

Analyse der Fragensituation bei Lycos iQ

PHILIPP NIEHUES
philipp_niehues@gmx.de

DIANA SCHULZ
dischu77@gmx.de

VANESSA SCHÄFER
vschaefer@web.de

BIANCA OMELAN
bomelan@googlemail.com

Im nachfolgenden Bericht wird die Fragequalität auf der Plattform des Frage-Antwort-Dienstes Lycos iQ untersucht. Ziel ist es herauszufinden, welche Fragen qualitativ geeignet sind, um in algorithmische Suchmaschinen mit eingebunden werden zu können. Es stellte sich hierbei die Frage, ob sich aus der Qualität der Antworten Rückschlüsse auf die Qualität der Fragen ziehen lassen. Hierfür wurde ein Antworten-Kriterienkatalog entwickelt, der eine subjektive Bewertung der Antworten zuließ. Anschließend wurden die dazugehörigen Fragen im Hinblick auf inhaltliche und formale Kriterien beleuchtet. Dabei wurden außerdem eine Kategorisierung nach Frageart und eine Verteilung der Fragen nach Themengebieten vorgenommen. Korrelationen zwischen Antwort- und Fragequalität konnten festgestellt werden. Aus der genauen Analyse ergibt sich unter anderem, dass sowohl Faktenfragen qualitativ gute Antworten generieren als auch dass die Antwortqualität von Fragen aus Themengebieten mit allgemeinem, freizeithlichen Interesse am besten ist und sich daher für eine Einbindung in eine Suche eignen. Der Bericht schließt mit einigen Empfehlungen zur Erhöhung der Frage- und damit auch Antwortqualität.

1. EINLEITUNG

Der Erfolg von Lycos iQ hängt sicherlich zu großen Teilen auch davon ab, wie hilfreich die Antworten auf die gestellten Fragen von Gästen und Community-Mitgliedern sind. Gute Antworten sind zum einen für die Mitglieder der Community besonders hilfreich und damit ein Anreiz, sich weiter auf der Website zu engagieren, zum zweiten eignen sich sicherlich nur richtige und gute Antworten zur Einbindung in andere Suchformen. Wird also die Frage gestellt, wie Frage-Antwort-Dienste algorithmische Suchen ergänzen können, kommt man im nächsten Schritt zu der Frage, welche Antworten und dazugehörige Fragen gut genug sind, eine algorithmische Suche zu ergänzen, und wie diese guten Frage-Antwort-Paare aus der Masse der Daten zu extrahieren sind.

Der Ansatz dieses Beitrages ist dabei folgender: Wenn ein User in einer klassischen Suche eine Frage formuliert (in diesem Zusammenhang wird die Eingabe eines einzelnen Suchbegriffes ebenfalls als stark verkürzte Frage angesehen), so ist eine Beantwortung dieser Frage mit Hilfe der Antworten aus den Frage-Antwort-

Diensten erstmal nur dann möglich, wenn man die passenden Antworten zu der Suchanfrage auch finden kann. Hierzu wäre es sinnvoll, nach Übereinstimmungen zwischen der gesamten Suchanfrage des Users und den Fragen innerhalb von Lycos iQ zu suchen - und nicht nach einem bestimmten Stichwort innerhalb der gesamten Datenbasis - und diese passenden Fragen dem User dann als Ergänzung zu seiner Suchanfrage anzubieten.

Dieses Vorgehen setzt allerdings einen Zusammenhang zwischen der Qualität der Frage und der Qualität der zugehörigen Antworten voraus, da dem User wie bereits erwähnt ja nur qualitativ hochwertige Antworten weiterhelfen und man von der Ausgestaltung der Frage auf die Qualität der Antworten schließen können müsste.

Weiterhin müsste auch die Frage selbst qualitativ hochwertig formuliert sein, damit sie zum Beispiel vom Suchenden auch direkt als verständliche, nützliche Frage identifiziert werden kann.

Ziel dieses Beitrages ist es daher, Zusammenhänge zwischen der Qualität der Fragen und der Qualität der zugehörigen Antworten aufzuzeigen und Kriterien zu entwickeln, nach denen gute Fragen, also gut formulierte Fragen mit qualitativ hochwertigen Antworten, von den nicht geeigneten Fragen separiert werden können. Außerdem sollen anhand dieser Kriterien Vorschläge entwickelt werden, wie eine Anleitung des Users mit dem Ziel der Verbesserung der Qualität von Fragen und Antworten erreicht werden könnte.

2. THESEN

Als Grundlage für die oben genannten Überlegungen dienen folgende Thesen, die im Weiteren untersucht und verifiziert werden sollen:

- Zur Einbindung in andere Suchformen eignen sich nur Fragen mit guten und richtigen Antworten.
- Zur Einbindung in andere Suchformen eignen sich nur Fragen, die bestimmten qualitativen Merkmalen entsprechen.
- Es besteht eine Korrelation zwischen formalen und inhaltlichen Merkmalen einer Frage.
- Es besteht eine Korrelation zwischen formalen und inhaltlichen Merkmalen der Frage auf der einen Seite und der Qualität ihrer Antworten auf der anderen Seite.
- Geeignete Fragen lassen sich daher im Allgemeinen anhand bestimmter qualitativer Merkmale bestimmen.

3. METHODIK UND VORÜBERLEGUNGEN

3.1 Thema der Fragen

Die vorliegenden Ergebnisse stützen sich im Wesentlichen auf eine statistische Analyse der verfügbaren Datenbasis, das heißt auf eine Analyse einer bestimmten Anzahl von Fragen und Antworten mit Hilfe definierter Kriterien und einer anschließenden Auswertung der Ergebnisse. Hierzu wurde eine Stichprobe von 250 Fragen (247 vollständige Datensätze) mit dazugehörigen 393 Antworten ausgewählt

und von einem Probanden-Team ausgewertet. Dabei wurde darauf geachtet, dass bei der Auswahl der Stichprobe alle von der Arbeitsgruppe „Analyse der Antwort-Qualität“ definierten Themengebiete einigermaßen gleichmäßig abgedeckt wurden (vgl. Abb. 5.1). Eine genaue Beschreibung dieser Themengebiete findet sich im vorherigen Beitrag dieses Bandes „Untersuchung der Qualität der Antworten bei Lycos iQ und deren Einbindung in die algorithmische Suche“.

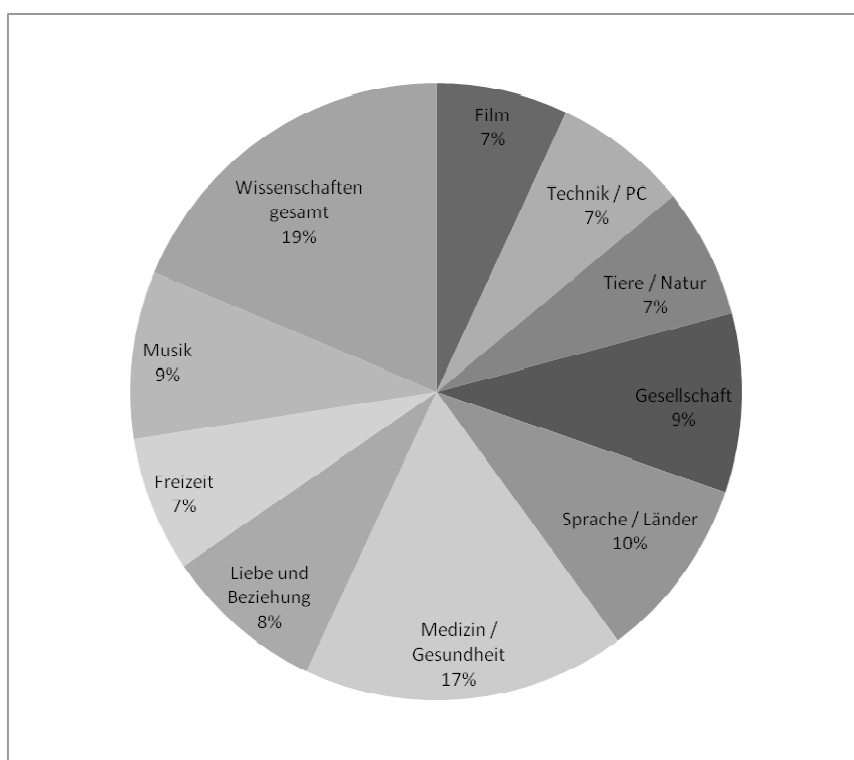


Abb. 5.1 Thematische Verteilung der Fragen

Die größere Anzahl der Fragen im Bereich Wissenschaften erklärt sich dadurch, dass hier verschiedene Wissenschaftliche Disziplinen (Politik, Geschichte, Literatur, etc.) zusammengefasst wurden.

3.2 Art der Fragen

Nach Auswahl der Fragen wurden diese nach ihrer Art kategorisiert. Bei der Einteilung lag die Annahme zu Grunde, eine bestimmte Frageart würde möglicherweise besonders qualitativ schlechte Antworten generieren, während eine andere Frageart dagegen mit auffällig qualitativ guten Antworten aufwarten könnte. Nach einer ersten Sichtung verschiedener zufällig gewählter Fragen zeigte sich, dass die Fragen in vier verschiedene Kategorien unterteilt werden können:

- **Faktenfragen**
Als Faktenfragen wurden alle Fragen definiert, auf die es auf jeden Fall eine eindeutig richtige Antwort geben muss, die in der Regel auch nachprüfbar ist.
Beispiele: Wo liegen die Osterinseln? Was ist eine Stammesfehde? Was ist der Unterschied zwischen einer GmbH und einer GmbH & Co KG?
- **Meinungsfragen**
Als Meinungsfragen galten alle Fragen, auf die es keine allgemein gültige Antwort geben kann. Die Fragen erlauben als Antwort eine subjektive Einschätzung und geben häufig die Möglichkeit zur Diskussion.
Beispiele: Gibt es Gott? Ist die Serie „Babylon 5“ empfehlenswert? Was ist besser...?
- **Lebensfragen**
Als Lebensfragen konnten Fragen bezeichnet werden, die als Antworten hilfreiche Tipps für den Alltag verlangen. Ob die Antworten auf solche Fragen wirklich sinnvoll sind und dem Fragesteller helfen, ist meist ohne entsprechenden Praxistest nicht direkt erkennbar.
Beispiele: Kann mir jemand ein günstigeres Mikro empfehlen? Wer kann mir da mal raten, wo ich Infos über ADHS herbekomme? Hat jemand Erfahrung mit...?
- **Spaßfragen**
Als Spaßfragen wurden Fragen bezeichnet, auf die der Fragesteller keine ernsthafte Antwort erwartet bzw. erwarten kann.
Beispiele: Warum gibt's hier bei Lycos iQ so viele Fragen? Sind alle Frauen dumm? Wie wird man immer glücklich?

Diese Einteilung der Fragen in vier verschiedene Arten basiert ebenfalls auf den Ergebnissen der Arbeitsgruppe „Analyse der Antwort-Qualität“, unterscheidet die Nicht-Faktenfragen aber noch differenzierter.

Um ein möglichst unverfälschtes Bild der quantitativen Verteilung der verschiedenen Fragearten zu bekommen, wurde eine zweite, größere Stichprobe ausgezählt. Für diese Stichprobe wurden 3259 Fragen ausgewertet und nach der Art der Frage kategorisiert. Dabei ergab sich die in Abbildung 5.2 dargestellte Verteilung:

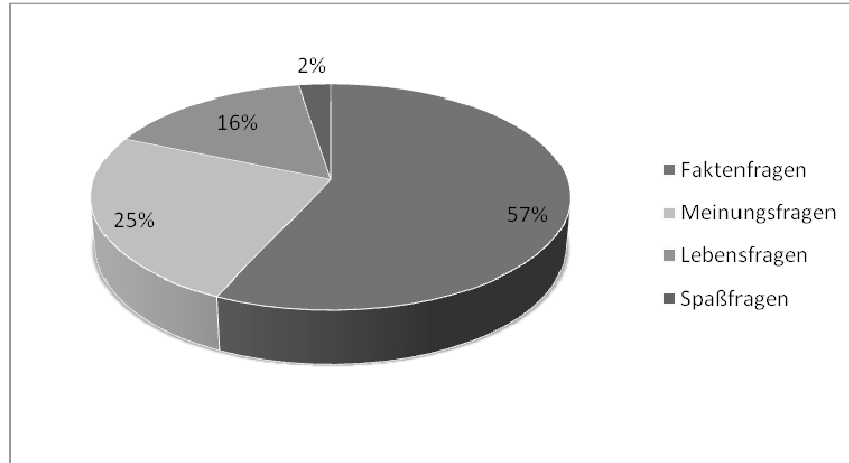


Abb. 5.2 Verteilung nach Art der Frage, große Stichprobe

Diese Verteilung stimmt allerdings bereits sehr gut mit der Verteilung innerhalb der kleineren Stichprobe überein (vgl. Abb. 5.3). Für die weiteren Analyseschritte werden daher die auf Grundlage der kleinen, 250 Fragen umfassenden Stichprobe erhobenen Daten verwendet.

