

DIRK LEWANDOWSKI

Informationskompetenz und das Potenzial der Internetsuchmaschinen

Preprint eines Buchkapitels, erscheint in:

Sühl-Strohmenger, W. (Hrsg.): Handbuch Informationskompetenz. Berlin: de Gruyter Saur, 2012

<http://www.degruyter.de/cont/fb/bb/detail.cfm?isbn=978-3-11-025473-0>

Schlüsselbegriffe

Suchmaschinen, Informationskompetenz, Google, Informationsverhalten

Abstract

Suchmaschinen sind für viele Nutzer *der* Zugang zu Informationen im Netz und oft fehlt ihnen die Kenntnis anderer Suchwerkzeuge. In diesem Kapitel wird das Thema Informationskompetenz anhand des Prozesses der Web-Suche (Auswahl des Suchwerkzeugs, Eingabe der Suchanfrage, Trefferselektion, Evaluierung des Treffers, Navigation/Suche innerhalb der Ergebnisquelle) sowie anhand des technischen Aufbaus von Suchmaschinen diskutiert.

Für welche Web-Recherchen braucht man Informationskompetenz?

Die Nutzung von Suchmaschinen setzt Informationskompetenz voraus, auch wenn es auf den ersten Blick nicht so scheinen mag. Was könnte einfacher sein, als eine Suchmaschine zu bedienen: Google aufrufen, Suchwörter eingeben, auf ein Ergebnis klicken ...

Vor allem dieser einfache Dreischritt der Recherche und die hohe Erfolgsquote machen es überhaupt nötig zu betonen, dass die Nutzung von Suchmaschinen Informationskompetenz voraussetzt: Im Allgemeinen sind Nutzer mit ihren Suchergebnissen zufrieden. Auch einige Schwierigkeiten im Verhältnis zwischen Information Professionals und Suchmaschinennutzern lassen sich aus diesem Umstand erklären: Es fällt den Information Professionals schwer zu verstehen, dass sich normale Nutzer nicht für Recherchestrategien interessieren; und die zufriedenen Suchmaschinennutzer sehen keinen Grund dafür, warum sie sich mit der Verbesserung ihrer (vermeintlich ausreichenden) Recherchefähigkeiten beschäftigen sollten.

Doch woraus resultiert diese Zufriedenheit mit den Suchergebnissen eigentlich? Um diese Frage zu beantworten, lohnt ein Blick auf die Verteilung der Suchanfragen nach Anfragetypen. Unter Anfragetypen versteht man vor allem das intendierte Ziel einer

Suchanfrage nach Zieldokument; allgemein wird nach Broder¹ in informationsorientierte, navigationsorientierte und transaktionsorientierte Suchanfragen unterschieden.

Mit navigationsorientierten Anfragen soll eine Seite (wieder)gefunden werden, die dem Benutzer bereits bekannt ist oder von der er annimmt, dass sie existiert. Beispiele sind die Suche nach Homepages von Unternehmen („Microsoft“) oder nach Personen („John von Neumann“). Solche Anfragen haben in der Regel ein richtiges Ergebnis. Das Informationsbedürfnis ist befriedigt, sobald die gewünschte Seite gefunden wird.

Bei informationsorientierten Anfragen ist das Informationsbedürfnis meist nicht durch ein einziges Dokument zu befriedigen. Der Nutzer möchte sich stattdessen über ein Thema informieren und liest deshalb mehrere Dokumente. Informationsorientierte Anfragen zielen auf jeden Fall auf statische Dokumente, nach dem Aufruf des Dokuments ist also keine weitere Interaktion auf der Website nötig, um an die gewünschten Informationen zu gelangen. Mit transaktionsorientierten Anfragen wird eine Website gesucht, auf der anschließend eine Transaktion stattfindet, etwa der Kauf eines Produkts, der Download einer Datei oder die Recherche in einer Datenbank.

