

Orakel oder Wissen über Wissen?

Yahoo als populäre Enzyklopädie des Cyberspace

Von Peter Haber

Die Enzyklopädien des anbrechenden 21. Jahrhunderts erscheinen nicht mehr in gedruckter Form, sondern präsentieren sich als Datenbanken im Internet. Sie offerieren Wissen zweiter Ordnung.

In den Pinakes, den Katalogen der alexandrinischen Bibliothek, hatte Kallimachos alle seine Schätze verzeichnet: Jede der schätzungsweise 400 000 bis 700 000 Buchrollen war vermerkt. Kallimachos liess in diesem Katalog nicht nur die Namen der Autoren, ihre Herkunft und ihre Fachgebiete aufschreiben, auch Hinweise auf Eigenarten der Schriftsteller oder ihre Todesumstände sollen in den Pinakes notiert worden sein. Die Pinakes des Kallimachos waren ein erstklassiges Arbeitsinstrument für die damalige Wissenschaft und zugleich ein getreues Abbild der bibliothekarischen Systematik des Museions.

Auch heute sind Pinakes ein wichtiges Instrument, um gesuchte Informationen aufzuspüren. Im Internet findet man gleich mehrere Projekte, die sich den Namen «Pinakes» zu eigen gemacht haben. Eines dieser Angebote besteht aus einer kleinen, aber feinen Zusammenstellung von spezialisierten Verzeichnisdiensten, von sogenannten Subject Gateways.¹ Solche Dienste sollen helfen, in der unübersichtlichen Masse von Information, die im World Wide Web angeboten wird, zu navigieren.

Und solche Instrumente sind nötig, denn die Informationsmenge wächst im Cyberspace mit einer völlig unkontrollierbaren Dynamik: Struktur und Technik des Netzes machen es praktisch unmöglich, die zur Verfügung stehende Informationsmenge auch nur annähernd genau zu quantifizieren. Das liegt vor allem an der technisch gegebenen Struktur des Netzes: «Das Internet» gibt es nämlich gar nicht. Als Internet wird vielmehr ein Konglomerat zahlreicher Teilnetze bezeichnet, die sich ständig verändern, vermehren oder auch wieder verschwinden. Es gibt keine Grenzen, kein Zentrum und auch keine wirkliche Kontrolle im Netz. «Das Internet» ist schlicht und einfach die Gesamtheit aller Geräte, die eine bestimmte Kommunikationstechnik (namens TCP/IP) verwenden.

Die Grösse des Internets gibt deshalb immer wieder zu Spekulationen Anlass. Vor einigen

Wochen geisterte die Nachricht durch die Fachpresse, das Web bestehe nicht aus ein bis zwei Milliarden Seiten, wie bisher angenommen, sondern aus mindestens 550 Milliarden Dateien, die allerdings nur zu einem Teil frei zugänglich seien. Das unsichtbare, unbekannte Netz erhielt den Namen «Deep Web». Ob nun diese Zahlen stimmen oder nicht: Die Informationsmenge im Internet ist ohne Hilfsmittel schon längst nicht mehr zu bewältigen. Man schätzt, dass das Web bereits heute mehr Informationen enthält als die grösste Bibliothek der Welt, die Library of Congress.

WISSEN ÜBER WISSEN

Nach Sprache, Schrift und Vervielfältigung scheint der Computer der vierte Auslöser einer tiefgreifenden Informationsrevolution in der Geschichte der Menschheit zu sein. Diese mediale Umwälzung zeitigt zahlreiche Folgen, die bis jetzt erst in Ansätzen untersucht worden sind. Als eine klare Tendenz scheint sich indes abzuzeichnen, dass Wissen zweiter Ordnung eine immer grössere Bedeutung erlangen wird. Dieses «Wissenswissen» ist die Kenntnis der richtigen Informationsbeschaffung oder «informiertes Wissen», wie es die Soziologin Nina Degele zutreffend genannt hat. Informiertes Wissen, so ihre These, gewinnt gegenüber den Wissensinhalten zunehmend an Bedeutung. Ein Blick in die Nutzungsstatistiken des Internets bestätigt diese These: Zu den beliebtesten Angeboten im Netz gehören die sogenannten Suchmaschinen und Verzeichnisdienste. Diese modernen Pinakes des Computerzeitalters sind die populären Anbieter von Wissen zweiter Ordnung geworden.

Die Geschichte unserer Kommunikationssysteme zeigt, dass es vor rund 500 Jahren, nach der Einführung des Buchdruckes, bereits eine sehr ähnliche Entwicklung gegeben hat. Mit seinen Innovationen löste Johannes Gutenberg nämlich eine Kettenreaktion aus: Zuerst einmal vereinfachte der Buchdruck mit beweglichen Lettern die Möglichkeit, Texte zu vervielfältigen. Gleichzeitig

erhöhte die neue Drucktechnologie mittelbar aber auch die Qualität der Texte, und sie beschleunigte die gesellschaftliche Kommunikation. Die Zahl der Bücher nahm in den folgenden Jahrzehnten stetig zu, und immer mehr Menschen schlossen sich dem «Kommunikationssystem Buch» an, wurden Teil des Typographeums, wie es der Kommunikationswissenschaftler Michael Giesecke genannt hat.