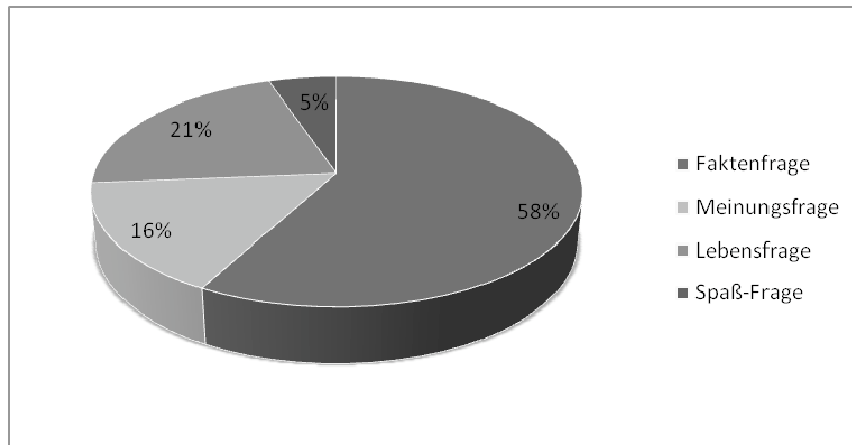


Abb. 5.3 Verteilung nach Art der Frage, kleine Stichprobe

Auffällig ist, dass mehr als die Hälfte aller Fragen Faktenfragen sind. Dies zeigt, dass es den Fragenden vor allem um eine konkrete Fragestellung geht. Auch der große Anteil der Lebensfragen, die meist nach einem Tipp zur Lösung eines bestimmten Problems suchen, bestätigt diesen Schluss. Obwohl gerade der Anteil der Meinungsfragen bei subjektiver Betrachtung manchmal höher erscheint, dient die Benutzung von Lycos iQ weniger dem Diskutieren kontroverser Themen oder dem Austausch von Meinungen zu einem bestimmten Thema.

Setzt man die Art der Frage in Abhängigkeit zu den verschiedenen Themengebieten, so ergibt sich die unten stehende Verteilung (vgl. Abb. 5.4). Das Ergebnis ist wenig überraschend, da der Anteil der Faktenfragen in Themengebieten, in denen eher eine Meinung als Fachwissen gefragt ist (z.B. „Liebe & Beziehung“, „Gesellschaft“, „Musik“), deutlich niedriger ist als in anderen Bereichen. Vor allem technische Themen wie „Technik & PC“ weisen einen sehr hohen Anteil an Faktenfragen auf. Dagegen fällt der Bereich „Wissenschaften“, dem man eigentlich einen recht hohen Anteil an Faktenfragen zubilligen würde, deutlich zurück, was damit begründet werden kann, dass hier verschiedenste wissenschaftliche Disziplinen zusammengefasst sind, die neben naturwissenschaftlichen auch gesellschaftliche oder politische Themen behandeln.

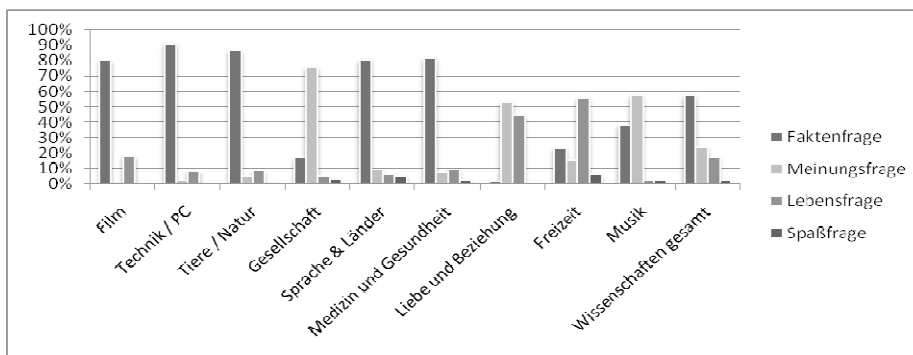


Abb. 5.4 Prozentualer Anteil der unterschiedlichen Fragearten in Abhängigkeit von den verschiedenen Themengebiete

4. ARBEITSCHRITTE

Die eigentliche Analyse gliederte sich in folgende Arbeitsschritte:

- **Erstellung eines Kriterienkatalogs zur Bewertung der Antworten**
Auf Grundlage der Ergebnisse der Arbeitsgruppe „Qualität der Antworten“ wurden die für die Qualität einer Antwort ausschlaggebenden Merkmale extrahiert und in Form eines Kriterienkataloges zusammengefasst.
- **Erarbeitung eines Punktesystems**
Um die Antworten bezüglich ihrer Qualität untereinander vergleichbar zu machen, wurde jedem Qualitätskriterium ein bestimmter Punktwert zugewiesen, der die Relevanz des Merkmales für die Qualität der Frage widerspiegelt.
- **Bewertung der Antworten**
Auf Grundlage des Punktesystems erfolgte eine Bewertung aller Antworten innerhalb der Stichprobe, so dass am Ende jeder Frage eine bestimmte Punktzahl als Indikator für die Qualität der Frage zugewiesen werden konnte.
- **Erarbeiten möglicher Qualitätskriterien für die Fragen**
Um die Qualität der Antworten in Relation zu den Fragen setzen zu können, wurde ein Kriterienkatalog entwickelt, der die Kategorisierung einer Frage nach bestimmten Merkmalen erlaubt.
- **Abbilden von Zusammenhängen**
Als zentraler Teil der Arbeit wurde im Anschluss versucht, Zusammenhänge zwischen einzelnen Frage-Kriterien und der Qualität der Antworten abzubilden, um so die für eine gute Frage ausschlaggebenden Merkmale zu extrahieren. Außerdem wurde versucht, Korrelationen der Merkmale untereinander herauszuarbeiten.
- **Formulierung von Empfehlungen**
Als letzter Teil wurden die Analyse-Ergebnisse auf die Fragestellung angewendet und in Form von Empfehlungen zur Einbindung von Fragen in anderen Suchformen und zur generellen Verbesserung der Fragequalität verdichtet.

Diese Arbeitsschritte werden im Folgenden eingehend erklärt.

5. KRITERIENKATALOG ZUR BEWERTUNG DER ANTWORTEN

Im Vorfeld musste zunächst geklärt werden, was eine qualitativ gute Antwort ist. Dafür wurden die Ergebnisse der Arbeitsgruppe zur Analyse der Antwort-Qualität herangezogen. Bei der Analyse der Antworten hat sich herauskristallisiert, welche Qualitätsmerkmale erfüllt sein müssen, um auf eine gute Antwort zu schließen. Im Folgenden werden die Merkmale einer guten Antwort ihrer Bedeutung nach (Rang) noch einmal kurz zusammengefasst, welche von der Arbeitsgruppe zur Analyse der Antwort-Qualität ermittelt wurden:

- Die Antwort muss richtig sein

- Die Antwort wurde mit Top vom Fragesteller bewertet, sie beinhaltet zusätzlich einen weiterführenden Link. Zudem hat der Antwortgeber in seinem eigenen Themengebiet geantwortet.
- Die Antwort wurde vom Fragesteller als hilfreich bewertet: Die Antwort zeichnet sich als Fließtext aus, beinhaltet zudem einen weiterführenden Link oder einen Wikipedia-Link.

Die Arbeitsgruppe zur Analyse der Antworten ermittelte ebenfalls die Ränge der Antwortgeber, die eine qualitativ gute Antwort gegeben hatten. Dieser Zusammenhang wird im Folgenden jedoch nicht berücksichtigt, da das Ergebnis verdeutlicht, dass ein „Einsteiger“ qualitativ bessere Antworten gibt als beispielsweise ein „Einstein“ oder „Bachelor“. Der Umkehrschluss muss demzufolge lauten, dass die Antwortgeber mit höherem Rang qualitativ schlechtere Antworten geben. Dies steht aber im Widerspruch zum Rang-Modell, denn der Antwortgeber hat mit aufsteigendem Rang im größeren Umfang qualitativ gute und richtige Antworten gegeben als die Antwortgeber der unteren Ränge. Aus diesem Grund wird der Rang des Antwortgebers nicht berücksichtigt.

5.1 Die Untersuchung der Antworten nach festgelegten Qualitätsmerkmalen

Das Analyse-Ergebnis der Arbeitsgruppe zur Analyse der Antworten lässt erkennen, welche Qualitätsmerkmale einer guten Antwort zugeordnet werden können. Die Arbeitsgruppe untersuchte jedoch die Qualität der Antworten ausschließlich im Hinblick auf die User-Bewertung. Es ist aber davon auszugehen, dass nicht alle User die Richtigkeit der Antworten zu ihrer gestellten Frage überprüfen. Deshalb musste eine Methode entwickelt werden, um die Qualität der Antworten subjektiv bewerten zu können.

Infolgedessen wurde eine Tabelle erstellt, welche die Qualitätskriterien der „Antworten-Arbeitsgruppe“ berücksichtigt, aber zudem eine Bewertung der einzelnen Antworten ermöglicht. Die Bewertung der Antworten wird mittels eines entworfenen Punktesystems vorgenommen. Jedem Kriterium wurden entsprechend der Bedeutung subjektiv beurteilt mehr oder weniger Punkte zugeteilt. Nachfolgend wird zunächst die Struktur der Tabelle vorgestellt (vgl. Abb. 5.5), auf Grund derer die Antworten untersucht wurden. Im Anschluss wird erläutert, wie die Antworten im Hinblick auf die festgelegten Qualitätskriterien bepunktet wurden.

Tabellenköpfe der einzelnen Spalten	Anmerkung
Fragen	Aufführung der Fragen aus der Stichprobe
Aus welchem Themengebiet stammt die Frage?	Die Unterteilung der Themengebiete in 10 grobe Kategorien erfolgte bereits durch die Antworten-Arbeitsgruppe. Nach deren Einteilung erfolgte die Zuweisung des Themengebietes.
Wie viele Antworten wurden gegeben?	Eintrag der Anzahl der gegebenen Antworten pro Frage
Wie hilfreich ist die Antwort?	Hier wurde eine subjektive Bewertung vorgenommen, inwiefern die Antworten tatsächlich hilfreich bei Lebens- und Erfahrungsfragen waren. Bei Faktenfragen wurde mittels Recherche geprüft, ob die Antworten tatsächlich richtig sind. Die Bepunktung dieses

	Qualitätskriteriums wird im nachfolgenden Punktesystem vorgestellt.
Wie viele Antworten sind richtig bzw. hilfreich?	Nur Eintrag der Antworten, die subjektiv gesehen tatsächlich hilfreich (bei Lebens- oder Erfahrungsfragen) oder richtig (bei Faktenfragen) sind.
Wie viele Antworten sind falsch, unverständlich oder unpassend?	Vornahme einer subjektiven Beurteilung: Bei Faktenfragen wurde die Anzahl der falschen Antworten eingetragen. Bei Lebens- und Erfahrungsfragen die Anzahl der Antworten, die nicht aussagekräftig oder unpassend sind.
Enthält die Antwort einen Fließtext?	Als Fließtext wird ein Text verstanden, der mehr als aus einem Wort besteht. Die einzutragende Anzahl der Antworten bezieht sich nur auf die Antworten, die zuvor subjektiv als hilfreich oder richtig bewertet wurden. Entsprechend wurden hier Punkte pro Antwort vergeben. Die Bepunktung wird nachfolgend vorgestellt.
Wie viele der hilfreichen oder richtigen Antworten bestehen ausschließlich aus einem Link?	Angabe der Anzahl von Antworten, welche die Frage lediglich durch einen weiterführenden Link beantwortet haben
Wie viele Antworten wurden durch die Angabe von Links unterstützt, dass heißt Fließtext und Link?	Eintrag der Antworten, die vom Fragesteller selbst formuliert wurden; ein weiterführender Link unterstützt die Antwort
Wie viele Antworten wurden insgesamt als Top bewertet?	Die Antworten wurden bezüglich der Beurteilung aus Fragestellersicht bewertet. Die dazu vorgenommene Bepunktung wird anschließend erläutert.
Wie viele Antworten wurden insgesamt als hilfreich bewertet?	Die Antworten wurden bezüglich der Beurteilung aus Fragestellersicht bewertet. Die Vergabe der Punkte wird anschließend erläutert.
Wie viele Antworten wurden weder als Top noch als hilfreich bewertet?	Der Fragesteller kann entweder eine Antwort nicht bewerten oder als nicht hilfreich bewerten. Hier sollten nur die Anzahl von Antworten eingetragen werden, die nicht vom Fragesteller bewertet wurden.
Wie viele Antworten wurden als nicht hilfreich bewertet?	Diese Antworten konnten in Lycos iQ durch einen Link eingeblenet werden. Die entsprechende Anzahl nicht hilfreicher Antworten sollte hier eingetragen werden.
Wie viele „gut“ Bewertungen bekamen die Antworten auf die Frage insgesamt?	Anderer User (nicht der Fragesteller) können die gegebenen Antworten als gute Antworten bewerten. Wurde eine Antwort als gut bewertet, wurde eine Bepunktung vorgenommen, die anschließend erläutert wird.
Nach wie vielen Minuten kam die schnellste Antwort?	Eintrag der Anzahl der Minuten, nach der die erste richtige oder hilfreiche Antwort dem Fragesteller zugegangen ist. Dabei wurde anschließend eine Bepunktung vorgenommen, die im Anschluss vorgestellt wird.
Hat der Antwortgeber in seinem Themengebiet geantwortet?	Wenn der Antwortgeber, wie aus dem Lycos-Profil ersichtlich, in seinem angegebenen Themengebiet geantwortet hat, wurden entsprechend dem folgenden Punktesystem Punkte vergeben.
Zu wie viel Prozent hat der	Die bisher hilfreichen Antworten sind ebenfalls im Lycos-Profil des

