

Das "Principle of Least Effort" in der Wissenschaft

Walther Umstätter

Vortrag gehalten im Rahmen des
Berliner Bibliothekswissenschaftlichen Kolloquiums (BBK)
am 7.10.2003 in der
Saur-Bibliothek des Instituts für Bibliothekswissenschaft.

Das didaktische Dilemma des Wissens:

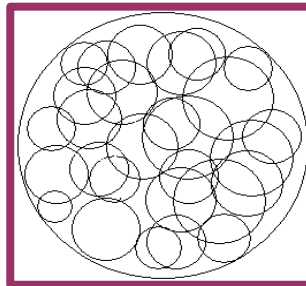
Was wir nicht wissen erscheint uns unbegreiflich.

Sobald wir Wissen erworben haben

erscheint es uns unbegreiflich ohne dieses Wissen zu sein.

Ein Beispiel hierfür ist das, was wir als Pseudo Maximum Entropy Principle bezeichnen können. In einer Universitätsbibliothek konnte über einen Zeitraum von einigen Wochen eine tägliche Nutzung der neusten Zeitschriften von 11% beobachtet werden. (Umstätter, W. und Rehm, M.: Bibliothek und Evolution. Nachr. f. Dok. 35 (6) S.237-249 (1984))

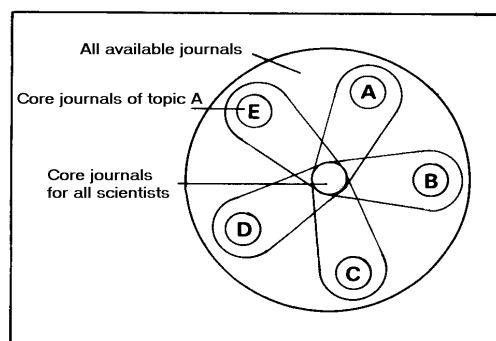
Dabei erschien zunächst völlig unverständlich, wie es dazu kommen konnte, dass die Nutzung am zweiten Tag sowohl 11% der Zeitschriften betraf, die schon am Vortag benutzt wurden, als auch 10% derer, die noch nicht benutzt worden waren, da allgemein bekannt ist, dass es Zeitschriften gibt, die sehr oft, und von vielen Wissenschaftlern gezielt benutzt werden, und solche, die fast niemand benutzt. Statt dessen entstand der Eindruck, als würden die Wissenschaftler rein randomisiert, d.h. zufällig in die Regale greifen.



Damit hat jede Zeitschrift scheinbar die gleiche Chance gelesen zu werden.

Der sich daraus ergebende Widerspruch lässt sich aber aufklären, wenn man annimmt, dass die Wissenschaftler sozusagen zwei Kräften folgen:

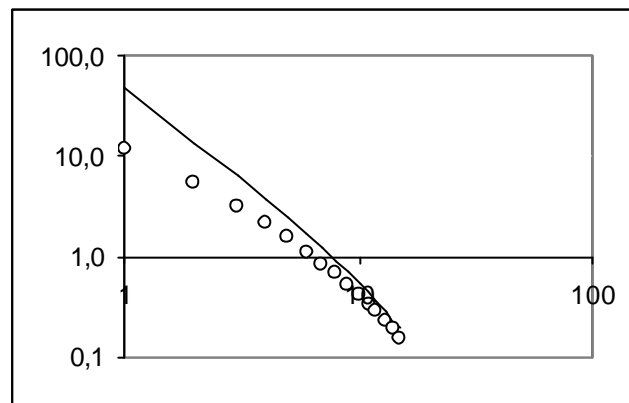
1. einer zentralisierenden Kraft
2. einer divergierenden Kraft



Beiden Kräften folgen sie interessanterweise in so ausgeglichener Form, dass sie gemeinsam das gesamte Gebiet wissenschaftlicher Publikationen ausgewogen überschauen. Hinsichtlich der zentralisierenden Kraft achten sie darauf, das zu lesen, was die Kolleginnen und Kollegen auch lesen, während sie gleichzeitig Themen suchen, auf denen sie sich spezialisieren können, weil sie niemand so gut, wie sie selbst besetzt.

Aus diesem Verhalten ergibt sich auch das Bradford's Law of Scattering, bei dem das jeweilige Spezialgebiet zum Zentrum wird, während alle anderen Zeitschriften als peripher anzusehen sind.

Bemerkenswert ist bei der Beobachtung der täglichen Zeitschriftennutzung auch noch, dass Zeitschriften wie NATURE oder SCIENCE, die allgemeinen Kernzeitschriften der Naturwissenschaft etwa jede 10te – 20te Nutzung betreffen. Das ergibt sich einerseits aus den Journal Citation Reports, in denen diese beiden Zeitschriften bei der Zitation in den anderen naturwissenschaftlichen Zeitschriften oft an 16ter oder 17ter Stelle stehen, und andererseits auch aus der Hochrechnung der Zeitschriftennutzung in der untersuchten Universitätsbibliothek.

