

Zeitschrift: Arbido
Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz
Band: 16 (2001)
Heft: 3

Artikel: Chancen und Risiken des Internets : der Mensch, das analoge Hindernis?
Autor: Hagmann, Jürg
DOI: <https://doi.org/10.5169/seals-769151>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 22.12.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

CHANCEN UND RISIKEN DES INTERNETS:

DER MENSCH, DAS ANALOGE HINDERNIS?

von Jürg Hagmann

Im Hinblick auf die kommende SVD-ASD-Arbeitstagung Anfang April 2001 in Ittingen möchte der Autor verschiedene im Verlauf des letzten Jahres publizierte Untersuchungen über das Internet vorstellen und besprechen, um Stoff und ein paar Gedanken in den Raum zu stellen, die sich um die Zukunft dieses Mediums drehen.

Die kürzlich in der NZZ besprochene Internet-«Bücherblase» (9.2.01, S.79)¹ hat gezeigt, dass diesbezüglich der Weisheit letzter Schluss bisher noch nicht gezogen werden konnte.

Der Internet-Boom sei nicht am Ende, und die gegenwärtige Krise sei nur eine vorübergehende Phase der Neuordnung, die die Grundlagen für neues solides Wachstum in dieser Branche schaffe:

Dies ist das Hauptergebnis der Ausgabe 2001 des Statusreports «Die Internet-Ökonomie – Strategien für die digitale Wirtschaft» des «European Communication Council (ECC)»^{1a}.

Der ECC-Report ist im Buchhandel erhältlich (wissenschaftlicher Springer-Verlag).

Im Kontrast zu dieser Studie malt Michael Mandel² eine allgemeine Rezession – als Folge der volkswirtschaftlichen Auswirkungen des Internets – an die Wand. Da sich in der neuen Ökonomie ein Teil der Forschung und Entwicklung (F+E) in Richtung Start-ups verlagert, die mit Risikokapital behaftet sind, steigt der Einfluss der Konjunktur auf das Innovationstempo: «Wenn die Aktienmärkte infolge einer Konjunkturdelle in die Knie gehen, dann werden die Venture Capitalists ihre Gelder nicht mehr so grosszügig an Jungunternehmen verteilen und das Innovationstempo nimmt ab. Damit hat sich die Rolle von Innovationen stark verändert: hat sie früher den Konjunkturzyklus eher gedämpft, verstärkt sie dessen Ausschläge heute.»

Sollte sich das Phänomen Internet wirklich als konjunkturdämpfend erweisen –

was ich kaum glaube –, so hiesse das indirekt, dass trotz einer globalen Zunahme der (digitalen) Informationsmenge nicht höhere Umsätze und Gewinne zu erzielen sind. «Quality is the king»: Diese etwas spekulative These geht davon aus, dass natürlich mehrheitlich Informationsschrott produziert wird, was aus der Erfahrung des Informationsspezialisten mehr oder weniger zutrifft (s. weiter unten).

Tatsache ist jedenfalls, dass die Informationsflut täglich anwächst. Die Menschheit produziert pro Jahr 1,5 Exabytes- oder 1,5 Milliarden Gigabytes-Daten. Das haben die Professoren Hal Varian und Peter Lyman von der School of Information Management and Systems an der Universität von California in Berkeley errechnet. Ihre Studie «How much information» (<http://www.sims.berkeley.edu/how-much-info/>) zeigt, welche gigantischen Datenmengen heute von der Menschheit produziert werden. Shakespeares gesammelte Werke entsprechen 30 Sekunden Video. Die Untersuchung berücksichtigt neben Gedrucktem und Computerharddisks erstmals den Ausstoss von Videos, DVDs, CDs, Kassetten, Rundfunk und Fotografie. «Was wir sehen, ist ein phänomenaler Zuwachs der Informationsproduktion», sagt Hal Varian, Co-Autor der Studie. «In den letzten zweieinhalb Jahren haben wir mehr Information erschaffen und gespeichert als in den 300 000 Jahren zuvor.» Möglich wurde dies durch die «Digitalisierung der Welt», wie sich Varian ausdrückt. 93 Prozent der Information, die 1999 produziert wurden, liegen nicht in Papierform vor, sondern sind digital auf Computerharddisks, CDs und DVDs gespeichert³. Die Studie

relativiert auch zwei Mythen: denjenigen des Internet-Wachstums und dass Content König sei im Internet^{3a}.

