

Abb. 2.11: Einstellungen und Werte

den frühen 70er-Jahren, sondern auch Reproduktionen und Neuheiten unter Maßgabe der Designvorstellungen nachgefragt.

## 3 SUCHEN UND FINDEN

### 3.1 Suchmethodik

#### 3.1.1 Daten = Informationen?

„Nie findet man das, was man sucht ...“ – Dies ist wohl der Eindruck der meisten Nutzer nach dem ersten Kontakt mit dem WWW. Nicht ohne Grund spricht die Informationswissenschaft vom Problem des „Information Search and Retrieval“, der Informationssuche und ihrer Wiedergewinnung. Wo findet man was und woher soll ich denn wissen, welche Suchmaschinen und Portale es gibt? –

Diese Fragen beantwortet der folgende Abschnitt oder besser: Zumindest erhalten Sie einen Einblick in Suchmethodik und Suchinstrumente. Ein allumfassendes Internetver-

*Einblick in Suchmethodik und Suchinstrumente*

Ein allumfassendes  
Internetverzeichnis  
existiert nicht

zeichnung existiert nämlich nicht. Selbst wenn dies möglich wäre, müsste zunächst definiert werden, was erfasst wird. Datenvolumen und Interpretationsmöglichkeiten sind zudem enorm groß.

*BEVOR ALSO EINE ZIELGERICHTETE SUCHSTRATEGIE FORMULIERT WERDEN KANN, IST ZUNÄCHST DAS SUCHOBJEKT ABZUGRENZEN.*

Was sind Daten und  
was macht ein Datum  
zur Information?

Zwar wurde bereits auf die Informationsflut und die Notwendigkeit einer Datenverarbeitung hingewiesen – doch was sind Daten und was macht ein Datum zur Information?

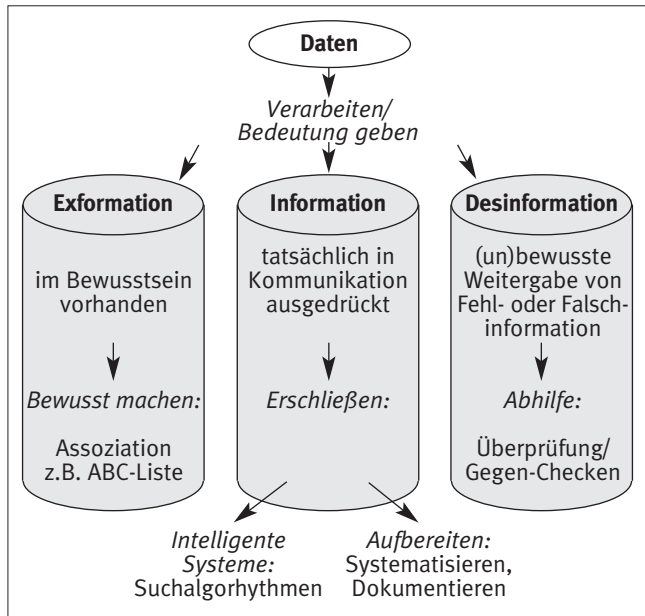


Abb. 3.1: Daten und ihre Formation

Erst die Verarbeitung  
von Daten gibt ihnen  
einen Sinn

Daten sind in einer Vielzahl vorhanden. Hierbei handelt es sich um alle erfassbaren Aspekte wie beispielsweise die derzeitige Zimmertemperatur, die Lichtintensität usw. Erst die Verarbeitung von Daten gibt ihnen einen Sinn; z.B. „Es ist zu kalt und zu dunkel, um in diesem Raum zu arbeiten“. Dabei werden Daten nicht immer bewusst verarbeitet.

Je nach Verarbeitungskontext lassen sich folgende Arten von Daten unterscheiden:

- Viele Daten befinden sich verarbeitet im Bewusstsein, als Erfahrung, erlerntes oder angeborenes abgespeichertes Wissen. Diese Daten können unter Bezug auf Nørretranders auch als EXFORMATION bezeichnet werden (Nørretranders 1994, S. 148). Über Exformationen verfügt jeder (Suchende). Assoziationen helfen, diese ins Bewusstsein zu rücken, um Suchanfragen zu konkretisieren.
- INFORMATIONEN sind bewusst verarbeitete, kommunizierte und individuell systematisierte Daten. Die Systematisierung verfolgt jedoch gewöhnlich ein bestimmtes Ziel, welches bei der Aufnahme und Weiterverarbeitung von Information bekannt sein sollte. Beispielsweise ist die Aussage: „Das Glas ist halb leer/voll“ faktisch richtig, jedoch unter einer Prämisse systematisiert. –

Eine derartig systematische Aufarbeitung kann auch durch so genannte intelligente Systeme vorgenommen werden. Die Verarbeitungsprogramme der Suchmaschinen und Datenbanken stellen einen ersten Schritt zum Ersatz der menschlichen Kognition dar (Lem 1996).

Dabei ist zu berücksichtigen, dass Verzerrungen zwischen Sender und Empfänger die individuelle Verarbeitung von Daten zur Information erschweren: Das Gemeinte ist nicht zwangsläufig das von anderen Verstandene. Dazwischen stehen die Übermittlungswege, z.B. Kataloge.

Jede Systematisierung gibt Information und reduziert zugleich den Informationsgehalt durch die Fokussierung auf Schlüsselbegriffe, Kernaussagen usw. –

Das Gemeinte (verarbeitete Daten) wird durch den Informationsträger geprägt, an den Sender vermittelt und in dessen Verarbeitung zum Verstandenen. Dabei fließt bei Einschaltung von Intermediatoren deren subjektives Wissen über den Meinungsgegenstand ein, z.B. durch manuelles Katalogisieren. Semantische Veränderungen sind dabei möglich (z.B. Branchenvolumen = Branchenumsatz), wodurch die ursprüngliche Information verändert wird.