Nun werden Suchmaschinen für alle drei Anfragetypen eingesetzt. Sieht man sich die Verteilung der Anfragetypen an, so zeigt sich, dass alle drei Typen nennenswerte Anteile am Gesamt der Suchanfragen haben. Für navigationsorientierte Anfragen sind eindeutig bestimmbare richtige Ergebnisse bekannt, und für Suchmaschinen bereitet es keine allzu großen Probleme, diese Anfragen korrekt zu beantworten². Da auch für die transaktionsorientierten Anfragen leicht festgestellt werden kann, ob das gewünschte Ziel der Transaktion mit Hilfe des von der Suchmaschine ausgegebenen Ergebnisses erreicht werden kann, stellt sich die Frage nach einer Bewertung des Suchergebnisses im klassischen Sinn nur für die informationsorientierten Anfragen. Nur hier muss zwischen verschiedenen, das Informationsbedürfnis unterschiedlich befriedigenden Dokumenten unterschieden werden. Aber auch bei den informationsorientierten Anfragen muss weiter eingeschränkt werden: Auch hier lassen sich viele Informationsbedürfnisse mit einem einzigen, vom Nutzer als „gut genug“ empfundenen Dokument befriedigen.

¹ Broder, Andrei: A taxonomy of web search. In: ACM Sigir forum 36 (2002). S. 3–10.

² Lewandowski, Dirk: The retrieval effectiveness of search engines on navigational queries. In: ASLIB Proceedings 61 (2011). S. 354-363.

Wenn man nun von einem Anteil informationsorientierter Anfragen von 45 Prozent ausgeht³ und dann noch die Anfragen abzieht, die aus Nutzersicht mit einem einzigen Dokument beantwortet werden können, so ergibt sich eine Erklärung dafür, warum Nutzer mit ihren Suchen in der Regel zufrieden sind und keine Notwendigkeit sehen, ihre Recherchekenntnisse anzuzweifeln oder gar nach einer Verbesserung derselben zu streben.

Festzuhalten ist also, dass Informationskompetenz im Sinne einer Recherchestrategie nur in einer Minderzahl der Web-Recherchen nötig ist. Ein Nutzer, der nur die Homepage eines Unternehmens oder ein Faktum sucht, wird auch, ohne sich Gedanken über eine Suchstrategie zu machen, zum Erfolg kommen. Für die Vermittlung von Informationskompetenz ist es daher von großer Bedeutung, diejenigen Fälle herauszuarbeiten, in denen tatsächlich strategisch an eine Web-Recherche herangegangen werden muss.

Die Frage nach der Informationskompetenz im Bereich der Internetsuchmaschinen stellt sich auf zweierlei Weise: Auf der einen Seite geht es um die Kompetenzen, die erworben werden können, um den eigenen Suchprozess zu optimieren; auf der anderen Seite geht es um die Kenntnis des Aufbaus und der technischen Funktionsweise der Web-Suchmaschinen. Basis der Diskussion sind die von Benno Homann⁴ beschriebenen Standards der Informationskompetenz:

„Der informationskompetente Student

1. bestimmt Art und Umfang der benötigten Informationen.
2. verschafft sich effizienten und effektiven Zugang zu den benötigten Informationen
3. evaluiert Informationen und seine Quellen kritisch und integriert die aus- gewählten Informationen in sein Wissen und sein Wertesystem
4. nützt Informationen effektiv sowohl als Individuum als auch als Gruppenmitglied, um ein bestimmtes Ziel zu erreichen
5. versteht viele der ökonomischen, rechtlichen und sozialen Streitfragen, die mit der Nutzung von Informationen zusammenhängen und er hat Zugang und nutzt die Informationen in einer ethischen und legalen Weise.“

Bezogen aus Suchmaschinen sind vor allem die ersten drei Punkte zu diskutieren. Die beiden letzten Aussagen beziehen sich auf das allgemeine Umfeld und sind nicht suchmaschinenspezifisch auszulegen.

³ Lewandowski, Dirk: Query types and search topics of German Web search engine users. In: Information Services & Use 26 (2006). S. 261-269.

⁴ Homann, Benno: Standards der Informationskompetenz. In: Bibliotheksdienst (2002), 36(5), 625–638.

Informationskompetenz im Prozess der Websuche

Der Prozess der Websuche (Abb. 1) beginnt mit der Auswahl einer geeigneten Suchmaschine (1), darauf folgt die Eingabe der Suchanfrage (2). Die Ergebnisse auf der Trefferseite der Suchmaschine (Search Engine Results Pages, SERP) werden mehr oder weniger gründlich durchgesehen und ein geeigneter Treffer ausgewählt (3). Nach dem Aufruf des Treffers erfolgt dessen Evaluation auf Tauglichkeit (4), eventuell folgt eine weitere Navigation oder Suche innerhalb der Website, auf der der Treffer gefunden wurde (5).

