aber im Sinne des Usability Testing (Nielsen 1994) Problembereiche in der Nutzung der *Wikipedia* qualifiziert auf.

An der Studie nahmen 6 männliche und 6 weibliche Studenten im Alter von 20-30 Jahren teil, die überwiegend gute Computerkenntnisse aufwiesen sowie regelmäßig das Internet und insbesondere auch die *Wikipedia* nutzen. Jede der Versuchspersonen hatte 10 Aufgaben zu lösen, die jeweils mit einem Problem der Informationssuche verbunden waren. Die Aufgaben waren dabei so gestaltet, dass jede Aufgabe mehrere Lösungsansätze erlaubte, dass aber jeder der oben eingeführten Wissensorganisationsstrukturen bei mindestens einer der Aufgaben eine prominente Rolle zukam<sup>11</sup>. Bei drei Aufgaben wurden die Nutzer explizit aufgefordert, den alphabetischen Index oder das Kategoriensystem zu nutzen. Der Test wurde in der *Thinking-Aloud-Technik* durchgeführt, die Teilnehmer waren also gehalten, ihre Gedanken über ihre Beobachtungen und über ihr Vorgehen laut zu äußern. Der Test wurde auf Video protokolliert. Im Anschluss an den Test wurden Fragen zu Person, Nutzung von Internet und *Wikipedia* sowie zur Einschätzung des Tests durch einen Fragebogen erhoben.

Für eine ausführliche Darstellung des Tests sei auf Lehnerer (2006) verwiesen. Hier ist nur Raum für eine zusammenfassende Darstellung der Ergebnisse.

- Den meisten Versuchsteilnehmern war die Existenz von alphabetischem Index und Kategoriensystem trotz regelmäßiger Nutzung der *Wikipedia* unbekannt.
- Die Handhabung der Suchfunktion war weitgehend problemlos. Der Unterschied zwischen Artikel- und Volltextsuche war aber vielen Nutzern unklar, so dass es bei manchen Fragen auch Schwierigkeiten mit der Interpretation der Suchergebnisse gab.
- Eine prominente Rolle spielten interne Verweise bei der Aufgabenlösung. Sie wurden von den Lesern problemlos erkannt und erfolgreich in die Suchstrategie integriert.
- Die nahe liegende Lösung, den Filesharing-Clients (Frage s.u.) *Azureus* über den alphabetischen Index zu suchen, fand nur eine Versuchsperson. Die Sortierregeln führten zu den zu erwartenden Problemen.
- Den meisten Teilnehmern war der Unterschied zwischen Themenportalen und Kategorien unklar. Dies wurde durch wechselnde Bezeichner befördert (*Artikel nach Themen, Themenportale*). Die jeweils individuelle Gliederung der Portale führte zu Orientierungsschwierigkeiten. Auch die hierarchische Struktur der Portale wurde als schwer zu durchschauen kritisiert, obschon der generelle Ansatz der Themenportale gelobt wurde.
- Der Umgang mit dem Kategoriensystem erwies sich als der schwierigste Teil des Tests. Die Darstellung wurde als unübersichtlich, die inhaltliche Gliederung nicht immer als nachvollziehbar empfunden. 5 Testpersonen waren in der Lage, das Kategoriensystem eigenständig für eine Lösung (insbesondere Aufgabe 4) zu nutzen.

Der Test zeigt, dass ein Hauptproblem in der Nutzung der Wissensorganisationsstrukturen der *Wikipedia* darin liegt, dass die Nutzer ihre Existenz erst gar nicht bemerken. Waren die

---

<sup>11</sup> Die Fragen des Tests sind in leicht gekürzter Form im Anhang dieses Beitrags aufgeführt.

Strukturen konzeptualisiert, führte ihre Nutzung zu effizienteren und effektiveren Lösungswegen. Schwierigkeiten gab es aber immer wieder im Detail (Sortierung, inhaltliche Gliederung usw.). Schon durch vergleichsweise einfache Maßnahmen an der Benutzungsoberfläche (z.B. Vereinheitlichung von Bezeichnern) sollten schon signifikante Verbesserungen in der Nutzung der Strukturen der Wissensorganisation in der *Wikipedia* zu erzielen sein.