Reine Informationsspeicher hingegen verarbeiten die Information nicht weiter, sondern stellen sie lediglich dem Empfänger bereit.

*Arten von Daten*

*Erfahrung, erlerntes oder angeborenes abgespeichertes Wissen*

*Informationen als bewusst verarbeitete, kommunizierte und individuell systematisierte Daten*

*systematische Aufarbeitung durch intelligente Systeme*

*Das Gemeinte ist nicht zwangsläufig das von anderen Verstandene*

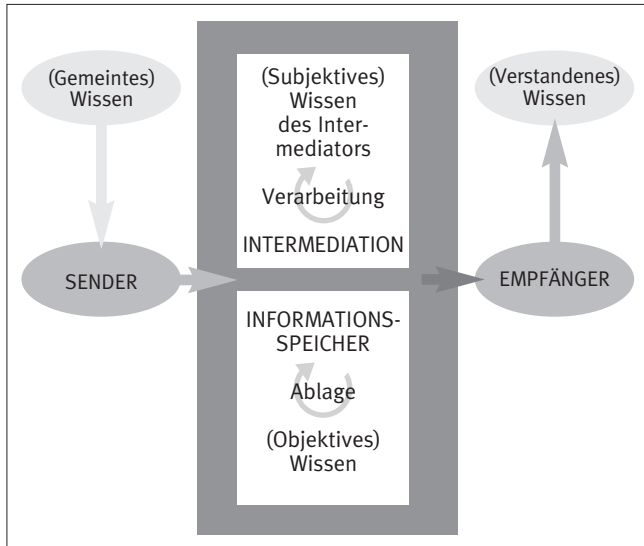


Abb. 3.2: Gemeintes und Verstandenes

*Desinformation durch (un-)bewusste Weitergabe von Fehl- oder Falschinformation*

- **DESINFORMATION** hingegen entsteht durch (un-)bewusste Weitergabe von Fehl- oder Falschinformation. Bei der Suche nach Informationen können diese durch die Überprüfung (Gegen-Checken) selektiert werden.

*Jede Suche wird durch spezifische Informationsbarrieren erschwert*

Suchen (SEARCH) und Finden (RETRIEVAL) gestaltet sich vor dem Hintergrund der aufgezeigten Facetten von Information als ein komplexer Prozess. Suchstrategie, Informationsaufbereitungsmöglichkeit und -bereitstellungsart sind daher voneinander abhängig.

Zunächst wird jede Suche unabhängig von der gewählten Suchstrategie durch spezifische Informationsbarrieren erschwert. Hierzu gehören beispielsweise rechtliche Restriktionen (z.B. Eigentums- und Urheberrechte), finanzielle Beschränkungen (z.B. Budget für professionelle Datenbanken), Zeitbarrieren (z.B. notwendig binnen 24 Stunden) oder politisch-ideologische Barrieren.

*intersubjektive Verzerrungen*

Dies kann als „Fehler erster Ordnung“ bezeichnet werden, da hierdurch intersubjektive Verzerrungen entstehen: Objektiv notwendige Informationen bleiben unzugänglich und damit unverarbeitet.

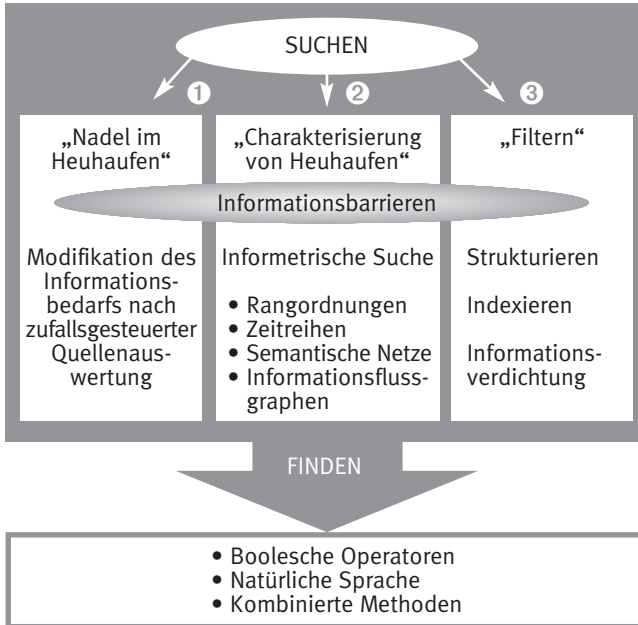


Abb. 3.3: Suchen und Finden

Die Suchstrategie ist abhängig von der Fähigkeit, den Informationsbedarf zu konkretisieren – und damit von Vorkenntnissen („Fehler zweiter Ordnung“). Dabei wird insbesondere bei problemorientierter Suche ein diachroner Einsatz der in Abbildung 3.3 dargestellten Methoden erfolgen.

- **„NADEL IM HEUHAUFEN“:** Hierbei wird ausgehend von einem in der Regel diffusen Informationsbedarf eher unsystematisch gesucht, indem Verweise von Informationsquellen auf andere genutzt werden. Ein vollständiger Überblick über alle relevanten Informationen wird mit dieser Methode nicht erreicht. Bezogen auf die Nutzung einer Suchmaschine entspricht diese Systematik der zufallsgesteuerten Auswahl einzelner, als Suchergebnis angezeigter Websites.
- **„CHARAKTERISIERUNG VON HEUHAUFEN“:** Durch systematische Aufbereitung, z.B. mittels Rangordnung oder semantischem Netz, der erhaltenen Informationen werden nicht näher analysierte Quellen charakterisiert. Dies entspricht

*mögliche Suchstrategien*

*unsystematische Suche, indem Verweise von Informationsquellen auf andere genutzt werden*

*Durch systematische Aufbereitung erhaltener Informationen nicht näher analysierte Quellen charakterisieren*

*Fachkenntnis innerhalb  
des Suchgebiets oder  
Einsatz intelligenter  
Suchsysteme*

*die Gesamtheit der  
verfügbaren Daten  
systematisieren*

*Informationsbedarf,  
Suchinstrument und  
Suchstrategie individuell  
aufeinander abstimmen*

beispielsweise einerseits einem Abbruch der weiteren Auswertung, wenn sich eine Informationsquelle als unseriös erweist und andererseits einer Aufnahme in die weitere Suchstrategie bei soliden Informationen. Eine derartige Sortierung setzt die Fachkenntnis innerhalb des Suchgebiets oder den Einsatz intelligenter Suchsysteme voraus. Beispielsweise unterstützt die Suchmaschine [www.kartoo.com](http://www.kartoo.com) diese Methodik durch die Visualisierung der gegenseitigen Verweise (Hyperlinks).

