

Diplomarbeit

# **Business Finder**

## **Unterstützung der regionalen Netzwerkbildung im www**

vorgelegt von  
**Benjamin Moos**  
Matr.-Nr.: 567792

**Betreut von:** Dipl. Inf. Tim Reichling  
**Erstprüfer:** Prof. Dr. Volker Wulf  
**Zweitprüfer:** Juniorprof. Dr. Volkmar Pipek

**Eingereicht am:** 20.09.2007

Diplomarbeit

# **Business Finder**

## **Unterstützung der regionalen Netzwerkbildung im www**

vorgelegt von  
**Benjamin Moos**  
Matr.-Nr.: 567792

**Betreut von:** Dipl. Inf. Tim Reichling  
**Erstprüfer:** Prof. Dr. Volker Wulf  
**Zweitprüfer:** Juniorprof. Dr. Volkmar Pipek

**Eingereicht am:** 20.09.2007

## **Eidesstattliche Erklärung**

Ich, Benjamin Moos, erkläre hiermit an Eides Statt, dass die vorliegende Diplomarbeit „Business Finder: Unterstützung der regionalen Netzwerkbildung im www“ von mir selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde. Die aus fremden Quellen übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder in gleicher noch in ähnlicher Form einer anderen Prüfungskommission vorgelegt oder veröffentlicht.

Siegen, den 20.09.2007

Benjamin Moos

# Danke

sage ich an dieser Stelle all den Personen, die dazu beigetragen haben, dass diese Arbeit realisiert werden konnte.

Mein Dank gilt insbesondere

- dem Betreuer der Arbeit, Tim Reichling, der mich fachlich beraten und unterstützt hat,
- Prof. Dr. Volker Wulf, der es mir ermöglichte, die Diplomarbeit innerhalb der Fachgruppe Wirtschaftsinformatik und Neue Medien zu schreiben,
- allen Mitarbeiterinnen und Mitarbeitern der Fachgruppe Wirtschaftsinformatik und Neue Medien für ihre Hilfsbereitschaft und ihre Unterstützung,
- den Beschäftigten der Abteilung Wirtschaftsförderung des Kreises Siegen-Wittgenstein Alexandra Schmitz-Kämpfer, Sabine Bartmann und Klaus Irle, die die Entwicklung des Business Finders im Kontext regionalpolitischer Maßnahmen ermöglichten,
- den Angestellten der Kommunalen Datenzentrale Westfalen-Süd in Siegen (KDZ) Georg Braß, Andreas Henning und Ralf Werthenbach, für ihr stetiges Bemühen, den Business Finder in das Umfeld der KDZ zu integrieren,
- den Interviewpartnern Helmut Bender, Thorsten Bölck, Helmut Föllmer, Hermann Grotmann, Jürgen Holdinghausen, Andre Klahold, Uwe Reinhold, Volker Sasse, Patrick Schulte, Oliver Schweißgut, Jan Soose, Marc Staiger, Rainer Steinke, Dr. Thomas Steffen, Karl-Eberhard Stötzel, Uwe Thomas, Dr. Dirk Thorsten Vogel, Karl-Helmut Wagener und Hans-Josef Weber, für die mir geopferte Zeit und die Bereitschaft, meine Fragen zu beantworten.



# Inhaltsverzeichnis

Abkürzungsverzeichnis .....	VI
Abbildungsverzeichnis .....	VII
Tabellenverzeichnis .....	VIII
<b>1 Einleitung .....</b>	<b>1</b>
<b>2 Stand der Forschung .....</b>	<b>4</b>
2.1 Regionale Netzwerke und deren Bedeutung .....	4
2.1.1 Definition des Begriffs „Cluster“ .....	4
2.1.2 Entwicklung und Förderung eines Clusters .....	6
2.1.3 Die Bedeutung von regionalen Netzwerken für die Wirtschaft .....	12
2.2 Aktionsforschung .....	17
2.2.1 Was ist „Action Research“? .....	17
2.2.2 Einfluss der Aktionsforschung in der Wirtschaftsinformatik .....	19
2.3 Die Rolle der IT in regionalen Netzwerken .....	21
2.4 Kataloge, Branchenbücher und Suchmaschinen .....	23
2.4.1 Internetbasierte Suchsysteme .....	24
2.4.2 Sonstige Suchsysteme und Informationsgewinnung .....	28
2.4.3 Funktionsweisen der Suchmaschinen .....	29
2.5 Einführung des ExpertFinding-Systems .....	34
2.6 Anwendung des ExpertFinding-Systems auf Unternehmen .....	36
<b>3 Forschungsfeld und Forschungsmethode .....</b>	<b>38</b>
3.1 Das Forschungsfeld .....	38
3.2 Methodisches Vorgehen .....	40
<b>4 Bedarfsanalyse .....</b>	<b>43</b>
4.1 Regionalität, Kooperation und Soziales Kapital .....	43
4.2 Erwartungen und Bedenken .....	44
4.3 Anforderungen .....	45
<b>5 Der Business Finder .....</b>	<b>47</b>
5.1 Konzept .....	47
5.2 Technische Grundlage: Das EF-System .....	48
5.2.1 Modularität .....	48
5.2.2 Profilerzeugung und Matching .....	49
5.3 Die Architektur des Business Finder .....	51
5.4 Implementierung .....	56
5.4.1 PHP-Funktionalitäten .....	56

5.4.2	<i>Erweiterungen des EF-Systems durch BF-Module</i> .....	59
<b>6</b>	<b>Evaluation</b> .....	<b>65</b>
<b>7</b>	<b>Zusammenfassung und Diskussion</b> .....	<b>70</b>
	<b>Literaturverzeichnis</b> .....	<b>74</b>
	<b>Anhang</b> .....	<b>81</b>
A	Interviewleitfäden.....	81
B	Lyz-Media-Breakfast Werbeflyer.....	85
C	Diagramme.....	86
D	CD-Rom .....	90

## Abkürzungsverzeichnis

AR	Action Research (engl. für Aktionsforschung/ Handlungsforschung)
ASP	Active Server Pages (serverseitige Scriptsprache für Webseiten)
BF	Business Finder
CIPM	Cluster Initiative Performance Model
EDV	Elektronische Datenverarbeitung
EF	ExpertFinding
FIT	Fraunhofer Institut Angewandte Informationstechnik
FuE	Forschung und Entwicklung
GUI	Graphical User Interface (Grafische Benutzerschnittstelle)
HTML	Hypertext Markup Language (Dateiformat für Internetseiten)
IDF	Inverse Document Frequency (inverse Dokumentenhäufigkeit)
IHK	Industrie- und Handelskammer
IKS	Informations- und Kommunikationssysteme
IKT	Informations- und Kommunikationstechnologie
IS	Informationssysteme
IT	Informationstechnologie
JSP	Java Server Pages (serverseitige Scriptsprache für Webseiten)
KMU	Klein- und mittelständische Unternehmen
LSI	Latent Semantic Indexing (engl. für schwache Bedeutungseinordnung)
OTE	Organisations- und Technikentwicklung
PHP	Hypertext Preprocessor (serverseitige Scriptsprache für Webseiten)
PPC	Pay-Per-Click
SMU	Small and medium units (engl. für KMU)
SSL	Secure Sockets Layer
TDM	Term-Dokumenten-Matrix
TF	Termfrequenz
URL	Uniform Resource Locator
VSM	Vector Space Model (engl. für Vektorielle Ähnlichkeitsfunktion)
WI	Wirtschaftsinformatik
WiFö	Wirtschaftsförderung
www	World Wide Web



## Abbildungsverzeichnis

Abbildung 1: Wertschöpfungskette Printmedien [angelehnt an Krät2001, S. 12] .....	5
Abbildung 2: Der Clusterlebenszyklus.....	7
Abbildung 3: Das Strategische Sechseck der Wirtschaftsförderung.....	9
Abbildung 4: Clusterinitiative Performance-Modell [nach Sölv2003, S. 25].....	11
Abbildung 5: Der Action Research Zyklus .....	19
Abbildung 6: Das STEPS-Modell [nach Floyd1989, S. 57] .....	20
Abbildung 7: Verhältnis zwischen statischen und dynamischen Webseiten.....	27
Abbildung 8: Beispiel des Vector Space Models .....	32
Abbildung 9: Entwicklungsprozess des Business Finder .....	41
Abbildung 10: Informationsgewinnung des Business Finder .....	47
Abbildung 11: Architektur des EF-Systems [nach Reic2005a, S. 65] .....	49
Abbildung 12: Die vier Komponenten des Business Finder .....	52
Abbildung 13: Matchingergebnis des Business Finder .....	54
Abbildung 14: Spezifische Einstellungen zur Profilgenerierung .....	55
Abbildung 15: Schlüsselwortprofil von Unternehmen B .....	55
Abbildung 16: Verarbeitung der Webservice-Rückgabe.....	58
Abbildung 17: Beispiel des Algorithmus zum Filtern interner Links .....	63
Abbildung 18: Gültige Links zu Unternehmensseiten der Region.....	68
Abbildung 19: Verhältnis von gültigen Links zu der Gesamtzahl der Links .....	69

## **Tabellenverzeichnis**

Tabelle 1: Verteilung von Medien-Unternehmen und Demografie.....	38
Tabelle 2: Auflistung der Interviewpartner .....	42
Tabelle 3: Platzierung bei verschiedenen Suchanfragen von Unternehmen L.....	67

# 1 Einleitung

*„Dieser zugleich unansehnliche und selbstbewusste Ort war die Stadt Grasse, seit einigen Jahrzehnten unumstrittene Produktions- und Handelsmetropole für Duftstoffe, Parfümeriewaren, Seifen und Öle: Guiseppe Baldini hatte ihren Namen immer mit schwärmerischer Verzückung ausgesprochen. Ein Rom der Düfte sei die Stadt, das gelobte Land der Parfumeure, und wer nicht seine Sporen hier verdient habe, der trage nicht zu Recht den Namen Parfumeur. [...] Nicht weniger als sieben Seifenkochereien machte Grenouille bei seinem Rundgang aus, ein Dutzend Parfümerie- und Handschuhmachermeister, unzählige kleinere Destillen, Pormadeateliers und Spezereien und schließlich einige sieben Händler, die Düfte en gros vertrieben.“ [Süsk1994, S. 211f]*

Die französische Stadt Grasse wird von Patrick Süskind in seinem Roman „Das Parfum“ als „das Rom der Düfte“ beschrieben. In der Stadt findet die Herstellung, die Verarbeitung und der Vertrieb von Parfümeriewaren statt. Was hier spezifiziert wird, ist die Ansiedlung von Unternehmen entlang einer Wertschöpfungskette. Diese Agglomeration ist ein ideales Fundament für das, was im Allgemeinen als *Cluster* bezeichnet wird. Neben der räumlichen Anordnung und der Anordnung der Unternehmen entlang einer Wertschöpfungskette beschreibt Süskind weitere typische Eigenschaften von regionalen Ballungszentren. So finden sich spezialisierte und gut ausgebildete Arbeitskräfte oder auch branchenfremde Betriebe, wie im angesprochenen Roman das Handwerk der Handschuhmacher, in regionalen Branchenzentren.

Cluster fördern die regionale Wirtschaft und bieten für teilnehmende Unternehmen wirtschaftliche Vorteile, so Michael E. Porter. Aus der Vielzahl an Referenzen [vgl. z.B. Saut2004, Wolf2002, Euro und Kete2006] auf Porters Publikationen ist zu schließen, dass er einer der Hauptvertreter von Clustertheorien ist. Ende des 20. Jahrhunderts trat zunehmend Kritik an seinen Theorien auf. Ursache hierfür ist die immer stärker werdende Präsenz des Internets und die zunehmende Digitalisierung. Mit immer schnelleren Kommunikationswegen und steigender Transparenz am globalen Markt vertreten Kritiker wie Larry Downes die Ansicht, dass räumliche Nähe in Zukunft eine untergeordnete Rolle spielt. Globalisierung, Digitalisierung und Deregulierung sind nach ihrer Auffassung die neuen Größen, die eine Wirtschaft vorantreiben [Down1998].

Diese Arbeit soll u.a. aufzeigen, dass neben rein wirtschaftlichen Faktoren auch soziale Faktoren dazu beitragen, einen lokalen Standort zu stärken. Es soll widerlegt werden, dass

Globalisierung und Digitalisierung zur Dezentralisierung führen. Außerdem wird begründet, dass diese Faktoren räumliche Nähe von Unternehmen begünstigen. Clustern ist eine große Bedeutung an wirtschaftlichen Nutzen und Innovationskraft zuzumessen. Wirtschaftsstandorte, in denen regionale Netzwerke existieren, sind gegenüber ihren Wettbewerbern im Vorteil. Deshalb gilt es diesen Vorzug zu erhalten und die Clusterbildung zu fördern.