Die äusserst empfehlenswerte und online erhältliche Analyse enthält zahlreiche Links zu Untersuchungen in den verschiedensten Gebieten der Informationswissenschaften. So etwa einen Hinweis auf die im Juli 2000 von der Firma BrightPlanet veröffentlichte Untersuchung «The Deep Web: Surfacing Hidden Value» über die Struktur und das Wachstum des WWW. (<http://www.completeplanet.com/Tutorials/DeepWeb/index.asp>). Die Studie hat durch empirische Messungen herausgefunden, dass der «unsichtbare», ausserhalb der Reichweite herkömmlicher Suchmaschinen liegende Teil des Web ungefähr 500-mal grösser ist als der allgemein zugängliche; d.h. mit andern Worten, dass z.B. die von Google abgesehen 1 Milliarde Webseiten nur 1/500 der Informationen darstellen, die im Prinzip im WWW vorhanden wären.

BrightPlanet verwendet den Begriff «Deep web» als Synonym zu «invisible web». Der durch Suchmaschinen frei zugängliche Inhalt wird als «Surface web» oder Oberflächen-Web bezeichnet. Der so genannte unsichtbare Teil besteht weitgehend aus gänzlich gesperrten oder

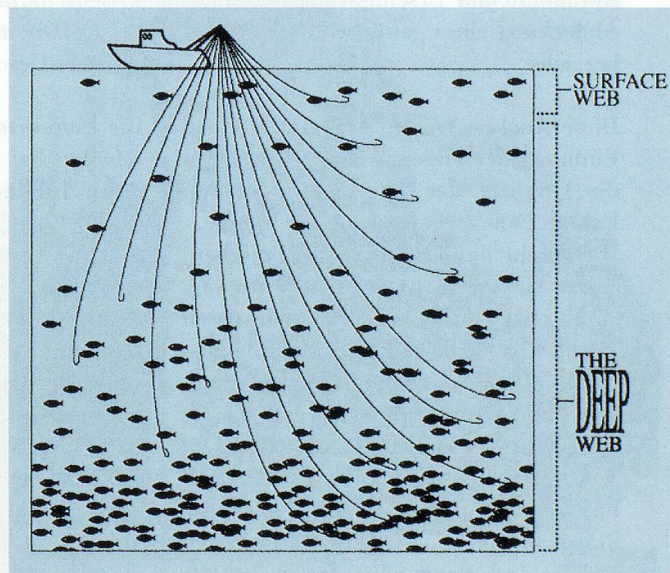


Tabelle: Total «Quality» Potential, Deep vs. Surface Web (The Deep Web, p. 27)

Search Type	Total Docs (million)	Quality Docs (million)
Surface Web		
Single Site Search	160	7
Metasite Search	840	38
Total Surface possible	1000	45
Deep Web		
Mega Deep Search	110,000	14,850
Total Deep possible	550,000	74,250
Deep v. Surface Web		
Improvement Ratio		
Single Site Search	688:1	2,063:1
Metasite Search	131:1	393:1
Total possible	655:1	2,094:1

kostenpflichtigen Informationen. Die 60 grössten Web Sites – darunter diejenigen der NASA und des amerikanischen Zentrums für Klimaforschung – enthalten rund 10% (oder 750 Terabytes) des «Deep Web».

BrightPlanet schätzt, dass der unsichtbare Webteil unter der Oberfläche aus rund 100 000 Websites besteht. Um auch einen bestimmten Teil aus diesem «Deep web» zugänglich zu machen, hat die Firma ein Tool entwickelt (*LexiBot*), das vorläufig die 600 relevantesten «unsichtbaren Web-Ressourcen» in die Suche einbezieht. Es ist geplant, die Suche auf 40 000 Sites auszudehnen, mit dem Fernziel, alle signifikanten Sites auszuwerten.

Die oben stehende Tabelle vergleicht die Suchwerte aus dem «Surface Web» mit denjenigen aus dem «Deep Web». «Quality Docs» impliziert eine Qualitätsrate, d.h., die verfügbare Anzahl Qualitätsdokumente ist eine Funktion ihrer Aktualität und Erschliessungstiefe bzw. Abdeckung eines gewünschten Themenbereichs.

Diese Analyse belegt, dass durch einen Einbezug des «Deep Web» in die Suche die Qualität der Ergebnisse um den Faktor 2000 verbessert werden könnte. Diese rein numerischen (theoretischen) Werte vermögen nicht darüber hinwegzutäuschen, dass der Zugang zu diesen «Deep sites» den kritischen Faktor ausmacht. Realistischerweise wird denn auch eingeräumt, dass der «Verbesserungsfaktor» höchstens 600 betrage. Werte, die nicht erstaunen, gibt es doch bereits im «Surface Web» massive Qualitätsprobleme. So wird etwa festgehalten (S.27), dass rund 44% des Webs von

1999 im Laufe des Jahres 1999 nicht mehr zugänglich waren und dass 45% der existierenden Sites unfertig, bedeutungslos oder trivial waren.