## 6 Zusammenfassung

Die beiden in diesem Artikel dargestellten Studien haben folgendes gezeigt. Die *Wikipedia* kann auch außerhalb der exzellenten Artikel in größeren thematischen Zusammenhängen verlässliche Information zur Verfügung stellen. Höhere Ansprüche an die Aufbereitung können dabei nicht gestellt werden, eine konsistente und vollständige Darstellung wird jedoch weitgehend erreicht. Qualitative Unterschiede zwischen den *Wikipedias* bestehen. Sie scheinen jedoch primär auf den unterschiedlichen Entwicklungsstand zurückzuführen zu sein. Klarheit können hier nur weitere Studien herbeiführen.

Hinsichtlich der Wissensorganisation ist festzustellen, dass die Nutzer der *Wikipedia* das System nur sehr unvollständig konzeptualisiert haben. In manchen Fällen führten irreführende Bezeichnungen sogar zu Fehlkonzertualisierungen. Hier wäre z.T. mit einfachen Maßnahmen Abhilfe zu schaffen. Erst dann kann fundiert untersucht werden, inwieweit die zuvor geäußerten dokumentarischen Bedenken hinsichtlich des Kategoriensystems von Bedeutung sind, oder ob dieser neue Systemtyp auch einen neuen, erfolgreichen Umgang mit bisher etablierten Strukturen der Wissensorganisation mit sich bringt.

### Anhang: Fragen des Nutzertests

1. Suche für Dein Vorstellungsgespräch Information über Geschäftsbereiche, Umsatz, Standorte usw. der Siemens VDO-AG!
2. Welche Kaninchenkrankheit wird durch Stechmücken übertragen?
3. Wie schnell ist der Transrapid? Benutze zur Suche das Kategoriensystem? Falls Du nicht fündig wirst, gehe über eine geeignetes Themenportal!
4. Selbst als Liebhaber klassischer Musik ist Dir von dem Namen eines bekannten Tenors nur der Anfangsbuchstabe des Nachnamens bekannt (K). Stelle eine Liste der Kandidaten zusammen! Suche anschließend eine Liste aller Orchester in Deutschland und finde heraus, wie viele in München beheimatet sind.
5. Versuche das taiwanesisches Unternehmen BenQ über das Kategoriensystem zu finden!
6. Du möchtest alle Artikel zum Thema Elektrotechnik finden. Wie gehst Du vor, wo suchst Du? Gibt es einen alternativen Lösungsansatz?
7. Du möchtest Deinen Auslandsaufenthalt in den USA vorbereiten und Dich über Regionen, Landschaften, Nationalparks, Sport, Politik, Wirtschaft informieren. Vor allem möchtest Du Artikel über die 10 größten Städte der USA lesen.
8. Du möchtest einen Artikel über einen Filesharing-Client lesen, von dem Du nur noch weißt, dass der Name mit *Azu* anfängt.

9. Du möchtest Dir einen 1er BMW kaufen. Wie viele davon wurden 2005 verkauft. Beginn beim alphabetischen Index. Zudem möchtest Du wissen, welche Fahrzeugtypen hinter der E90/91 Baureihe bei BMW stehen. Suche auch hier über den alphabetischen Index!
10. Welche Ländervorwahl hat Zypern? Welches Land hat die Vorwahl +993?

## Literatur

*Boyce, Charles*, 1996, *The Wordsworth Dictionary of Shakespeare*. Wordsworth Editions Ltd., Hertfordshire.

*Crystal, David* und *Crystal, Ben*, 2004, *Shakespeare's Words*, Penguin Books Ltd.

*Danowski, Patrick* und *Voss, Jakob*, 2005, *Das Wissen der Welt – die Wikipedia*. In: *Open Source Jahrbuch*. [http://www.opensourcejahrbuch.de/download/jb2005/chapter\\_06/osjb2005-06-05-danowskivoss](http://www.opensourcejahrbuch.de/download/jb2005/chapter_06/osjb2005-06-05-danowskivoss), zitiert am 23.3.2007.

*Emigh, William.*, und *Herring, Susan C.*, 2005, *Collaborative authoring on the web: A genre analysis of online encyclopedias*. Paper presented at the Proceedings of the Annual Hawaii International Conference on System Sciences. <http://ella.slis.indiana.edu/~herring/wiki.pdf>, zitiert am 6.1.2007.

*Gaus, Wilhelm*, 1995, *Dokumentations- und Ordnungslehre*, Springer.

*Giles, Jim*, 2005, *Internet encyclopaedias go head to head*. *Nature* 438, S. 900-901, <http://www.nature.com/nature/journal/v438/n7070/full/438900a.html>, zitiert am 6.1.2007.