*Kooperation* und *Konkurrenz* sind die zentralen Größen in einem Cluster. Die Kooperation resultiert aus der Spezialisierung der Unternehmen entlang einer Wertschöpfungskette. Aufgaben einzelner Unternehmen werden auf die für diese Aufgabe spezialisierten Unternehmen ausgelagert. Mit der Konzentration auf einen Aufgabenbereich können Unternehmen Produkte, Dienstleistungen und Prozessabläufe optimieren. Die räumliche Nähe im Cluster führt dazu, dass ansässige Unternehmen von solchen Innovationen schneller informiert werden als „verbindungslose“ Unternehmen. Konkurrierende Unternehmen müssen die eigenen Prozesse und Produkte verbessern, um am Markt bestehen zu können. Der hohe Informationsfluss führt zur Verbesserung der in einem Cluster teilnehmenden Unternehmen. Eine höhere Transparenz fördert somit die regionale Wirtschaft und sollte unterstützt werden.

Im Fokus dieser Arbeit steht die Transparenz in einem regionalen Cluster mit Hilfe von Informationstechnologie (IT) zu steigern und zu unterstützen. In Kooperation mit der Wirtschaftsförderung des Kreises Siegen-Wittgenstein und mit ansässigen Unternehmen der IT- und Medienbranche der Region Siegen sind die Anforderungen und der Bedarf an solch einem System evaluiert worden. Die Ergebnisse der Studie bilden die Grundlage für das zu entwickelnde System mit dem Namen *Business Finder* (BF). Die Basis bilden das in der Fachgruppe „Wirtschaftsinformatik und Neue Medien“ der Universität Siegen entwickelte ExpertFinding-System und das Branchenbuch der Wirtschaftsförderung, die Lüz-Media-Datenbank. Damit wird eine Verknüpfung von Branchenbuch und Recommender-System erreicht, welche die Vorteile beider Systeme nutzt. Das Resultat ist eine auf die Region Siegen bezogene Unternehmenssuchmaschine, die bei Suchanfragen dem Nutzer geeignete Unternehmen empfiehlt.

Im Rahmen dieser Arbeit wird weiter untersucht, ob das System die gewünschten Effekte der Clusterbildung und –förderung bewirkt.

Kapitel 2 beschreibt die Grundlagen dieser Arbeit. Es umfasst die Motivation und die Vorteile von regionalen Netzwerken sowie den Einfluss von IT auf Cluster. Die technischen Grundlagen des Business Finders werden erläutert. Neben der Analyse von bestehenden Suchsystemen wird ein Überblick über das ExpertFinding-System und die Übertragung dieses Systems von „Experten“ auf Unternehmen dargestellt.

In Kapitel 3 ist der Bezug zur Region Siegen-Wittgenstein und der ansässigen IT- und Medienbranche hergestellt. Dabei wird das methodische Vorgehen zur Entwicklung des Business Finders beschrieben. Die Ergebnisse der Gespräche mit den beteiligten Unternehmen werden in Kapitel 4 zusammengeführt und die Anforderungen an den Business Finder entwickelt.

In Kapitel 5 wird die Entwicklung des Business Finders dargestellt. Eine detaillierte Beschreibung der Anforderungen und deren technische Umsetzung sind Gegenstand dieses Kapitels.

Mit der Integration des Business Finders in die Umgebung der Wirtschaftsförderung und einer ersten Evaluation zum Nutzen des Systems befasst sich Kapitel 6.

Abschließend werden alle Ergebnisse in Zusammenhang mit der Unterstützung von Clustern durch Informationstechnologie zusammengefasst und Verbesserungspotentiale sowie Einsatzgebiete des Business Finders aufgezeigt.

## 2 Stand der Forschung

### 2.1 Regionale Netzwerke und deren Bedeutung

Regionale Netzwerke, auch *Cluster* genannt, gewinnen zunehmend in der Wirtschaft und Wissenschaft an Bedeutung. Während des Internetbooms und der Globalisierung haben regionale Kooperationen an Ansehen verloren. In den letzten Jahren hat sich die Auffassung durchgesetzt, dass gerade wegen der Globalisierung regionale Zusammenschlüsse von Bedeutung sind. Zu Beginn des 20. Jahrhunderts wurde die Auflösung von regionalen Branchenzentren durch Mechanisierung und Mobilität prognostiziert. Analog dazu, entstand zum Ende des vorigen Jahrhunderts eine ähnliche Sichtweise die die Globalisierung und Digitalisierung als Ursachen für Dezentralisierung sieht [Schi2003, S. 16]. Holger Schiele beschreibt die Euphorie, die die digitale Revolution Ende des 20. Jahrhunderts ausgelöst hat, ist aber der Meinung, dass Cluster auch in Zukunft von Bedeutung sind.

Cluster sind nicht die Universallösung wirtschaftlicher Probleme. Bieger und Scherer bezeichnen Cluster als „Modewort der Wirtschaftsförderung“ [Bieg2003, S. 9].

Dennoch werden regionalen Agglomerationen derzeit in Wirtschaft, Forschung und Politik beträchtliche Aufmerksamkeit geschenkt.

#### 2.1.1 Definition des Begriffs „Cluster“

Cluster (engl.), gleichbedeutend mit Anhäufung oder Ansammlung gleicher oder ähnlicher Einzelteile, steht im Zusammenhang mit regionalen Netzwerken für eine räumliche Ansammlung von Unternehmen und Instituten. Von dieser Grundlage ausgehend setzen sich Cluster aus einer Reihe von Akteuren zusammen, die laut Stefan Donhauser [Donh2006, S. 18] thematisch orientiert sind. Dabei verfolgen die Akteure in den Clustern das gemeinsame Ziel, die Innovationskraft der Unternehmen innerhalb der Cluster zu steigern, um den regionalen Standort im globalen Wettbewerb zu stärken. Durch kooperative Maßnahmen und Vernetzung wird versucht, diese Ziele zu verwirklichen. Michael E. Porter [Port2000, S 254] definiert Cluster als eine „geographisch benachbarte Gruppe von miteinander vernetzten Unternehmen und angegliederten Instituten eines speziellen Gebietes.“

*„A cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities. [...] Clusters take varying forms depending on their depth and sophistication, but most include endproduct or service companies; suppliers of*

*specialized inputs, components, machinery, and services; financial institutions; and firms in related industries.”*

Cluster unterscheiden sich in Tiefe und Details, weisen aber dennoch viele Gemeinsamkeiten auf. Ähnlich wie Cluster selbst existieren Definitionen des Begriffs, die sich in Detailfragen unterscheiden, aber in den Hauptmerkmalen übereinstimmen. Björn Sautter [Saut2004, S.66] kritisiert die Vielzahl an verschiedenen Definitionen und zielt damit besonders auf Wirtschaftsförderer und Wirtschaftsberater, die den Begriff Cluster selten klar definieren. Der Begriff „Cluster“ wird häufig als Sammelbegriff für unterschiedliche Konzepte und Maßnahmen verwendet. Sautter geht einen Schritt weiter und spricht in diesem Zusammenhang vom „Mythos“ Cluster. Damit beschreibt er die Annahme der Förderer und Berater, dass ein Cluster immer zur gesamtwirtschaftlichen Situation einer Region fördernd beiträgt und verweist auf die Vielzahl von Veranstaltungen und Tagungen zu Cluster-Strategien. Damit zeigt sich, dass für eine eindeutige Definition des Begriffs die Gemeinsamkeiten der verschiedenen Definitionen gefiltert werden müssen. Eine gute Basis bietet die von Porter oben erwähnte und in der Literatur oft verwendete Definition eines Clusters. Björn Sautter versucht, Porters Definition in einem Satz zu verwirklichen:

*„Zwei zentrale Merkmale definieren also ein Cluster: Die räumliche und sektorale Konzentration in einer Wertschöpfungskette.“*

Porters Erklärung findet sich u.a. in der Eigenschaft der sektoralen Konzentration wieder. Er nennt zwar in dieser Spezifizierung nicht explizit den Begriff der Wertschöpfungskette, hat diesen aber geprägt und beschreibt hier einzeln die Teilnehmer einer Wertschöpfungskette mit speziellen Anbietern, Produkten, Unternehmen, Maschinen und Abnehmern.

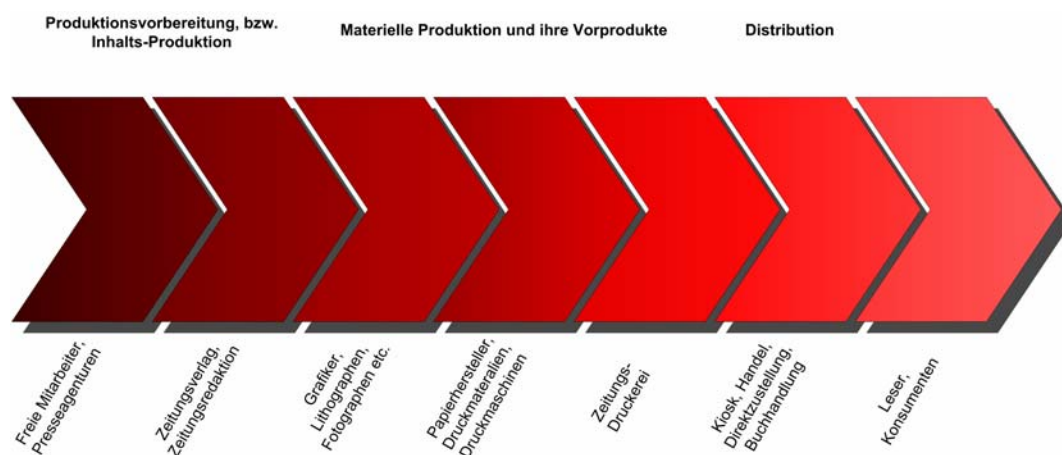


Abbildung 1: Wertschöpfungskette Printmedien [angelehnt an Krät2001, S. 12]

Die Teilnehmer der Wertschöpfungskette müssen nicht zwingend der gleichen Branche angehören. Abbildung 1 zeigt, dass entlang der Wertschöpfungskette für Printmedien auch branchenfremde Unternehmen, wie Hersteller von Papier und Druckmaschinen, ansiedeln.

Neben der sektoralen Konzentration ist die räumliche Konzentration für die Bildung eines Clusters Grundvoraussetzung. Dabei ist der Umfang des Clusters unbedeutend und besteht ohne räumliche Grenzen. Cluster können sich auf kommunaler, regionaler oder auf nationaler Ebene befinden [vgl. Saut2004, S.66].

Im Folgenden werden Cluster und synonym dazu „regionale Netzwerke“ als eine geographische Anhäufung von Akteuren entlang einer Wertschöpfungskette verstanden. Dabei verfolgen die einzelnen Akteure, die nicht zwingend der gleichen Branche angehören müssen, das gemeinsame Ziel, den regionalen Standort und somit das eigene Unternehmen zu stärken.

## **2.1.2 Entwicklung und Förderung eines Clusters**

### *2.1.2.1 Entwicklungsprozess eines Clusters*

Cluster entstehen durch unvorhersehbare zufällige Ereignisse und durchlaufen einen langfristigen Entwicklungsprozess [Saut2004, S67]. Die Versuche der regionalen Wirtschaftsförderung, neue Cluster ins Leben zu rufen, scheinen überflüssig. Bisher ist noch kein Fall bekannt, bei dem alleine durch regionalpolitische Maßnahmen ein Cluster initiiert wurde [vgl. Saut2004, S. 67]. Selbst Silicon Valley, das wohl bekannteste und erfolgreichste Mikroelektronik-Cluster, ist durch eine Reihe von Zufällen entstanden. Begonnen hat die Entwicklung des Valleys mit der Vermietung der Ländereien der Stanford University. Diese recht günstigen Gewerbeflächen führten zu den ersten Ansiedlungen heute renommierter Unternehmen. Durch das Hinzukommen des Nobelpreisträgers William B. Shockley und dessen im Alter zunehmender Verfolgungswahn führten dazu, dass acht seiner talentiertesten Mitarbeiter eigene Unternehmen gründeten. Eine ähnliche Entwicklung wiederholte sich in den darauf folgenden Jahren. Wissenschaftler verloren den Einfluss in ihren Unternehmen und gründeten neue Firmen [vgl. Male1998].

Die „Geburt“ eines regionalen Netzwerks kann nach Porter [Port1998a, S. 84] durch weitere Aspekte bedingt sein. Neben den erwähnten Zufällen entstehen Cluster an „historischen“



Orten, durch spezielle Anforderungen der Region, durch die Einflüsse eines bestehenden Clusters oder durch ein bzw. zwei innovative Unternehmen.

Das niederländische Transportcluster findet z.B. seine Wurzeln in Rotterdam. Der Seehafen und die vielen Wasserwege stützten die Entstehung des Clusters. Finnlands Schadstoffproblem führte zur Bildung eines Umweltclusters und in San Diego (U.S.A.) entstand aus dem Luftfahrtcluster ein Cluster für Golf-Equipment. Die technische Verbesserung von Metall in der Luftfahrtsindustrie bildete den neuen Grundstock für die Entwicklung von Golfschlägern.