Im Schaubild wird deutlich, dass der Prozesse der Websuche nicht geradlinig verlaufen muss, sondern dass es an verschiedenen Stellen zum Abbruch oder zur Neuorientierung kommen kann.

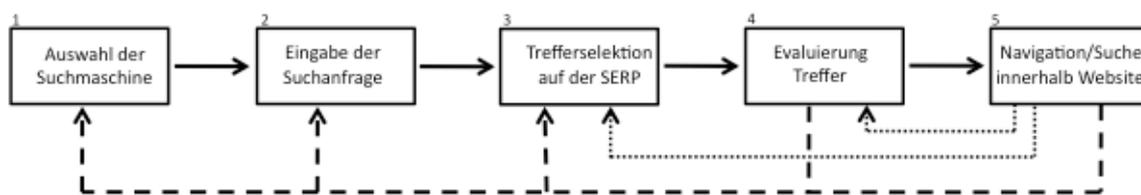


Abbildung 1: Prozess der Websuche

Anhand der Stufen des Prozesses der Websuche soll im Folgenden die für die effektive und effiziente Nutzung der Suchmaschinen nötige Informationskompetenz diskutiert werden.

Dabei wird das allgemeine Nutzerverhalten skizziert, um dann aufzuzeigen, welche Informationskompetenz auf der jeweiligen Stufe erforderlich wäre.

Schon auf der ersten Stufe, der Auswahl einer geeigneten Suchmaschine, zeigt sich erhebliches Potential. Nicht nur wird von vielen Nutzern die Suche im Web mit einer Suche in der Suchmaschine Google gleichgesetzt, sondern Google wird sogar als der geeignete Einstiegspunkt für jegliche Art von Recherche angesehen.⁵ Schon zwischen den allgemeinen Web-Suchmaschinen (Universalsuchmaschinen) gibt es eine größere Auswahl, die allerdings von den überwältigenden Marktanteilen von Google überdeckt wird. Die alleinige Verwendung einer Suchmaschine wird oft mit der (scheinbar) überragenden Qualität der Ergebnisse begründet. Allerdings erfüllt die Konsultation einer weiteren Suchmaschine zumindest noch einen weiteren Zweck, nämlich das Einholen einer „zweiten Meinung“. Die

⁵ Zur Bedeutung von Suchmaschinen für den Wissenserwerb s.a.: Lewandowski, Dirk: Informationen finden im Netz. In: UNESCO heute (1/2008).

Suchmaschinen unterscheiden sich nicht nur in ihren Datenbeständen, sondern gravierender noch hinsichtlich ihrer Ergebnisreihung.⁶ So entsteht durch die Recherche in einer weiteren Suchmaschine eine größere Vielfalt der Ergebnisse. Außerdem bedeutet selbst eine überlegene Qualität der Treffer im Vergleich zu anderen Suchmaschinen nicht, dass auch für jede Anfrage das bestmögliche Ergebnis gefunden wird.⁷ So lässt sich manche vermeintlich schwierige Recherche lösen, indem einfach die gleiche Suchanfrage bei einer anderen Suchmaschine eingegeben wird. Denkt man aber an spezifischere Recherchen, so ist die Auswahl einer geeigneten Spezialsuchmaschine⁸ – hier soll noch gar nicht von weiterführenden Fachdatenbanken gesprochen werden – nötig.

1. Eine kompetente Quellenauswahl muss also anhand zweier Fragen erfolgen: Welcher Typ von Suchwerkzeug ist für die Recherche geeignet? Welche individuelle Suchmaschine ist für die Recherche geeignet?
2. Die Eingabe der Suchanfrage erfordert vom Nutzer die Kompetenz, sein Informationsbedürfnis in eine geeignete Suchanfrage zu „übersetzen“. Auf der einen Seite sind Informationsbedürfnisse jedoch oft nicht klar definiert oder können sich im Verlauf einer Recherche verändern⁹, andererseits fällt es oft schwer, das Informationsbedürfnis in einer Suchanfrage auszudrücken. Studien zeigen, dass die in die Web-Suchmaschinen eingegebenen Suchanfragen in der Regel kurz sind und oftmals zu unspezifisch, um zu passenden Ergebnissen zu führen.¹⁰ Die Suchmaschinen haben auf dieses Verhalten reagiert, indem sie die Suchanfragen in den Kontext des Suchprozesses stellen und durch diese zusätzlich gewonnenen Informationen zu besseren Ergebnissen leiten.¹¹ Das kompetente Formulieren von Suchanfragen will gelernt sein. Allerdings sollte hier nicht vergessen werden, dass bei vielen Informationsbedürfnissen auch einfache Suchanfragen zum Ziel führen. Im Rahmen von Schulungen sollte deshalb klar gemacht werden, in welchen Fällen sich komplexere Suchanfragen überhaupt eignen und dies entsprechend durch Beispiele belegt werden.