Mit dieser Entwicklung stieg auch der Bedarf an Navigationshilfen, die durch das Informationsmeer geleiten, das in erster Linie aus gedruckten Traktaten, Pamphleten und Büchern bestand; schon bald waren Bibliographien und Enzyklopädien in grosser Zahl auf dem Markt erhältlich. Mit dem Internet ist es noch einmal einfacher geworden, Informationen zu «veröffentlichen». Und erneut explodiert die publizierte Informationsmenge und steigt die Nachfrage nach Orientierungshilfen: Es vergeht kaum ein Tag, an dem nicht irgendwo im Internet ein neues Portal oder ein neuer Suchdienst seine Dienste anböte. Der unbestrittene Klassiker und Marktführer in diesem Milliardenmarkt ist der 1994 gegründete Dienst Yahoo² – in den kurzlebigen Zeitbegriffen des Cyberspace bereits ein Veteran.

ONTOLOGIE DER YAHOO-DATENBANK

Der Kern von Yahoo ist nicht eine Suchmaschine, sondern ein Verzeichnisdienst. Ein solches Angebot lässt sich mit einem Telefonbuch vergleichen: Nachgewiesen werden nicht einzelne Seiten des Netzes, sondern die verschiedenen Angebote. Wenn eine Firma ihre Produkte auf hundert verschiedenen Seiten im Netz beschreibt, so gibt es in Yahoo trotzdem nur einen einzigen Eintrag. Die Verzeichnisse von Yahoo wurden anfangs von einer zentralen Redaktion bearbeitet, mittlerweile sind weltweit zwei Dutzend lokaler Dienste mit zum Teil eigenen Redaktionen hinzugekommen.

Verzeichnisse sind also immer das Ergebnis intellektueller Arbeit. Irgendjemand muss die Adressen auswählen, begutachten und einer Kategorie zuordnen. Dies ist auch der grosse Unterschied zu den Suchmaschinen, bei denen Roboter automatisch alle Seiten im Web, die sie finden, vollständig indizieren und in eine grosse Datenbank eingeben. In einer Suchmaschine wie zum Beispiel Altavista³ oder Google⁴ sind dann unter Umständen alle hundert Seiten eines Netzangebotes einzeln verzeichnet, doch es gibt keine Klassifizierung der Inhalte.

Verzeichnisse wie Yahoo, könnte man sagen, haben die Funktionen populärer Enzyklopädien

übernommen. Alles Wissen, das im *enkyklios paideia* des Cyberspace, im – virtuellen – Kreis des Wissens, gespeichert ist, soll mit Hilfe des Yahoo-Suchbaumes erschlossen werden. Doch wie kann das funktionieren? Die grösste Enzyklopädie der Welt soll es auf unvorstellbare 22 877 Bände gebracht haben: Das Werk wurde 1407 in China vollendet. Wie viele Einträge bei Yahoo gespeichert sind, wird geheim gehalten. Schätzungen gehen aber von mehr als einer Million aus. Auch die Grundsätze der Systematik zählen zu den Geschäftsgeheimnissen der Firma, so erfährt man auf Anfrage. Die «Ontologie der Datenbank», erklärt eine hörbar gereizte Yahoo-Sprecherin, sei nicht für die Öffentlichkeit bestimmt.

Diese Geheimnistuerei entspricht nicht gerade der geistigen Tradition moderner Enzyklopädien: Als die berühmte «Encyclopédie» von Jean le Rond d'Alembert und Denis Diderot erschien, schickte d'Alembert dem Werk einen «Discours préliminaire» voraus, in welchem er die verschiedenen Konzeptionen der Wissenschaftssystematik ausführlich erläuterte. Mit ihrem Werk verfolgten Diderot und d'Alembert nämlich mehrere Ziele: Zum einen wollten sie das damals bekannte Wissen der Welt versammeln und klassifizieren. Mehr noch wollten sie aber auch das bekannte Wissen ersetzen durch ein besseres Wissen. Die «Encyclopédie» war ein Kind der Aufklärung. Doch das Projekt entfaltete schon bald eine Eigendynamik: Spätestens seit den fulminanten Recherchen des amerikanischen Historikers Robert Darnton wissen wir, dass die «Encyclopédie» zu einem der grössten Spekulationsobjekte der Mediengeschichte wurde. In einem wilden Gestrüpp von illegalen und halblegalen Geschäften ging es im Verlaufe der späteren Auflagen weder um Wissen noch um Aufklärung – Geld allein stand im Mittelpunkt des gigantischen Geschäftes im Dreieck zwischen Paris, Neuenburg und Marseille.