Nach ihrer Entstehung durchlaufen Cluster weitere Entwicklungsphasen [Saut2004, S. 67 und Port1998b, S. 237-240]. Während die Anfänge eines Clusters mehr oder weniger durch Zufälle bedingt sind, so Porter 1998 [Port1998b, S 240], ist die Entwicklung vorhersehbarer. Porter unterscheidet die drei Stufen *Geburtsphase*, *Entwicklungsphase* und *Schrumpfungsphase* beim Lebenszyklus eines Clusters (vgl. Abbildung 2). In der Literatur lassen sich unterschiedliche Beschreibungen der Entwicklungsstufen finden. Diese sind an die von Porter beschriebenen Stufen angelehnt. Dabei wird meist eine von Porters Phasen unterteilt. Beispielsweise wird die Entwicklungsphase in eine „Entstehungsphase“ und eine „Wachstumsphase“ aufgeschlüsselt [vgl. Kete2006, S. 11]. Eine etwas größere Abänderung bietet die European Commission. In ihrem Bericht von 2002 [Euro2002, S. 15] stellen sie einen typischen Verlauf für die Entstehung und Entwicklung eines regionalen Clusters dar. Dabei wird der Entwicklungsprozess in sechs Abschnitte unterteilt. In Abbildung 2 zeigt sich, dass auch diese Stufen zu dem von Porter entwickelten 3-Stufen-Modell passen.

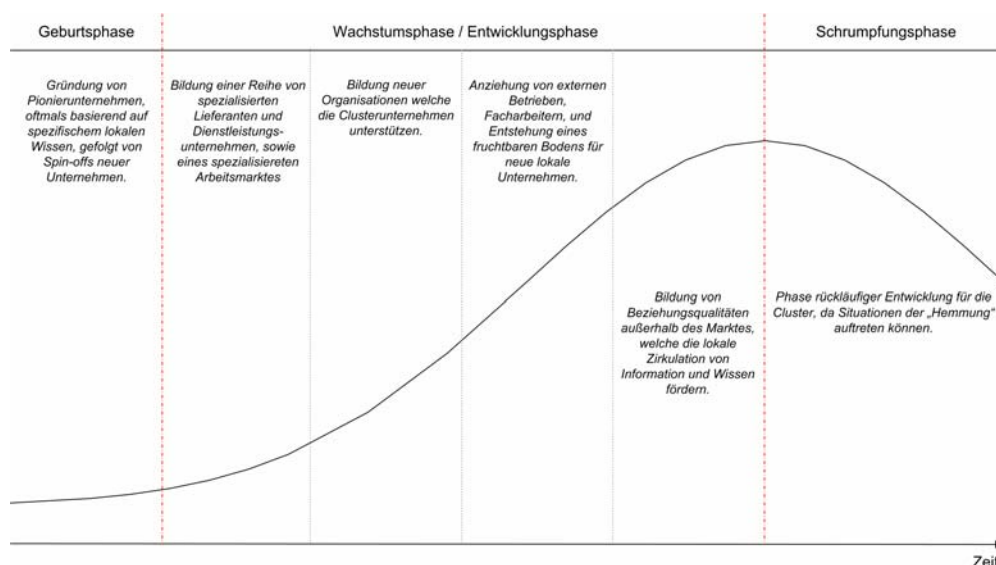


Abbildung 2: Der Clusterlebenszyklus

Für ein funktionierendes Cluster ist es wichtig, dass ein selbstlaufender Prozess entsteht. Je mehr Institute und Unternehmen die Wichtigkeit des Clusters erkennen, umso mehr spezialisierte Organisationen entstehen. Ist eine kritische Masse erreicht, steigt die Bedeutung für Organisationen, Investoren und Unternehmen, die zunächst nicht dem regionalen Zentrum angehören. Dieses unter dem Begriff „Netzeffekt“ bekannte Phänomen lässt sich mit Metcalfe's Gesetz<sup>1</sup> erklären: Der Wert eines Netzwerkes steigt mit dem Quadrat der Anzahl seiner Nutzer. Allgemeiner formuliert: Je mehr Akteure dem Netzwerk angehören, umso größer sein Nutzen. Je größer der Nutzen, desto größer die Anziehungskraft auf neue Akteure. Das heißt: Regionale Cluster müssen eine bestimmte Anzahl von teilnehmenden Unternehmen erreichen, um erfolgreich zu werden, so die Ergebnisse einer Studie der European Commission. Die Vielzahl unterschiedlicher Cluster geben aber keine Rückschlüsse auf eine perfekte Anzahl an Akteuren. Festzuhalten bleibt, dass mit steigender Anzahl an Unternehmen auch das Interesse fremder Investoren, Unternehmen und wirtschaftlichen Organisationen am Cluster wächst. Dies beinhaltet Spezialisierungen in Produktionen, Expertisen oder etwa im Finanzsektor. Das Hinzukommen fremder Teilnehmer, stellt Porter fest, kann das Wachstum eines Clusters beschleunigen. Der Begriff „Teilnehmer“ zielt nicht ausschließlich auf Organisationen und Unternehmen; auch einzelne Experten steigern den Wert des Clusters oder werden von diesem angezogen. 1960 wurde Medtronic, ein teilnehmendes Unternehmen des Medizin-Clustes in Massachusetts, auf eine Veröffentlichung von zwei Ingenieuren aufmerksam. Die Veröffentlichung handelte von der Entwicklung eines speziellen Herzschrittmachers. Medtronic arbeitete zur gleichen Zeit an einer ähnlichen Entwicklung. Die beiden Entwickler erkannten die Chance, ihre Idee besser mit einem renommierten Unternehmen aus einem wachsenden Cluster vermarkten zu können [vgl. Port1998b, S. 242].

### 2.1.2.2 Förderungsmaßnahmen

Cluster benötigen nach Porter [Port1998a, S. 85] mehrere Jahrzehnte, um ihren wirtschaftlichen Vorsprung zu entwickeln. Dies ist ein Grund, warum der politische Versuch, Cluster zu gründen, nicht funktioniert. Entsprechend lange kann es dauern, bis die Schrumpfungsphase einsetzt. Fördermaßnahmen während der Wachstumsphase wirken wie der Turbolader eines Autos und helfen Clustern, sich schneller zu entwickeln [vgl. Sölv2003, S. 12]. Welche Maßnahmen zur Förderung stattfinden, ist abhängig von dem jeweiligen

---

<sup>1</sup> Metcalfe, Robert (\* 1946)

Cluster und dem Entwicklungsstatus des Clusters. Maßnahmen, die bei einem Cluster zum Erfolg geführt haben, können bei anderen Clustern das genaue Gegenteil bewirken. Erfolgreiche regionale Wirtschaftsförderung basiert auf dem Verständnis für die Stärken und Schwächen einer Region [Saut2004, S.70].

Zunächst zeigt sich, dass eine große Zahl von Einflussfaktoren auf die Entscheidungen und die Ausrichtung der Wirtschaftsförderung einwirkt. Nach Stefan Leuninger und Holger Held [Leun2003, S. 162] existieren sechs Faktoren, welche die Orientierung und Durchführung der Wirtschaftsförderung beeinflussen (vgl. Abbildung 3).



**Abbildung 3: Das Strategische Sechseck der Wirtschaftsförderung**

Die genannten Faktoren beziehen sich allgemein auf wirtschaftsfördernde Maßnahmen. Deshalb können sie ebenfalls auf Förderungsversuche für Cluster angewandt werden. Das Wichtigste für die Wirtschaftsförderung ist dabei die zu fördernde *Zielgruppe*. Im Clusterfall sind das die einzelnen Akteure des Clusters. Je nach Art, Größe und Standort muss ein unterschiedliches Angebot an Beratung, Service und Unterstützung zur Verfügung gestellt werden. Ebenfalls lokale *politische Akteure* wirken auf die Entscheidungen der Wirtschaftsförderung ein. Dabei treten oftmals kommunikative Probleme auf, so Leuninger und Held. Weitere *konkurrierende Standorte* führen dazu, dass es notwendig ist, fördernde Maßnahmen abzuändern, um sich am globalen Markt etablieren zu können. Selbst innerhalb des Clusters ist eine Absprache mit weiteren gegebenenfalls *konkurrierenden wirtschaftsfördernden Instituten* wichtig. Umwelteinflüsse, speziell sich *ändernde Trends* in

Technologie und Innovation aber auch die Globalisierung führen zu sozioökonomischen und konjunkturellen Veränderungen. An diese Veränderungen muss sich nicht nur das Cluster anpassen, sondern auch die *Personen der Wirtschaftsförderungen*. Kenntnisse über z.B. europäische Förderprogramme und neue Rechtslegungen werden abverlangt. Initiativen der Clusterbildung sollten diese Faktoren berücksichtigen.

Nach Sölvell et. al. [Sölv2003, S. 13] finden Länder, in denen bereits Clusterinitiativen bestehen, einen leichteren Zugang zu neuen erfolgreichen Initiativen. Als Clusterinitiative werden organisierte Bemühungen verstanden, die das Wachstum sowie die Wettbewerbsfähigkeit fördern [Sölv2003, S. 31]. Sölvell et. al. haben die Förderungsmöglichkeiten in einem Modell vereint und es mit „Cluster Initiative Performance Model“ (CIPM) betitelt. Dieses Modell besteht aus drei antreibenden Größen: Dem sozialen, politischen und ökonomischen Umfeld, den Zielen der Clusterinitiative sowie dem Entwicklungsprozess der Clusterinitiative. Diese drei Kräfte wirken auf den Erfolg der Clusterinitiative, der vierten Größe im CIPM. Jede der vier Komponenten setzt sich aus einzelnen Faktoren zusammen [vgl. Sölv2003, S.9], die im Folgenden beschrieben werden.

Das Geschäftsumfeld wird vom lokalen und nationalen Umfeld beeinflusst. Diese Beeinflussung findet sich in den Unternehmensphilosophien sowie den regionalen kulturellen Einstellungen und Bemühungen bezüglich Clusterpolitik wieder [vgl. Euro2002, S. 44]. Für die Initiative spielt die Vergangenheit des Clusters, die Konkurrenz, die Wettbewerbsfähigkeit, der Technologiestand sowie Angebot und Nachfrage eine wichtige Rolle. Diese Punkte treffen eine Aussage über die Stärken des regionalen Netzwerks und müssen entsprechend auf Förderungsmaßnahmen abgestimmt werden.

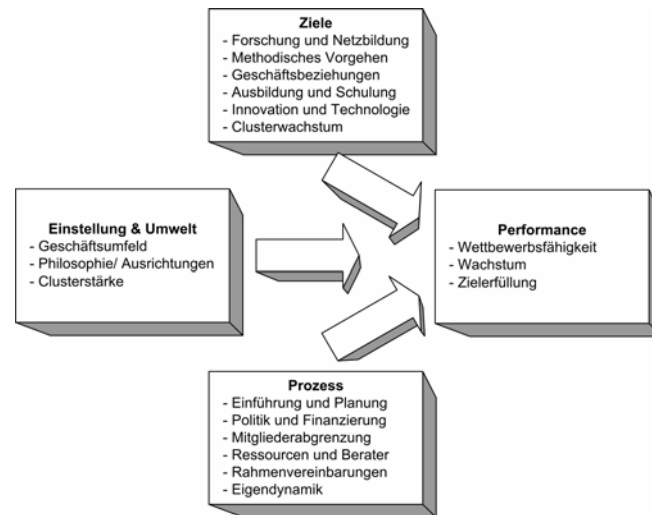


Abbildung 4: Clusterinitiative Performance-Modell [nach Sölv2003, S. 25]

Nach Abbildung 4 lassen sich die *Ziele* in sechs Kategorien unterteilen. Netzbildung gehört bei allen Clusterinitiativen zu den häufigsten Unterstützungsmaßnahmen. Deshalb wird vielfach der Begriff des „regionalen Netzwerkes“ mit einem Cluster gleichgesetzt. Dabei geht es um „Information-Sharing“ innerhalb des Clusters, Bekanntmachung durch Berichte oder das Veranstellen von Seminaren, um die Kooperation zu fördern. Die Stärkung der Verbindung zu Forschungseinrichtungen und politischen Instituten gehört ebenfalls zu den häufigeren Handlungen einer Clusterinitiative. Seltener finden Maßnahmen im technischen Bereich statt. Viele Aktivitäten zielen auf die Expansion eines Clusters. Maßnahmen mit Werbung und Veröffentlichungen sollen potenziellen Akteuren die Wichtigkeit und Marktrelevanz des Clusters demonstrieren.

Bei der Durchführung einer Clusterinitiative zeigt sich bereits bei der Einführung und Planung, ob der Initiator aus dem wirtschaftlichen Sektor des Clusters kommt oder politische Maßnahmen hinter der Initiative stehen. Entscheidungen über die Finanzierung und politischen Anweisungen sowie die Grenzen des Clusters werden getroffen. Jede Clusterinitiative benötigt eigene organisatorische Ressourcen und entsprechende Berater. Ähnlich wie bei einem Cluster selbst ist es notwendig, dass die Förderung eine gewisse Eigendynamik entwickelt und zum Selbstläufer wird [vgl. Saut2004, S. 70].

Die Leistungsfähigkeit einer Clusterinitiative ist nach Abbildung 4 durch drei Größen bestimmbar: Erstens kann durch Innovation und technischen Fortschritt die Auswirkung der Initiative gemessen werden, zweitens ist die Wachstumsrate des Clusters eine wichtige Größe und drittens werden durch die Kontrollen der Zielerfüllung die Aussagen über den Erfolg oder Misserfolg der Clusterinitiative getroffen.