⁶ Spink, Amanda, Bernard J. Jansen, Chris Blakely u. Sherry Koshman: A study of results overlap and uniqueness among major web search engines. In: *Information Processing & Management* 45 (2006). S. 1379–1391; Bar-Ilan, Judit: Comparing rankings of search results on the Web. In: *Information Processing & Management* 41 (2005). S. 1511-1519.

⁷ Lewandowski, Dirk: The retrieval effectiveness of web search engines: considering results descriptions. In: *Journal of Documentation* 64 (2008). S. 915-937.

⁸ Lewandowski, Dirk: Spezialsuchmaschinen. In: Lewandowski, Dirk (Hrsg.), *Handbuch Internet-Suchmaschinen*. Heidelberg: AKA 2009, S. 53-69.

⁹ Marchionini, Gary: Exploratory search: from finding to understanding. In: *Communications of the ACM* 49 (2006). S. 41–46. (ACM.)

¹⁰ Höchstötter, Nadine u. Martina Koch: Standard parameters for searching behaviour in search engines and their empirical evaluation. In: *Journal of Information Science* 35 (2009). S. 45.

¹¹ Lewandowski, Dirk: Query Understanding. In Lewandowski, Dirk (Hrsg.), *Handbuch Internet-Suchmaschinen 2: Neue Entwicklungen in der Web-Suche*. Heidelberg: Akademische Verlagsgesellschaft AKA 2011, S. 55-75.

3. Zur Trefferselektion werden dem Nutzer Surrogate („Abstracts“) der eigentlichen Treffer zur Verfügung gestellt, welche üblicherweise aus dem Titel des Dokuments, einer kurzen Beschreibung sowie der URL des Dokuments bestehen. Weitere Informationen werden von den Suchmaschinen teilweise hinzugefügt.

Während die Surrogate einen Einfluss auf die Selektion innerhalb der Trefferdarstellung haben, ist diese jedoch vor allem durch die Position der Treffer beeinflusst. Besonders weit vorne stehende Treffer, insbesondere die im sichtbaren Bereich (also Treffer, die sichtbar sind, ohne dass heruntergescrollt werden müsste¹²), werden bevorzugt, außerdem besonders hervorgehobene Treffer.

Eine kompetente Trefferauswahl setzt voraus, dass nicht nur einer der ersten Treffer angeklickt und die Recherche damit als beendet betrachtet wird. Allerdings ist auch hier zu beachten, dass ein Informationsbedürfnis (auch wenn es sich um eine informationsorientierte Suchanfrage handelt) oft schon durch ein einziges Dokument befriedigt werden kann. In Schulungen sollte klar gemacht werden, in welchen Fällen tatsächlich eine Sichtung mehrerer Dokumente oder die Sichtung mehrerer Ergebnisseiten der Suchmaschinen sinnvoll ist.

Der Interpretation und Evaluierung der Suchergebnisseite (hier im Gegensatz zu den eigentlichen Trefferdokumenten) sollte besondere Aufmerksamkeit entgegengebracht werden. Nutzer sollten erstens grundsätzliche Kenntnisse über das Zustandekommen des Rankings haben (s.u.), zweitens in der Lage sein, die in den Surrogaten enthaltenen „Vorab-Informationen“ zu interpretieren, und drittens um die Möglichkeiten der Manipulation von Trefferseiten (sei es durch Suchmaschinenbetreiber oder externe Anbieter durch Methoden der Suchmaschinenoptimierung) wissen.

4. Die Evaluierung eines individuellen Ergebnisses erfolgt im Durchschnitt innerhalb weniger Sekunden. Nutzer blicken auf das Dokument und lassen sich vom „information scent“ leiten, d.h. sie suchen nach Anhaltspunkten dafür, dass sich in diesem Dokument für sie relevante Informationen finden. Lassen sich solche Anhaltspunkte nicht finden, kehren sie zur Trefferliste der Suchmaschine zurück und wählen ein anderes Dokument aus. Eine Evaluierung der Ergebnisse nach Kriterien wie Verlässlichkeit oder faktischer Korrektheit erfolgt meist nicht, in dieser Hinsicht vertrauen die Nutzer der „Empfehlung“ durch die vermittelnde Suchmaschine.