- „FILTERN“: Durch Strukturieren, Indexieren (z.B. Thesaurus, Klassifikation, Textwortmethode, Zitationsindexierung) oder Informationsverdichtung (Zusammenfassungen bzw. Abstracts, Topic Maps, MindMaps, Ontologie) wird die Gesamtheit der verfügbaren Daten systematisiert. Webkataloge versuchen, diesem Suchanspruch gerecht zu werden. Jedoch erfassen diese in der Regel nicht alle verfügbaren Websites.

Die Grundproblematik des Suchenden stellt sich ebenso dem Informationsverarbeitenden bei der Bereitstellung von Informationen. Informationsbedarf, Suchinstrument und Suchstrategie müssen daher individuell aufeinander abgestimmt werden.

Erschwerend beim (Wieder-)auffinden von Informationen (Information Retrieval) kommt hinzu, dass dem Suchenden häufig ein unnatürliches Verhalten aufgedrängt wird. Statt die natürliche Sprache zu nutzen, erfordert die Nutzung eines Großteils der Suchinstrumente eine Reduktion auf wenige Termini (Schlagwortsuche). Dabei sind die Anstrengungen, Suchanfragen in natürlicher Sprache zu gestalten, rund 40 Jahre alt: Gerard Salton (1962), Roger Kent Summit, Karen Sparck-Jones (1968) und Gerhard Lustig (1968) gehören zu den Vorreitern in der Entwicklung natürlichsprachiger Systeme. Die 1966 von Roger Kent Summit gegründete Datenbank Thomson Dialog ([www.profound.com](http://www.profound.com)) unterstützt diese Retrieval-Technik.

### **3.1.2 Schlagworte: Suchformulierung**

Suchinstrumente im Internet greifen auf die Systematisierung von Daten zurück. Da nur wenige Auskunftswerkzeuge die

## SUCHMETHODIK

natürliche Sprache (z.B. „Wie groß ist der Markt für Süßwaren?“) unterstützen, muss auf Schlagworte zurückgegriffen werden. Es bietet sich daher die Kodierung des Informationsbedarfs in treffende Termini an. Dabei gilt insbesondere bei populären Themen:

*WENIGER (IM SUCHERGEBNIS) IST MEHR! – GENERIEREN SIE MEHRERE SCHLAGWORTE, UM DIE TREFFERQUOTE AUF EINE ÜBERSCHAUBARE ZAHL EINZUGRENZEN.*

Die Generierung geeigneter Schlagworte kann erfolgen durch

- Assoziationen, also gedankliche Verknüpfungen, sowie
- Einordnung in über- oder untergeordnete Themen.

*Generierung geeigneter Schlagworte*

### ASSOZIATION

Wenn beispielsweise Vergleichsdaten für die relativen Distributionskosten einzelner Vertriebskanäle als Vergleichsmaßstab für das eigene Unternehmen gesucht werden, so bietet

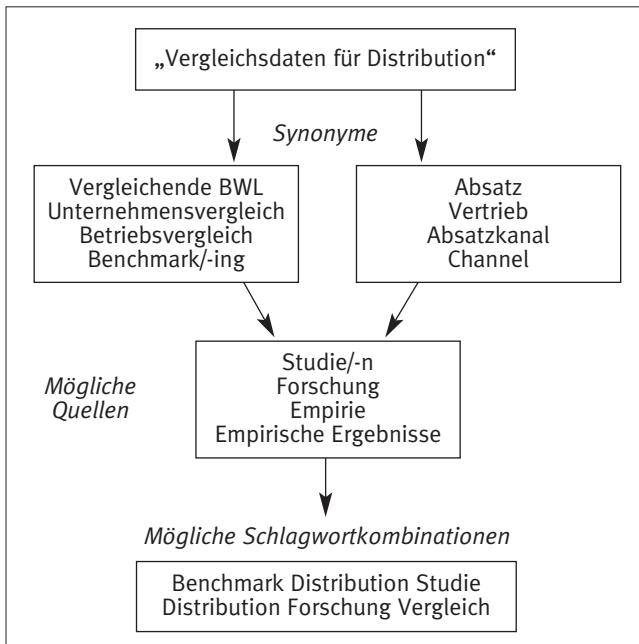


Abb. 3.4: Schlagwortgenerierung

sich zunächst die Entwicklung einer Assoziationskette oder eines Brainstorming mit synonymen oder verwandten Begriffen zum Informationsbedarf an (siehe Abb. 3.4).

*Bildung von ABC-Listen,  
um geeignete Schlag-  
worte zu finden*

Die Assoziationstechnik kann durch ABC-Listen verfeinert werden:

- Versuchen Sie Wörter zum Thema mit den Anfangsbuchstaben von A bis Z zu finden.
- Disziplinieren Sie sich! Tragen Sie Assoziationen mit eindeutigen Themenbezug zusammen!
- Versuchen Sie dabei, ausschließlich auf Fachtermini zurückzugreifen!