Es ist zu berücksichtigen, dass es sich bei allen Förderungsmaßnahmen um Bemühungen handelt, ein bereits bestehendes Cluster zu fördern. Ein neues Cluster „künstlich“ ins Leben zu rufen, ist nach Auffassung der Wissenschaft bisher nicht gelungen. Michael E. Porter ist der Meinung, dass die Regierung und der private Sektor nicht versuchen sollen, neue Cluster zu kreieren. Sie sollen sich auf die Unterstützung existierender und entstehender Cluster konzentrieren [Port1998a, S. 89]. Weiter ist Porter der Auffassung, dass es sich immer lohnt, ein Cluster zu unterstützen, egal welchem Wirtschaftszweig es angehört oder wie alt es ist. Er gründet seine These darauf, dass Cluster meistens die Produktivität und Löhne steigern. Die Förderungsmaßnahme selbst ist aber nur durch Marktstudien und nicht durch Politik entscheidbar ist.

### **2.1.3 Die Bedeutung von regionalen Netzwerken für die Wirtschaft**

Wie oben beschrieben gewinnen regionale Standorte trotz oder gerade wegen der Globalisierung an Bedeutung. Porter bezeichnet in [Port1998b, S. 236] dieses Phänomen mit „The Location Paradox“.

Warum siedeln Unternehmen eines bestimmten Sektors in der gleichen Region, wenn schnelle Kommunikationswege und Transportmöglichkeiten Nähe unbedeutend erscheinen lassen? Geringere Transport- und Reisekosten reichen nicht aus, um dies befriedigend zu beantworten. Stefan Donhauser [Donh2006, S. 18] beschreibt die Anstöße zur Zentralisierung durch die Steigerung der Innovationskraft, somit der Produktivität und des Wirtschaftswachstums. Dabei legt Donhauser zugrunde, dass Unternehmen miteinander kooperieren und somit diese Entwicklung hervorgerufen wird. In der Literatur lassen sich zwei Kategorien feststellen: Die ökonomisch monetären- und die sozialen Ansätze. Während die ökonomischen Ansätze zum Teil gegen regionale Agglomerationen sprechen, zeigt die soziale Sichtweise die Bedeutung von Lokalität in sozialen Netzwerken auf. Dennoch sind beide Ansätze im Zusammenhang zu betrachten. So kann z.B. der gute persönliche Kontakt zu einem Lieferanten Geschäftsprozesse optimieren, die dadurch zu Kostensenkung führen. Michael E. Porter [Porter2000, S. 259] unterscheidet zwischen drei wesentlichen Faktoren, die zu einem Wettbewerbsvorteil führen:

- Steigerung der statischen Produktivität durch
  - Zugang zu spezialisierten Vorprodukten und ergänzenden Gütern,
  - Zugang zu spezialisierten Facharbeitern,
  - Zugang zu Wissen und Information,
  - Zugang zu Instituten und öffentlichen Gütern und

- das Bilden von Anreizen.
- Steigerung der Innovation und der Produktivität.
- Bieten eines Anreizes, neue Unternehmen zu gründen und das Cluster zu expandieren.

Unter dem Begriff der Produktivität wird im Allgemeinen eine wirtschaftliche Größe verstanden, die das Verhältnis zwischen der Ausbringungsmenge eines Faktors und der dazu erbrachten Einsatzmenge beschreibt. In der Auflistung zeigt sich der Einfluss nicht monetärer Größen auf die statische Produktivität. Aber auch die Steigerung von Innovation und das Bilden von Anreizen verweisen auf soziale Einflüsse. Beziehungen zwischen Menschen führen zu Wissensaustausch, Kommunikation, Kooperation und Konkurrenzverhalten. Im Folgenden wird der Einfluss dieser Größen auf Innovation und Produktivität bzw. die Wirtschaftlichkeit beschrieben.

### *2.1.3.1 Wissensaustausch unter den Akteuren*

Ein wichtiger Punkt auf dem Weg in die Dienstleistungsgesellschaft ist der Informations-, Wissens- und Ideenaustausch [vgl. Vorwort Lothar Späth in Schi2003, S. 9]. Aber entgegen der Annahme, dass durch das Internet und der damit verbundenen Digitalisierung der Kommunikationswege persönliche Zusammentreffen unwichtig werden, zeigt sich, dass nicht alles Wissen digital übermittelt werden kann.

Im Allgemeinen werden zwei Arten von Wissen unterschieden: Explizites und implizites Wissen [vgl. Pola1946, S. 21-41]. Bei explizitem Wissen handelt es sich um ausgesprochenes, formuliertes und dokumentiertes Wissen, welches in verschiedenen Formen wie Datenbanken, Schriften, Plänen, Karten usw. vorliegen kann. Eine Person, die aufgrund von Erfahrungen, Geschichte, Praxis und des geführten Lebens Wissen im Sinne von „Know-how“ aufgebaut hat, verfügt über sogenanntes implizites Wissen [Will1998, S. 12]. Diese Person muss nicht die Kenntnis um ihr Wissen haben. Deshalb muss sie dieses Wissen auch nicht erklären können. Folglich ist implizites Wissen schwer dokumentier- und kontrollierbar. Jeder einzelne Mitarbeiter eines Unternehmens besitzt „Know-how“, dessen er sich entweder bewusst ist oder nicht. In persönlichen Kontakten überträgt er dieses Wissen - bewusst oder unbewusst - an seine Mitmenschen. Dies kann innerhalb eines Unternehmens geschehen, z.B. in einer Meister - Lehrlingsbeziehung, bei der der Meister sein angesammeltes Wissen dem Lehrling vermittelt. Interessanter und für die Unterstützung von lokalen Unternehmenszentren ist aber der Kontakt zu Personen anderer Unternehmen. Durch räumliche Nähe entwickeln sich

persönliche Kontakte, z.B. bei Freizeitaktivitäten, in Pausen oder beim Mittagessen im Schnellimbiss. Die in einer Organisation als „Flurgespräche“ bekannte Kommunikation hat positive Auswirkungen auf einen Ballungsraum.

Zusätzlich zum übertragenen Wissen werden Informationen über Innovationen weitergegeben. So werden neue Ideen schneller verbreitet und die ansässigen Unternehmen und Angestellten profitieren davon. Gespräche zwischen Mitarbeitern fördern die Innovation und Produktivität.

Für die Unternehmen bedeutet dies, dass sie spezialisierte Arbeitskräfte aus dem „Labour-Pool“ [Schi2003, S. 24] ziehen können und durch die Fähigkeiten der Arbeitskräfte schneller und leichter lernen, neue innovative Technologien zu entwickeln und zu nutzen [vgl. Port1998a, S. 83]. Isolierte Unternehmen erfahren später als Clusterunternehmen von Innovationen. Die rein geografischen Effekte werden in Europa zusätzlich durch nationale Grenzen verstärkt [Schi2003, S. 24]. „Inselunternehmen“, die auf sich alleine gestellt sind, haben somit einen Nachteil gegenüber ihren Wettbewerbern, da sie für Innovation alleine verantwortlich sind. Der Lernfaktor von anderen Unternehmen, die ähnliche Probleme bearbeiten, ist bei Inselunternehmen geringer. Ein ideales Umfeld für Innovationen ist eine räumliche Ansammlung von Organisationen, in der eine hohe Transparenz herrscht. Unternehmen kennen sich gegenseitig, haben Kenntnisse über die Innovationen anderer Akteure und können beurteilen, ob eine andere Organisation vertrauenswürdig ist [Schi2003, S. 19].

### 2.1.3.2 Soziales Kapital

Für eine gute Verteilung von Informationen ist eine gewisse Vertrauensbasis zwischen den Akteuren notwendig [vgl. Port1998, S. 81]. Als Beispiel sei hier das „Verbundprojekt der Region Siegen-Wittgenstein und Olpe“ angeführt. Teilnehmer des Projekts waren ebenfalls Gesprächspartner der Bedarfsanalyse (Kapitel 4). Die einzelnen Akteure bestätigten in Gesprächen unabhängig voneinander, dass durch den überschaubaren persönlichen Rahmen ein gestärktes Vertrauen zu den anderen Akteuren existiert. Der persönliche Kontakt zu den Teilnehmern spielt dabei eine wichtige Rolle, um eine Vertrauensbasis aufzubauen. Aussagen von Akteuren in späteren Interviews bestätigen das Vertrauensverhältnis zwischen den Teilnehmern des Verbundprojektes.

Im Zusammenhang mit Vertrauen und Netzwerken wird zunehmend der Begriff des *Sozialen Kapitals* in der Literatur verwendet [vgl. Eule2006, S. 13]. Seine erste Verwendung geht auf Lyda Judson Hanifan zu Beginn des 20. Jahrhunderts zurück. Diese Definition findet in der



Forschung jedoch keine weitere Verwendung. [vgl. Müll2006, S. 35]. Der häufige Gebrauch des Begriffs führte zu einer Vielzahl an Definitionen [vgl. Eule2006, S. 15]. Loury definiert Sozialkapital als die Menge der Ressourcen, die in Familienbeziehungen und in sozialen Organisationen der Gemeinschaft enthalten sind und die die kognitive oder soziale Entwicklung eines Kindes oder Jugendlichen fördern. [vgl. Eule2006, S. 12f] Bourdieu führt hingegen den „Sozial-Kapital-Begriff“ neben kulturellem- und ökonomischem Kapital als eine von drei Kapitalformen ein. Dabei handelt es sich um nutzbare Größen, die auf der Zugehörigkeit zu einem Netzwerk beruhen [Bour1997, S. 63]. Robert D. Putnam sieht im Sozial-Kapital-Begriff dagegen einen Wert, z.B. Freunde, Bekannte oder Familie, der in „Krisensituationen“ zum eigenen Vorteil genutzt werden kann. Weiter überträgt Putnam diese Annahme auf ganze Netzwerke, die Vorteile aus dem sozialen Beziehungsgeflecht ziehen [Putn2001, S19f].

Es lassen sich grundsätzlich zwei Betrachtungsweisen des Sozialkapitalbegriffs unterscheiden, für die jeweils Bourdieu und Putnam stehen. Nach Bourdieu ist soziales Kapital „eine individuell nutzbare Ressource“, während Putnam in dem Begriff das Zusammenspiel von sozialen Netzwerken und ganzen Gesellschaften sieht [vgl. Eule2006, S.16].

In einem Punkt stimmen die Definitionen dennoch überein: Das Erschaffen von sozialen Netzwerken ist mit Bemühungen verbunden, die den späteren Fortbestand sichern [vgl. Beck2004, S. 333]. Ein Beziehungsnetzwerk ist durch die gezielte Schaffung und Erhaltung sozialer Kontakte, die einen unmittelbaren Nutzen versprechen geprägt [Bour1997, S. 65]. Welcher Ansatz auch gewählt wird: Die Bedeutung des Sozialen Kapitals für die Wirtschaft und somit für regionale Ballungsräume bleibt unumstritten.

Das Konzept des Sozialen Kapitals und des Vertrauens hilft, eine Erklärung zu finden, warum bedeutende ökonomische Aktivitäten dazu tendieren, sich in regionalen Netzen zu vereinigen.

*„It [Social Capital] becomes more valuable as the process of globalization continues; it is not equally available in all communities; it cannot be purchased or transferred; and it is difficult to imitate or replicate. Trust, as a component of social capital, helps overcome market failures or reduce the level of market costs for firms in densely related networks, by supporting stable and reciprocal exchange relationships among them” [Wolf2002, S. 11].*

Soziales Kapital ist demnach, ähnlich implizitem Wissen, ein schwer transportierbares Gut. Es lässt sich weder vervielfältigen noch kopieren und trägt somit zur Bedeutung von regionalen Agglomeration bei. Nicht nur aus diesem Grund ist Soziales Kapital für

Ballungsräume wichtig. Soziales Kapital führt durch das gesteigerte Vertrauen zu einer höheren Bereitschaft sich gegenseitig zu unterstützen, Verbindungen (Geschäftsbeziehungen) einzugehen und, wie bereits im Kapitel 2.1.3.1 beschrieben, Wissen auszutauschen. Die gestärkten Beziehungen zwischen den Akteuren tragen schließlich durch eine vertraute und offenere Kommunikation zur Optimierung von Produktions- und Innovationsmöglichkeiten bei.

### 2.1.3.3 Konkurrenz und Transparenz

Am Ende der Wertschöpfungskette werden Anbieter oftmals durch anspruchsvolle Kunden dazu gezwungen, ihre Produkte zu verbessern. Dieser Effekt lässt sich entlang der Wertschöpfungskette verfolgen. Somit sind alle Unternehmen, die an der Wertschöpfungskette beteiligt sind, in den Innovationsprozess einbezogen [vgl. Port1998, S. 83]. Die Unternehmen können durch die Nähe zum Markt schneller agieren und entsprechend schneller auf die Wünsche ihrer Kunden eingehen. Die Transparenz im Netzwerk führt somit zu Fortschritt und Entwicklung neuer Technologien.

Neben dem Effekt des stärkeren Informationsaustausches steigt innerhalb des Clusters durch die Vielzahl ähnlicher Unternehmen und durch die Transparenz der Konkurrenzdruck.