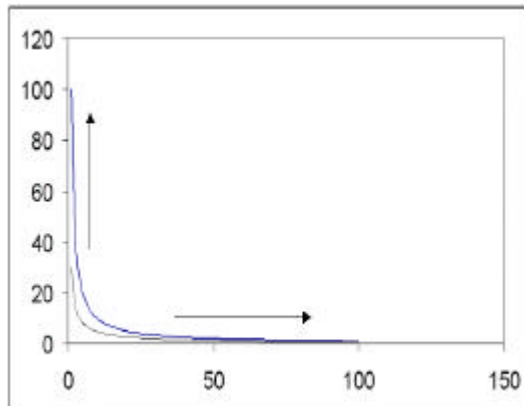


Hier zeigte sich, dass die häufig, also täglich mehrmals genutzten Zeitschriften, zu niedrige Werte aufweisen, weil die Untersuchungen nur die Feststellung erlaubten, ob eine Zeitschrift im Laufe eines Tages benutzt wurde oder nicht.

Geht man allerdings von 7 Benutzungen pro Tag aus, bei einer Nutzungsdauer von etwa 30 Minuten und einer Öffnungszeit von 12 Stunden, so erhält man ein Verhältnis von etwa 3 zu 1. Über die Warteschlangentheorie, die im Prinzip einer Poissonverteilung folgt, lässt sich damit abschätzen, wie groß die Wahrscheinlichkeit ist, dass ein Benutzer die Bibliothek betritt und sich daran gehindert sieht, seine Informationsbedürfnisse abzudecken.

Erfahrungsgemäß muss dieser sogenannte Satisfaction Level unterhalb von rund 40% liegen, da die Nutzer anderenfalls der Bibliothek fernbleiben. Der Aufwand die Bibliothek zu besuchen lohnt sich somit nicht mehr, wenn man feststellt, dass etwa jeder zweite Informationswunsch unbefriedigt bleibt. Dies kann aber schon mit einer Zeitschriftennutzung von durchschnittlich 11% erreicht werden, wenn man die starke Kumulation dieser Nutzungen auf bestimmte Zeitschriften berücksichtigt. Hinzu kommt, dass die Nutzung einer Bibliothek nicht über den Tag völlig gleich verteilt ist und der Eingang neuer Zeitschriften teilweise bekannt ist. So wird beispielsweise jede Tageszeitung oder auch eine Wochenzeitung wie Die ZEIT am Tag ihres Erscheinens weitaus häufiger nachgefragt als danach.

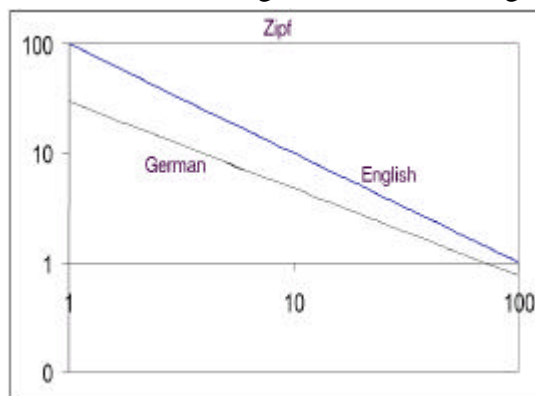
Eine interessante Parallele zu der Benutzung von Zeitschriften finden wir im Zipfschen Gesetz. (Zipf, G.K.: Human Behavior and the Principle of Least Effort, Hafner, New York (1949)). Diese zunächst von Estoup 1916 veröffentlichte statistische Regelmäßigkeit bei der Verteilung von Worten in unserer natürlichen Sprache, hatte Zipf in seiner Dissertation näher untersucht und später in seinem berühmten Buch bekannt gemacht. Dabei identifizierte er zwei Kräfte, die zur hyperbolischen Verteilung der Wörter führt. Sortiert man alle Wörter eines Textes nach Häufigkeit, und trägt auf der Ordinate diese Häufigkeit auf, und auf der Abszisse ihre Rangreihenfolge, so erhalten wir eine Hyperbel der folgenden Form.



Zipf nannte die Kraft in Richtung der Ordinate „Force of unification“ und die andere „Force of diversification“. Die erste führt dazu, dass wir bestimmte Wörter wie „the“, „it“, „is“ etc. so oft benutzen, dass beispielsweise jedes 10te - 20te Wort in einem englischen Text „the“ ist. Gleichzeitig nutzen wir aber in einem Text extensiv seltene Worte, die im gesamten weiteren Text nicht mehr vorkommen. Sie erscheinen in englischen Texten ebenfalls an jeder 10ten - 20ten Stelle, statistisch betrachtet. Die Hyperbel ist im Englischen somit etwa symmetrisch.