Für das obige Beispiel könnte eine mögliche ABC-Liste folgende Elemente enthalten. Dabei finden sich – je Kenntnisstand – Fachtermini neben gewöhnlichen Substantiven und Verben.

<b>A</b>	Absatz, Absatzkanal, Ähnlichkeit, Akquisition, ACD (Automated Call Distribution)
<b>B</b>	Benchmark (als Vorbild in der Distribution), Bahn (Transport), Balanced Scorecard
<b>C</b>	Channel, Category Management
<b>D</b>	Daten, Distribution, Distributionsgrad (als Ausdruck der Lieferbereitschaft)
<b>E</b>	Erfassung, Empirie (zur Distribution), Efficient Replenishment, Efficient Consumer Response
<b>F</b>	Forschung (der Distributionsstrukturen)
<b>G</b>	Geschwindigkeit, Güterverkehr
<b>H</b>	Handel
<b>I</b>	Index (zur Distribution), Indizes
<b>J</b>	Just-in-time
<b>K</b>	Kennzahl, Kooperation(-sgrad), Kosten (je Einheit), Kapitalbindung
<b>L</b>	Logistik, Lieferbereitschaft, Lagerreichweite
<b>M</b>	Masse
<b>N</b>	Netz (Verteilungsnetz, Lagernetz usw.)

<b>O</b>	Operator (für die Logistik zuständiger Mitarbeiter)
<b>P</b>	Produktivität
<b>Q</b>	Qualität, Qualitätsvergleich
<b>R</b>	Rentabilität, Rücklaufquote
<b>S</b>	Similarity (Ähnlichkeit), Supply Chain Management
<b>T</b>	Testmarkt, Termineinhaltung
<b>U</b>	Umschlag(-skennziffern)
<b>V</b>	Vertrieb, Verfügbarkeit im (Einzel-)Handel, Verteilen
<b>W</b>	Win-Win-Gemeinschaft (bei Kooperation zwischen Herstellern und Handel), Weg
<b>X</b>	?
<b>Y</b>	Yield (Ertrag, Ertragsmanagement in Distribution)
<b>Z</b>	Zentrale, Zulieferer, Zwischenlager

Abb. 3.5: ABC-Liste zu „Vergleichsdaten in der Distribution“

Wird nun die ABC-Liste nach den beiden Hauptkriterien „Vergleich“ und „Distribution“ geordnet, ergeben sich sinnvolle Hinweise auf Schlagwortkombinationen. Häufig zeigen sich Verbindungen zu verwandten Themen, deren Kenntnis man sich zuvor nicht bewusst war. ABC-Listen eignen sich insbesondere dann, wenn wenig, jedoch einiges zum Thema bekannt ist (Birkenbihl 2003, S. 53).

Dabei kommt der Kombination allgemeiner und fachspezifischer Ausdrücke eine besondere Bedeutung zu: Nicht jedes Dokument basiert ausschließlich auf dem „Fachchinesisch“. Durch die Ergänzung um verwandte Begriffe aus der Alltagssprache erzielt man regelmäßig eine höhere Trefferquote. Für das genannte Beispiel ergeben sich allein bei sechs ausgewählten (Meta-)Suchmaschinen folgende Trefferzahlen:

- „Vergleichsdaten +Distribution“: 239 (www.google.de), 234 (www.kartoo.com), 106 (www.metager.de), 29 (www.vivismo.com), 19 (www.ixquick.com), 5 (Altavista);
- „Benchmark +Studie +Distribution“: 3.580 (www.google.de), 1.010 (www.kartoo.com), (www.vivisimo.com), 746 (www.altavista.de), 642 (www.ixquick.com), 134 (www.metager.de).

*sinnvolle Hinweise auf Schlagwortkombinationen*

*Ergänzung um verwandte Begriffe aus der Alltagssprache*

Die Trefferquoten variieren um ein Vielfaches

Die Trefferquoten variieren um ein Vielfaches. Dabei bedeutet ein angezeigtes Suchergebnis keinesfalls „Treffer“ im Sinne des Informationsbedarfs. Die Qualität der Suchergebnisse steht in keinem linearen Zusammenhang mit der Quantität.

So liefern die jeweils ersten zehn Treffer auf den sechs ausgewählten Suchmaschinen übereinstimmend folgende Ergebnisse:

- Anleitung für eine Treiberinstallation unter dem Betriebssystem Linux,
- Benchmarkingstudie der Uni Hannover zum Gesundheitssystem,
- Benchmarkingstudie der GfK zur Distribution der Top 100 Markenartikel in Europa,
- Benchmarkingstudie der Bundesanstalt für Milchforschung zum Distributionsgrad von Milchprodukten.

Unter den ersten Treffern waren also drei zutreffende Dokumente und eines ohne Bezug zum Thema. Dies kann aus dem Gebrauch derselben Termini in unterschiedlichen Zusammenhängen resultieren.

In der EDV wird beispielsweise als „distribution“ bzw. „distributed computing“ eine verteilte Datenverarbeitung bezeichnet, weshalb die Treiberinstallation zum Ergebnisreis gehört.

*VERSUCHEN SIE, DIE JEWEILIGE PROBLEMSTELLUNG IN ÜBERGEORDNETE THEMENKOMPLEXE EINZUORDNEN UND GENERIEREN SIE TREFFENDE SCHLAGWORTE.*

Eine zu starke Differenzierung kann jedoch die Konsequenz haben, dass keine Ergebnisse angezeigt werden. Die Kombination von fünf Fachtermini zum dargestellten Beispiel (hier: Benchmark, Studie, Absatzkanal, Distribution, Vergleich) führt zum Ergebnis „keine Treffer“.