*„Competition in today’s economy is far more dynamic. Companies can mitigate many input-cost disadvantages through global sourcing, rendering the old notion of comparative advantage less relevant. Instead, competitive advantage rests on making more productive use of inputs, which requires continual innovation. [...] Without vigorous competition a cluster will fail.” [Port1998a, S. 78]*

Michael E. Porter hat erkannt, dass Kooperation und Konkurrenz innerhalb eines Clusters nebeneinander existieren können. Vielmehr notiert er, dass beides für ein erfolgreiches Cluster notwendig ist. Der Spruch „Konkurrenz belebt das Geschäft“ spiegelt die Bedeutung des Wettbewerbsdrucks wider. Durch bessere Transparenz wird den Unternehmen bewusst, welche innovativen Fortschritte die Konkurrenz unternimmt. Um nicht aus dem Markt gedrängt zu werden, ist es deshalb notwendig, im Fluss der Innovationen „mit zu schwimmen“ und selbst Verbesserungen der eigenen Produktivität vorzunehmen. Dadurch führen Cluster schneller zu neuen Ideen und zu technischem, organisatorischem und anderem Fortschritt.

Die Ansiedlung der Unternehmen entlang der Wertschöpfungskette ergibt eine ausgeprägte Arbeitsteilung. Unternehmen im regionalen Netzwerk gliedern Aufgaben auf entsprechende spezialisierte Anbieter aus [vgl. Schi2003, S. 21]. Verbindungslose Unternehmen müssen

hingegen alle Aufgaben alleine bewerkstelligen. Dadurch ergeben sich Maschinenfehlzeiten, ein schlechterer Zugang zum Markt und ein geringerer Informationsfluss, so dass diese Unternehmen unwirtschaftlicher arbeiten als die Konkurrenz aus dem Cluster. Clusterunternehmen kennen die Preise und Kosten ihrer Zulieferer und können somit Druck ausüben. Die Anbieterunternehmen sind gezwungen, durch die Vielzahl an Anbietern wirtschaftlicher zu arbeiten, um die Preise zu senken und ihre eigene Produktivität zu verbessern. Der Wettbewerbsdruck in einem Cluster führt somit zu höherer Qualität bei Zwischenprodukten und trägt zur Qualität des Endproduktes bei. Für die Anbieter führen die hohen Qualitätsansprüche zum Zwang, neue technische Möglichkeiten zu entwickeln, um den Ansprüchen der Kunden gerecht zu werden. Diese Innovation trägt dazu bei, dass sich diese Unternehmen am globalen Markt etablieren können.

## 2.2 Aktionsforschung

Es ist bisher bewusst darauf verzichtet worden, regionale Cluster als Agglomeration von Unternehmen zu sehen. Erstens, weil neben dem regionalen Aspekt der sektorale Aspekt nicht außer Acht gelassen werden darf, und zweitens, weil über die wirtschaftliche Seite hinaus Politik und Wissenschaft für die Entwicklung des Clusters mitverantwortlich sind. Eine Kooperation zwischen diesen Akteuren ist demnach unumgänglich. Wie findet sich aber die Wissenschaft in der Wirtschaft ein, wo praktizierende Unternehmer kritisieren, dass die aktuelle Wissenschaft mehr und mehr an Relevanz für die Wirtschaft verliert [vgl. Susm1978, S. 582]? Begründet ist diese Annahme durch den geringen Anteil der Forschungseinrichtungen, die sich aktiv am Praxisalltag von Unternehmen beteiligen. Um dem entgegenzuwirken und der Wissenschaft einen näheren Bezug zur Praxis zu geben, wird *Aktionsforschung* betrieben.

### 2.2.1 Was ist „Action Research“?

„Action Research“ (AR) hat in der deutschen Sprache eine Anzahl von Synonymen. So finden sich in der Literatur neben der eingängigen „Aktionsforschung“ Übersetzungen wie „Handlungsforschung“ und „Tat-Forschung“. Die erste Verwendung des Begriffs geht auf Kurt Lewin im Jahre 1946 zurück [Susm1978, S. 586]. Sieben Jahre später schreibt er:

*„Die für die soziale Praxis erforderliche Forschung lässt sich am besten als eine Forschung im Dienste sozialer Unternehmungen oder sozialer Technik kennzeichnen. Sie ist eine Art Tat-Forschung (,action research'), eine vergleichende Erforschung der Bedingungen und Wirkungen verschiedener*

*Formen des sozialen Handelns und eine zu sozialem Handeln führende Forschung“ [Lewi1953, S. 280].*

Parallel zu Lewin entstand in Großbritannien am Tavistock-Institut ein weiterer Zweig, der sich mit einer ähnlichen Problematik beschäftigte. Neben seiner Studie über Tavistock hat Rapoport 1970 seine eigene Definition aufgestellt:

*„Action research aims to contribute both to the practical concerns of people in an immediate problematic situation and to the goals of social science by joint collaboration within a mutually acceptable ethical framework“ [Rapo1970, S. 499].*

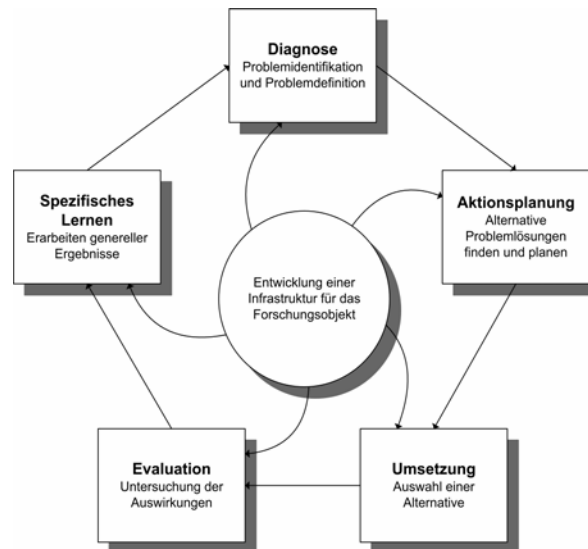
Auf die freie Marktwirtschaft übertragen beteiligt AR sowohl Forschung als auch Unternehmen an den Untersuchungsergebnissen. Susman und Evered [Susm1978, S. 588] fügen diesen beiden Perspektiven die Entwicklung selbsthelfender Kompetenzen als dritte Komponente hinzu. Dies soll der Entwicklung von Eigeninitiative dienen, um so den Teilnehmern des Forschungsobjektes die Weiterentwicklung auch ohne den Forscher zu ermöglichen.

AR darf nicht als eine umfassende Forschungsmethode gesehen werden, sondern als ein Ansatz, der verschiedene Vorgehensweisen beinhaltet. Die häufigste dieser Methoden ist die *teilnehmende Handlungsforschung*. Die Hauptaufgaben des Aktionsforschers liegen dabei in der Untersuchung der komplexen sozialen Prozesse. Diese lassen sich am Besten analysieren, indem der Forscher selbst am Forschungsobjekt beteiligt ist und Änderungen innerhalb des Systems herbeiführt. Die Auswirkungen dieser Änderungen werden vom System und von ihm registriert und weiter beobachtet [Bask1999, S.2].

In diesem Zusammenhang erstellen Susman und Evered den Zyklus der Aktionsforschung (Abbildung 5). Dieser ist in die fünf Phasen Diagnose, Aktionsplanung, Umsetzung, Evaluation und spezifisches Lernen unterteilt. Der Zyklus wird während eines Forschungsprojektes mehrmals durchlaufen. Nach der Identifikation eines Problems wird vom Forscher gemeinsam mit dem Client-System<sup>2</sup> in einer Planungsphase nach möglichen Lösungen gesucht. Die Auswirkungen der durchgeführten Maßnahme werden beobachtet und untersucht. Als letzter Schritt in einem Zyklus werden Schlüsse aus den Ergebnissen gezogen.

---

<sup>2</sup> Mit Client-System bezeichnen Susman und Evered [Susm1978, S. 588] das soziale System, in dem Probleme mittels dem AR-Ansatz gelöst werden sollen.



**Abbildung 5: Der Action Research Zyklus**

In der Aktionsforschung werden verschiedene Ansätze praktiziert die in einem Punkt übereinstimmen: Der Forscher wird in das Client-System einbezogen. Dadurch ist es ihm möglich, an der Praxis teilzunehmen, einen Einblick ins System zu gewinnen und eine praxisorientierte Forschung und Problemlösung zu betreiben. Dadurch grenzt sich die Aktionsforschung von traditionellen Forschungsmethoden ab [vgl. Lewi1953, S. 280].

### 2.2.2 Einfluss der Aktionsforschung in der Wirtschaftsinformatik

Die Wirtschaftsinformatik (WI) ist ein interdisziplinärerer Forschungsbereich. Sie setzt sich nicht nur aus den Bereichen Wirtschaft und Informatik, wie dem Wortstamm nach zu vermuten ist, zusammen, sondern beinhaltet zusätzlich die Bereiche Hardware-Engineering, Informations- und Kommunikationstechnik (IKT), Organisationswissenschaft sowie ethnographische und soziologische Ansätze. Am deutlichsten wird dies bei der Interaktion von Organisationen mit Informations- und Kommunikationssystemen (IKS) [Bask1999, S. 6]. So arrangieren IKS Menschen und Technologie, um wirtschaftliche Aufgabenstellungen zu lösen [Fran1998, S. 75]. Der wichtigste Punkt der Aktionsforschung ist dabei, die Organisationen als Ganzes zu betrachten. Eine Aufteilung in soziale Aspekte, wie die Organisation selbst und ihre Informationstechnologie, genügt nicht, um ausreichend Informationen über die Organisation als Ganzes zu gewinnen [Bask1999, S 2].

Die Softwareentwicklung hat in den letzten 40 Jahren eine entsprechende Veränderung durchlaufen. 1970 hat Royce erstmals das Wasserfall-Modell vorgestellt [Royc1970, S. 389]. Dieses statische Modell, das „wasserfallmäßig“ von einer Phase in die nächste führt, hat lange

Zeit Anklang in der Softwareentwicklung gefunden. Später wurde es um verschiedene Aspekte, z.B. der Möglichkeit in die jeweils vorige Phase zurückzukehren, erweitert.

Im Jahre 1988 stellt Barry Boehm das Spiralmodell vor [Boeh1988, S. 64]. Boehms Modell durchläuft in seinem Entwicklungsprozess mehrere Phasen, die spiralförmig angeordnet sind, um die Möglichkeit, zu früheren Phasen, zurückzukehren zu verdeutlichen. Die größte Schwäche beider Modelle liegt in der Konzentration auf Artefakte, die aus jeder Phase hervorgehen müssen, um in die nächste Phase einzutreten. Diese dokumentenbasierte Sichtweise erklärt das Scheitern vieler Softwareprojekte in der Vergangenheit. Benutzer werden oft nur zu Beginn der Entwicklung und selten in späteren Phasen in den Entwicklungsprozess integriert. Weiter haben Entwickler bei diesen Modellen eine geringe Einsicht in die Prozessabläufe der Organisationen. Dieser Punkt ist mit der Kritik an den traditionellen Forschungsmethoden zu vergleichen. Abhilfe soll das von Floyd entwickelte STEPS-Modell (Abbildung 6) schaffen. Dieses prozessorientierte Vorgehensmodell sieht den Entwicklungsprozess für Entwickler und Nutzer als einen Lernprozess.

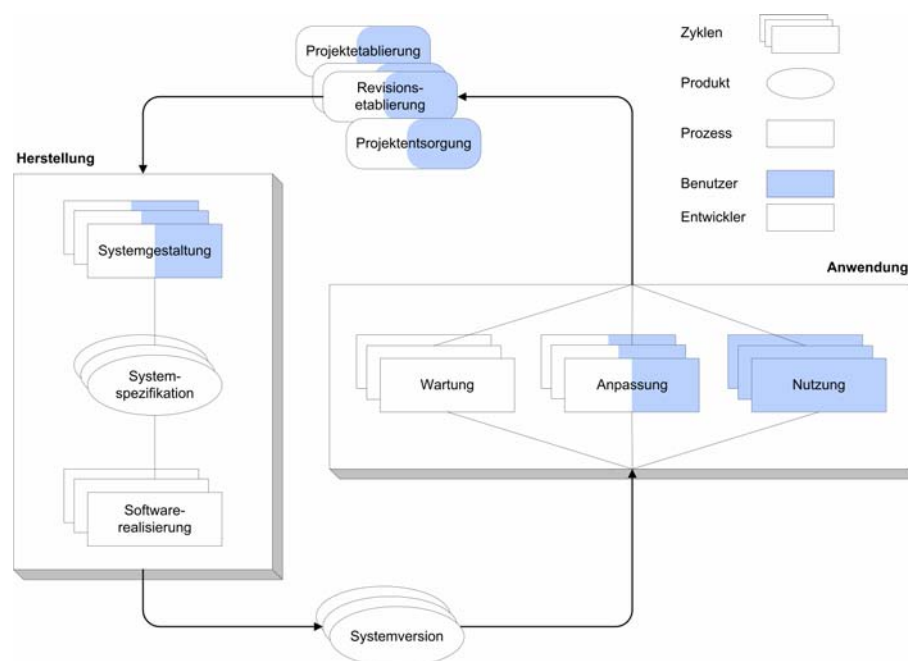


Abbildung 6: Das STEPS-Modell [nach Floyd1989, S. 57]

Eingebettet in das Umfeld der *integrierten Organisations- und Technikentwicklung* (OTE) ist es das Ziel, ein System zu entwickeln, welches auf die Bedürfnisse der Benutzer eingeht und weiterhin anpassbar bleibt [vgl. Kahl1997, S. 44, Wulf1995, S. 60].