*Hammwöhner, Rainer; Fuchs, Karl-Peter; Kattenbeck, Markus und Sax, Christian*, 2007, *Qualität der Wikipedia. Eine vergleichende Studie*. Angenommen für das Internationale Symposium Informationswissenschaft, ISI 2007, in Köln.

*IEEE 2004*, *Inspec*, <http://www.iee.org/publish/inspec/about/>, <http://www.iee.org/publish/support/inspec/document/thes/>, zitiert am 1.4.2007.

*Kermode, Frank*, 2001, *Shakespeare's Language*, Penguin Books Ltd.

*Kuhlen, Rainer*, 2005, *Wikipedia – Offene Inhalte im kollaborativen Paradigma – eine Herausforderung auch für Fachinformation*. *Forschung und Lehre*, Nr. 10, S. 546-548.

*Lanier, Jason*, 2006, *Digitaler Maoismus. Kollektivismus im Internet, Weisheit der Massen, Fortschritt der Communities? Alles Trugschlüsse*. <http://www.sueddeutsche.de/kultur/artikel/306/78228/article.html>, zitiert am 29.6.2006.

*Lehnerer, Sabrina*, 2006, *Wissensorganisation in der Online-Enzyklopädie Wikipedia*, Magisterarbeit im Fach Informationswissenschaft an der Universität Regensburg.

*Lih, Andrew*, 2004, *Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource*. In: *Proceedings of the 5th International Symposium on Online Journalism*. <http://jmsc.hku.hk/faculty/alih/publications/utaustin-2004-wikipedia-rc2.pdf>, zitiert a, 6.1.2007.

*Molz, Johannes* und *Burghardt, Manuel*, 2006, *Shakespeare's Werk in der Wikipedia. Eine bilinguale Untersuchung zu Umfang und Qualität von Wikipediaartikeln in einem abgegrenzten Themenbereich, Ausarbeitung zum Projektseminar „Hypermedia“*.

*Jakob Nielsen*, 1994, *Usability Engineering*. Morgan Kaufmann.

Rühle, Alex, 2006, Im Daunenfedergestöber. Wikipedia-Fälschungen. sueddeutsche.de, <http://www.sueddeutsche.de/kultur/artikel/631/90541/>, zitiert am 6.1.2007.

Schlieker, Christian, 2005, Explorative Untersuchung von Wissen in kollektiven Hypertexten, Diplomarbeit, Fachbereich 08, Soziologie, Universität Bremen.

Stvilia, Besiki; Twidale, Michael B.; Gasser, Les; und Smith, Linda C., 2005a, Information Quality Discussions in Wikipedia. In: S. Hawamdeh (Ed.), *Knowledge Management: Nurturing Culture, Innovation, and Technology - Proceedings of the 2005 International Conference on Knowledge Management*. Charlotte, NC: World Scientific Publishing Company. S 101-113. <http://mailer.fsu.edu/~bstvilia/papers/qualWiki.pdf>, zitiert am 6.1.2007.

Stvilia, Besiki; Twidale, Michael B.; Gasser, Les; und Smith, Linda C., 2005b, Assessing information quality of a community-based encyclopedia. In: *Proceedings of the International Conference on Information Quality - ICIQ 2005*. Cambridge, MA. S. 442-454. <http://mailer.fsu.edu/~bstvilia/papers/quantWiki.pdf>, zitiert am 6.1.2007

Voss, Jakob, 2005, Informatrische Untersuchungen an der Online-Enzyklopädie Wikipedia, Magisterarbeit im Fach Bibliothekswissenschaft, Institut für Bibliotheks und Informationswissenschaft, Humboldt-Universität zu Berlin.

Voss, Jakob, 2006, Collaborative Thesaurus Tagging the Wikipedia Way, <http://arxiv.org/abs/cs/0604036>, zitiert am 26.3.2007.

Wiegand, Dorothee, 2007, Entdeckungsreise, Digitale Enzyklopädien erklären die Welt, c't, Magazin für Computer und Technik, Nr. 6, S. 136-145.

### **Zitierte Artikel der Wikipedia**

Wikipedia, 2007, Startseite der Wikipedia, <http://www.wikipedia.org/>, zitiert am 6.1.2007.