Antwortgeber bisher hilfreiche Antworten gegeben?	Antwortgebers ersichtlich. Je nach prozentualem Anteil wurden nach dem Punktesystem entsprechend Punkte verteilt.
Summe Relevanz	Punkte aller Antworten, die subjektiv als hilfreich oder richtig bewertet wurden
Summe Fließtext	Punkte aller Antworten mit einem Fließtext
Summe Links	Punkte aller Antworten mit weiterführenden Links
Summe Bewertung	Punkte aller Antworten, die vom Fragesteller als Top oder hilfreich bewertet wurden oder wenn keine Bewertung vom Fragesteller erfolgte
Summe „gut“ Bewertung	Punkte aller Antworten, die von anderen Usern als gute Antwort beurteilt wurden
Summe Themengebiet	Punkte aller Antworten, bei denen der Antwortgeber nach seinem angegebenen Themengebiet geantwortet hat
Summe bisher hilfreiche Antworten des Antwortgebers	Punkte aller Antworten, die erreicht werden konnten, wenn der Antwortgeber bisher hilfreiche Antworten gegeben hat
Summe Geschwindigkeit	Punkte aller Antworten nach Beurteilung der Geschwindigkeit der ersten richtigen bzw. hilfreichen Antwort
Gesamtsumme	Gesamtpunktzahl aller erreichten Punkte in den einzelnen Qualitätskriterien
Wert pro Frage	Gesamtsumme dividiert durch die Anzahl der gegebenen Antworten zu einer Frage

Abb. 5.5 Untersuchungskriterien zur Bewertung der Antworten

5.2 Die Bewertung der Antworten nach den Qualitätskriterien anhand des Punktesystems

Die Beurteilung der Qualität der Antworten erfolgte subjektiv nach einem entwickelten Punktesystem. Die einzelnen Qualitätskriterien, die bezogen auf das Punktesystem bewertet wurden, werden nachfolgend noch einmal aufgeführt. Nicht jedem Qualitätskriterium wurde dieselbe maximale Punkteanzahl zugeordnet. Aus welchen Gründen hier differenziert wurde, wird im Anschluss erörtert.

5.2.1 Erstes Qualitätskriterium: *Wie hilfreich ist die Antwort?*

Punktesystem:

sehr hilfreich = 50 Punkte pro Antwort
bedingt hilfreich = 25 Punkte pro Antwort
nicht hilfreich = 0 Punkte pro Antwort

Hier sollte überprüft werden, ob die Antwort für den Fragesteller hilfreich ist, wenn es sich um eine Erfahrungs- oder Lebensfrage handelt. Lag eine Faktenfrage vor, sollte die Richtigkeit mittels Recherche beurteilt werden.

Bei allen Fragearten wurde zudem beurteilt, ob eine sprachlich verständliche Formulierung der Antwort vorlag. Qualitativ war die Antwort auch bei allen Fragearten höher zu bewerten, wenn die Antwort eine Erklärung enthielt oder ausreichend begründet wurde.

Die Bepunktung erfolgte subjektiv, was auch notwendig war. Denn bei der Betrachtung der jeweiligen Antworten zu einer Frage wurde festgestellt, dass eine als Top, hilfreich oder gut beurteilte Frage durch den Fragesteller oder andere User nicht automatisch richtig oder hilfreich war.

Gerade bei Faktenfragen nimmt die richtige Beantwortung der Frage einen besonders großen Stellenwert ein, wenn davon auszugehen ist, dass diese Frage auf der Lycos-Suche zukünftig mit eingebunden werden soll. War die Antwort falsch, wurden daher keine Punkte vergeben.

Bei Lebens- oder Erfahrungsfragen wurde die Qualität der Antwort danach beurteilt, inwiefern diese hilfreich für den Fragesteller war. Eine Antwort war nur bedingt hilfreich beurteilt worden, wenn zwar ein Ratschlag gegeben wurde, jedoch nicht begründet wurde, warum dieser Ratschlag das Problem hilfreich lösen könnte.

Beispiel:

Wie bekomme ich einen Rotweinfleck aus meinem Teppich?

Antwort 1: Streu Salz drauf!

→ Ratschlag wurde zwar gegeben, aber keine Begründung oder Erklärung dazu.

Daher nur bedingt hilfreich = 25 Punkte.

Antwort 2: Wenn der Rotweinfleck frisch ist, streu sofort Salz darauf. Nicht damit sparsam sein, denn Salz zieht den Rotwein aus dem Teppich heraus.

→ Ratschlag wurde zudem auch begründet, daher sehr hilfreich = 50 Punkte.

Des Weiteren war die Antwort qualitativ höher, wenn der Antwortgeber die Beantwortung selbst formuliert hat und diese nicht nur durch einen angegebenen Link indirekt gegeben wurde.

Wie die Antwort-Arbeitsgruppe belegt hat, wird eine gute Antwort auch durch einen weiterführenden Link unterstützt. Aus diesem Grund wurden die angegebenen Links der Antwortgeber auf ihre Richtigkeit kontrolliert bzw. beleuchtet, inwiefern diese die Antwort hilfreich unterstützen. War das der Fall, konnte die Antwort zusätzliche Punkte erhalten (siehe Qualitätskriterium Nr. 2).

Bei dieser Bewertung konnte eine Antwort nach dem Punktesystem auch die höchste Punktzahl erreichen. Es ist davon auszugehen, dass der Fragesteller sich auf die Richtigkeit der Antworten verlässt, wenn es sich um Faktenfragen handelt. Eine besonders kritische Beurteilung dieser war daher unerlässlich.

5.2.2 Zweites Qualitätskriterium: Erhält die Antwort einen weiterführenden Link?

Punktesystem:

Ja = 10 Punkte pro Antwort
Nein = 0 Punkte

Wie bereits angemerkt, wurde die Qualität der weiterführenden Links unter Punkt 1 beurteilt. Wenn ein Link die Antwort hilfreich unterstützt hat, konnten zusätzlich 10 Punkte pro Link vergeben werden.

5.2.3 Drittes Qualitätskriterium: Erhält die Antwort einen Fließtext?

Punktesystem:

Ja = 10 Punkte
Nein = 0 Punkte

Eine Antwort besteht aus einem Fließtext, wenn sie aus mehr als einem Wort besteht. Die Qualität der Antwort erhöht sich, wenn diese auch für den Fragesteller begründet wird.

Beispiel:

Frage: Kennt jemand den St. Georgsorden?

Positiv: Ein Fließtext liegt vor, wenn der Antwortgeber begründet, wobei es sich bei diesem St. Georgsorden handelt.

Negativ: Wenn der Antwortgeber beispielsweise nur mit „Ja“ oder „Nein“ geantwortet hat.

5.2.4 Viertes Qualitätskriterium: Wie wurde die Antwort vom Fragesteller bewertet?

Punktesystem:

Top = 20 Punkte pro Antwort
Hilfreich = 10 Punkte pro Antwort
Keine Bewertung = 5 Punkte pro Antwort

Dieses Kriterium sollte der unter Punkt 1 vorgenommenen subjektiven Beurteilung entgegenstehen. Jedoch ist die maximale Punktzahl geringer als unter Punkt 1. Denn, wie bereits angemerkt wurde, war eine durch den Fragesteller als Top oder hilfreich bewertete Antwort nicht automatisch richtig. Wurde, wie das Punktesystem deutlich macht, die Antwort vom Fragesteller als Top bewertet, erhielt die Antwort die doppelte Punktzahl als eine hilfreiche Antwort. Eine Antwort konnte aber auch noch 5 Punkte dazu gewinnen, wenn sie überhaupt nicht vom Fragesteller bewertet wurde. Hintergrund dieser Punktevergabe ist, dass der Fragesteller eine

Antwort auch als nicht hilfreich beurteilen kann. Diese Antworten erhielten dann entsprechend keine Punkte.

5.2.5 Fünftes Qualitätskriterium: Wie oft wurde die Antwort von anderen Usern als „gut“ bewertet?

Punktesystem:

Für jede „Gut-Bewertung“ = 5 Punkte

Wenn eine Antwort von anderen Usern als „gut“ bewertet wurde, ist davon auszugehen, dass diese nicht nur für den Fragesteller qualitativ wertvoll ist. Die geringere Bepunktung liegt darin begründet, dass in erster Linie der Fragesteller mit der Antwort zufrieden gestellt sein sollte.

5.2.6 Sechstes Qualitätskriterium: Zu welchem Prozentsatz haben die Antwortenden bisher hilfreiche Antworten gegeben?

Punktesystem:

- > 95 % = 10 Punkte*
- > 90 % = 5 Punkte*
- < 90 % = 0 Punkte*

Bei diesem Kriterium wurde die Verlässlichkeit der bisherigen hilfreichen Antworten durch die Antwortgeber bewertet. Die bisher abgegebenen hilfreichen Antworten konnten im Lycos-Profil eingesehen werden. Hat der Antwortgeber bisher mehr als 95 Prozent aller Fragen hilfreich bzw. richtig beantwortet, wurden 10 Punkte vergeben. Lag die Quote zwischen 90 und 95 Prozent, so konnten 5 Punkte erreicht werden. Bei unter 90 Prozent der bisher hilfreichen bzw. richtigen Antworten wurden hingegen keine Punkte verteilt.

Zwar werden bei dieser Bepunktung die vergangenen gegebenen Antworten beurteilt, aber mit dieser zusätzlichen Punktevergabe differenziert sich heraus, in welchem Themengebiet besonderes viele richtige Antworten bisher gegeben wurden.

5.2.7 Siebtes Qualitätskriterium: Hat der Antwortende in seinem Themengebiet geantwortet?

Punktesystem:

- Ja = 10 Punkte*
- Nein = 0 Punkte*

Wie die Antwort-Arbeitsgruppe bei der Analyse der Antworten herausgefunden hat, sind Antworten von höherer Qualität, wenn der Antwortgeber aus seinem eigenen Themengebiet geantwortet hat, denn die Beantwortung im eigenen Themengebiet verstärkt die Glaubwürdigkeit der gegebenen Antwort. Daher wurden für jede

Antwort zusätzlich 10 Punkte vergeben, wenn der Antwortgeber entsprechend seinem Lycos-Profil in seinem Themengebiet geantwortet hat.

Eine höhere Punktevergabe wurde hier vernachlässigt, da die „Antworten-Gruppe“ herausgefunden hat, dass durchaus gute Antworten gegeben werden, wenn der Antwortgeber nicht im eigenen Themengebiet geantwortet hat.

5.2.8 *Achtes Qualitätskriterium: Dauer bis zur ersten richtigen Antwort*

Punktesystem:

<i>Schnell</i>	< 60 Minuten = 10 Punkte
<i>Mittel</i>	< 400 Minuten = 5 Punkte
<i>Langsam</i>	> 400 Minuten = 0 Punkte

Dieses Zeit-Kriterium wurde aufgenommen, weil davon auszugehen ist, dass eine Antwort für den Fragenden besonders hilfreich ist, wenn sie möglichst schnell gegeben wird. Die Einteilung der obigen Zeitbereiche wurde subjektiv vorgenommen und stützt sich auf die durchschnittliche Wartezeit bis zur ersten richtigen Antwort. In einigen Fällen kann eine erste richtige Antwort bereits nach 10 Minuten eintreffen, in anderen Fällen kam die erste richtige Antwort aber erst nach mehreren Stunden. Außerdem muss hierbei berücksichtigt werden, dass dem Antwortgeber auch die Möglichkeit gegeben sein sollte, sich ausführlich mit der Beantwortung einer Frage auseinanderzusetzen. Gerade im Hinblick auf die Erfahrungs- oder Lebensfragen ist die Antwort zumeist umfangreicher als bei einer Faktenfrage.