Nutzer werden beim STEPS-Modell in den Herstellungsprozess einbezogen beziehungsweise Entwickler in den Anwendungsprozess [vgl. Wulf1995, S. 60]. „Mit der Verknüpfung

evolutionärer und partizipativer Aspekte der Softwareentwicklung mit Erkenntnissen der modernen Organisationsentwicklung in der Integrierten Organisations- und Technikentwicklung (OTE) soll der starken Verflechtung von Organisation und Technik Rechnung getragen werden“ [Kahl1997, S. 45; Kahl1996, S.101]. Dabei zeigen sich die Auswirkungen der Aktionsforschung auf den Softwareentwicklungsprozess und IKS stellen somit das Resultat menschlichen Handelns dar und sind demzufolge sozial konstruiert [Fran1998, S. 75]. OTE ist infolgedessen ein Ansatz, der die beteiligten Akteure aktiv in die Gestaltung von Organisation und Technik einbezieht, indem sie in die Weiterentwicklung von Prozessmodellen integriert werden [Kahl1997, S. 45, Wulf1995, S. 60]. Verfolgt man die beschriebenen Ansätze der OTE, sind die Parallelen zur Aktionsforschung zu erkennen. OTE stellt folglich einen AR-Ansatz für die Softwareentwicklung dar. Ende des 20. Jahrhunderts gewinnt die Aktionsforschung mehr und mehr an Bedeutung für Informationssysteme (IS) [Bask1999, S. 1].

### **2.3 Die Rolle der IT in regionalen Netzwerken**

In Kapitel 2.1.2 wurde die Beschreibung der technischen Unterstützung von Clustern bewusst außer Acht gelassen. Ein Grund ist, dass in der einschlägigen Literatur nur wenige Ansätze behandelt werden [Amen1992, S. 174]. Die Fälle, in denen der Technikbegriff genannt ist, zielen meist auf das elektronische Paradoxon, das Globalisierung und Digitalisierung räumliche Nähe bedeutend werden lässt. Kapitel 2.1 hat diesen Ansatz bereits ausreichend erläutert. Wenn es aber selbstverständlich ist, dass Cluster von IT unterstützt wird, stellt sich die Frage, warum in Literatur und Clusterinitiativen nichts darüber zu finden ist? Ein technischer Aspekt, den alle Cluster gemein haben, ist die Internetpräsenz [vgl. Sölv2003, S. 30]. Die Webseiten repräsentieren die Cluster am globalen Markt, dienen aber nicht dazu, Beziehungen zwischen Akteuren innerhalb des Clusters zu stärken. Dies soll nicht die Effektivität dieser Plattformen schmälern. Es ist durchaus wichtig, den Bekanntheitsgrad eines Clusters durch das Internet zu unterstützen. Hier wird allerdings eine „aktiverere“ Förderung gesucht, eine Förderung die einzelne Akteure des Clusters einbezieht.

Durch die zunehmende Digitalisierung nähern sich die Prozessabläufe der Unternehmen einander an. Standardisierte Software in verschiedenen Bereichen wie Workflow, E-Business oder Customer-Relationship-Management (CRM) bilden das Grundgerüst dieser Prozessketten. Umso wichtiger ist es für Unternehmen, sich durch die „nicht-digitalen“ Abläufe von anderen Unternehmen zu unterscheiden [vgl. Schi2003, 94-96]. Diese nicht

digitalen Abläufe spiegeln die sozialen Beziehungen wider, die in Ballungsräumen entstehen. In dieser Hinsicht fördern die genannten Systeme das Bilden von regionalen Netzwerken. Autoren, die sich mit diesem Thema befasst haben und untersuchten, wie IKT auf Cluster wirkt, waren u.a. Jed Kolko [Kolk2001] und Anthony J. Venables [Vena2001].

Carole Maignan [Maig2003, S.11] führt diese und weitere Autoren auf, um die Einflüsse der IKT zusammenzufassen. Sie ist der Auffassung, dass keine allgemeingültigen Aussagen getroffen werden können, da die Untersuchungen zu spezifisch sind. Dennoch werden Cluster durch diese Systeme technisch unterstützt. Da die Effekte aber nicht bewusst herbeigeführt werden, sondern zufällig entstanden sind, fördert E-Business und die untersuchte IKT passiv die Bildung von regionalen Netzwerken. In diesem Zusammenhang kann deshalb nicht von „aktiven“ Förderungsmaßnahmen gesprochen werden.

Krätke und Scheuplein [Krät2001, S. 4] schlagen ein Softwaresystem zur Identifizierung und Analyse von Clustern vor. Das System soll die regionalpolitischen Institute unterstützen, Entscheidungen über politische Maßnahmen zu treffen. Die Anforderungen, die sie an das System stellen, sind die sachliche und räumliche Abgrenzung eines Clusters, das Aufzeigen sämtlicher interner Verflechtungen der Akteure sowie die Möglichkeit, regionsinterne oder überregionale Vergleiche anzuzeigen. Ein System zu entwickeln, welches ausschließlich der Wirtschaftsförderung dient, kann nicht der Sinn einer technischen Unterstützung von Ballungsräumen sein. Dennoch können die letzten beiden Punkte aufgegriffen werden. Wie sich aus den im Rahmen dieser Arbeit durchgeführten Interviews zeigt, sind Verflechtungen zwischen Unternehmen, meist als Referenzen bezeichnet, für Unternehmer von großer Bedeutung. Eine Seite ist, andere Unternehmen aufgrund ihrer Referenzen intuitiv zu bewerten, die andere Seite führt dazu, dass über Referenzen Kontakte zu neuen Kunden geknüpft werden.

Vergleichende oder empfehlende Systeme, in der Literatur als Recommender-System bezeichnet, gewinnen in vielen Bereichen und besonders in der IKT an Bedeutung. Die Bezeichnung Recommender-System geht auf Paul Resnick und Hal Varian zurück [Resn1997]. Sie beschreiben, dass der Alltag des Menschen von „Empfehlungen“ geprägt ist. Somit liegt es nahe, dies auch auf Funktionen in der IKT umzusetzen. Eines haben diese Systeme gemeinsam: Ein Benutzer muss eine Anfrage stellen und bekommt daraufhin eine Empfehlung [vgl. Reic2005a, S. 62]. Kundenspezifische Internetkaufhäuser empfehlen Produkte, Expertensuchmaschinen filtern den richtigen Ansprechpartner heraus oder



Internetsuchmaschinen bieten geeignete Webseiten zu einer gestellten Suchanfrage. Die Bedeutung der Recommender-Systeme wird auf diese Weise in vielen Bereichen sichtbar. Einer der Hauptaspekte von Clusterinitiativen ist die Netzwerkbildung und damit die Wahrnehmung anderer Unternehmen in der Region. Es liegt nahe, genau dies zu unterstützen. Nach Porter ist es notwendig, den Akteuren eines regionalen Netzwerkes die Möglichkeit zu geben, zu erkennen, welche weiteren Akteure am Cluster teilnehmen [Port1998a, S. 319]. Leuninger und Held sehen eine durch ein IS realisierte Adress- und Kommunikationsplattform [Leun2003, S. 166], als notwendig, geben aber keine weitere Aufklärungen bezüglich möglicher Systeme.

Diese Ausführungen zeigen eine Möglichkeit, IKT in den Unterstützungsprozess einzubeziehen und gehen der Frage nach, mit welcher IT es möglich ist, ein Cluster zu fördern. Einige brauchbare Ansätze, die nicht direkt mit der Unterstützung regionaler Vernetzung in Verbindung gebracht wurden, könnten hilfreich sein, diese Frage zu beantworten. Eine Bewusstseinssteigerung, andere Unternehmen wahrzunehmen und diese Unternehmen vergleichen zu können, forderten Krätke und Scheuplein. Das technische Informationssystem muss die Möglichkeit bieten, ansässige Unternehmen präsentieren und vergleichen zu können. Um eine zweckmäßige Bewertung abzugeben, ist eine Sammlung von Daten über die Unternehmen notwendig. Eine reine Adressdatenbank ist nicht geeignet, um „Experten“-Unternehmen zu speziellen Gebieten aufzufinden. Die hohe Bedeutung von Recommender-Systemen in der wissenschaftlichen Literatur und der Praxis bietet einen ersten Ansatz für ein „clusterunterstützendes“ System.

## **2.4 Kataloge, Branchenbücher und Suchmaschinen**

Ende des 20. Jahrhunderts hat mit der anwachsenden Verbreitung des Internets und seiner steigenden Userzahlen auch der Informationsgehalt stark zugenommen. Ein zentraler Punkt des Internets besteht in der Aufgabe, zu fast allen Themen die entsprechenden Informationen bereitzustellen [vgl. Höls2002, S. 13]. Um der Informationsflut Herr zu werden, ist es notwendig, Maßnahmen zu ergreifen. Gewisse Abhilfen schaffen Suchmaschinen und Webkataloge. 90 Prozent aller Suchen im Internet beginnen mit bekannten Suchsystemen [Glög2003, Vorwort]. Auch wenn der derzeitige Trend etwas anderes zeigt, so gibt es im Internet eine große Menge an unterschiedlichen Suchsystemen. Einige davon sind Suchmaschinen wie Google, andere stellen Informationen katalogweise zur Verfügung oder sind thematisch beschränkt. Speziell zur Suche im Internet lassen sich Webkataloge,

Suchmaschinen und Metasuchmaschinen unterscheiden. Branchenbücher stellen einen Sonderfall dar. Sie dienen nicht der Suche nach Internetpräsenzen, sondern der reinen Suche nach Unternehmen.

## **2.4.1 Internetbasierte Suchsysteme**

### *2.4.1.1 Webkataloge*

Ein Webkatalog ist eine themenorientierte Konglomeration von Internetadressen [vgl. Ellw2002, S. 82]. Die einzelnen Adressen werden durch Mitarbeiter des Katalogs persönlich geprüft, bewertet und sortiert. Die Mitarbeiter entscheiden über eine mögliche Eintragung oder eine Ablehnung der Adresse [vgl. Glög2003, S. 2]. Dadurch ist eine höhere Qualität als bei „normalen“ Suchmaschinen, die eine Zusammenstellung von Links automatisch durchführen, zu erwarten. Der Nachteil ergibt sich aus demselben Grund. Durch die manuelle Bearbeitung der Einträge ist nur ein wesentlich kleinerer Datenbestand möglich als bei Suchmaschinen.

Die Ergebnisse werden im Allgemeinen in einer strukturierten Form nach Themenbereichen präsentiert. In einer Baumstruktur kann man durch immer spezialisierendere Links von der Wurzel, dem übergeordneten Thema, zum entsprechenden Blatt des Baumes, dem Link, zur eigentlichen Webseite gelangen. Oftmals bieten Webkataloge noch eine Suchfunktion an. Im Gegensatz zu Suchmaschinen bezieht sich diese nicht auf eine Suche im Internet, sondern auf die bereitgestellten Links im Katalog. Eine themenübergreifende Suche wird auf diese Weise ermöglicht.

### *2.4.1.2 Suchmaschinen, Meta-Suchmaschinen und sonstige Suchmaschinen*

Neben Webkatalogen stellen Suchmaschinen den zweiten Grundtyp von Suchmöglichkeiten im Internet dar. Suchmaschinen durchsuchen selbstständig das World Wide Web (www) nach Informationen und bilden eine Indexierung über die durchsuchten Internetseiten. Sie erzeugen ihren Datenbestand somit automatisch. Ende 2003 ergab eine Studie der Bertelsmann-Stiftung, dass 91 Prozent aller deutschen Internetsurfer Suchmaschinen nutzen [vgl. Erlh2005, S.13 und 14]. Die Bedeutung von Suchmaschinen für die Recherche im Internet ist somit klar. Die wesentlichen Unterschiede zwischen einer Suchmaschine und einem Webkatalog liegen in der automatisierten Datenbeschaffung, der automatisierten Dokumentenanalyse und –bewertung, dem automatisierten Aufbau und der Verwaltung von Datenstrukturen sowie der automatisierten Berechnung von Relevanzen bei Suchanfragen [vgl. Glög2003, S. 4].

Suchmaschinen stellen zum einen ein konkretes Suchsystem und zum anderen einen Sammelbegriff dar. Sie lassen sich in „normale“ *Suchmaschinen*, *Meta-Suchmaschinen*, *Payed Listing Suchmaschinen*, *Cluster-Suchmaschinen* und *Deep-Web-Suchmaschinen* untergliedern. Während die ersten drei Arten bekannte Verfahren aufweisen, zeichnen sich Cluster- und Deep-Web-Suchmaschinen durch neuere Ansätze aus.

### **„Normale“ Suchmaschinen**

Suchmaschinen funktionieren auf die oben beschriebene Art und Weise. Sie durchsuchen große Teile des Internets, um ihren Datenbestand zu vergrößern und aktuell zu halten. Als Beispiele für Suchmaschinen sei hier Google und Altavista genannt.