Nötig ist hier die Kompetenz zur Bewertung des individuellen Suchergebnisses und der Quelle (Website), in der dieses enthalten ist. Auch hier ist darauf hinzuweisen, dass die

¹² Siehe dazu: Höchstötter, Nadine u. Dirk Lewandowski: What users see – Structures in search engine results pages. In: Information Sciences 179 (2009). S. 1796-1812.

Fakten- und Quellenprüfung nicht in allen Fällen durchgeführt werden muss, und es sind entsprechende Fälle zu beschreiben.

5. Wird in dem Ergebnisdokument die gesuchte Information nicht gefunden, wird häufig auf der Website weiter navigiert bzw. die site-interne Suche verwendet. Wird die Suche verwendet, ergeben sich die gleichen Fragen wie unter 2. genannt; wird innerhalb der Website navigiert, muss der Nutzer in der Lage sein, die grundlegende Navigation zu verstehen. Nötig ist eine Vermittlung grundlegender Konzepte der Site-Navigation sowie die Kenntnis der Stärken und Schwächen site-interner Suchmaschinen.

Kenntnisse über den Aufbau und die Funktionsweise von Suchmaschinen

Nicht in den Suchprozess mit einbezogen ist in dieser Darstellung die eigentliche Verarbeitung der Suchanfrage und die Produktion des Ergebnisses durch die Suchmaschine (beides zwischen Schritt 2 und Schritt 3). Hierbei handelt es sich um die „black box“ Suchmaschine: Es ist für den Nutzer nicht transparent, was in diesem Schritt „in der Suchmaschine“ passiert. Um allerdings Suchstrategien entwickeln und Ergebnisse adäquat bewerten zu können, ist eine grundlegende Kenntnis der Funktionsweise von Suchmaschinen nötig.

Für Nutzer ist die Kenntnis des grundlegenden Aufbaus einer Web-Suchmaschine (Abb. 2) wichtig, da nur aus dieser Kenntnis die Stärken und Schwächen von Suchmaschinen erkannt werden können. Insbesondere ist hier auf die Erstellung des Index hinzuweisen; dieser wird durch das Crawling erstellt, wodurch nicht-crawlbare Dokumente (vor allem Inhalte von Datenbanken) nicht erfasst werden können (das sog. „Invisible Web“), andererseits können die Crawler nicht jedes Dokument stets aktuell erfassen. Dies führt dazu, dass aktuelle Dokumente teils nicht über die Web-Suchmaschinen zu finden sind und daher direkt die jeweilige Quelle (welche wiederum über eine Suchmaschine recherchiert werden kann) aufgesucht werden muss.

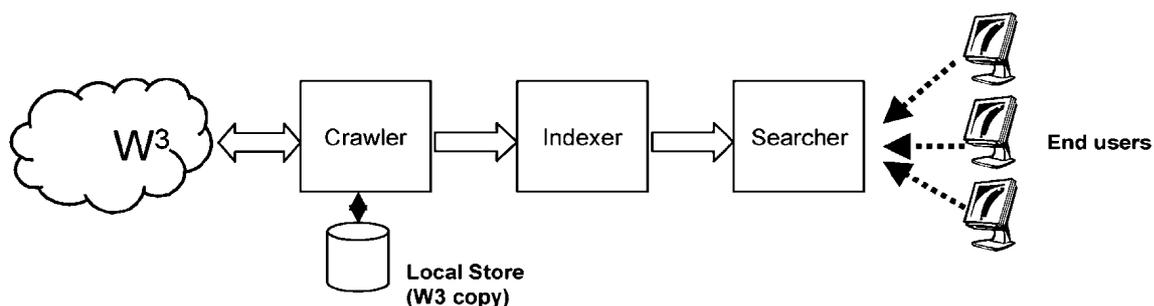


Abbildung 2: Aufbau einer Web-Suchmaschine¹³

Im Prozess der Indexierung werden die Dokumente so aufbereitet, dass sie über die Benutzerschnittstelle gefunden werden können. Informationen, die in diesem Prozess nicht berücksichtigt werden (beispielsweise vom Betreiber einer Website erstellte, umfangreiche Metadaten), können auch nicht recherchiert werden.