Die Studie von Lawrence und Giles aus der Zeitschrift *Nature*⁴ schätzte, dass der Abdeckungsgrad des Web durch die besten Suchmaschinen in nur einem Jahr von 32% (1998) auf 16% (1999) zurückging.

Bei solchen Zahlen und Überlegungen kommen doch die einen oder andern Zweifel auf über die Zuverlässigkeit der Informationsbeschaffung über das Internet.

Ich denke, es gehört eben zu den Risiken, dass die schier unbegrenzten Möglichkeiten der Informationsverbreitung teilweise überschätzt werden. Man darf sich vorläufig noch nicht einer zu grossen Illusion einer sofortigen und transparenten Vermittlung insbesondere von Fachinformationen und speziell von wissenschaftlichen Texten hingeben. Wie gesagt, wir stehen erst am Anfang der Internet-Geschichte.

Die *Schweizerische Zeitschrift für Geschichte* hat solche Fragen kürzlich in ihrer Jubiläumsnummer exemplarisch für die Geisteswissenschaften diskutiert⁵. Im Aufsatz mit dem Titel «*Quelques réflexions rapides sur le passage des revues du papier au numérique*» (J.M. Baldner) wird dargelegt, dass zwar das Internet neue Diffusionskanäle öffnet, aber zugleich auch die schwierige Frage nach der Identität der Zeitschriften stellt. Die entsprechenden Fragestellungen sind weit komplexer als die Optionen Modernität oder Tradition, deren Argumente ähnlich unerschöpflich und un-

entscheidbar sind. Wichtige Identitätsfragen stellen sich in der Regel jenseits des Mediums. Es sind analoge Fragen angesichts der natürlichen analogen Begrenztheit, aber auch Unvergleichbarkeit menschlichen Tuns.

Wie sind die Beziehungen zwischen AutorInnen und LeserInnen, wie steht es um die AutorInnenrechte und die Anerkennung? Welches sind die Ziele und Orientierungsmuster, Themen und Methoden, Schreib- und Argumentationsweisen, was ist mit Mehrsprachigkeit, Finanzen oder Verteilkanälen (*marché de l'abonnement à la carte*)?

Es kann hier nicht darum gehen, die Errungenschaften und auch Zukunftschancen des Internets zu minimieren, sondern nur darum, die digitalen Vorteile stets in Bezug auf einen Gesamtkontext und/oder Gesamtnutzen zu betrachten.

Der Mensch darf nicht nur als analoges Hindernis betrachtet werden. Letztlich geht es um die Fortsetzung der intellektuellen Fähigkeit, authentisch zu schreiben (und nicht nur Bestehendes zu komponieren), um Texte zu verstehen und über diesen Prozess Identität herstellen zu können.

«*A la fin, lu sur écran ou imprimé, le texte ne se différencie pas dans sa linéarité argumentative. (...) Un des enjeux forts de la numérisation (...) se moule dans la réceptivité.*»⁶ Diskutieren wir doch mal solche Fragen im Zusammenhang mit der geforderten Online-Version von ARBIDO.

¹ Endet die neue Wirtschaft in einer Depression? Technologie und Risikokapital als destabilisierende Faktoren, NZZ 9.2.01, S.79

^{1a} Vgl. NZZ v. 16.2.01, S.69, «Nun beginnt das normale Internet-Geschäft». Ergebnisse einer Studie des ECC

² Michael Mandel, *The Coming Internet Depression*, New York 2000 (Basic Books), 19.20 \$

³ Vgl. auch Die Sinnflut, O. Zihlmann, in *Sonntagszeitung*, 19.11.2000

^{3a} As Odlyzko (2000) puts it, «communication (connectivity by e-mail), not content, is the killer app.», vgl. A. Odlyzko (AT&T Labs), Content is not king, <http://www.research.att.com/~amo>; und A. Odlyzko, Internet growth: Myth and reality, use and abuse, <http://www.research.att.com/~amo>

⁴ S. Lawrence and C.L. Giles, «Accessibility of Information on the Web», *Nature* 400: 107–109, July 8, 1999

⁵ SZG Vol.50, 2000, Nr.4

⁶ «*Quelques réflexions rapides sur le passage des revues du papier au numérique*» (J.M. Baldner), in: SZG, Vol. 50, 2000, Nr. 4, S. 511–514