Wikipedia, 2007a, Portalseite der Wikipedia, [http://de.wikipedia.org/wiki/Portal:Wikipedia\\_nach\\_Themen](http://de.wikipedia.org/wiki/Portal:Wikipedia_nach_Themen), zitiert am 10.5.2007

Wikipedia, 2007b, Listen in der Wikipedia, <http://de.wikipedia.org/wiki/Wikipedia:Listen>, zitiert am 24.3.2007

Wikipedia, 2007c, Formatvorlagen, <http://de.wikipedia.org/wiki/Kategorie:Wikipedia:Formatvorlage>, zitiert am 26.3.2007.

Wikipedia, 2007d, Textbausteine, <http://de.wikipedia.org/wiki/Kategorie:Vorlage:Kategorie>, zitiert am 26.3.2007.

Wikipedia, 2007e, Simplified Ruleset, [http://en.wikipedia.org/wiki/Wikipedia:Simplified\\_Ruleset](http://en.wikipedia.org/wiki/Wikipedia:Simplified_Ruleset), zitiert am 26.3.2007.

Wikipedia, 2007f, Hilfe, <http://de.wikipedia.org/wiki/Kategorie:Wikipedia:>, zitiert am 26.3.2007.

Wikipedia, 2007g, Bewertungsbausteine, <http://de.wikipedia.org/wiki/Wikipedia:Bewertungsbausteine>, zitiert am 26.3.2007.

Wikipedia, 2007h, Featured Article Criteria, <http://de.wikipedia.org/wiki/Wikipedia:Bewertungsbausteine>, zitiert am 26.3.2007.



Wikipedia, 2007i, Wiki Project Elizabethan theatre,  
[http://en.wikipedia.org/wiki/Wikipedia:WikiProject\\_Elizabethan\\_theatre](http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Elizabethan_theatre), zitiert am 29.3.2007.

Wikipedia, 2007j, Diskussion: Kategorien,  
[http://de.wikipedia.org/w/index.php?title=Wikipedia\\_Diskussion:Kategorien/Archiv1](http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Kategorien/Archiv1),  
[http://de.wikipedia.org/w/index.php?title=Wikipedia\\_Diskussion:Kategorien/Archiv2](http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Kategorien/Archiv2),  
[http://de.wikipedia.org/w/index.php?title=Wikipedia\\_Diskussion:Kategorien/Archiv3](http://de.wikipedia.org/w/index.php?title=Wikipedia_Diskussion:Kategorien/Archiv3), zitiert am  
31.3.2007

Wikipedia, 2007k, Hilfe: Kategorien, <http://de.wikipedia.org/wiki/Hilfe:Kategorien>, zitiert am  
31.3.2007

Wikipedia, 2007l, Benutzte aber nicht angelegte Kategorien,  
[http://de.wikipedia.org/wiki/Spezial:Gew%C3%BCnschte\\_Kategorien](http://de.wikipedia.org/wiki/Spezial:Gew%C3%BCnschte_Kategorien), zitiert am 31.3.2007

Wikipedia, 2007m, Meistbenutzte Kategorien,  
[http://de.wikipedia.org/wiki/Spezial:Meistbenutzte\\_Kategorien](http://de.wikipedia.org/wiki/Spezial:Meistbenutzte_Kategorien), zitiert am 31.3.2007

Wikipedia, 2007n, Kategorien, <http://de.wikipedia.org/wiki/Spezial:Kategorien>, zitiert am 31.3.2007

Wikipedia, 2007o, Wikipedistik, <http://de.wikipedia.org/wiki/Wikipedia:Wikipedistik>, zitiert am  
10.5.2007

### **Kontakt zum Autor:**

Prof. Dr. Rainer Hammwöhner  
Philosophische Fakultät IV – Sprach- und Literaturwissenschaften  
Institut für Medien-, Informations-, und Kulturwissenschaft  
Lehrstuhl für Informationswissenschaft  
Universität Regensburg  
93040 Regensburg

### **Bitte diesen Artikel wie folgt zitieren:**

Hammwöhner, Rainer (2007), Qualitätsaspekte der Wikipedia. In: Stegbauer, Christian / Schmidt, Jan / Schönberger, Klaus (Hrsg.): Wikis: Diskurse, Theorien und Anwendungen. Sonderausgabe von kommunikation@gesellschaft, Jg. 8. Online-Publikation: [http://www.soz.uni-frankfurt.de/K.G/B3\\_2007\\_Hammwoehner.pdf](http://www.soz.uni-frankfurt.de/K.G/B3_2007_Hammwoehner.pdf)