Qualitätskriterium	Punktesystem
Wie hilfreich ist die Antwort?	Sehr hilfreich = 50 Punkte Bedingt hilfreich = 25 Punkte Nicht hilfreich = 0 Punkte
Enthält die Antwort einen weiterführenden Link?	Ja = 10 Punkte Nein = 0 Punkte
Enthält die Antwort einen Fließtext?	Ja = 10 Punkte Nein = 0 Punkte
Wie wurde die Antwort vom Fragesteller bewertet?	Top = 20 Punkte Hilfreich = 10 Punkte Keine Bewertung = 5 Punkte
Wie oft wurde die Antwort von anderen Usern als „gut“ bewertet?	Für jede „gut“ Bewertung = 10 Punkte
Zu welchem Prozentsatz haben die Antwortenden bisher hilfreiche Antworten gegeben?	> 95 % = 10 Punkte > 90 % = 5 Punkte < 90 % = 0 Punkte
Hat der Antwortende in seinem Themengebiet geantwortet?	Ja = 5 Punkte Nein = 0 Punkte
Dauer bis zur ersten richtigen Antwort	Schnell < 60 Minuten = 10 Punkte Mittel < 400 Minuten = 5 Punkte Langsam > 400 Minuten = 0 Punkte

Abb. 5.6 Überblick des Punktesystems

6. METHODIK: FRAGEKATALOG

Nachdem 393 Antworten nach dem Punktesystem bewertet und qualitativ eingeordnet wurden, musste herausgefunden werden, ob sich aus der Art und Qualität einer Frage möglicherweise Rückschlüsse auf die Qualität der Antworten ergeben. Zu diesem Zweck war es erforderlich, die 247 Fragen, die zu den bereits bewerteten Antworten gehörten, näher zu betrachten und zu analysieren.

Für die systematische Untersuchung der Fragen war es notwendig, sich vorab Kriterien zu überlegen, mit deren Hilfe es theoretisch möglich sein könnte, die Qualität einer Frage zu beurteilen. Dabei lag bei den meisten Kriterien eine Hypothese zu Grunde, wie sich das gewählte Kriterium auf die Qualität der Antworten und Fragen auswirken könnte. Rückschlüsse, welches Kriterium sich tatsächlich für eine Unterscheidung der Fragen gut oder schlecht eignet, können aber erst nach Auswertung aller Daten getroffen werden.

Folgende Kriterien wurden vorab festgelegt:

6.1 Formale Kriterien

Es wurde davon ausgegangen, dass eine qualitativ gute Antwort von formalen Kriterien einer Frage abhängen könnte. Bei der Betrachtung formaler Merkmale sollte das inhaltliche Verständnis einer Frage unberücksichtigt bleiben. Es wurden folgende formale Kriterien, nach denen eine Frage bewertet werden kann, festgelegt:

- **Zeichnet sich die Frage durch eine gute Rechtschreibung aus?**

Jede Frage wurde auf Schwächen in der Rechtschreibung untersucht. Die Untersuchung ließ nur zwei Bewertungen zu: „Gut“ bedeutete, das Kriterium wurde beachtet, „Schlecht“ zeigte, das Kriterium wurde nicht erfüllt.

- **Wurde die Grammatik beachtet?**

Jede Fragen wurde auf Schwächen in der Grammatik untersucht. Die Untersuchung ließ nur zwei Bewertungen zu: „Gut“ bedeutete, das Kriterium wurde beachtet, „Schlecht“ zeigte, das Kriterium wurde nicht erfüllt.

Negativ-Beispiele: Wie entsteht eine Tsunami? Was ist von den Film "Lissi und der wilde Kaiser" zu halten? Gibt es schon Trailers? Welches ist die Hauptstadt von Mongolei?

- **Entspricht die Satzstellung einer Frage?**

Der Aufbau der Fragestellung wurde separat untersucht. Dabei sollte herausgefunden werden, ob eine Frage als solche richtig formuliert wurde. Es erschien wichtig, anhand der Satzstellung sofort erkennen zu können, dass es sich um eine konkrete Frage handelt. Das Kriterium ließ nur zwei Bewertungen zu: „Gut“ bedeutete, das Kriterium wurde beachtet, „Schlecht“ zeigte, das Kriterium wurde nicht erfüllt.

Beispiele für eine schlechte Fragestellung: Suche Infos zu Personen, Hauptstadt Boliviens, Igel zugelassen, Definition für die Bereicherung.

6.2 Inhaltliche Kriterien

Es wurde vermutet, dass eine qualitativ gute Antwort von inhaltlichen Kriterien einer Frage abhängen könnte. Deshalb wurden die Fragen nach Merkmalen untersucht, die die inhaltliche Qualität einer Frage bewerten. Bei der Betrachtung dieser Merkmale sollte ausschließlich das inhaltliche Verständnis einer Frage berücksichtigt werden. Aus diesem Grund wurden folgende inhaltliche Kriterien festgelegt:

- **In welchem Sprachstil ist die Frage formuliert?**

Es wurde untersucht, ob die Frage in verständlichem Hochdeutsch formuliert oder ob Umgangssprache verwendet wurde. Es lag die Annahme zu Grunde, durch eine umgangssprachliche Formulierung der Frage würde es zu Missverständnissen kommen, welche falsche Antworten hervorrufen könnten.

Beispiel für eine Formulierung in Umgangssprache: Wie krieg ich's gebacken?

- **Werden Fremd- oder Fachwörter verwendet?**

Jede Frage wurde auf Verwendung von Fremdwörtern untersucht. Als Fremd- und Fachwörter sollten alle Wörter verstanden werden, die aus einer fremden Sprache oder aus einer Fachsprache übernommen wurden und deren Bedeutung im Allgemeinen nur schwer erfasst werden kann. Um Verständnis für die Frage zu erlangen, muss ein Antwortender die Fremd- und Fachwörter kennen oder die Bedeutung des spezifischen Begriffes nachschlagen. Es lag die Annahme zu Grunde, je weniger die Frage durch eine „schwere“ Sprache allgemein verständlich ist, desto schlechter könnten auch die Antworten ausfallen.

Beispiele: Wird Drive Defrag der TuneUp Utilities 2008 taugen? Was sind für euch die besten Psychedelic Hits?

- **Animiert die Fragestellung zu kurzen Antworten?**

Es wurde untersucht, ob die Fragestellung zu einer wenig ausführlichen Antwort animiert, obwohl davon auszugehen ist, dass der Fragesteller eigentlich mehr Informationen wünscht. Die Antworten könnten dann aus einem Wort bestehen wie „Ja“ oder „Nein“, „Gut“ oder „Schlecht“ und damit für den Fragesteller als unbefriedigend bzw. schlecht bewertet werden.

Beispiele: Kann mir jemand sagen, wo Senegal liegt? Glaubt ihr, es geht beim VfB wieder aufwärts in nächster Zeit? Wie findet ihr den Film ...?

- **Ist die Frage logisch verständlich?**

Es wurde untersucht, ob die Frage logisch verständlich formuliert wurde. Als logisch galt eine Frage, wenn aus ihr alle Informationen hervorgingen, die nötig waren, um eine eindeutige Antwort zu erhalten.

Beispiele für eine solche unlogische Frage: Kann mir jemand ein Buch empfehlen? Warum brennen immer meine Augen? Fabeltiere?

Um als logisch eingeordnet zu werden, musste die Frage außerdem eine inhaltliche Logik aufweisen. Es musste möglich sein, eine sinnvolle Antwort auf die Frage zu finden.

Beispiele für eine unlogische Frage: Gibt es Weihnachten 2020 Schnee? Was ist der mathematische Wert des Gottes?

- **Provoziert die Frage zu einer unpassenden Antwort?**

Es wurde untersucht, ob die Frage Antwortende dazu reizen könnte, eine ausfallende, verärgerte Antwort zu geben. Als provokativ galt eine Frage, wenn sie

Gelegenheit gab, Klischees zu bestätigen oder Tatsachen voraussetzte, die nicht vorauszusetzen waren.

Beispiele für provokative Fragen: Warum singen Schwarze besser als alle anderen? Warum sind Scientologen gefährlich? Wieso müssen sich Muslimas unter einem Kopftuch verstecken?

- **Wird die Ausführungsbox für weitere Erklärungen zur Frage genutzt?**
Es wurde untersucht, ob die Frage im anschließenden Zusatzfeld näher beschrieben wurde. Durch das Zusatzfeld wird dem Antwortenden die Möglichkeit gegeben, ein besseres Verständnis für den Sachverhalt zu erlangen, über den der Fragesteller etwas in Erfahrung bringen möchte. Es lag daher die Annahme zu Grunde, durch Ergänzungen im Zusatzfeld könnten bessere Antworten erzielt werden.
Beispiel?
- **Ist die Frage zeitlich begrenzt?**
Es wurde untersucht, ob die Frage einer zeitlichen Begrenzung unterliegt.
Beispiel: Wann findet das nächste Madonna-Konzert statt?

6.3 Schwierigkeitsgrad

Es erfolgte eine Einteilung nach dem Wissen, das ein Antwortender möglicherweise benötigt, um auch tatsächlich eine gute Antwort geben zu können. Der Wissensgrad wurde abhängig vom Rechercheaufkommen beurteilt. Es sollte folgende Annahme überprüft werden: Je speziellere Kenntnisse die Beantwortung einer Frage erfordert, desto schlechter könnten die Antworten ausfallen. Dabei wurde eine Einteilung in drei verschiedene Schwierigkeitsgrade vorgenommen: Allgemeinwissen, Fachwissen und Spezialwissen.

- **Kann ein Antwortender die Frage mit Allgemeinwissen beantworten?**
Es wurde untersucht, ob die gute Beantwortung der Frage nur allgemeines Wissen voraussetzt. Solche Fragen können von vielen Personen aus ihrem allgemeinen Kenntnisstand heraus oder aber durch eine schnelle und einfache Recherche beantwortet werden.
Beispiel für eine Frage, deren Beantwortung Allgemeinwissen benötigt: Wo steht das Empire State Building?
- **Kann nur ein Antwortender die Frage beantworten, der mindestens über Fachwissen verfügt?**
Es wurde untersucht, ob die gute Beantwortung der Frage Fachwissen voraussetzt. Solche Fragen können nicht von allen Personen aus ihrem allgemeinen Kenntnisstand heraus beantwortet werden. Ein Antwortender benötigt fachliche Kenntnisse über den Sachverhalt, muss sich mit dem Themengebiet schon im Vorfeld auseinandergesetzt haben oder für eine gute Antwort intensiv recherchieren.
Beispiel für eine Frage, deren Beantwortung Fachwissen benötigt: Hat Regina Halmig verdient gewonnen?
- **Kann nur ein Antwortender die Frage beantworten, der über Spezialwissen verfügt?**
Es wurde untersucht, ob die gute Beantwortung der Frage Spezialwissen voraussetzt. Solche Fragen können nur von wenigen Personen beantwortet werden,

die sich mit einer speziellen Thematik bereits ausgiebig beschäftigt haben. Die Fragen beinhalten oft bereits spezielle Fachwörter, über deren Bedeutung sich nur Antwortende mit speziellen Kenntnissen im Klaren sind.

*Beispiel für eine Frage, deren Beantwortung Spezialwissen benötigt:
Was ist besser, GoZilla oder FlashGet?*

6.4 Sonstige Kriterien

Nach der Feststellung formaler und inhaltlicher Kriterien der Fragen und der Berücksichtigung des Schwierigkeitsgrades zur Beantwortung einer Frage wurden sonstige weitere Kriterien festgelegt:

- **Wie viele Bonuspunkte wurden ausgesetzt?**

Es wurde untersucht, wie viele Bonuspunkte ein Fragesteller für die gute Beantwortung seiner Frage aussetzt. Dabei lag die Annahme zu Grunde, dass ein Fragesteller, der eine richtige und zufrieden stellende Antwort erwartet, eine höhere Anzahl Bonuspunkte als „Belohnung“ aussetzt als jemand, der es nur als Zeitvertreib und Spaß ansieht, Fragen bei Lycos iQ zu stellen. Des Weiteren gab es die Mutmaßung, ein Fragesteller, der Wert auf eine besonders gute Antwort legt und damit höhere Bonuspunkte aussetzt, würde sich gleichzeitig auch um eine formal und inhaltlich richtige Fragestellung bemühen.

- **Anzahl der Möglichkeiten, um eine Frage zu beantworten**

Es wurde untersucht, ob zu einer Frage mehrere richtige Antworten möglich sein könnten. Dabei lag die Annahme zu Grunde, dass bei Meinungs-, Lebens- und Erfahrungsfragen mehrere richtige Antworten vorstellbar sind – bei annähernd gleich bleibender Qualität. Bei Faktenfragen dagegen ist größtenteils nur eine richtige Antwort möglich, so dass dem Fragesteller doppelte und damit mehrfach dieselben richtigen Antworten zugehen. Diese Doppelungen könnten jedoch zu einer insgesamt schlechteren Beantwortung einer Frage führen, da ein Fragesteller, der bereits die eine richtige Antwort erhalten hat, die Antworten der "Nachahmer" wohl nicht mehr als nützlich empfinden würde. In einigen Fällen könnte es aber durchaus möglich sein, dass eine Faktenfrage mehrfach richtige Antworten generiert. Zum Beispiel, wenn es sich um eine allgemeine Personenfrage handelt, wie „Wer war Albert Schweitzer?“ könnten Antworten wie: "Theologe", "Orgelkünstler", "Musikforscher", "Philosoph" und "Arzt" in Frage kommen. In diesen Fällen könnte die unterschiedliche Beantwortung zu Erhöhung der Qualität der Antwort führen, da dem Fragesteller ein umfangreicheres Wissen zu seiner Frage vermittelt worden ist.