POPULÄRE NAVIGATIONSHILFE

Ähnlich dürfte es auch vor einigen Jahren David Filo und Jerry Yang ergangen sein. Die beiden jungen Forscher sassen in Stanford an ihren Dissertationen im Fach Electric Engineering, als der Boom im Internet ausbrach. Im Frühjahr 1994 begannen sie damit, ihre persönlichen Lieblingsadressen im Internet als Liste auch anderen Netzbesuchern zur Verfügung zu stellen. Noch im gleichen Jahr professionalisierten sie diesen Dienst und bauten ein Datenbanksystem auf, das zu allen möglichen (und später auch unmöglichen) Themen Netzadressen verzeichnete. Yahoo war geboren. Der seltsame

Name soll dem Vernehmen nach für «Yet Another Hierarchical Officious Oracle» stehen: noch ein aufdringliches hierarchisches Orakel. Heute, nur wenige Jahre später, ist Yahoo der grösste Verzeichnisdienst im World Wide Web. Die Firma erzielte allein im zweiten Quartal dieses Jahres Einnahmen von über 270 Millionen Dollar.

Die Grundstruktur von Yahoo hat sich in den letzten Jahren kaum verändert. Auf der obersten Ebene kann der Netzbesucher zwischen 14 Kategorien auswählen. Mit der Struktur der *artes liberales*, welche d'Alembert als Grundlage diskutierte, haben die Einteilungen von Yahoo nur noch wenig gemein: Es gibt zwar vier Kategorien, die auf wissenschaftliche Themen verweisen, doch die Einteilung bleibt an der Oberfläche. Wer sich im deutschen Verzeichnis über die Grosskategorie «Geisteswissenschaften» in die Kategorie «Philosophie» klickt, hat wiederum 14 Unterbegriffe zur Auswahl, darunter «Bildung und Ausbildung», «elektronische Texte», «Firmen», «Organisationen» und «Verzeichnisse». Nur gerade «Ethik» und «Logik» sind als eigentliche philosophische Disziplinen mit eigenen Kategorien (aber nur ganz wenigen Einträgen) vertreten. Es wird schnell klar: Yahoo ist keine Enzyklopädie für die Fachwelt, sondern eine populäre Navigationshilfe für Netzreisende.

Ein Eintrag in Yahoo ist für die Autoren eines Web-Angebotes indes von existenzieller Bedeutung. Nur wer über die grossen Suchdienste gefunden wird, kann auf die Dauer genug Netzbesucher auf seine Seiten locken und so seine Werbeflächen zu Geld machen.

Das Problem der Orientierungslosigkeit der meisten Surferinnen und Surfer können Verzeichnisse wie Yahoo letztlich nicht lösen: Die Auswahl wird immer willkürlich und kaum nachvollziehbar bleiben. Solange Kriterien und Strukturen von Verzeichnisdiensten nicht offengelegt sind, stellen solche Dienste keine glaubwürdigen Informationsquellen dar. Die Dynamik des Netzes hat aber in den letzten Monaten neue Formen der

enzyklopädischen Wissensvermittlung hervorgebracht: Auf der einen Seite sind es sogenannte Wissensportale wie zum Beispiel Xipolis⁵ («Ihre Bibliothek des Wissens»), wissen.de⁶ oder Questia,⁷ die eine Vielzahl von bestehenden Informationsangeboten bündeln und unter einer einheitlichen Oberfläche anbieten (werden). Nicht selten sind es Verlage, die solche Dienste ins Netz stellen; ein Teil des Angebotes ist zumeist auch kostenpflichtig.

DIE ZUKUNFT DES WISSENSWISSENS

Auf der anderen Seite entstehen im Netz immer mehr der eingangs erwähnten «Subject Gateways». Diese Dienste versuchen, die Idee von Yahoo aufzugreifen, wollen dabei aber ganz anderen Qualitätsansprüchen genügen. Subject Gateways beschränken sich auf klar definierte Fachbereiche und versehen die einzelnen Einträge mit zahlreichen Zusatzinformationen. So findet man Antworten auf Fragen wie: Von wem stammt das Angebot? Wie aktuell ist es? Welche Zusatzdienste werden geboten? Diese Gateways werden nicht selten von grossen wissenschaftlichen Bibliotheken und Forschungseinrichtungen gepflegt, die auf diese Weise ihre Kompetenz im Bereich der inhaltlichen Erschliessung von Printmedien auf die neuen Anforderungen des Cyberspace anwenden. Mit grossem Erfolg, wie die jüngsten Debatten zeigen: Vor einigen Monaten ist ein von der EU finanziertes Programm angelaufen, das die Vernetzung dieser Gateways zum Ziel hat. Es scheint, als ob die Qualität der populären Enzyklopädien, die im Internet abrufbar sind, sich in nächster Zukunft verbessern könnte.

¹ <http://www.hw.ac.uk/libWWW/irn/pinakes/pinakes.html>

² <http://www.yahoo.com> und <http://www.yahoo.de>

³ <http://www.altavista.com> oder
<http://ragingsearch.altavista.com>

⁴ <http://www.google.com>

⁵ <http://www.xipolis.net>

⁶ <http://www.wissen.de>

⁷ <http://www.questia.com>