#### EINORDNUNG IN ÜBER- ODER UNTERGEORDNETE STRUKTUREN

Kein Begriff steht für sich allein. Die Einordnung von Schlagworten in über- und untergeordnete Strukturen kann nach verschiedenen Aspekten erfolgen, z.B.:

- **FACHSYSTEMATIKEN** wie die Betriebswirtschaft im Rahmen der Wirtschaftswissenschaften, die nach den funktionalen

Aspekte der Hierarchisierung

## SUCHMETHODIK

(z.B. Beschaffung, Produktion, Marketing, Steuern, Rechnungswesen) und institutionalen Aspekten (z.B. Handel, Banken, Versicherungen) geordnet sind;

- WIRTSCHAFTSZWEIGE wie beispielsweise Land- und Forstwirtschaft, Produktion, Bauwirtschaft, Dienstleistungen;
- GEOGRAFISCHE GEBIETE wie Staaten, Bundesländer, Kommunen usw.;
- RESSOURCENTYPEN wie Berichte von Forschungseinrichtungen, Unternehmen, Non-Government Organizations (NPOs), Privatpersonen, Tagespresse usw.

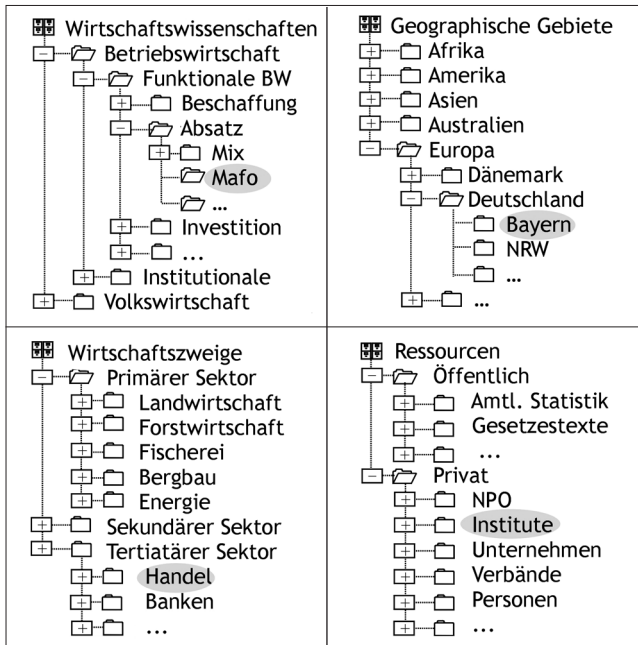


Abb. 3.6: Einordnung in über- und untergeordnete Strukturen

Bei der Einordnung des Informationsbedarfs in über- und untergeordnete Strukturen sind HOMONYME zu berücksichtigen.

Ein und dasselbe Wort kann, etwa in verschiedenen Fachdisziplinen, unterschiedliche Bedeutungen haben. So bedeutet „Yield“ beispielsweise sowohl Ertrag als auch Resultat im Rahmen der Informatik – das Yield Management befasst sich

Homonyme  
berücksichtigen

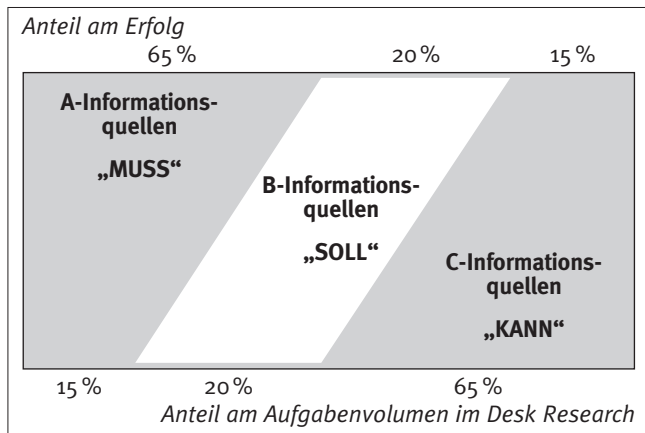
hingegen mit einer Ertragssteigerung auf der Basis einer zeitlichen Preisdifferenzierung. Die Begriffe „Segment“ bzw. „Segmentgröße“ werden nicht nur in der Betriebswirtschaft, sondern auch in der Informatik genutzt und stehen dort – im Gegensatz zum „Teilmarkt“ – für Elemente eines TCP-Datenrahmens.

*KOMBINIEREN SIE DAHER HOMONYME SUCHBEGRIFFE MIT VERWANDTEN TERMINI.*

**PROIRITÄTENSETZUNG**

*Erkenntnisse des Pareto-Prinzips nutzen*

Bei der Suche nach Informationen können die Erkenntnisse des Pareto-Prinzips genutzt werden. Dieses besagt, dass rund 20 Prozent der aufgewendeten Zeit für rund 80 Prozent der Ergebnisse verantwortlich ist.



*Abb. 3.7: ABC-Prinzip im Internet Desk Research*

*Klassifikation von Informationen*

Dabei steht die Klassifikation von Informationen bzw. Informationsquellen für:

- A-Informationsquellen: Informationen, die beschafft werden müssen bzw. Quellen, die befragt werden müssen (KERNINFORMATIONEN);
- B-Informationsquellen: Informationen, die bereitgestellt bzw. Quellen, die genutzt werden sollten (ZUSATZINFORMATIONEN);

## SUCHMETHODIK

- C-Informationsquellen: Informationen, die idealerweise bereitgestellt bzw. Quellen, die genutzt werden könnten (RANDINFORMATIONEN).

Im Rahmen einer Bestandsaufnahme zum Thema „Existenzgründungsstimmung in Deutschland“ könnte beispielsweise folgende Priorität gesetzt werden: A-Quellen wie Verbandsinformationen, Informationen des Instituts für Mittelstandsforschung, der Wirtschaftsförderungsgesellschaften oder empirische Erhebungen, B-Quellen wie Statistiken der Arbeitsagentur und C-Quellen wie Erfahrungsberichte von Existenzgründern in verschiedenen Medien.