Kriterium	Untersuchung	Antwort
Einteilung nach Art der Frage:	Um was für eine Art Frage handelt es sich?	Faktenfrage
		Meinungsfrage
		Lebensfrage
		Spaßfrage
Formale Kriterien:	Zeichnet sich die Frage durch eine gute Rechtschreibung aus?	Ja/Nein
	Wurde die Grammatik beachtet?	Ja/Nein
	Entspricht die Satzstellung einer Frage?	Ja/Nein
Inhaltliche Kriterien:	In welchem Sprachstil ist die Frage formuliert?	Hochdeutsch/ Umgangssprache
	Werden Fremd- oder Fachwörter verwendet?	Ja/Nein
	Animiert die Fragestellung zu kurzen Antworten?	Ja/Nein
	Ist die Frage logisch verständlich?	Ja/Nein
	Provoziert die Frage zu einer unpassenden Antwort?	Ja/Nein
	Wird das Zusatzfeld für weitere Erklärungen zur Frage genutzt?	Ja/Nein
	Ist Frage zeitlich begrenzt?	Ja/Nein
Wie viele User bewerteten die gestellte Frage als „Gute Frage“?	Anzahl der Gut-Bewertungen durch die User	
Kriterien in Bezug auf den Fragesteller:	Wie viele Bonuspunkte wurden ausgesetzt?	Anzahl der Punkte
	Welchen Status hat der Fragesteller?	Status des Fragestellers
Einteilung nach Schwierigkeitsgrad:	Welches Wissen benötigt jemand, um die Frage zu beantworten?	Allgemeinwissen
		Fachwissen
		Spezialwissen

Abb. 5.7 Überblick aller Kriterien

7. METHODIK DER AUSWERTUNG

Die für die Analyse notwendigen Daten wurden manuell von einem Probanden-Team erhoben, indem Fragen und Antworten mit Hilfe der definierten Kriterien subjektiv beurteilt wurden. Anschließend erfolgte eine statistische Auswertung der Daten, um Zusammenhänge aufzuzeigen.

Da Zusammenhänge der einzelnen Antwort-Merkmale untereinander bereits im Beitrag der Arbeitsgruppe zur Analyse der Antwort-Qualität dargestellt wurden (vgl. den vorherigen Beitrag in diesem Band), sollen hier nur die Kriterien erläutert werden, die in direktem Zusammenhang mit der jeweils zugehörigen Frage stehen. Als Grundlage dient also lediglich die nach oben erklärtem Muster errechnete Punktzahl pro Frage, die die Antwortqualität widerspiegelt.

8. ERGEBNISSE

8.1 Verteilung der Antwortqualität

Nach Auswertung aller 393 Antworten ergab sich eine durchschnittliche Punktzahl pro Frage von 70,09 Punkten. Eine Antwort durchschnittlicher Qualität würde also eine Punktzahl von 70,09 Punkten erreichen. Dabei waren maximal 120 Punkte zuzüglich je 5 Punkten pro vergebener Gut-Bewertung zu erreichen. Die in Abb. 5.9 dargestellte Verteilung ergibt sich aus der Ordnung der Fragen nach folgender Skala:

Sehr gut:	> 100
Gut:	100 <> 80
Mittel	80 <> 60
Schlecht	60 <> 40
Sehr schlecht	< 40

Abb. 5.8 zeigt exemplarisch die Punkteverteilung für eine Frage mit sehr guter Antwortqualität und eine sehr schlecht bewertete Frage.

Qualitätskriterien	Was bedeutet der Begriff „Halb-schmarotzer“ in der Pflanzenwelt?	Auf welche Beweise begründet sich die Behauptung, Jesus sei in einer Werkstatt geboren?
Wie hilfreich ist die Antwort?	100	25
Enthält die Antwort einen weiterführenden Link?	0	10
Enthält die Antwort einen Fließtext?	20	10
Wie wurde die Antwort vom Fragesteller bewertet?	40	0
Wie oft wurde die Antwort von anderen Usern als „gut“ bewertet?	10	15
Zu welchem Prozentsatz haben die Antwortenden bisher hilfreiche Antworten gegeben?	20	0
Hat der Antwortende in seinem Themengebiet geantwortet?	20	10
Dauer bis zur ersten richtigen Antwort	10	10
Summe	220	80
Anzahl der Antworten	2	5
Wert der Frage	110	16
Bewertung	Sehr gut	Sehr schlecht

Abb. 5.8 Exemplarische Bewertung einer Frage mit sehr guter und einer Frage mit sehr schlechter Antwortqualität

Abb.5.9 zeigt, dass ca. ein Drittel aller Fragen keine brauchbaren Antworten generieren. Angesichts der fehlenden qualitativen Kontrolle der Antworten in der Community ist dieser Wert positiv zu bewerten, vor allem da der Anteil der

mindestens guten Antworten deutlich höher ist als der Anteil der höchstens schlechten Antworten. Allerdings wurden auch nur 14 Prozent der Antworten von den Probanden als sehr gut und 23 Prozent als gut bewertet. Setzt man nun voraus, dass sich, wie eingangs erwähnt, ausschließlich gute und sehr gute Antworten zur Ergänzung klassischer Suchformen eignen, reduziert sich die Menge der potenziell geeigneten Antworten auf ca. ein Drittel der Gesamtmenge.

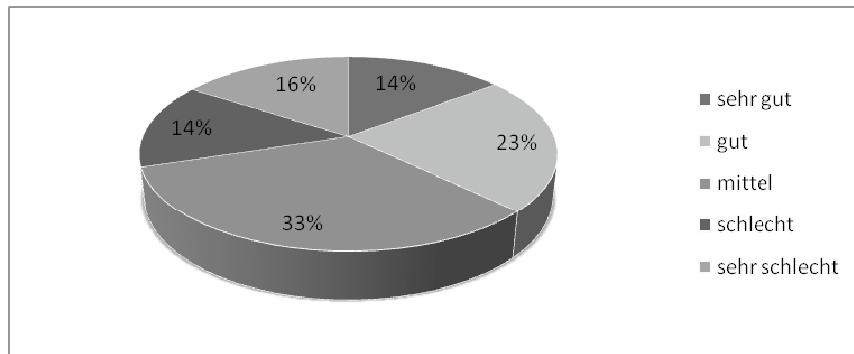


Abb. 5.9 Verteilung der Fragen nach ihrer Qualität

Betrachtet man die thematische Verteilung der Antwortqualität, so ergibt sich die in Abb. 5.10 dargestellte Verteilung. Auffällig sind hier vor allem die sehr gute Qualität der Antworten im Bereich „Filme“ und das gleichzeitig schlechte Abschneiden der Wissenschaftsbereiche. Hier liegt der Schluss nahe, dass Bereiche mit hohem allgemeinen Interesse (auch „Freizeit“, „Sprachen/Länder“) und niedrigem benötigten Fachwissen eine besonders gute Antwortqualität generieren. Symptomatisch für die spezialisierten Bereiche dagegen ist der Bereich „Computer & Technisches“, dessen Fragen, die in einem Technik-affinen Umfeld gestellt werden, nur eine durchschnittliche Antwortqualität von ca. 67 Punkten erreichten.

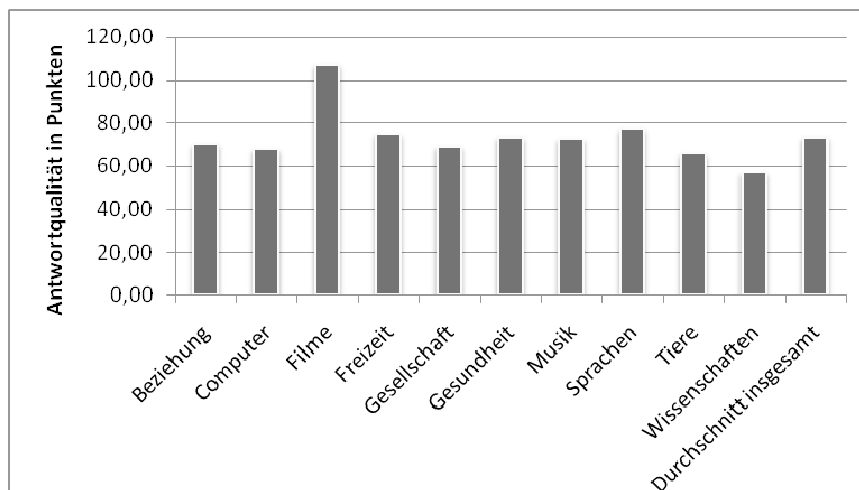


Abb. 5.10 Durchschnittliche Antwortqualität pro Themengebiet, n=247

Verteilt man die Antwortqualität nun auch nach der Art der Frage, erhält man eine wenig überraschende Verteilung, die den Faktenfragen die beste Antwortqualität zubilligt. Dies dürfte vor allem in der Art der Antworten begründet liegen, da Antworten auf Faktenfragen meist präziser und weniger strittig sind als Antworten, die eine einzelne Meinung widerspiegeln (vgl. Abb. 5.11).

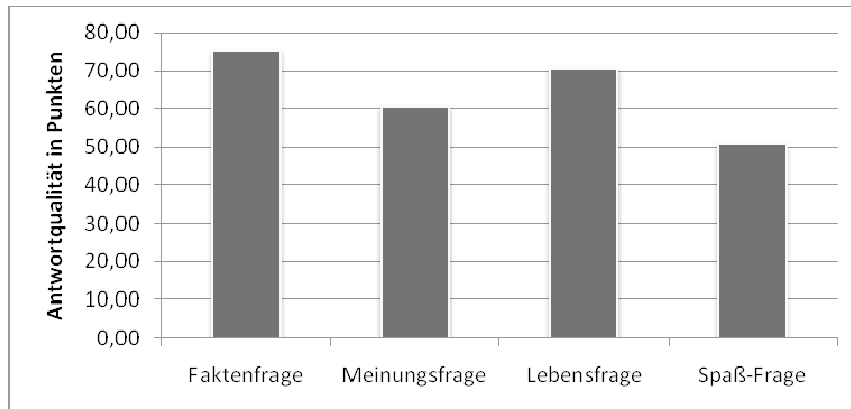


Abb. 5.11 Durchschnittliche Antwortqualität pro Art der Frage, n=247

Als erstes Ergebnis ist an dieser Stelle also festzuhalten, dass Fragen dann besonders gute Antworten generieren, wenn sie ernst gemeint sind und als Antwort keine Meinung verlangen. Themengebiete von hohem allgemeinen Interesse und eher geringem benötigten Fachwissen schneiden besser ab als schwierigere, spezielle Themen. Faktenfragen aus dem Bereich „Film“ erreichen so auch tatsächlich im Schnitt eine Antwort-Punktzahl von 86,08 Punkten, während Spaßfragen aus dem Bereich „Wissenschaften“ nur auf eine Punktzahl von 35,13 kommen.

Die Art der Frage und das Thema, in dem die Frage gestellt wurde, lassen also bereits erste Schlüsse über die zu erwartende Qualität der Antworten zu. Vor allem der Zusammenhang zwischen Antwortqualität und Thema der Frage ist dabei auf Grund der einfachen Möglichkeit einer späteren, automatisierten Selektion der guten Fragen interessant.

8.2 Schwierigkeitsgrad

Wenn man sich nun dem Zusammenhang zwischen der Qualität der Antworten und der analysierten Merkmale zuwendet, lässt sich das oben beschriebene Ergebnis auch mit der folgenden Grafik belegen (vgl. Abb. 5.12). Die Prozentzahlen spiegeln dabei jeweils den Anteil der Fragen mit positiver Bewertung wider. Wird der Anteil der Fragen, der mit Allgemeinwissen zu beantworten ist, in einem Themengebiet mit 90 Prozent angegeben, so bedeutet dies, dass die Frage: „Kann ein Antwortender die Frage mit Allgemeinwissen beantworten?“ in 90 Prozent der Fälle mit „Ja“ beantwortet wurde.

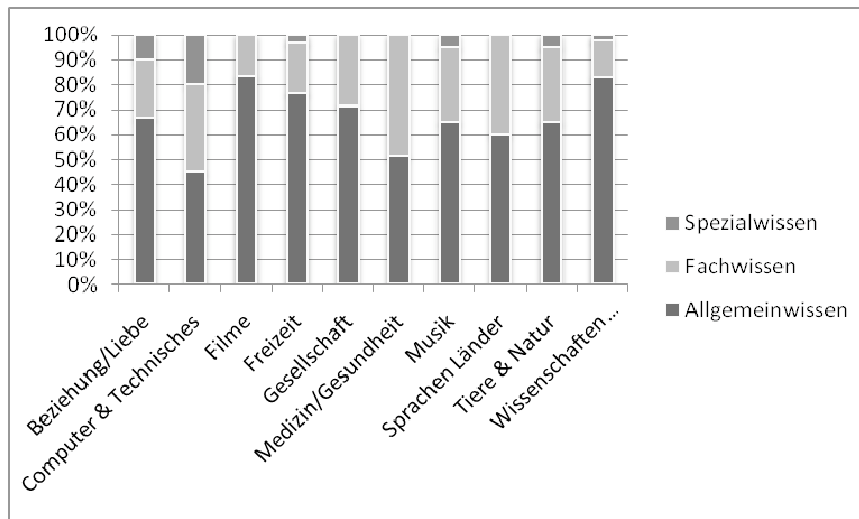


Abb. 5.12 Schwierigkeitsgrad der Fragen in Abhängigkeit vom Themengebiet

Es wurde analysiert, welches Wissen der Antwortende zur Beantwortung der Frage benötigt. Hier zeigt sich ein Zusammenhang zwischen dem Anteil der Fragen, die man als Anfänger beantworten kann, und der oben aufgezeigten Antwortqualität pro Themengebiet. Je höher der Anteil an Allgemeinwissensfragen, desto besser die Antwortqualität.