Wir können somit schreiben: Die Häufigkeit H ist gleich K / x^1
 K ist dabei eine Konstante, die sich aus der Zahl der insgesamt ausgezählten Wörter ergibt.
 Während x dem jeweiligen Rang in der Häufigkeitsverteilung entspricht.

Da Hyperbeln bei doppelt logarithmischer Darstellung als lineare Verteilung erscheinen



erkennt man den Unterschied zur deutschen Sprache in dieser Abbildung wesentlich deutlicher. Dabei zeigt sich, dass einerseits ein Wort wie „the“ im Deutschen als „der“, „die“ oder „das“ nur zu einem Drittel so häufig vorkommen kann, die Gerade somit wesentlich niedriger beginnt, und

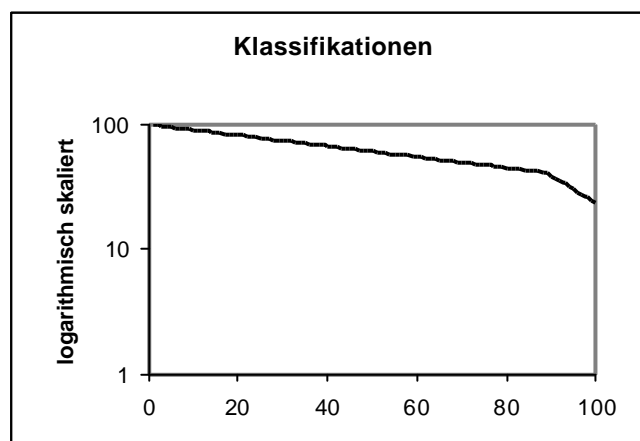
dass andererseits die Gerade einen flacheren Verlauf hat, weil das Deutsche sehr viel mehr Wörter zusammen schreibt, so dass die Zahl seltener Wörter wesentlich größer ist. Wir sehen somit auch hier eine Ausgeglichenheit zwischen zwei Kräften, obwohl sich kaum ein Mensch darüber bewusst sein dürfte, dass er sich um diese Ausgeglichenheit bei jedem neuen Satz bemüht.

Diese Ausgeglichenheit, die sich im Deutschen durch den Exponenten $x^{-0,8}$ und im Englischen durch x^{-1} manifestiert, ergibt sich genau genommen aus der Syntax der Sprache bzw. aus der interdisziplinären Abhängigkeit der Wissenschaftsdisziplinen untereinander.

Wir erkennen somit, dass sowohl die Wissenschaftler bei der Auswahl von Zeitschriften, als auch jeder Text erzeugende Mensch, diesem Prinzip des minimierten Aufwandes erfahrungsgemäß und eher algorithmisch folgt, und dass erst die Wissenschaft bewusst zu machen vermag, was hier die Triebkraft ist.

Es ist bei genauer Betrachtung in beiden Fällen die Informationstheorie, die uns diese Phänomene verständlich macht, da das „principle of least effort“ eigentlich nichts anderes ist als ein ausgewogenes Verhältnis von Information und Redundanz, was Zipf noch nicht realisiert hat, da die Informationstheorie von Shannon, Weaver u.a. erst danach ihren Siegeszug antrat.

Bei den hier betrachteten Funktionen handelt es sich um das sogenannte Potenzgesetz oder „power law“, das in zahlreichen verschiedenen Ausprägungen eine Reihe von Phänomenen zu beschreiben vermag. Das reicht von der horizontalen Geraden mit konstanter Gleichverteilung bis zu fast senkrecht abfallenden Funktionen, bei denen die Hochzahl 3 und größer sein kann.



Dabei ist allerdings darauf zu achten, ob eine Verteilung erst bei doppellogarithmischer Auftragung linearisiert wird, oder schon bei halblogarithmischer Auftragung, wie es für alle e-Funktionen (gestrichelte Linie) der Fall ist.

Ein solcher Übergang von einer hyperbolischen zu einer logarithmischen Funktion tritt ein, wenn die „force of unification“ und die „force of diversification“ jeweils abgeschwächt wird, bzw. Gegenkräfte wirksam werden. Das ist der Fall bei Thesauri und Klassifikationen.

Bevor wir allerdings auf diese Problematik kurz eingehen, soll noch ein grundsätzliches Phänomen des „principle of least effort“ angesprochen werden, das der „Didaktischen

Reduktion“. Unter diesem Prinzip wird zunächst etwas sehr einfaches angesprochen, die Notwendigkeit, in der Didaktik komplexe Probleme auf überschaubare Zusammenhänge herunter zu bringen. Im Prinzip ist das der selbe Vorgang, den wir bei der Modellierung durchlaufen, in der wir alle vernachlässigbaren Elemente abstrahieren.

In unserem Zusammenhang hier hat aber die Didaktische Reduktion noch einen sehr fundamentalen Aspekt, den wir auch als „Biologisches Lernen“ im Gegensatz zum herkömmlichen didaktischen Lernen, bezeichnen können.