Steve Lawrence stellt 1999 fest, dass Suchmaschinen Webseiten unterschiedlich indexieren, neue Seiten erst nach Monaten aufnehmen und eine einzelne Suchmaschine nicht mehr als 16 Prozent des Internets abdeckt [vgl. Lawr1999, S.1]. Unterschiedliche Suchmaschinen decken mitunter unterschiedliche Bereiche ab. Bei Suchanfragen ist es deshalb hilfreich, mehrere Suchmaschinen zu verwenden. Suchmaschinen, die dies automatisieren, sind Meta-Suchmaschinen.

### **Meta-Suchmaschinen**

Meta-Suchmaschinen verfügen über keinen eigenen Datenbestand. Vielmehr nutzen sie die Datenbestände von anderen Suchmaschinen und Webkatalogen. Das Kernstück einer Meta-Suchmaschine, wie z.B. MetaGer oder MetaCrawler, ist ein Programm, das Suchanfragen an mehrere verschiedene Suchdienste weiterleitet, die erhaltenen Resultate zusammenführt und nach eigenen Kriterien gewichtet [Glög2003, S. 8]. So trivial dies zunächst erscheinen mag, so einfach ist die Funktionalität einer Meta-Suchmaschine nicht. Ellwein führt zwei wichtige Nachteile auf [Ellw2002, S. 75]. Der erste Punkt bezieht sich auf umfassendere Funktionalitäten der „normalen“ Suchmaschinen. Fast jede Suchmaschine bietet für sich spezifische Operationen an, die über logische Verknüpfungen hinausgehen. Meta-Suchmaschinen können diese Funktionalitäten nicht voll ausnutzen. Punkt zwei betrifft die Wiederaufarbeitung und Präsentation der zurückgelieferten Ergebnisse [Erlh2005, S. 29]. Häufig werden die Beurteilungen von den ursprünglichen Seiten übernommen. Aufgrund der heterogenen Bewertungsalgorithmen und der Qualität einzelner Suchmaschinen ist dieses Verfahren suboptimal [vgl. Erlh2005, S. 30].

### **Payed Listing Suchmaschinen**

Die Bezeichnung „Suchmaschine“ ist im Hinblick auf die Definition eigentlich unzutreffend. Beim Payed-Listing-Verfahren werden Rangpositionen bei verschiedenen Suchsystemen gekauft [Glög2003, S.9]. Ein Grundbestandteil dieses „Suchdienstes“ ist das Pay-Per-Click-Verfahren (PPC). Die Rangpositionen werden an den Höchstbietenden verkauft. Alle Nachfolgegebote werden entsprechend in der Rangfolge darunter angesiedelt. Geboten wird auf den Preis, pro Klick auf die eigene Webseite. Für dieses Verfahren sind Monitoring und Controlling-Verfahren zur Linkverfolgung notwendig. Wenige Suchmaschinen und Webkataloge verfügen über eine Technologie, die dieses Verfahren unterstützt. Aus diesem Grund arbeiten sie mit externen Dienst Anbietern, die PPC im Versteigerungsverfahren realisieren, zusammen [Glög2003, S. 11].

### **Cluster-Suchmaschinen**

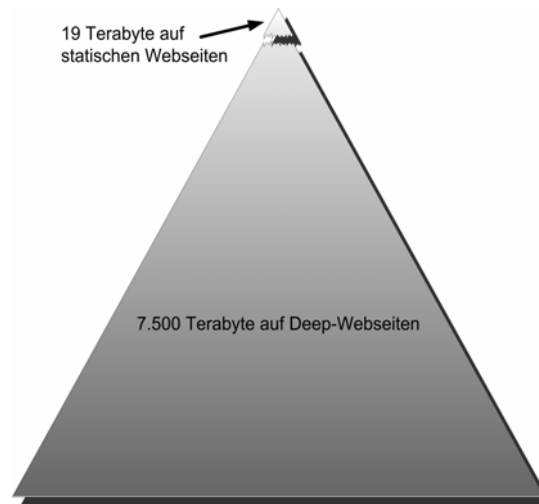
Auch wenn die Wortübereinstimmung verwirrend ist: Cluster-Suchmaschinen haben nicht das Ziel, regionale Netzwerke ausfindig zu machen. Sie können als eine besondere Art der Meta-Suchmaschinen gesehen werden [Ellw2002, S. 78]. Der Unterschied zu Meta-Suchmaschinen besteht in der Datenaufbereitung. Durch Analyse der URLs, Titelzeilen und Kurzbeschreibungen wird versucht, die Ergebnisse thematisch zu gruppieren. So soll dem Nutzer der Zugang zu den Informationen erleichtert werden. Vivisimo und Teoma sind Beispiele für zwei Cluster-Suchmaschinen.

### **Deep-Web Suchmaschinen**

Ein besonderer Nachteil herkömmlicher Suchmaschinen liegt in der Auslassung dynamischer Webseiten. Seiten, die in PHP, JSP, ASP oder Ähnlichem geschrieben sind, erzeugen ihren Inhalt je nach Anfrage dynamisch. Dieser Inhalt kann von Suchmaschinen, die sich über eingebettete Links von einer Webseite zur anderen hangeln, nicht gefunden werden. Hier beginnt die Aufgabe der Deep-Web Suchmaschinen, diese dynamischen Inhalte zu erfassen. [Ellw2002, S. 81]. Michael Bergman nennt verschiedene Gründe, warum die Tiefen des Internets so wichtig sind. Die Datenmenge der statischen Webseiten bildet nur die „Spitze des Eisbergs.“ Dynamische Webseiten, so Bergman, fassen nochmals etwa das 400 bis 500 Fache der herkömmlichen Webseiten (vgl. Abbildung 7). Ausgangspunkt für diese Aussage ist eine Studie, die er in Kooperation mit BrightPlanet<sup>3</sup> im März 2000 durchgeführt hat.

---

<sup>3</sup> <http://www.brightplanet.com/>



**Abbildung 7: Verhältnis zwischen statischen und dynamischen Webseiten**

Nach Bergman ist das Deep-Web der am schnellsten wachsende Bereich für Informationen. Diese Informationen sind höchst relevant und verweilen ungefunden in Datenbanken. Außerdem erzeugen Deep-Webseiten um 50 Prozent mehr „Traffic“ als ihre statischen Konkurrenten.

Deep-Web-Suchmaschinen werden allerdings selten auf externen Servern betrieben und setzen die Installation eines Programms auf dem eigenen Computer voraus [vgl. Ellw2002, S.81]. Erschwerend kommt hinzu, dass viele dieser Maschinen kostenpflichtig sind, so zum Beispiel CiteLine.

#### 2.4.1.3 Branchenbücher

Branchenbücher, auch Gelbe Seiten, Yellow-Pages, Telefonbuch oder Business Listing genannt, heben sich von den genannten Suchmaschinen ab. Im Gegensatz zu Ihnen durchsuchen Branchenbücher keine Internetpräsenzen. Sie stellen im Allgemeinen eine Datenbank bereit, in der Unternehmen oder Benutzer ihre Daten zu einer oder mehreren Branchen/Themen eintragen können. Ein deutschlandweites Branchenbuch ist z.B. GelbeSeiten und für die Region Siegen Si-Ri24.

Ähnlich wie Webkataloge werden die Unternehmen nach Branchen sortiert und dem Nutzer präsentiert. Viele Branchenbücher verfügen über eine eigene Suchmöglichkeit. Diese bezieht sich auf die eingetragenen Daten und nicht auf mögliche Webseiten der Unternehmen. Ein großer Vorteil der Branchenbücher liegt in der „geringen“ Datenmenge, die sie anbieten. Nutzer, die gezielt nach Unternehmen einer bestimmten Branche suchen, werden auf diese Weise vor der Informationsflut geschützt. Allerdings mit dem Nachteil, dass Unternehmen

nur zu spezifischen Branchen und im Idealfall zu einigen Stichworten gefunden werden. Die Zahl der Branchenbücher hat mittlerweile ähnliche Ausmaße wie die der Suchmaschinen angenommen. Manche sind auf einen bestimmten Sektor bezogen, andere liefern nur die nötigsten Kontaktdaten, sind dafür aber im Hinblick auf Branchen und die Anzahl an Unternehmen wesentlich umfassender; wieder andere beschränken sich auf einen Raum oder eine Region.

#### **2.4.2 Sonstige Suchsysteme und Informationsgewinnung**

Aktuell geht der Trend zu den Wikis (hawaiianisch für „schnell“). Dies sind dynamische Internetseiten, die von dem Leser nicht nur gelesen, sondern auch bearbeitet werden können. Vorreiter dieses Konzeptes ist die Internetseite „Wikipedia“, die eine umfassende Enzyklopädie zu fast jedem Thema aufgebaut hat. Der Grundgedanke dabei ist dem User eine möglichst umfassende und verständliche Enzyklopädie anzubieten [vgl. Ellw2002, S. 89].

In Kapitel 2.3 wurde bereits der Begriff der Recommender-Systeme erläutert. Mit zunehmender Bedeutung finden sich diese Systeme in immer mehr Bereichen wieder. Internetsuchmaschinen (wie oben definiert) dienen als Beispiel solcher Systeme. Eine weitere Verbreitung finden diese Systeme in so genannten Personen- und Expertensuchmaschinen die nicht nur nach Namen, sondern auch nach möglichen Kompetenzen suchen [Ellw2002, S.97]. Oft sind es abgeschlossene Systeme für einen bestimmten Bereich oder eine Organisation. Das in Kapitel 2.5 angesprochene ExpertFinding-System ist ein solches System.

Im Internet geht der Trend zu Personendatenbanken, in denen sich die Nutzer selber eintragen und ihre Daten verwalten. Wichtig bei diesen Systemen ist, dass eine „kritische Masse“ an Benutzern erreicht wird. Die Internetplattform „Xing“, ehemals „Open-BC“, hat dies als erstes System erfolgreich umgesetzt und findet täglich neue Nutzer. Interessante Optionen bei diesen Systemen sind die Möglichkeiten, Verbindungen zu anderen Personen und Kontakten zu knüpfen und visuell darzustellen. Die Plattform Xing wirbt damit, dass alle Menschen auf der Welt über „sechs Ecken“ miteinander verbunden sind und zeigt bei Personensuchen direkte Verbindungswege über andere Personen an. Diese Personensuchportale sind entgegen der Expertensysteme reine Nachschlagewerke bzw. Kontaktpflegeportale. Eine Empfehlung oder Bewertung zu Suchanfragen findet in der Regel nicht statt.

Es gibt eine Reihe weiterer spezialisierter Suchmaschinen. Ähnlich einem Branchenbuch dienen sie nur einem bestimmten Zweck. Bibliotheken zum Auffinden von Büchern und

wissenschaftlichen Texten, Atlanten, die Kartenmaterial bereitstellen oder Routenplaner, die Adressen finden, anzeigen und ganze Routen berechnen, sind nur ein Teil der im Web verfügbaren Suchmaschinen. Eine etwas andere Möglichkeit der Informationsgewinnung zu einem bestimmten Thema ist das Eintragen in eine Mailingliste oder das Beantragen eines Newsletters. Zu einem bestimmten Thema werden an die angemeldeten Mitglieder wöchentlich oder monatlich E-Mails, so genannte Newsletter mit Informationen zu diesem Thema, versendet.

Weitere Informationsquellen im Internet sind Foren und Newsgruppen. Sie nutzen das Wissen der angemeldeten Teilnehmer für Fragen anderer Anwender. Aus gestellten Fragen ergeben sich oft Diskussionen zwischen den einzelnen Akteuren, so dass auf diese Weise ein bestimmtes Thema ausführlich behandelt wird. Diese Informationen sind im Internet abrufbar.

### **2.4.3 Funktionsweisen der Suchmaschinen**

#### *2.4.3.1 Indexierungsverfahren*

Die Aufgabe der automatischen Dokumentenanalyse besteht darin, Begriffe bzw. Deskriptoren aus einem Text zu extrahieren und einem Index zusammenzustellen. Dabei wird den Begriffen entsprechend ihres Auftretens ein unterschiedliches Gewicht zugeordnet. [Salt1983, S. 57 und S. 65]. Bereits hier erkennt man die Bedeutung der Indexierung für Recommender-Systeme, die auf einer Dokumentenanalyse basieren. Einzelnen Begriffen wird ein Wert gemäß der Häufigkeit seines Auftretens zugewiesen, der die Grundlage für spätere Empfehlungen bietet. Auf dieser Annahme, belegt durch Textanalysen verschiedener Gebiete, beruhen die meisten automatischen Indexierungsverfahren.

Die Annahme, die einem der ersten Algorithmen zu Grunde gelegt wird, besagt, dass das Auftreten repräsentativer Begriffe im mittleren Häufigkeitsbereich am höchsten ist. Eine besondere Aufgabe ist die Festlegung von Schwellenwerten, zwischen denen sich die Deskriptoren befinden. Eine zu starke Eingrenzung führt zum Informationsverlust [vgl. Fran1997, S.193]. Nicht alle hochfrequenten Begriffe müssen Stoppworte<sup>4</sup> sein. Durch das Wegfallen dieser „falsch erkannten“ Stoppworte können Informationen bezüglich eines Dokumentes verloren gehen.

---

<sup>4</sup> Als Stoppworte werden Worte bezeichnet, die bei einer Indexierung keine Beachtung finden. Allgemein sind dies Artikel, Konjunktionen, Präpositionen und Negationen.