Der Anteil der Fragen, für die Spezialwissen benötigt wird, ist insgesamt extrem gering. Eine Ausnahme bildet hier der Bereich „Medizin/Gesundheit“, was aber nicht sonderlich überrascht. Interessant ist, dass im Bereich „Filme“, dem Bereich mit den besten Antworten, die leichtesten Fragen gestellt werden. Dazu passt, dass alle Fragen, die man als Anfänger beantworten kann, im Schnitt eine Antwort-Punktzahl von 73,03 erreichen, während die Fragen für Fortgeschrittene nur eine Antwortqualität von 66,24 und die Fragen für Spezialisten sogar nur eine Antwortqualität von 62,35 erreichen. Der Schwierigkeitsgrad der Frage scheint also direkten Einfluss auf die Qualität der Antworten zu haben, auch wenn die Unterschiede weniger gravierend ausfallen, als das eventuell zu erwarten gewesen wäre.

Dies erklärt die Ausnahmestellung des Bereiches „Filme“ bei der Antwortqualität in Teilen. Da allerdings auch der Bereich „Wissenschaften“ nur relativ einfache Fragen aufweist, müssen neben der Schwierigkeit der Frage auch andere Faktoren wie zum Beispiel das Interesse am Thema der Frage eine Rolle spielen.

Stellt man den Schwierigkeitsgrad in Abhängigkeit von der Fragen-Art dar, so lässt sich die oben gewonnene Erkenntnis nicht bestätigen (vgl. Abb. 5.13). Die Verteilung zeigt zwar, dass die qualitativ eher schlechter beantworteten Meinungsfragen auch den höchsten Schwierigkeitsgrad aufweisen, die qualitativ deutlich besser beantworteten Faktenfragen weisen jedoch einen nahezu ebenso hohen Schwierigkeitsgrad auf. Die Spaßfragen mit dem logischerweise niedrigsten Schwierigkeitsgrad generieren sogar die schlechtesten Antworten.

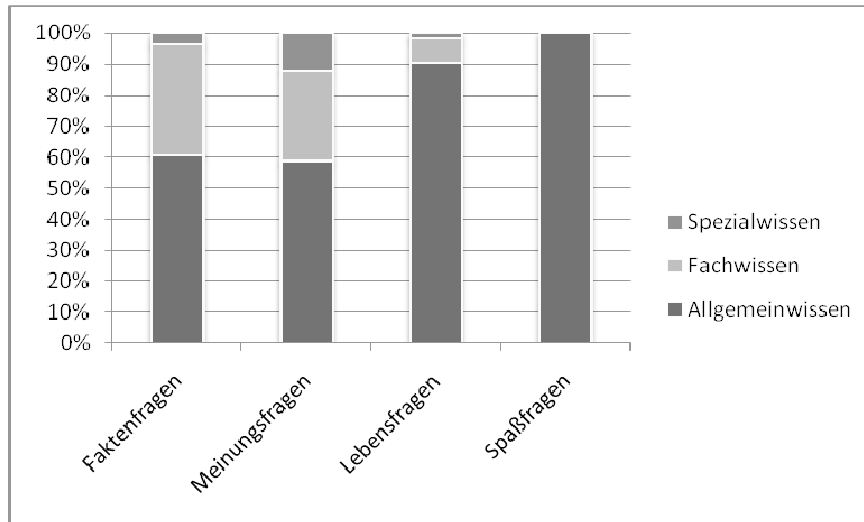


Abb. 5.13 Schwierigkeitsgrad der Fragen in Abhängigkeit von der Art der Frage

Setzt man den Schwierigkeitsgrad in direkte Abhängigkeit zur Qualität der Antworten, so zeigt sich, dass gute Antworten zwar tendenziell einen etwas höheren Anteil an einfachen Fragen aufweisen, der Zusammenhang aber recht gering ist (vgl. Abb. 5.14). Die Qualität der Antworten lässt sich also nicht allein durch den Schwierigkeitsgrad der Frage beeinflussen. Je nach Art oder Thema der Frage variiert die Antwortqualität teils erheblich.

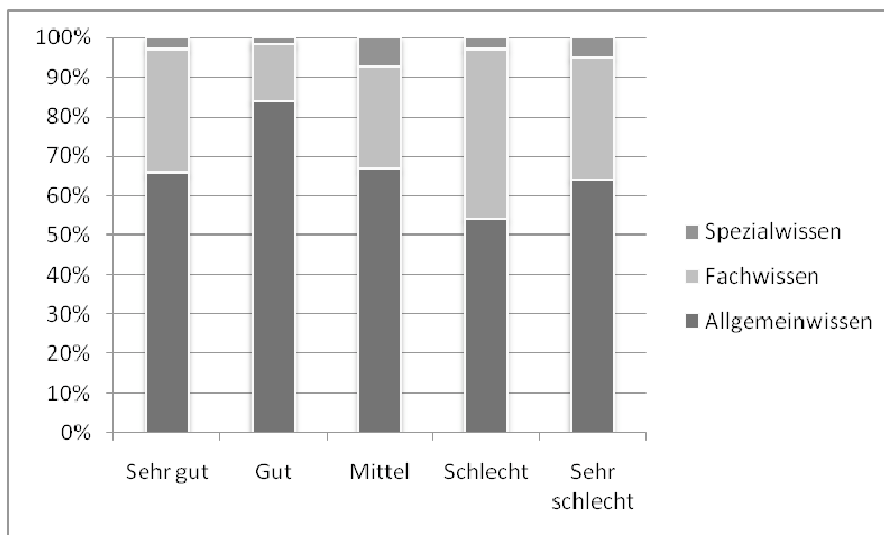


Abb. 5.14 Schwierigkeitsgrad der Fragen in Abhängigkeit von der Antwortqualität

8.3 Anzahl der Antworten

Ein weiteres analysiertes Merkmal, das zwar nicht in die Berechnung der Antwortqualität eingeflossen ist, aber durchaus von Bedeutung sein könnte, ist die Anzahl der Antworten, die pro Frage gegeben wurden. Stellt man die Anzahl der Antworten und die Antwortqualität für die einzelnen Themen dar (vgl. Abb. 5.15), so sieht man, dass sich die Anzahl der Antworten eher antiproportional zur Antwortqualität verhält. Dieser Zusammenhang ist zwar sehr schwach ausgeprägt, aber trotzdem lässt sich festhalten, dass die Qualität der Antworten bei Fragen mit vielen Antworten tendenziell schlechter ist als bei Fragen, auf die nur wenig User antworten. Ein Vergleich mit Abbildung 5.14 zeigt, dass der Schwierigkeitsgrad der Frage keinen direkten Einfluss auf die Anzahl der Antworten hat.

Innerhalb der Themen generieren Fragen aus dem Themenbereich „Sprachen/Länder“ die meisten Antworten pro Frage. Die geringste Anzahl Antworten pro Frage wird im Themenbereich „Computer & Technisches“ generiert.

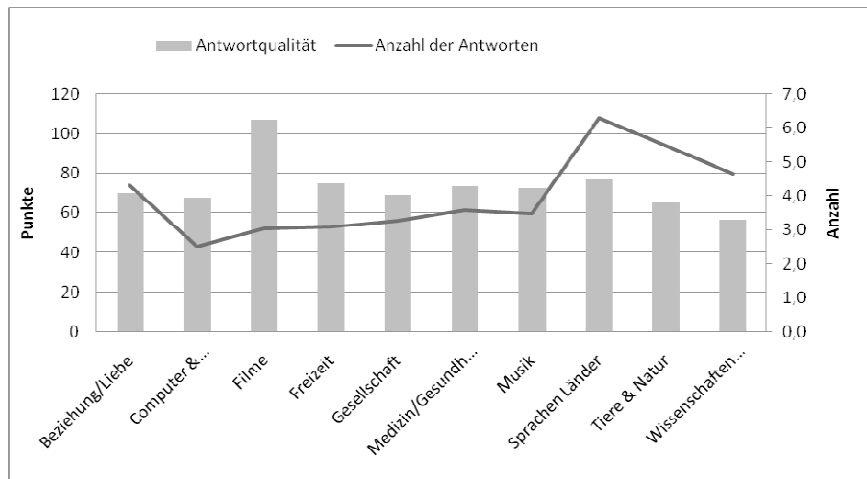


Abb. 5.15 Durchschnittliche Antwortqualität und durchschnittliche Anzahl der Antworten pro Frage in den Themengebieten.

Interessant ist, dass die Anzahl der Antworten nicht nur vom Thema, sondern auch von der Art der Frage abzuhängen scheint. Faktenfragen, die wie bereits gezeigt, die beste Antwortqualität generieren und auch insgesamt den größten Anteil der Fragen ausmachen, generieren am wenigsten Antworten pro Frage (vgl. Abb. 5.16). Die aufgestellte Hypothese, bei Faktenfragen, also bei Fragen mit meist nur einer Antwortmöglichkeit, könnte die Anzahl der Antworten geringer sein, scheint sich hier zu bewahrheiten. Grund hierfür könnte sein, dass sich bei nur einer Antwortmöglichkeit schnell Dopplungen in den Antworten ergeben und die User daher keinen Anreiz für weitere Antworten sehen.

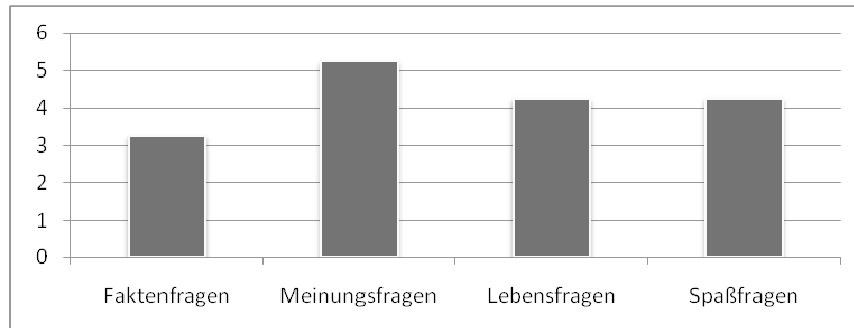


Abb.5.16 Anzahl der Antworten pro Frage-Art

8.4 Bonuspunkte

Stellt man die Anzahl der durchschnittlich ausgesetzten Bonuspunkte pro Frage und die durchschnittliche Antwortqualität für die verschiedenen Themen dar, so zeigt sich, dass die Themen mit hoher Antwortqualität im Schnitt weniger Bonuspunkte pro Frage aufweisen. Die Anzahl der Bonuspunkte verhält sich dabei teilweise sogar antiproportional zur Antwortqualität. Dies zeigt sich vor allem im Themenbereich „Filme“ und wiederum im Themenbereich „Computer & Technisches“ (vgl. Abb. 5.17).

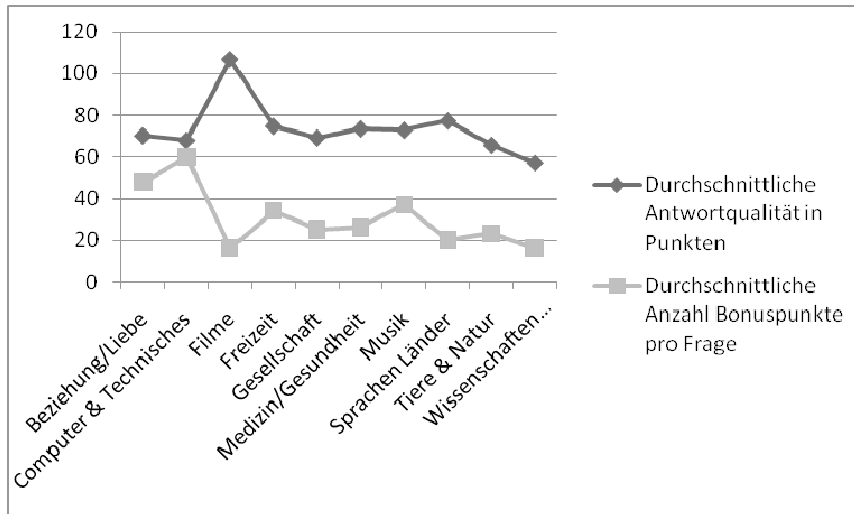


Abb. 5.17 Bonuspunkte für Fragen in Abhängigkeit zur Antwortqualität in den Themenbereichen

Die Anzahl der ausgesetzten Bonuspunkte scheint erstaunlicherweise also keine bzw. negative Auswirkungen auf die Antwortqualität zu haben. Eine mögliche Erklärung hierfür wäre, dass viele Bonuspunkte zu unüberlegten Antworten einladen, mit denen versucht wird, trotz mangelnder Sachkenntnis Punkte für das eigene Konto

zu sammeln. Sollte diese Hypothese zutreffen, müsste sich aber eine Korrelation zwischen der Anzahl der Antworten pro Frage und der Höhe der Bonuspunkte ergeben. Abbildung 5.18 zeigt, dass eine solche Korrelation nicht besteht. Eine andere Erklärung wäre, dass Fragesteller für schwierige Antworten tendenziell mehr Bonuspunkte aussetzen, diese Fragen von den Usern im Durchschnitt aber schlechter beantwortet werden. Hierauf würde auch der deutlich antiproportionale Zusammenhang zwischen Bonuspunkten und Antwortqualität bei Fragen aus dem Bereich „Filme“ hindeuten. Die negative Auswirkung der Bonuspunkte auf die Antwortqualität, bzw. der Zusammenhang zwischen Bonuspunkten und Antwortqualität wäre aber noch näher zu untersuchen.