Während sich die meisten didaktischen Ansätze darum bemühen, Wissen möglichst systematisch zu vermitteln, erfolgt das Lernen aller Lebewesen auf einem anderen Grundprinzip, das wir beim Erwerb unserer Sprache sehr schön beobachten können. So lernt ein Kind täglich durchschnittlich drei Worte hinzu, deren Auswahl selbstverständlich dem Zipfschen Gesetz folgt. Das sind etwa zehntausend verschiedene Wörter, die je nach Sprachgebrauch über zehn Jahre erworben und wiederholt werden. Bei Menschen mit eher einfacher Bildung übersteigt die Zahl dagegen den Wert von Tausend Wörtern nur wenig.

Während also einerseits permanent Worte gelernt und wiederholt werden, die im täglichen Sprachgebrauch auftreten, tauchen immer wieder solche auf, die gerade dadurch auffallen, dass sie uns neu und damit besonders informativ erscheinen. Wir filtern somit zwei Gruppen von Wörtern aus der natürlichen Sprache heraus, hochredundante und hochinformativ Wörter. Bei den hochredundanten können wir von „Redundanzlernen“ sprechen, da wir sie nur zur Kenntnis nehmen, weil sie immer wieder auftauchen.

Bei den hochinformativen Wörtern können wir von Sensationslernen sprechen, weil wir ihr Auftauchen als besonderen Reiz der Sinne ansehen können. Dieses Sensationslernen ist vergleichbar mit Erlebnissen, die wir im Extremfall nur ein einziges Mal im Leben erfahren haben, uns aber für den Rest unseres Daseins merken.

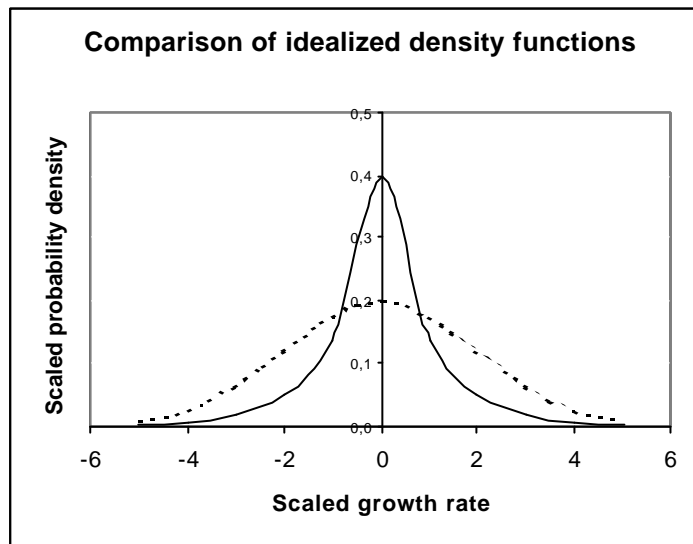
Alle anderen Ereignisse, die entweder zu wenig redundant oder zu wenig informativ sind, verrauschen mehr oder minder rasch in der Vergesslichkeit.

Es ist somit kein Zufall, dass die Relation Information \times Redundanz = Konstante, in der Schreibweise $I = K / R$, eine Hyperbel ergibt, bei der K als Schwellenwert oder Filterkonstante anzusehen ist. K wird je nach Nachrichtenflut angehoben oder gesenkt. Insofern setzen sich Lebewesen selten einer wirklichen Informationsüberflutung aus. Je mehr Zeichen auf sie einströmen, desto stärker erhöhen sie die Filterkonstante, bis hin zum Schließen der Augen, Ohren oder Nase.

Die oft wiederholte Ansicht, dass wir heute einer besonders hohen Informationsflut ausgesetzt sind, ist im Prinzip falsch, da wir vielmehr im zunehmend harten Wettbewerb um die jeweils beste Information diese gezielt suchen. Unsere vermeintliche Überflutung von außen ist somit eigentlich eher ein gezielt erzeugter Sog von Innen, den wir uns nur leisten können, weil wir immer bessere Filter zu Informationsüberflutung besitzen. Das reicht von den Abstellknöpfen bei Radio, Fernseher oder Computer, über die Bücher und Zeitschriften, die wir nur auf Wunsch aus dem Bücherschrank nehmen, bis hin zu den eigenen vier Wänden, in die wir nur gebetene Gäste lassen.

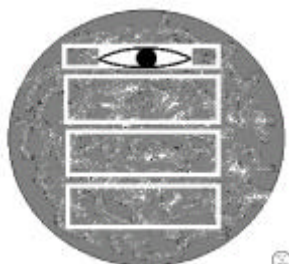
Aus Sicht der Informationstheorie kommt noch hinzu, dass unsere Vorfahren vor Jahrmillionen weitaus weniger Wissen über diese Welt besaßen und daher einer Informationsüberflutung weit mehr ausgesetzt waren, während wir heute sehr viel Information in eine Art a priori Redundanz umgewandelt haben. Wir wissen in weiten Bereichen was morgen geschieht, weil wir dies selbst so planen.

