

Wikipedia in großen Suchmaschinen

Die großen Suchmaschinen Google und Yahoo setzen beide auf die Inhalte des freien Nachschlagewerks Wikipedia. Sie wollen die Inhalte bei entsprechenden Anfragen prominent platziert vor der eigentlichen Trefferliste anzeigen. Google verwendet sie auch, um an die Suchmaschine gestellte Fragen zu beantworten. Damit gewinnt die nicht unumstrittene Freiwilligen-Enzyklopädie weiter an Bedeutung.

Google hat im April die aus dem US-Angebot bekannte Funktion der Begriffsdefinitionen für die deutschsprachige Seite vorgestellt. Leitet man eine Anfrage mit "definiere" ein, so wird an erster Stelle (noch vor der regulären Trefferliste) eine Definition des entsprechenden Begriffs angezeigt. Zwar werden andere Quellen aus dem offenen Web eingebunden, es zeigt sich aber, dass die Wikipedia-Einträge an erster Stelle platziert sind. Es wird jeweils der erste Abschnitt aus dem Wikipedia-Artikel präsentiert, der vollständige Text kann über einen Link abgerufen werden. Klickt man die Überschrift "Definitionen von ... im Web" an, bekommt man, wenn vorhanden, weitere Treffer aus anderen Quellen angezeigt. Ein Beispiel zeigt die Bandbreite der Definitionen: Die Suche nach einer Definition für Glück ergibt neben dem ausführlichen Wikipedia-Artikel unter anderem Treffer von skat.com ("Faktor, der beim Skatenspiel nicht wegzudenken ist") und von jesu-botschaft.de. Hier muss sich Google (wieder einmal) den Vorwurf gefallen lassen, nicht die Autoritäten, also die hochwertigsten Quellen einzubeziehen, sondern irgendwelche, die (vermeintlich) etwas zum Thema zu melden haben.

Eine weitere Einbindung der Wikipedia-Quellen wird bereits auf Googles US-Site getestet: Stellt man der Suchmaschine eine Frage (und nicht wenige Suchmaschinennutzer tun dies tatsächlich!), so bekommt man mit etwas Glück direkt eine Antwort angezeigt, zum Beispiel auf die Frage "Who is David Bowie?" die Antwort: "is a British rock and roll musician, actor, and artist who has had a profound influence on rock". Unter jedem Treffer ist ein Link auf die Quelle angegeben. Die Informationen kommen nicht nur aus der Wikipedia, sondern zum Beispiel auch aus dem World Fact Book (bei geographischen Anfragen) oder von Seiten wie "Pub Quiz Help". Die aus den Webseiten gezogenen Exzerpte sind von unterschiedlicher Qualität: Fragt man nach (populären) Zahlenangaben wie etwa der Höhe

des Mount Everest, funktioniert es gut, bei den Fragen nach Personen wird manchmal ein etwas unpassender Teil des Artikels angezeigt, so dass man sich auf den vollständigen Artikel weiterklicken muss.

Bislang geht Yahoo einen anderen Weg als Google. Gibt man in der US-amerikanischen Version Begriffe ein, die auf ein klassisches "Nachschlage-Bedürfnis" hindeuten, wird ähnlich wie bei Google mit den Definitionen ein "Shortcut" angezeigt (Password 2/2004 und 11/2004). Die Quellen dort sind allerdings direkt in das Yahoo-Angebot eingebaut, d.h. man verlässt nicht das Yahoo-Portal, wenn man sich die vollständigen Artikel ansehen möchte. Und die Quellen wirken wesentlich vertrauenerweckender: Die Enzyklopädie-Einträge kommen aus der Columbia Electronic Encyclopedia. Sucht man nach Ländernamen, so bekommt man Treffer aus dem World Fact Book, "the U.S. government's complete geographical handbook".

Nun hat Yahoo angekündigt, die Einträge aus der Wikipedia bei solchen Anfragen prominent platziert anzeigen zu wollen. Bisher sind allerdings nur einzelne Beispiele zu sehen: Gibt man beim französischen Ableger von Yahoo einen Ländernamen ein, so wird bereits ein Treffer aus der Wikipedia angezeigt. Des Weiteren stellt Yahoo den Betreibern der Wikipedia Serverplatz zur Verfügung. Es wird jedoch betont, dass dies nicht das Ende der Gespräche zwischen Wikipedia und Google bedeute.

Wie gut ist Wikipedia? Sollten die Suchmaschinen nicht besser auf hochwertige Inhalte setzen?