Für Internet Desk Research bedeutet dies:

*BEGINNEN SIE IHRE SUCHE DORT, WO SIE MIT GROSSER WAHRSCHEINLICHKEIT DIE WESENTLICHEN INFORMATIONEN ERHALTEN. FÜR RAND- UND ZUSATZINFORMATIONEN KÖNNEN ENTSPRECHEND AUFWÄNDIGERE SUCHSTRATEGIEN EINGESETZT WERDEN.*

Dies setzt die Kenntnis von Einstiegsquellen wie themenbezogenen Portalen, Verbandsseiten, Datenbanken oder E-Zines voraus.

*Die Kenntnis von Einstiegsquellen ist unerlässlich*

### 3.1.3 Trailing / Surfen

Das Internet ist ein riesiger Informationsspeicher. Die Grundidee, insbesondere für das WWW, ist der HYPERLINK. Dieser ermöglicht die Anlage von Verweisen, welche Pfade (TRAILS) im Sinne assoziativer Sprünge von einem Dokument zum anderen generieren. Das Gleiten (SURFEN) von einem Dokument zum anderen ist dadurch möglich (siehe Abb. 3.8).

*HYPERLINKS ermöglichen das Surfen im Netz*

Die Überlegung stammt aus den 20er-Jahren des 20. Jahrhunderts, als sich Forscher Gedanken über ein Wissensmanagement machten. Unter dem Titel „As we may think“ veröffentlichte Vannevar Bush später, im Jahr 1945, einen zukunftsweisenden Artikel im ATLANTIC MONTHLY. Er stellte darin ein Konzept für die assoziative Verknüpfung dokumentbasierten Wissens vor: Die Grundidee für HYPertext MARKUP LANGUAGE (HTML), welche das WWW erst ermöglichte.

*assoziative Verknüpfung dokumentbasierten Wissens*

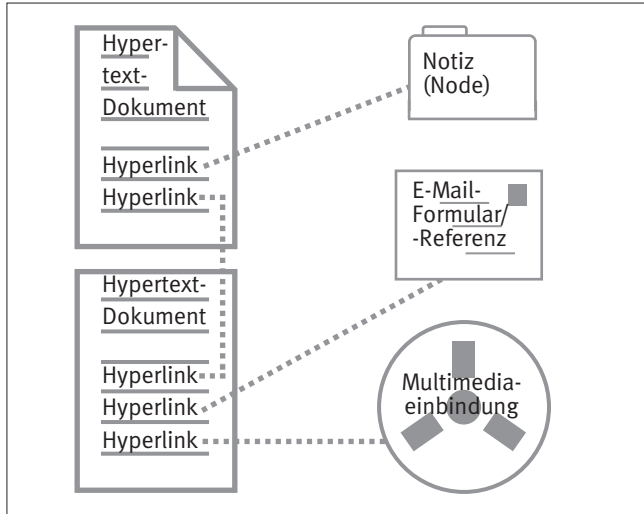


Abb. 3.8: Hyperlink-Prinzip

te. Das als MEMEX (Memory Extender) bezeichnete System sollte der Anlage von Notizen und Verweisen zu anderen Dokumenten dienen. Hierdurch hätten die zur damaligen Zeit bereits großen Informationsmengen durch den Einzelnen besser erschlossen werden können. Das aufgrund der vorherrschenden Technologie mechanisch konstruierte System blieb jedoch ein Entwurf und wurde nie realisiert.

Das System der Gedankenpfade (TRAILS) bzw. der Verweise (LINKS) wurde jedoch realisiert: Im World Wide Web. Das Surfen von einer Website zur anderen entspricht dabei einem ungesteuerten Schneeballprinzip.

Die „Internetarbeiter“ verarbeiten durch den Verweis auf andere Websites ihr Wissen und stellen es jedem interessierten Nutzer zur Verfügung. Wenn also eine Website über keine oder nur wenige Hyperlinks zu anderen Dokumenten verfügt, sagt dies nichts über Existenz und Anzahl weiterer Websites zum Thema aus.

### 3.1.4 Suchoperatoren

Die meisten Suchinstrumente erlauben die Verwendung von BOOLESCHEN OPERATOREN. Diese, nach dem britischen Ma-

Das Surfen entspricht  
einem ungesteuerten  
Schneeballprinzip

BOOLESCHE  
OPERATOREN

## SUCHMETHODIK

thematiker und Logiker George Boole (1815-1864) benannten logischen Kombinationen AND, OR und NOT erlauben die Verknüpfung von Suchbegriffen mittels Worten oder mathematischer Zeichen.

Dabei können bei webbasierten Suchinstrumenten in der Regel folgende Kombinationen verwendet werden:

- AND bzw. & bzw. + bzw. ANY: Die mit AND verknüpften Begriffe müssen im Dokument enthalten sein und dürfen höchstens zehn Wörter voneinander entfernt sein (Schnittmenge);
- OR bzw. | bzw. ANY: Einer der mit OR verknüpften Begriffe muss im Dokument enthalten sein;
- NOT bzw. ! bzw. -: Mit dem Minuszeichen bzw. mit AND NOT oder OR NOT verknüpfte Begriffe dürfen nicht im Dokument enthalten sein;
- NEAR bzw. ~: Die mit NEAR verknüpften Begriffe müssen im Dokument enthalten sein und dürfen höchstens zehn Wörter voneinander entfernt stehen;
- \* bzw. ? , \$ oder #: Sternchen, Joker oder Fragezeichen sind Stellvertreter für einzelne oder mehrere Zeichen, während die Raute für ein einzelnes Zeichen steht. Hierdurch können konjugierte oder deklinierte Varianten des Suchbegriffs gefunden werden (sog. Fragmentierung);
- „“ bzw. PHRASE: Die Verwendung von Anführungszeichen erlaubt die Suche nach dem genauen Begriff (Phrasen), welche im Suchergebnis exakt enthalten sein müssen;
- Das Leerzeichen (blank) zwischen zwei Suchbegriffen wird automatisch als AND gewertet;
- XOR liefert die Exklusionsmenge.