Nur am Rand sei hier noch erwähnt, dass das fundamentale Naturgesetz der Physik, dass das Verhältnis von Beschleunigung (Acceleration A), Kraft (Force F) und Masse (Mass M) beschreibt, auch einer Hypertel entspricht, wenn in $A = F / M$ die Kraft F konstant gehalten wird. Dies ist bei Wachstumsprozessen interessant, bei denen man wiederholt festgestellt hat (Amaral, L.A.N., Gopikrishnan P, Plerou, V., Stanley H.E. u.a.) dass die Streuung in der Wachstumsgeschwindigkeit (nicht die Wachstumsgeschwindigkeit selbst) mit fortgeschrittenem Wachstum geringer wird.



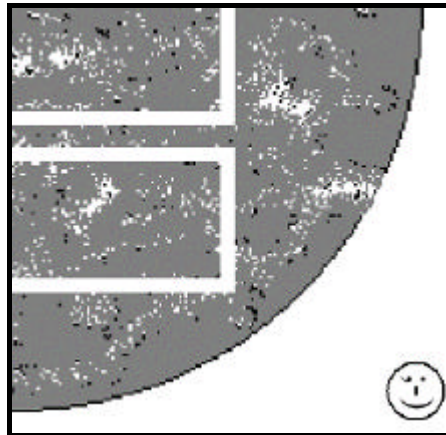
Die Tatsache, dass die Varianz V im Verhältnis $V = 1 / G^{0,2}$ zur Größe eines Unternehmens beispielsweise abnimmt, führt dazu, dass wir als Verteilung aller sich in einem Markt befindlichen Unternehmen keine echte Gaußverteilung (gestrichelte Linie) vorfinden, sondern eine Verteilung mit Kurtosis (durchgezogene Linie). Damit zeigen Wachsende Systeme eine charakteristische Variante der Gaußverteilung, die allerdings nur bei Systemen auffallen dürfte, die stark divergierende Größen erreicht haben und somit in ihrer Verteilung dem „power law“ folgen. Das ist durchaus bemerkenswert, weil wir noch sehen werden, dass bei solchen Verteilungen das Wachstum nicht gleich verteilt ist.

Bezüglich der Didaktischen Reduktion bzw. des Biologischen Lernens ist noch ein Beispiel interessant, aus dem Bereich des Sehens. Wenn wir das folgende Bild betrachten,



so sehen wir zunächst eine graue Fläche, die wir als Redundanz eines Grauwertes ansehen können, oder auch als informationstheoretisches Rauschen, das übergangen werden kann. Darüber liegen Rechtecke, die eher durch ihre Redundanz auffallen und ein aus Kreissegmenten bestehendes Gebilde, das eher durch seine Singularität auffällt.

Bei genauerer Betrachtung, bzw. Senkung unserer Filterkonstanten, würden wir im Rauschen weitere Strukturen erkennen.



Wobei noch auffällt, dass wir alle herausgefilterten Informationen und Redundanzen mit Mustern vergleichen, die wir als eine Art Archetypen im Gehirn gespeichert haben. So erkennen wir wiederholt Augen bzw. ganze Gesichter, im Mond, auf Schmetterlingsflügeln, oder auch in diesem Testbild, obwohl die Abstraktion erheblich ist.

Dieser Abgleich hat viel gemeinsam mit dem Abgleich von Texten mit Thesaurusbegriffen im Bibliotheks- und Dokumentationswesen.

Daraus erkennen wir, erstens, dass wir alles unterhalb eines bestimmten Schwellenwertes als Rauschen betrachten, zweitens, dass unsere Erkennung von Strukturen im Prinzip nichts anderes ist, als die Registrierung von Redundanzen, und dass damit drittens Informationen als Ausnahmen solcher Redundanzen sensationell wirken.

Für viele der Strukturen haben wir bereits vorgefertigte Mustererkennungen, die wir aus phylogenetischer oder auch aus ontogenetischer Erfahrung heraus durchführen.

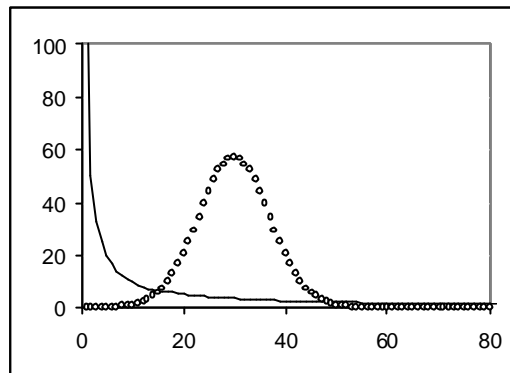
Was wir als Redundant erkennen ist allerdings davon abhängig, auf welcher Erkenntnisebene wir uns befinden. So sind alle Wörter unserer Sprache, die wir im Bibliotheks- und Dokumentationswesen, als Stopwörter kennen, im höchsten Maße redundant. Dies gilt aber nicht auf der Bedeutungsebene.