Von besonderer Bedeutung für das Verständnis der Suchmaschinen sind grundlegende Kenntnisse, wie Suchmaschinen die Dokumente in der Trefferliste anordnen (Ranking). Damit ist nicht gemeint, dass Nutzer die Rankingverfahren im Detail kennen müssen, sondern nur, dass sie verstehen, welche Art von Dokumenten besonders gute Chancen auf einen hohen Trefferplatz haben. Insbesondere ist hier ein Verständnis der Qualitätsmessung der Suchmaschinen, welche vor allem auf der Basis der Popularität von Dokumenten erfolgt, zu nennen. Insgesamt lässt sich das Suchmaschinen-Ranking durch vier Gruppen von Rankingfaktoren erklären: Textspezifische Faktoren, linktopologische Verfahren, Aktualität und Lokalität.¹⁴

Zusammengefasst kann gesagt werden, dass die Nutzung von Suchmaschinen zwar intuitiv erfolgt und in vielen Fällen ohne nennenswerte Informationskompetenz auch zum Erfolg führt, bei ernstzunehmenden Recherchen jedoch Informationskompetenz auf unterschiedlichen Ebenen nötig ist. Es konnte gezeigt werden, dass es auf allen Ebenen des Suchprozesses Anknüpfungspunkte für die Vermittlung von Informationskompetenz gibt und diese entsprechend in Schulungen berücksichtigt werden sollten. Doch weder der Suchprozess allein noch allein das Verständnis des technischen Aufbaus von Suchmaschinen kann hier zum Erfolg führen, sondern nur ein Verständnis beider Bereiche.

Das Potential der Suchmaschinen

Suchmaschinen werden hinsichtlich der Entwicklung von Informationskompetenz oft kritisch betrachtet. Der Vorwurf lautet, sie würden sich so stark dem einfachen Nutzerverhalten anpassen, dass Nutzer auch in anderen Informationssystemen keine Notwendigkeit mehr

¹³ Risvik, Knut Magne u. Michelsen, Rolf: Search engines and web dynamics. In: Computer Networks 39 (2002). S. 289–302. (Elsevier.)

¹⁴ Lewandowski, Dirk: Der OPAC als Suchmaschine. In: Bergmann, Julia u. Patrick Danowski (Eds.), Handbuch Bibliothek 2.0. München: De Gruyter Saur 2010, S. 87-107. detaillierte Darstellung des technischen Aufbaus von Suchmaschinen findet sich in Lewandowski, Dirk: Web Information Retrieval: Technologien zur Informationssuche im Internet. Frankfurt am Main: Deutsche Gesellschaft f. Informationswissenschaft u. Informationspraxis 2005.

sähen, Recherchestrategien zu entwickeln und sich mit der Formulierung von Suchanfragen zu beschäftigen.

Allerdings bieten die Suchmaschinen auch Chancen hinsichtlich der Informationskompetenz: Anhand der alltäglichen Web-Recherche kann Nutzern gezeigt werden, wie sie ihr Rechercheergebnis mit Hilfe von Suchstrategien, der Verwendung spezieller Befehle und Operatoren, einer kompetenten Ergebnisselektion, usw. verbessern können. Dadurch können sie erkennen, dass solche Strategien auch in anderen Informationssystemen angewendet werden können oder gar müssen. Sind die Strategien bereits von den Suchmaschinen bekannt, so lassen sich in anderen, schwieriger zu bedienenden Informationssystemen leichter Erfolge erzielen.

Kurzbiographie

Dirk Lewandowski ist Professor für Information Research & Information Retrieval an der Hochschule für Angewandte Wissenschaften Hamburg. Davor war er unabhängiger Berater im Themenbereich Suchmaschinen und Information Retrieval sowie Lehrbeauftragter an der Universität Düsseldorf. Dr. Lewandowskis Forschungsinteressen sind Web Information Retrieval, Qualitätsfaktoren von Suchmaschinen, das Rechercheverhalten der Suchmaschinen-Nutzer sowie die gesellschaftlichen Auswirkungen des Umgangs mit den Web-Suchmaschinen. Zu seinen Veröffentlichungen gehören neben den Büchern "Web Information Retrieval" und "Handbuch Internet-Suchmaschinen" (bislang zwei Bände) zahlreiche Aufsätze, die in deutschen und internationalen Fachpublikationen veröffentlicht wurden.