Diese Operatoren werden im VERY EASY RODENT ORIENTED NETWORK INDEX TO COMPUTERIZED ARCHIVES (VERONICA) des Gopher-Systems ebenso wie in Suchmaschinen, Datenbanken und Verzeichnissen im WWW akzeptiert.

Die Booleschen Operatoren beziehen sich auf die Logik der Mengenlehre. Wenn beispielsweise Musikinformationen zu den Stilrichtungen „Country“ und „Western“ gesucht werden, ergeben sich bei Booleschen Operatoren folgende Mengenbeschreibungen:

*Verknüpfung von Suchbegriffen mittels Worten oder mathematischer Zeichen*

*Kombinationen bei webbasierten Suchinstrumenten*

*Logik der Mengenlehre*

Beispiele für den unterschiedlichen Einsatz von Suchoperatoren

- „Country AND Western“ liefert die **SCHNITTMENGE**, also Dokumente, welche sowohl „Country“ als auch „Western“ enthalten, z.B. Mike Bladely oder Johnny Cash;
- „Country AND OR Western“ fragt nach der **VEREINIGUNGSMENGE**. Die Treffer enthalten alle Dokumente, welche entweder „Country“ oder „Western“ oder beides enthalten, z.B. Bono, Johnny Cash;
- „Country AND NOT Western“ erzeugt die **EXKLUSIVMENGE**, welche alle Dokumente mit „Country“, jedoch nicht mit „Western“ umfasst, z.B. Carter Family, Pete Seeger, Hank Williams;
- „Country XOR Western“ bezeichnet die **EXKLUSIONSMENGE**, also alle Dokumente die entweder „Country“ oder „Western“ enthalten, jedoch keinesfalls beide Bezeichnungen, z.B. Pete Seeger, Chris Goss.

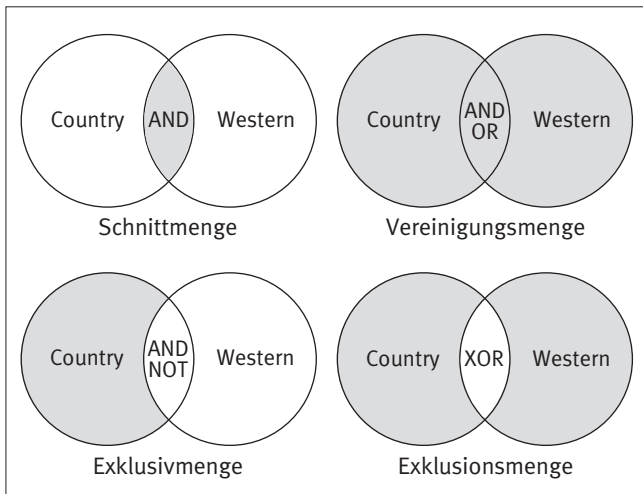


Abb. 3.9: Boolesche Operatoren

Trefferquote durch Nutzung mehrerer Operatoren auf eine relevante Restmenge reduzieren

Die Trefferquote lässt sich durch Nutzung mehrerer Operatoren auf eine relevante Restmenge reduzieren.

Wenn beispielsweise Daten zum Markt für Naturkosmetik gesucht werden, die das Wort „Ringelblumenextrakt“ enthalten, jedoch nicht das Wort „Edelweißextrakt“, müsste die Suchanfrage lauten „Naturkosmetik AND Ringelblume AND NOT Edelweiß“.

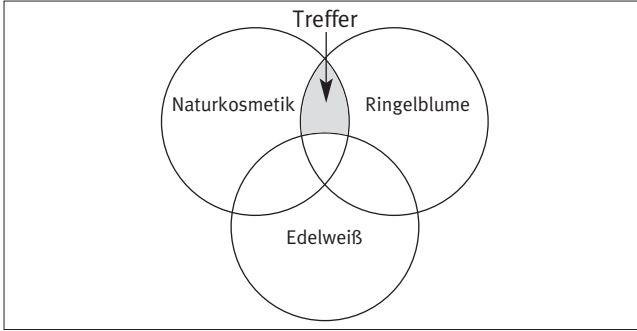


Abb. 3.10: Beispiel für eine Suchanfrage „Naturkosmetik AND Ringelblume AND NOT Edelweiß“

Als besonders hilfreich bei der Nutzung von (Meta-)Suchmaschinen erweist sich das Minuszeichen, da hiermit bestimmte Kriterien ausgegrenzt werden können. Dies können beispielsweise Dateitypen, Länderkennung (Top Level Domains, TLD) oder Schlagworte sein.

Beispielsweise werden Informationen zum Thema „Kundenzufriedenheit“ über die Suchmaschine www.vivisimo.com gesucht. Um Websites der Softwareanbieter mit so genannten CRM-Produkten auszugrenzen, wird das Minuszeichen verwendet. Die Suchanfrage liefert 58 Ergebnisse im deutschsprachigen Web.

Mit dem Minuszeichen lassen sich bestimmte Kriterien ausgrenzen

Abb. 3.11: Verwendung boolescher Operatoren

- Das Schlagwort „Kundenzufriedenheit“ liefert hingegen 205 Treffer und damit dreimal mehr potenzielle Quellen, welche auf ihren Informationsgehalt geprüft werden müssen.