Wenn wir beispielsweise Homonyme wie Bank, Plasma, Stift, Weide etc. und speziell Polyseme, wie Blatt, Buch, Haus, Kraft, Pferd etc. betrachten, so ist klar, dass ihr jeweiliger Bedeutungswandel sich immer aus ihrer Syntax heraus ergibt.

Informationstheoretisch ist das Wort „es“ also hoch redundant, während es semiotisch als Homonym angesehen werden muss, das je nach Kontext für Atom, Buch, Geld, Haus, Kaninchen, Kind, Sein, Thema, Universum, Wissen etc. stehen kann.

Hier bestätigt sich abermals die wichtige Feststellung Warren Weavers von 1949, „information must not be confused with meaning“.

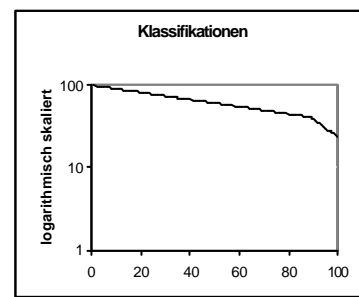
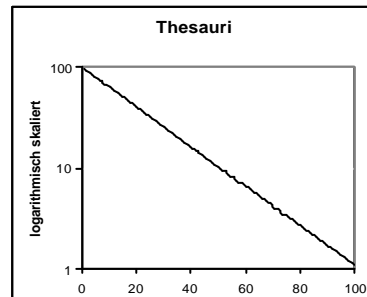
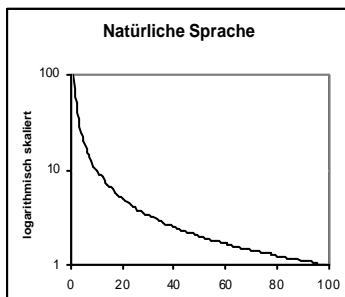
Wenn wir mit der Funktion $I = K / R$ eine Relation zwischen den beiden Zipfschen Kräften unification und diversification geschaffen haben, so ist damit auch den Übergangsbereich zwischen den beiden Extremen, hoch redundant und hoch informativ, festgelegt. Er wurde insbesondere von dem Pionier der Informationswissenschaft, dem Emigranten Hans Peter Luhn in seinem sogenannten „automatic abstracting“, das eigentlich ein automatisches exzerpieren war, zur Texterkennung ausgewählt.



Luhn hatte in den 50er und 60er Jahren des letzten Jahrhunderts bei IBM das Selective Dissemination of Information (SDI) ebenso, wie das Keyword in Context (KWIC) entwickelt.

Hinsichtlich des „automatic abstracting“ erkannte er, dass Stoppwörter ohne wesentlichen Verlust ebenso beseitigt werden können, wie all die Worte, die in einem Text nur ein einziges mal vorkommen. Als charakteristisch für einen solchen Text kann man aber wahrscheinlichkeitstheoretisch die Worte herausfiltern, die informativ und gleichzeitig redundant genug sind, um die Sätze herauszufiltern, die für diesen Text typisch sind. Dies Verfahren gleicht sehr stark dem, das Studierende anwenden, wenn sie bestimmte Sätze in Studentexten markieren.

Auch die Erzeugung von Thesauri und Klassifikationen beruht grundsätzlich auf dem gleichen Prinzip, in dem man die Hyperbel der Wortverteilung natürlicher Sprache von beiden Achsen her staucht. Wie bereits erwähnt geht damit die Lineare der doppelt logarithmischen Darstellung in eine Gerade der halb-logarithmische Darstellung über: Man könnte auch von einem Wechsel vom hyperbolischen Verlauf zu einem Halbwertszeitverlauf sprechen.



Damit zeigen sich bei halblogarithmischer Darstellung für die drei Fälle: natürliche Sprache, Thesauri und Klassifikationen, entsprechend des Zipfschen Rankings, drei charakteristische Kurvenverläufe, die zur Typologie genutzt werden können.

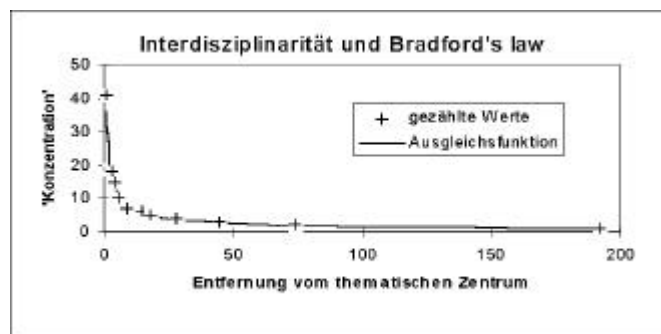
Die Bibliothekswissenschaft beobachtet schon seit langem in der schlichen Erschließung eine Tendenz zur Redundanzdifferenzierung, die mit den Überschriften Klassifikation des Wissens, dokumentarische Erschließung des Wissens und Wissensorganisation zu charakterisieren sind.