Deutlicher ist der Zusammenhang zwischen der Anzahl der Bonuspunkte und der Anzahl der positiven Bewertungen, die die Antworten zu einer Frage insgesamt bekommen. Hier führen hohe Bonuspunkte offensichtlich zu vielen Gut-Bewertungen. Die extrem hohe Anzahl der Bonuspunkte im Bereich der Spaßfragen, die für den nicht ernst gemeinten Charakter dieser Fragen spricht, kann den Wert der positiven Bewertungen allerdings nicht in gleichem Maße steigern. Warum dieser Zusammenhang zwischen der Anzahl der Gut-Bewertungen und der Höhe der Bonuspunkte besteht, lässt sich nur spekulieren. Möglich wäre, dass User, die viele Bonuspunkte aussetzen, die erhaltenen Antworten einfach auch aktiver bewerten als User, die auf das Aussetzen von Bonuspunkten weniger Wert legen.

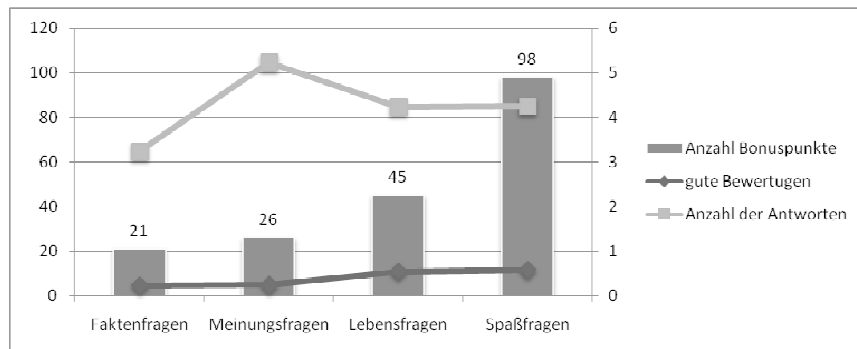


Abb. 5.18 Korrelationen: Bonuspunkte, Gut-Bewertungen, Antwortenanzahl

8.5 Gut-Bewertungen

Neben dem Zusammenhang zwischen Gut-Bewertungen und der Anzahl der Bonuspunkte ist auch der Zusammenhang zwischen der Anzahl der Gut-Bewertungen und der Antwortqualität interessant. Mit Ausnahme der guten Antworten ist die Anzahl der Gut-Bewertungen bei allen Qualitätsstufen recht ähnlich, wenn auch bei sehr schlechten Fragen erwartungsgemäß am niedrigsten (vgl. Abb. 5.18). Trotzdem scheint auch die Bewertung der Antworten durch die User kein geeignetes Kriterium zur Bestimmung guter Antworten zu sein. Dieser Eindruck wird durch Abbildung 5.19 noch verstärkt. Hier zeigt sich, dass sich die Anzahl der guten Bewertungen in Abhängigkeit von der Art der Frage eher antiproportional zur tatsächlichen Antwortqualität verhält. Dieses Ergebnis wird dabei natürlich durch den zu Grunde gelegten Kriterienkatalog ermöglicht, der die Qualität der Antworten nicht nur in

Abhängigkeit von der User-Bewertung, sondern vor allem auf Grund einer durch unabhängige Probanden vorgenommenen Bewertung einordnet.

Eine gute Bewertung ist also kein geeignetes Indiz für eine gute Antwort. Im Gegenteil scheint eine gute Bewertung vor allem bei den Faktenfragen und den Meinungsfragen eher auf eine geringere Qualität der Antworten hinzuweisen. Dieser Zusammenhang ist auf Grund der gänzlich anderen Erwartungshaltung besonders überraschend.

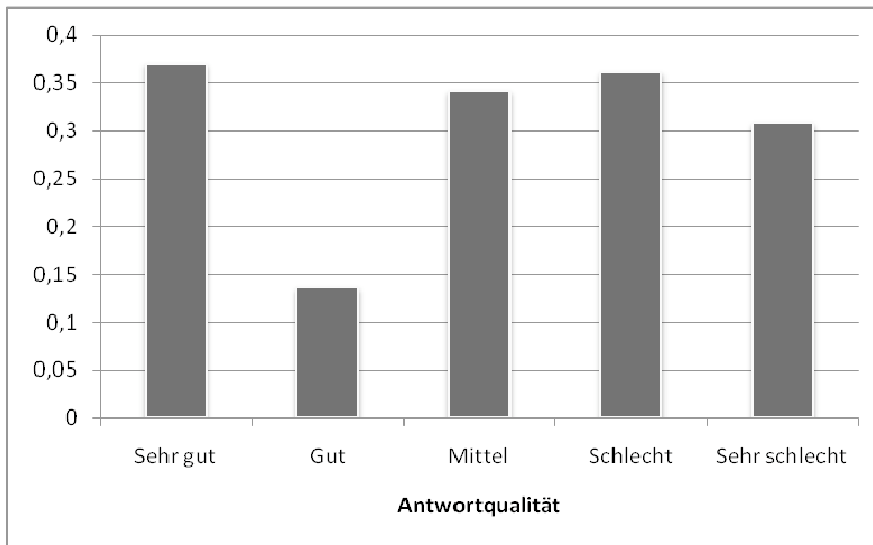


Abb. 5.18 Anzahl der Gut-Bewertungen pro Frage in Abhängigkeit von der Antwortqualität, n=247

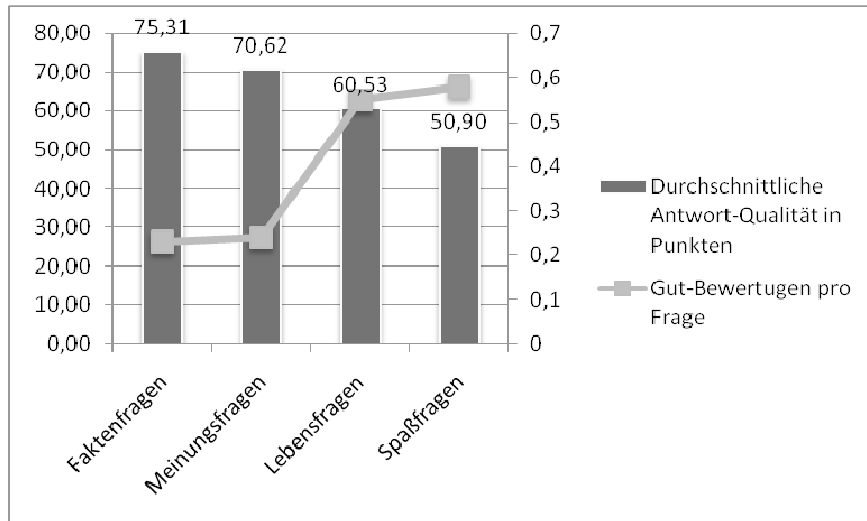


Abb. 5.19 Anzahl der Gut-Bewertungen und Antwort-Qualität in Abhängigkeit von der Art der Frage, n=247

8.6 Formale Merkmale

Stellt man die Qualität der formalen Merkmale in Abhängigkeit von der Antwortqualität dar, so lässt sich kein genereller Zusammenhang feststellen (vgl. Abb. 5.20). Die Abbildung stellt dabei die anteiligen Fragen dar, die für das jeweilige formale Merkmal positiv bewertet wurden.

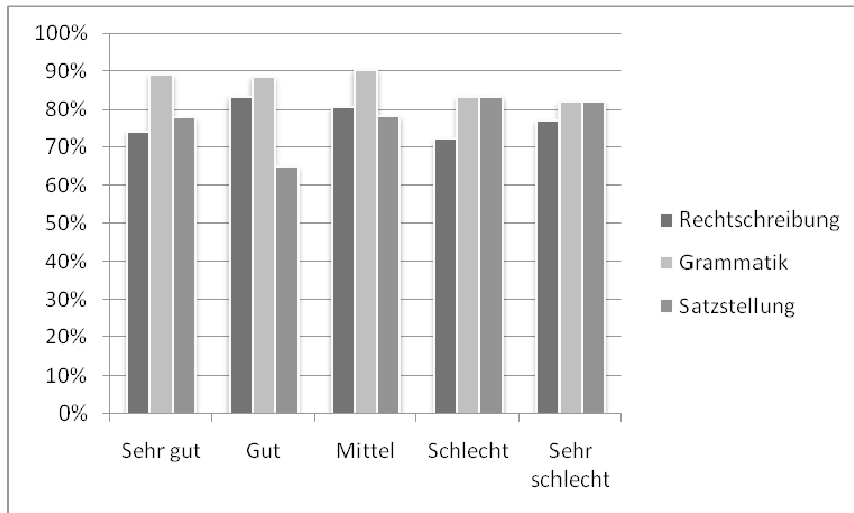


Abb. 5.20 Formale Merkmale in Abhängigkeit zur Antwortqualität

Tendenziell erscheint es zwar so, dass eine gute Grammatik zumindest eine schlechte Antwort unwahrscheinlicher macht, ein Kriterium zur Beurteilung der Qualität einer Frage lässt sich so aber sicherlich nicht ableiten. Auch die Rechtschreibung hat offensichtlich keinen Einfluss auf die Antwortqualität.

Allerdings lässt sich ein Zusammenhang zwischen der formalen Qualität einer Frage als Ganzes und der Antwortqualität in Abhängigkeit vom Thema der Frage zumindest ansatzweise erkennen (vgl. Abb. 5.21). Dies wird vor allem in den Bereichen „Filme“ und „Sprachen/Länder“ deutlich. Hier wirken sich gute formale Merkmale also positiv auf die Antwortqualität aus.

Der insgesamt geringe Einfluss der formalen Gestaltung der Frage auf die Qualität der Antworten ist bei genauerer Überlegung aber auch nicht weiter verwunderlich. Es ist davon auszugehen, dass ein User, wenn er die Frage verstanden hat, bei einer formal schlecht gestalteten Frage genauso „klug“ und richtig antworten dürfte, wie bei einer formal sehr gut gestalteten Frage.

Weiterhin ist festzustellen, dass die formalen Merkmale bei Fragen aus dem Bereich „Sprachen/Länder“ deutlich besser sind als zum Beispiel im Bereich „Computer & Technisches“.

Interessant ist auch der relativ deutliche Zusammenhang zwischen der Qualität der formalen Merkmale und der Anzahl der Antworten. Offensichtlich generieren gut gestaltete Fragen mehr Antworten als Fragen mit formalen Mängeln. Dies impliziert zwar keine Aussage über die Qualität, da eine hohe Anzahl von Antworten aber

sicherlich recht positiv auf den fragenden User wirkt, ist dieses Kriterium für ein späteres Tutorial (vgl. Kap. 10) von Bedeutung.

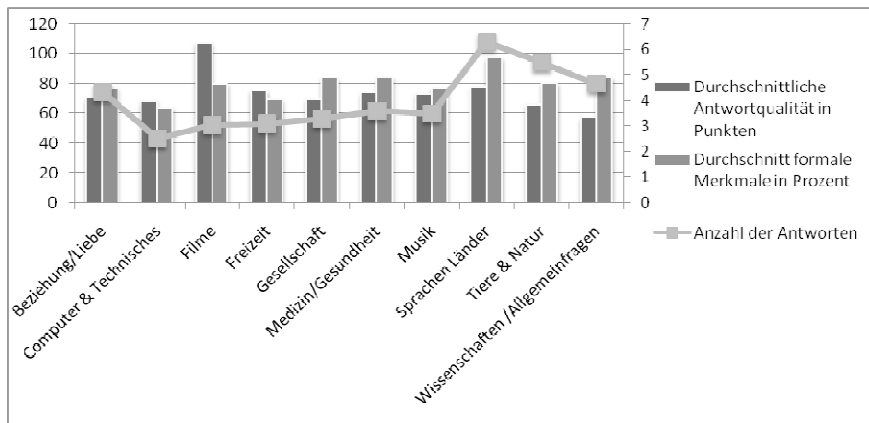


Abb. 5.21 Formale Merkmale der Frage und deren Antwortqualität in Abhängigkeit vom Themengebiet

Positiv anzumerken ist insgesamt die durchgehend hohe formale Qualität der Fragen, die in einer Online-Community, denkt man an den Sprachstil in vielen Internet-Foren, so nicht zu erwarten war.