The screenshot shows the Vivísimo search interface. At the top, there are navigation links for 'firma', 'produkte', 'lösungen', 'kunden', 'demos', 'partner', and 'presse'. A search bar contains the text 'Kundenzufriedenheit' and a dropdown menu is set to 'das Web'. To the right of the search bar are buttons for 'Suche', 'Erweitert', and 'Hilfe!'. Below the search bar, there are radio buttons for 'weltweit' and 'deutschsprachiges Web'. The main content area displays 'Clustering Ergebnisse' with a list of categories: 'Kundenzufriedenheit (205)', 'Firma (38)', 'Hersteller und Lieferanten (38)', 'Marketing (19)', 'Kundenbindung (13)', 'Consulting (16)', 'Marktforschung (12)', 'Hohe Kundenzufriedenheit (11)', 'Impressum, Einführungstour Schritt für Schritt Zinsen (11)', 'Software (7)', and 'Oberstes (10)'. Below this list is a search input field labeled 'Suchbegriff'. The search results section shows the first four results, each with a title, a brief description, and source information.

Abb. 3.12: Ergebnisse ohne Eingrenzung durch boolesche Operatoren

Option „in den Ergebnissen/Clustern“ suchen

- Darunter befinden sich elf Seiten, die sich ausschließlich mit CRM oder Software befassen. Diese Sortierung erfolgt durch die Option „in den Ergebnissen/Clustern“ suchen und wird von vielen Suchinstrumenten angeboten.

This screenshot shows the same Vivísimo search interface as the previous one, but with the search results filtered. The search bar still contains 'Kundenzufriedenheit'. The 'Clustering Ergebnisse' list on the left is now sorted by relevance, showing categories like 'Consulting (16)', 'Marktforschung (12)', 'Hohe Kundenzufriedenheit (11)', 'Impressum, Einführungstour Schritt für Schritt Zinsen (11)', 'Software (7)', 'Oberstes (10)', 'Werkstatt (6)', 'FUCUSA, Beratungsunternehmen für Mitarbeiter- und Kundenorientierung (4)', 'Care (4)', 'Qualitätsmanagement (6)', 'Analysen (6)', 'Training (6)', 'Artikel (3)', 'Mystery Shopping (2)', 'Controlling (3)', and 'Communication (3)'. The search results section shows the first three results, which are more specific to CRM and software, such as 'pki Software für Online Umfragen und Marktforschung', 'CFM - CFM - Center for Financial Services Marketing & Management', and 'Qlist.de - Der kostenlose Newsletter Versand'.

Abb. 3.13: In Suchergebnissen suchen

Eine Verfeinerung bieten die erweiterten Verknüpfungen, welche jedoch nicht von jedem Suchinstrument unterstützt werden. Nähere Hinweise zur Akzeptanz der Operatoren finden Sie in den jeweiligen Hilfeseiten.

Folgende Operatoren werden beispielsweise von wenigen Suchdiensten unterstützt:

- **ADJ (Adjacent):** Als Ergebnis nur Treffer, in welchen beide oder mehrere Suchbegriffe nebeneinander stehen.
- **BEFORE:** Das Ergebnis enthält nur Treffer, in welchen der eine Suchbegriff vor dem anderen erscheint. Auf diese Weise lassen sich Websites ausschließen, welche eine Aufzählung von Stichworten in den Keywords oder im (un-)sichtbaren Text (Body) enthalten, die keinen Bezug zum dargestellten Inhalt aufweisen.
- **FAR (wichtig):** Zwei durch FAR verknüpfte Begriffe stehen nicht mehr als 25 Worte auseinander. Diese Verknüpfung empfiehlt sich insbesondere dann, wenn Kombinationen der Schlagworte mit unterschiedlicher Aussage möglich sind. Beispielsweise könnte die Kombination „Kind“ und „Fahrrad“ folgende Inhalte als Treffer liefern: Fahrradfahren erlernen, Verkehrssicherheit, Verkehrserziehung, Warentest zu Kinderfahrrädern, Fahrradwanderung für Familien usw. – Wird jedoch Information zum Thema „Fahrradgröße bzw. Rahmenhöhe für Kinder“ gesucht, empfiehlt sich die Verknüpfung durch den Operator FAR.
- **ALLCAPS, NOPS, CAPS:** Alle Buchstaben, Kleinbuchstaben bzw. nur Großbuchstaben am Wortbeginn. Damit können Abkürzungen für Organisationen oder Techniken von Namen getrennt werden: beispielsweise „VERONIKA“ als Suchsystem versus „Veronika“ als Name.
- **PLURAL, SINGULAR:** Liefert Ergebnisse nach Deklination: „PLURAL JOB“ zeigt auch Ergebnisse zu „Jobs“ an.

*Verfeinerung durch erweiterte Verknüpfungen, die aber nicht jedes Suchinstrument unterstützt*

*erweiterte Operatoren*

## PRAXIS

### Merker: Suchoperatoren

- Formulieren Sie Ihren Informationsbedarf so exakt wie möglich;
- prüfen Sie die Suchbegriffe auf Homonyme;

- Ordnen Sie die Begriffe nach logischen Verknüpfungen;
- Schnittmengen (AND; +) enthalten die Ergebnisse, welche beide oder mehrere Suchbegriffe umfassen;
- Vereinigungsmengen (AND OR; ANY; |) enthalten alle Treffer mit einem oder mehreren der Suchbegriffe;
- Exklusivmengen (AND NOT; + -) liefern ausschließlich die Ergebnisse eines oder mehrerer Suchbegriffe ohne die anderen.
- Exklusionsmengen (XOR) umfassen alle Treffer, in welchen beide oder mehrere Suchbegriffe separat enthalten sind, jedoch nicht gemeinsam.
- Informieren Sie sich über weitere mögliche Suchoperatoren auf der jeweiligen Hilfeseite.

## 3.2 Informationsquellen

### 3.2.1 Websites

„Website“ steht für „Ort im Netz“ wie eine Seite im Buch, welches in einem Regal steht. Um das Indexieren und Katalogisieren von Websites als Voraussetzung eines erfolgreichen Suchmaschineneinsatzes verstehen zu können, ist ein Blick hinter die sichtbaren Kulissen notwendig:



Abb. 3.14: Sichtbare Information und unsichtbarer HTML-Code