Dabei haben Bibliotheken zunächst das Wissen in eine möglichst monohierarchische Klassifikation gebracht.

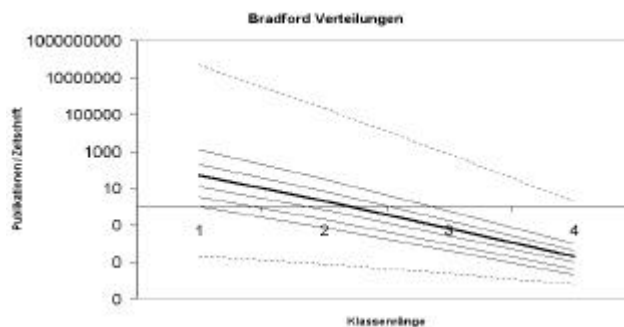
Dokumentationen haben dagegen schon sehr viel kleinere polyhierarchische Äquivalenzklassen erzeugt.

Während das Wissensmanagement nun möglichst exakt auch noch zwischen der Bedeutung von scheinbaren Synonymen und Homonymen differenziert.

Wenn eingangs festgestellt wurde, dass sich Wissenschaftler einerseits möglichst ausgewogen auf ihr zentrales Spezialthema konzentrieren und andererseits die Entwicklungen in der allgemeinen Wissenschaft zu beobachten versuchen, so liegt das natürlich an der interdisziplinären Vernetzung des gesamten Wissens dieser Welt. Wobei die Bibliothekswissenschaft in Bradford's Law of Scattering eine sehr schöne und eindeutige Funktion für die Beschreibung dieser Interdisziplinarität besitzt.



Die Funktion macht deutlich, warum wir auch heute noch auf unseren Spezialgebieten den Überblick behalten können, weil die thematischen Kerne sich sehr stark auf einige wenige Zeitschriften konzentrieren. Daraus ergibt sich auch, warum wir weiterhin eine stetig steigende Zahl an Zeitschriften haben, die direkt proportional zur Zahl der Spezialgebiete steigt.

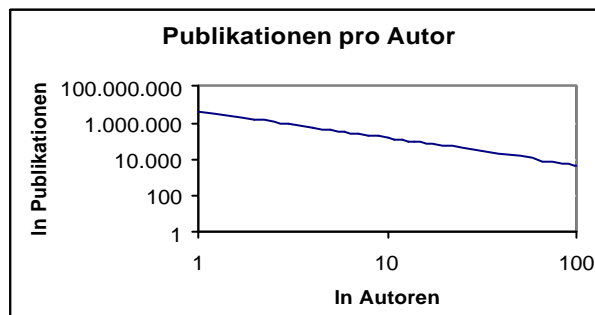


Betrachtet man das Bradford's Law of Scattering als ein Grundsätzliches Phänomen der Wissenschaft, sozusagen von der ersten Publikation eines neuen Spezialgebietes bis hin zur gesamten Wissenschaft, so erhält man bei halblogarithmischer Auftragung eine Schar von Geraden. Hinsichtlich dieser Geraden ist anzumerken, dass die Klassen nach Bradford im Verhältnis $n^0 ; n^1 ; n^2 ; n^3 ; n^4$ anwachsen.

Die beiden gestrichelten Linien zeigen die jeweiligen Begrenzungen auf, während die hervorgehobene Gerade im Mittelfeld die Entstehung einer neuen Zeitschrift markiert. Dies geschieht etwa bei der Größenordnung von ~200 - 300 Aufsätzen/Jahr.

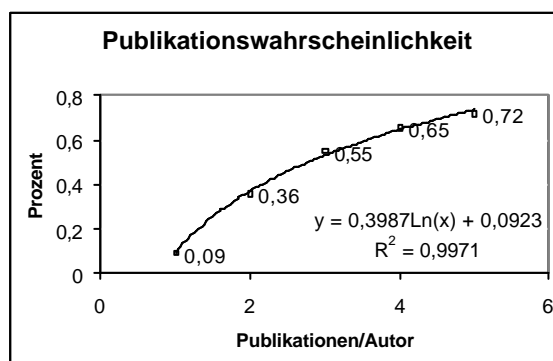
Die wiederholt zu machende Beobachtung, dass Wissenschaftler schon bei einem Aufkommen von etwa hundert Publikationen eine neue Zeitschrift begründen möchten, führt dazu, dass auf zwei Neugründungen im allgemeinen eine Pleite folgt, so dass wiederholt vom Zeitschriftensterben gesprochen wird, obwohl dieser Markt weiterhin eine ungebrochene Verdopplungsrate von etwa 20 Jahren aufweist, auch wenn immer mehr Zeitschriften nur noch elektronisch erscheinen.