Mit der sehr umfangreichen Wikipedia bietet sich für die Suchmaschinenbetreiber eine Möglichkeit, durch die manuelle Einbindung einer einzigen Quelle vermeintlich alle Fragen kompetent abzudecken. Der bisher zumindest bei Yahoo betriebene technische Aufwand, verschiedene Quellen (für die man als Suchmaschinenbetreiber vielleicht sogar bezahlen muss!) einzubinden und trotzdem nicht auf jede Frage eine schnelle Antwort liefern zu können, kann gespart werden. Letztlich stellt sich allerdings die Frage nach der Qualität der Treffer aus Wikipedia.

Die Wikipedia wird von Freiwilligen erstellt, wobei es sich nicht um eine geschlossene Gruppe handelt. Vielmehr wird jeder Internet-Nutzer dazu eingeladen, sich an der

Erstellung der Enzyklopädie zu beteiligen. Artikel werden gemeinschaftlich erarbeitet und sollen so dem Ideal nach mit der Zeit immer besser werden. Gibt es aktuelle Entwicklungen, so können diese direkt in die Artikel eingebunden werden.

Es ist aber zu fragen, ob die Artikel durchweg die von einem Nachschlagewerk zu erwartende Qualität liefern können. Sicherlich gibt es in der Wikipedia sehr gute Artikel. Allerdings gibt es auch einen großen Anteil von schlechten, unsystematischen und fragwürdigen Eintragungen. Ein Blick auf das "Portal Bibliothek, Information, Dokumentation" zeigt exemplarisch einige Schwachstellen: Zunächst ist das Portal unvollständig. Das ist verständlich und wäre nicht besonders schlimm, wenn dies für den Nutzer immer ersichtlich wäre. Zwar ist dies klar, wenn zu einem Begriff kein Artikel angezeigt wird. Artikel, die sich noch in einem "frühen Stadium" befinden (und deshalb noch nicht oder nur marginal überarbeitet wurden), sind jedoch nicht als solche gekennzeichnet. Sie können deshalb, insbesondere wenn Suchmaschinen an prominenter Stelle auf sie verweisen, in ihrer Qualität leicht mit Einträgen aus echten Enzyklopädien verwechselt werden. Viele Nutzer werden ihnen damit eine vergleichbare Autorität zumessen.

Ein weiterer Kritikpunkt betrifft den unsystematischen Aufbau der Wikipedia. Wieder ein Beispiel aus dem Portal BID: Ein Artikel zum Booleschen Retrieval ist nicht vorhanden, wohl aber zum Erweiterten Booleschen Retrieval. Dieser besteht hauptsächlich aus einer Ansammlung von Formeln und hat damit wenig mit einem enzyklopädischen Artikel zu tun.

Eine Enzyklopädie sollte etwas über die Bedeutung der behandelten Themen aussagen. Klassisch bekommen die wichtigeren Themen längere Artikel, die weniger wichtigen entsprechend kürzere. Die Wikipedia gibt hier keine Orientierung: Alle Themen scheinen gleich bedeutend zu sein, zumindest drückt sich die Bedeutung nicht in der Länge aus. So ist der Artikel über den Computerspiel-Helden Super Mario, bei dem man sich ja schon generell fragen darf, was er in einer Enzyklopädie zu suchen hat, ca. 30.000 Zeichen lang, der über Gerhard Schröder nur 17.000.

Was bedeutet die zumindest zu hinterfragende Qualität der Wikipedia und ihre Einbindung in die Suchmaschinen für den Suchmaschinen-Nutzer? Der Kernpunkt wurde

Fortsetzung auf Seite 29.

Wikipedia in großen Suchmaschinen

Fortsetzung von Seite 21.

bereits erwähnt: Die Nutzer werden der Wikipedia eine Autorität beimessen, die diese nicht hat. Die Suchmaschinen unterstützen mit der Einbindung die Annahme, die Wikipedia könne die Qualität und Orientierung einer echten Enzyklopädie bieten. Die Suchmaschinen wären besser beraten, verstärkt auf die Einbindung geprüfter, hochwertiger Quellen zu setzen. Einen solchen Ansatz verfolgt MSN mit der Einbindung der hauseigenen Encarta-Enzyklopädie. Deren Nutzung ist immerhin für jeden Nutzer zwei Stunden lang kostenlos, dann muss bezahlt werden.

*Dirk Lewandowski,
dirk.lewandowski@uni-duesseldorf.de,
www.durchdenken.de/lewandowski*