8.7 Inhaltliche Merkmale

Welche inhaltlichen Kriterien sind wesentlich für eine gute Frage? Diese Frage versuchen wir mit der folgenden Auswertung zu beantworten (vgl. Abb. 5.22). Einige Rückschlüsse lassen sich hier auf den ersten Blick feststellen. Zum einen senkt der Verzicht auf eine provokative Wortwahl die Gefahr einer schlechten Antwort deutlich, zum anderen sind fast alle Fragen mit sehr guten Antworten in sich logisch und ein Großteil der Fragen mit schlechter Antwortqualität lassen eben diese Logik vermissen.

Der Anteil der Fragen mit mehr als einer Antwort-Möglichkeit ist bei den sehr schlechten Antworten deutlich höher als bei den übrigen Qualitätsstufen.

Die anderen untersuchten Kriterien scheinen in keinem direkten Zusammenhang mit der generierten Antwortqualität zu stehen. Erstaunlicherweise unterscheiden sich weder der Anteil der Fremd- und Fachwörter noch der Anteil der zeitlich begrenzten Fragen bei Antworten unterschiedlicher Qualität. Wie auch schon bei den formalen Kriterien scheint die Antwortqualität von der Gestaltung der Frage kaum abzuhängen. Nur inhaltliche Merkmale, die das Verstehen der Frage erschweren oder zu einer schlecht bewerteten Antwort provozieren, wirken sich auch auf die Qualität der Antworten aus.

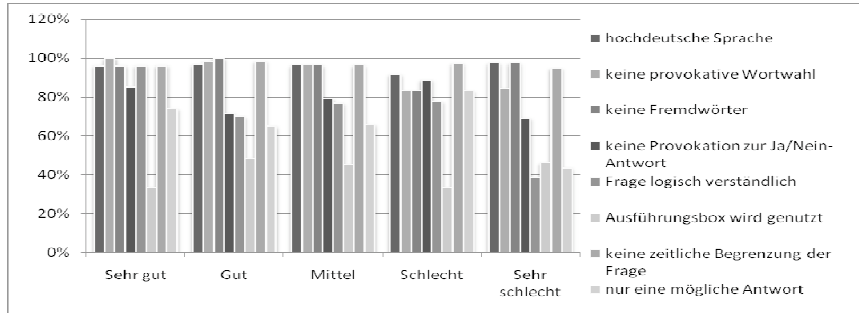


Abb. 5.22 Anteil der Fragen in Abhängigkeit zur Antwortqualität, die das jeweilige Kriterium erfüllt

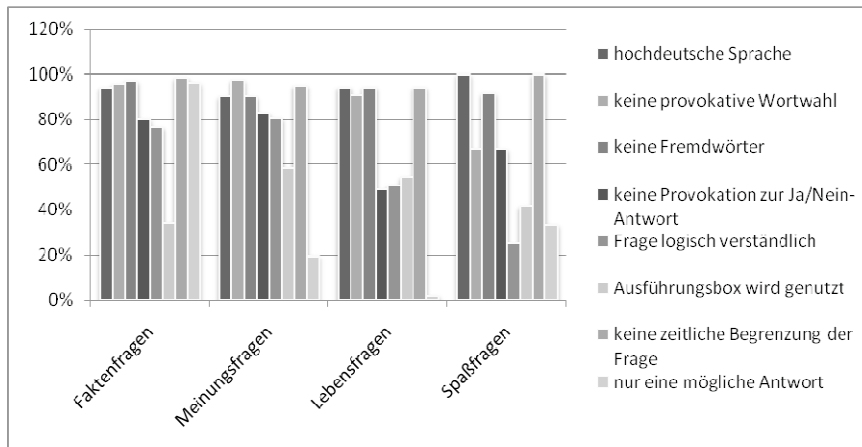


Abb. 5.23 Anteil der Fragen in Abhängigkeit zur Art der Frage, die das jeweilige Kriterium erfüllt

Ordnet man die Ausprägung der unterschiedlichen Kriterien den verschiedenen Fragearten zu, so bestätigt sich erneut die hohe Qualität der Faktenfragen (vgl. Abb. 5.23). Auffällig ist hier auch noch einmal, dass fast ausschließlich Faktenfragen nur eine Antwort zulassen. Zu dieser Bestimmtheit der Frageart passt auch, dass die Ausführungsbox bei den Faktenfragen am wenigsten gebraucht wird. Auch der geringe Anteil von Fragen ohne provokative Wortwahl bei den Spaßfragen passt genau ins Bild/erfüllt die Erwartungen (aber warum?).

9. EINBINDEN VON FRAGEN AUS LYCOS IQ IN ANDERE SUCHFORMEN

Anfänglich haben wir die Frage gestellt, inwieweit sich Fragen aus Lycos iQ zur Ergänzung anderer Suchformen eignen. Nach Abschluss der Untersuchung ist festzustellen, dass sich Spaßfragen und eventuell auch Meinungsfragen auf Grund der niedrigeren Qualität nicht für eine Einbindung in andere Suchformen eignen.

Faktenfragen und Lebensfragen sind zwar von leicht unterschiedlicher Qualität, weisen aber insgesamt eine höhere Qualität auf. Generell stellt sich aber die Frage, wie man die Frageart automatisch ermitteln soll, so dass dies nicht einziges Unterscheidungsmerkmal sein kann. Hier sind technische Grenzen gesetzt, die eine komplette Unterscheidung nach der Frageart verhindern. Trotzdem ist es sinnvoll, bevorzugt auf Faktenfragen zurück zu greifen. Eine Möglichkeit diese zu identifizieren wäre zum Beispiel das Anlegen von Wortlisten. Typisch für Faktenfragen sind zum Beispiel Fragewörter wie „Wo“, „Wann“ und „Wer“, während „Wie“ eher bei Lebensfragen auftauchen dürfte. Mit entsprechender Analyse könnte man eine solche Liste erweitern und die Fragen einigermaßen nach ihrer Art filtern.

Hier stellt sich aber auch die Frage, wie interessant Faktenfragen als Bereicherung einer herkömmlichen Suche sind, da sich diese meist zufriedenstellend über algorithmische Suchen beantworten lassen. Interessanter für den User wäre daher vielleicht gerade eine Auswahl der kontrovers diskutierten Fragen und der persönlichen Meinungen, um eine andere Art der Fragen-Beantwortung anzubieten. Hier müsste man aber eine etwas geringere Qualität der Fragen in Kauf nehmen. Dies stellt eine Grundsatzentscheidung dar, die von uns so nicht beantwortet werden kann.

Ein weiteres Kriterium ist das Thema der Frage. Wir haben gezeigt, dass die Antwortqualität in Themen von allgemeinem, freizeitleichem Interesse am besten ist. Dementsprechend sollten bevorzugt Fragen aus diesen Bereichen verwendet werden, während Bereiche, die ein spezielles Fachwissen erfordern, ausgeklammert werden sollten. Passend dazu haben wir gezeigt, dass Fragen, die mit normalem Allgemeinwissen zu beantworten sind, tendenziell die etwas besseren Antworten generieren. Da dies aber zwischen den einzelnen Fragearten und Themen variiert, lässt sich hieraus kein wirklich geeignetes Kriterium ableiten.

In dem Umstand, dass Fragen mit geringem benötigten Fachwissen die deutlich besseren Antworten generieren, liegt allerdings auch ein Problem, da die Lycos iQ-Antworten dann nicht zur Lösung komplizierterer Fragen beitragen können.

Inhaltlich kann gesagt werden, dass geeignete Fragen keine provokative Wortwahl enthalten und zudem logisch aufgebaut sein sollten. Vor allem die Wortwahl könnte anhand weiterer Stoppwortlisten, die die Fragen nach beleidigenden Ausdrücken oder bestimmten Redewendungen durchsuchen, überprüft werden.

Weitere Kriterien, die eine Frage disqualifizieren, sind eher Merkmale, die sich direkt auf der Lycos iQ-Plattform ablesen lassen und die sich daher gut zu einer automatischen Selektion der Fragen eignen. Ein Indiz für qualitativ schlechte Fragen sind so interessanter Weise nach unserer Untersuchung eine hohe Anzahl positiver Bewertungen und eine hohe Anzahl ausgesetzter Bonuspunkte, wobei die Anzahl der positiven Bewertungen in Teilen von der Anzahl der Bonuspunkte abhängt. Die Korrelation zwischen der Anzahl der ausgesetzten Bonuspunkte und der Qualität der Antworten sowie der Anzahl der positiven Bewertungen und der Qualität der Antworten ist also tendenziell negativ.

Zusammengefasst kann also gesagt werden, dass eine Frage, die sich zur Ergänzung einer herkömmlichen Suche eignet, aus einem Themengebiet stammen sollte, das kein spezielles Interesse und Fachwissen erfordert. Außerdem sollten Fragen ausgeklammert werden, die Worte und Formulierungen enthalten, die typisch für Provokationen und Beleidigungen sind. Letztendlich bieten sich eher Fragen mit

einer niedrigen Anzahl positiv bewerteter Antworten und wenigen Bonuspunkten an, da diese tendenziell qualitativ bessere Antworten generieren.

10. TUTORIAL

Bei der Entwicklung eines Tutorials muss man bedenken, dass man einen User beim Einstellen einer Frage eventuell etwas lenken, ihm die Art und das Thema der Frage aber nicht vorschreiben kann. Schließlich soll jede Frage zulässig sein. Einige Kriterien, die die Qualität von Fragen und Antworten verbessern können, bieten sich aber dennoch an, um in Form einer kleinen Hilfe oder eines Tutorials präsentiert zu werden. Folgende Formulierungen wären dabei denkbar:

- Achten Sie auf eine ordentliche formale Gestaltung Ihrer Fragen. Formulieren Sie die Frage im ganzen Satz und achten Sie auf die Rechtschreibung. So erhöhen Sie die Wahrscheinlichkeit, dass Sie viele hilfreiche Antworten auf Ihre Frage erhalten.
- Verzichten Sie auf jegliche provokative Wortwahl oder einen provokativen Charakter ihrer Frage. Beleidigungen oder Angriffe auf Personen oder Gruppen sind nicht nur unerwünscht, sondern führen auch eher zu schlechten Antworten.
- Formulieren Sie Ihre Frage logisch. Fragen ohne innere Logik werden deutlich öfter sehr schlecht beantwortet als logisch formulierte Fragen.
- Achten Sie darauf, dass der Sinn Ihrer Frage den anderen Mitgliedern der Community sofort klar wird.
- Seien Sie sparsam mit Bonuspunkten. Auch bei besonders schwierigen Fragen führt eine hohe Anzahl Bonuspunkte nicht zu einer besseren Antwort.

Jeder Tipp sollte außerdem durch ein entsprechendes, aussagekräftiges Beispiel veranschaulicht werden, an dem der Fragesteller sich orientieren kann. Ideal wäre hier die Anlage eines Tutorials als kleine Videosequenz.

11. FAZIT

Im Rahmen der Analyse konnten wir allgemeine Zusammenhänge zwischen Fragen und Antworten aufzeigen. Einige Korrelationen waren dabei recht deutlich, andere eher schwach und einige der erhofften Zusammenhänge konnten so auch nicht belegt werden. Vor allem Zusammenhänge zwischen inhaltlichen und formalen Kriterien und der Antwortqualität ließen sich nur schwer aufzeigen. Überraschend ist, dass sich die formalen Kriterien zwar insgesamt schwach auf die Qualität der Antworten auswirken, schlechte Formalien die Lycos iQ-User aber nicht wirklich vom Beantworten einer Frage abzuhalten scheinen. Dafür entspricht das in allen Bereichen gute Abschneiden der Faktenfragen unseren Erwartungen.

Ein interessanter Ansatzpunkt für weitere Untersuchungen bietet die Frage, warum die Antwortqualität bei Fragen, die kaum Fachwissen benötigen, am besten ist. Das populäre Schlagwort „Wisdom of the crowd“ im Gegenteil zu dem Wissen einzelner Experten scheint hier eine Rolle zu spielen. Herauszufinden wäre also, ob einfache Fragen aus Themen, die viele Menschen interessant finden, von einer Menge

besser beantwortet werden können als von einzelnen Experten und ob umgekehrt einzelne Experten mit spezielleren Fragen besser umgehen können als die Masse.

Durch eine deutliche Erhöhung des Stichprobenumfangs und eine noch feinere Aufspaltung der einzelnen Kriterien und Themenkategorien könnte eine weiterführende Untersuchung einige Ergebnisse sicherlich noch vertiefend darstellen. Diese Analyse ist daher eher als erster Überblick und als Anreiz für tiefere, speziellere Untersuchungen zu verstehen. Trotzdem hoffen wir, dass die Ergebnisse sich bei zukünftigen Entscheidungen und Verbesserungen an der Lycos iQ-Plattform als nützlich erweisen könnten.