Beobachtungen dieser Art können heute bereits dazu herangezogen werden, ein interessantes Modell der Wissenserzeugung aufzubauen. So können wir abschätzen, dass heute etwa 10 Mio. Wissenschaftler und Wissenschaftlerinnen aktiv tätig sind, und dass dies etwa 80% aller Wissenschaftler auf dieser Welt sind. Somit können wir von insgesamt 12 Mio. wissenschaftlichen Autoren in der Welt ausgehen, von denen 36% bisher nur eine Veröffentlichung hervorbrachten. Das sind etwa 4,4 Mio.



Die Abnahme folgt hier der Potenz -1,48 und lässt damit eine gute Abschätzung dafür zu, wie wahrscheinlich es ist, als Wissenschaftler einen, zwei, drei, etc. Publikationen zu veröffentlichen. Diese Wahrscheinlichkeit nimmt von 36%, über 20%, 13%, 9% etc. ab.

Anders betrachtet nimmt aber die Wahrscheinlichkeit zu, einen weiteren Aufsatz zu publizieren.



Das ist der Hintergrund für die wiederholt gemachte Behauptung, dass das Publizieren, ebenso wie das zitiert werden, einem Matthäuseffekt unterliegt. Aus dieser Entwicklung heraus lässt sich grob abschätzen, dass die Chance eine erste wissenschaftliche Publikation zu erzeugen bei nur 9% liegt, während die Wahrscheinlichkeit den 16ten Aufsatz dieser Art zu verfassen schon 90% erreicht. Autoren die in den Bereich , und die während .

Eigentlich ist es nicht sehr sinnvoll hier von einem Matthäuseffekt zu sprechen, da trotzdem jede neue Arbeit und auch jedes neue damit verbundene Zitat mit neuen Anstrengungen verbunden ist. Hinzu kommt, dass der viel zitierte Matthäuseffekt im Science Citation Index nicht nachweisbar ist, weil Literatur im Laufe der Zeit immer mit einer Halbwertszeit von etwa 5 Jahren abnehmend und nicht zunehmend zitiert wird.

Der Hauptgrund dafür, dass auch bekannte Aufsätze nicht dadurch, dass sie zitiert werden, immer öfter zitiert werden, liegt in der kompensatorischen Uncitedness 3, wie sie Garfield nannte. Danach werden bekannte Aufsätze mit zunehmendem Bekanntheitsgrad immer weniger zitiert. Man nennt sie nur noch nach ihrem Namen, wie z.B. das Haber-Bosch-Verfahren, die Fehlingsche Zuckerbestimmung oder die Doppelhelix von Watson und Crick. Eine auffällige Ausnahme bildet die Proteinbestimmung von Lowry, Rosebrough, Farr, und Randall von 1951, mit 250,000 Zitationen. Sie bildet den Zitationsrekord, der vermutlich nur darauf basiert, dass die dabei eingesetzte Folin Reagenz mit dem Verweis auf das Original ausgeliefert wird, so dass alle Autoren, diese Referenz nur abzuschreiben brauchen.

Simkin, M.V. und Roychowdhury, V.P. (Read before you cite. *Preprint*, <http://xxx.lanl.gov/abs/cond-mat/0212043>, (2002). Und Copied citations create renowned papers? *Preprint*, <http://xxx.lanl.gov/abs/cond-mat/0305150>, (2003)) haben ein "model of random-citing scientists" entwickelt, in dem sie behaupten, dass "the majority of scientific citations are copied from the lists of references used in other papers.". Danach sollen nur 22% der zitierten Publikationen von den Zitierenden gelesen werden. Diese Aussage ist vermutlich aus mehreren Gründen falsch, weil die Autoren vorwiegend viel zitierte Publikationen untersucht haben, weil sie die Zahl der Selbstzitationen, die nicht selten allein bei 20-30% liegen, unberücksichtigt ließen und weil sie nicht unterschieden, wie viel Arbeiten nicht über Zitationen sondern über die Durchsicht der neusten Literatur referenziert werden.

Anmerkung:

Dieser Text ist in deutsch abgefasst und gibt damit deutschsprechenden Wissenschaftlerinnen und Wissenschaftlern die Chance, die sich daraus ergebenden weiteren Erkenntnisse rascher zu erfassen, als Menschen, die die deutsche Sprache nicht verstehen.

Dass die Zahl derer, die diese Ausführungen verstehen und weiterzuentwickeln vermögen, sich damit erheblich verringert, gegenüber entsprechenden Ausführungen in englisch, wird in Kauf genommen, da es dem Grundprinzip entspricht, den Studierenden und Institutsangehörigen neuste eigene Erkenntnisse möglichst rasch anzubieten.