Ein anderer Punkt, auf dem dieser erste Algorithmus seine Funktionalität gründet, ist die Annahme der absoluten Häufigkeit. Angenommen, in einer Fachbibliothek über Maschinenbau werden ausschließlich Texte zu diesem Thema abgelegt. Es ist anzunehmen, dass in jedem dieser Texte mindestens einmal das Wort „Maschine“ genannt wird. Somit kann „Maschine“ kein Deskriptor für einen dieser Texte sein. Er gibt zwar den Inhalt eines Dokumentes wieder, hilft aber nicht dabei, die Texte voneinander zu unterscheiden. Infolgedessen muss das Auftreten eines Wortes in Relation zu der Gesamtzahl des Wortes in allen Texten gesehen werden. Ein Begriff muss häufig in einigen wenigen Dokumenten enthalten sein und darf nur relativ selten in der Gesamtzahl der Dokumente vorkommen. In der Literatur finden sich verschiedene Gewichtungsfunktionen [Salt1983, S. 68], so z.B. das Verfahren der inversen Dokumenthäufigkeit (IDF). In diesem Fall wird die Gewichtung auf Begriffe mit einer niedrigen Dokumenthäufigkeit gelegt [Salt1983, S.69]. Im Zusammenhang mit der Termfrequenz  $TF$  wird  $IDF$  vornehmlich bei der automatischen Indexierung verwendet. Die  $TF$  gibt dabei einen Hinweis auf die Bedeutung eines Terms  $k$  für ein Dokument.

Die Entwicklung der Indexierungsverfahren ist in den letzten 20 Jahren stetig vorangeschritten. Entsprechend weisen viele der Verfahren von heute Ähnlichkeiten mit den früheren auf und gewinnen durch Recommender-Systeme wieder zunehmend an Bedeutung. Grundsätzlich sind bei allen Algorithmen fünf Schritte durchzuführen [vgl. Fran1997, S.]:

1. Separation von Sätzen aus dem Dokument für weitere Indexierung.
2. Entfernen aller Stoppworte.
3. Finden und indexieren von Wortkombinationen in den Sätzen.
4. Indexieren von Einzelwörtern die im Satz auftauchen.
5. Erzeugen eines Dokumentenprofils.

Neben den grundlegenden Ansätzen gibt es eine Reihe von fortgeschrittenen Modellen und Verfahren. Dazu zählen Clusterverfahren, das Verarbeiten von Zitaten sowie das Vector Space Modell (VSM). Clusterverfahren versuchen die Menge der Deskriptoren in bestimmte Begriffsklassen aufzuteilen, um so die Dokumente in Sachgebietenklassen einzuordnen (vgl. Abschnitt Cluster-Suchmaschinen in Kapitel 2.4.1.2). Das Verarbeiten von Zitaten wird in Kapitel 2.4.3.3 verwendet und dort näher erläutert.



### 2.4.3.2 Vektorielle Ähnlichkeitsfunktion und schwache Bedeutungseinordnung

In der Zeit der objektorientierten Softwareentwicklung wird davon ausgegangen, dass alles durch Objekte beschreibbar ist. Die Objekte besitzen Eigenschaften (Attribute), die sie identifizieren [vgl. Salt1983, S. 213]. So stellt eine Flasche ein Objekt dar. Die Attribute, die sie besitzt, lassen sich durch Öffnungszustand, Füllstand und Glasart beschreiben. Die Eigenschaften werden durch Werte beschrieben: Geöffnet/Geschlossen, Voll/Leer, Weißglas/Buntglas. Die Objekte im VSM werden durch die vorhandenen Dokumente repräsentiert. Ihre Eigenschaften sind die aus Ihnen extrahierten Deskriptoren. Jeder der Deskriptoren besitzt einen Wert für die Häufigkeit *Freq* seines Auftretens in diesem Dokument *i*. Die Dokumente können folglich in einem mehrdimensionalen semantischen Raum als Vektoren *Dok* dargestellt werden:

$$Dok_i = \begin{pmatrix} Freq_{i1} \\ Freq_{i2} \\ \vdots \\ Freq_{in} \end{pmatrix} \quad Dok_j = \begin{pmatrix} Freq_{j1} \\ Freq_{j2} \\ \vdots \\ Freq_{jn} \end{pmatrix} \quad \dots \text{ mit } n \in \mathbb{N}, \text{ der Gesamtzahl aller Deskriptoren}$$

Alternativ können die einzelnen Vektoren zu einer Matrix der Größe  $n \times m$  zusammengefasst werden, in der die Spalten  $m$  Dokumente und die Zeilen die vollständige Anzahl von  $n$  Deskriptoren repräsentieren. Diese Matrix wird Term-Dokumenten-Matrix genannt (TDM). Die einzelnen Achsen in dem semantischen Raum entsprechen den einzelnen Deskriptoren. Die Häufigkeit der einzelnen Deskriptoren entspricht dem Lot des Vektors auf diese Achse. Ein Beispiel soll das VSM verdeutlichen. Es geht vereinfachend von zwei Dokumenten  $i$  und  $j$  aus. Der semantische Raum wird durch drei Deskriptoren aufgebaut<sup>5</sup> und zwar aus den Begriffen *Vector*, *Space* und *Model* (vgl. Abbildung 8). In Dokument  $i$  taucht das Wort *Vector* 12-mal, das Wort *Space* 6-mal und das Wort *Model* 7-mal auf. Für Dokument  $j$  gelten in gleicher Reihenfolge der Begriffe die Werte 8, 11 und 4. Mit der Häufigkeit der Begriffe lassen sich die Lagen der Vektoren im semantischen Raum beschreiben:

$$Dok_i = \begin{pmatrix} 12 \\ 6 \\ 7 \end{pmatrix} \quad Dok_j = \begin{pmatrix} 8 \\ 11 \\ 4 \end{pmatrix}$$

<sup>5</sup> Auch dies ist eine starke Vereinfachung der Wirklichkeit. Textanalysen ergeben im Allgemeinen eine Vielzahl an Deskriptoren. Um aber eine visuelle Darstellung zu ermöglichen wurde diese Approximation gewählt.

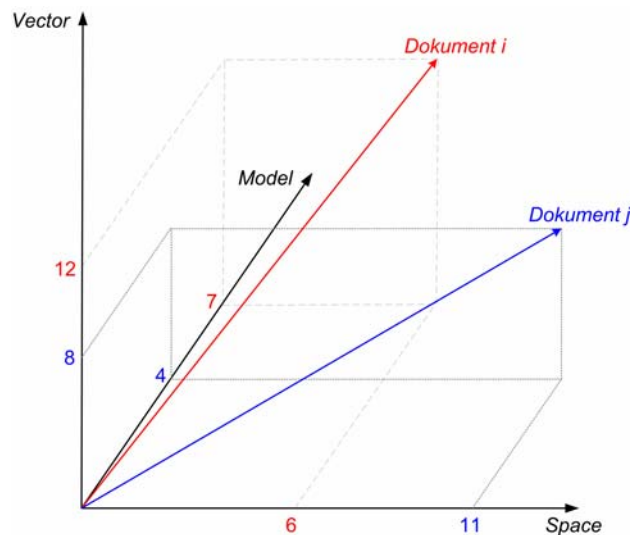


Abbildung 8: Beispiel des Vector Space Modells

Um die Ähnlichkeit der Dokumente zu bestimmen, wird die Lage der Vektoren zueinander verglichen. Das Bilden des Skalarproduktes bzw. das Bestimmen des Winkels zwischen den Vektoren stellt nur eine Auswahl verschiedener mathematischer Möglichkeiten dar. Salton und McGill geben z.B. fünf Funktionen zur Ähnlichkeitsbewertung an [Salt1983, S. 215]. Alle Funktionen haben eins gemeinsam: Ihre Werte nehmen mit steigender Übereinstimmung zu [Salt1983, S. 216]. Der inverse Fall ist allerdings auch möglich wie die Variante der Winkelbestimmung zeigt. Je geringer der Winkel zwischen zwei Vektoren ist, desto größer ist deren Übereinstimmung. Ein Deskriptor kann in einem Dokument beliebig oft ( $Freq \geq 0$ ) vorkommen. Im Falle des VSM existiert folglich nur der mathematisch positive Raum. Für das Beispiel lässt sich der Winkel  $\alpha$  aus dem Skalarprodukt der beiden Vektoren dividiert durch das Produkt aus den Längen beider Vektoren berechnen:

$$\cos(\alpha) = \frac{Dok_i \cdot Dok_j}{\|Dok_i\| \times \|Dok_j\|} \Rightarrow \frac{\begin{pmatrix} 12 \\ 6 \\ 7 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ 11 \\ 4 \end{pmatrix}}{\sqrt{12^2 + 6^2 + 7^2} \times \sqrt{8^2 + 11^2 + 4^2}} = \frac{12 \times 8 + 6 \times 11 + 7 \times 4}{15,13 \times 14,18}$$

$$\approx 0,89 \Rightarrow \underline{\underline{\alpha = 27,67^\circ}}$$

Auf VSM aufbauend wird das *Latent Semantic Indexing*-Verfahren (LSI) im Folgenden beschrieben. LSI ist ein rein mathematisches Verfahren, welches darauf abzielt, die TDM zu optimieren. Dabei verfolgt LSI das Ziel, Synonyme und Polysemie aufzudecken und zusammenzufassen. In der TDM bedeutet dies, einzelne Zeilen zusammenzufassen oder zu eliminieren. Begriffe werden zu Konzepten zusammengefasst. Das Wort „Auto“ ist z.B. ein

Konzept für die Begriffe Kraftfahrzeug, Personenkraftwagen, Untersatz und Flitzer. Vorteil der durch LSI angestrebten Möglichkeiten ist, Dokumente zu finden, in denen synonyme Begrifflichkeiten auftauchen. Reichling, Schubert und Wulf sehen einen entscheidenden Vorteil in der Optimierung der TDM. Ausgehend von der Annahme, dass viele Dokumente unterschiedlichster Bereiche indexiert wurden, ergeben sich verschiedene Deskriptoren, die nur in einer Reihe von Dokumenten auftauchen. Die Folge sind viele Nullfelder innerhalb der Matrix [vgl. Reic2005a, S.64]. Diese Nullen vergrößern die Matrix unnötig. Mit LSI wird versucht, dieses Problem zu lösen.

Hauptaspekt des Verfahrens ist die Singulärwertzerlegung. Dabei wird die TDM mathematisch in drei Matrizen umgewandelt:  $TDM = U \cdot S \cdot V^T$ . Die Matrix S richtet sich nach dem frei wählbaren Wert  $k$ . S ist eine Diagonalmatrix die mit  $k$  die Anzahl der neuen künstlichen Konzepte beschreibt [vgl. Batt2005, S. 39]. Die Wahl von  $k$  ist von entscheidender Bedeutung. Eine zu kleine Wahl führt zu einer zu starken Approximation, während die Wahl eines großen  $k$  keine deutliche Optimierung hervorbringt. Deerwester, Dumais und Harshman sehen in ihrer Veröffentlichung zum LSI-Verfahren [Deer1990, S. 21] einen Wert zwischen 50 und 150 als realistisch, fügen aber hinzu, dass sie keine Untersuchung unternommen haben, dies zu bestätigen.

LSI besitzt somit die Fähigkeit, Synonyme zu Konzepten zusammenzufassen. Die Probleme mit Polysemie werden dagegen nicht gelöst [vgl. Deer1990, S. 21].

#### 2.4.3.3 PageRank

1997 stellen die Entwickler Sergey Brin und Lawrence Page die Beta-Version einer weiteren Suchmaschine namens „Google“ der Öffentlichkeit zur Verfügung, welche die Welt der Suchmaschinen in den folgenden Jahren revolutionieren sollte. Die Indexierung der Suchmaschine basiert auf dem PageRank-Verfahren. Die Annahme, die dieser Methode zugrunde liegt, beruht auf dem Zählen von Literaturverweisen und Zitaten in wissenschaftlichen Arbeiten [vgl. Brin1997]. Bei diesem Verfahren wird davon ausgegangen, dass bedeutende wissenschaftliche Arbeiten häufig in anderen Arbeiten verwendet werden und dass bedeutende Arbeiten ebenfalls nur auf signifikante Texte verweisen. Diesen Grundgedanken übertrugen Brin und Page auf Internetseiten. Bedeutsame Seiten werden häufig von anderen Webseiten aus verlinkt und bedeutende Seiten selbst werden keine Links auf unbedeutende Webseiten beinhalten [Brin1997]. Das Problem, das sich ergibt, ist die Unbekanntheit der Stellen, die auf einen Text verweisen. Bei der Analyse einer Webseite

kann nicht festgestellt werden, von welchen Seiten aus sie verlinkt ist (Backlinks), wohl aber auf welche Seiten sie verlinkt. Page et al. [Page1998, S. 3] bemessen dem ersten Faktor, die Wichtigkeit der Seite, von der aus der Backlink erfolgt, eine große Bedeutung zu. Sie untermauern ihre These, dass ein reines Zählen der Backlinks alleine nicht aussagekräftig ist mit dem Beispiel eines Links von der Yahoo-Homepage. Dieser Link habe, so Page et al. eine größere Bedeutung, als ein Link auf die gleiche Stelle von einer dubiosen anderen Webseite. Damit ist folgende Beschreibung für PageRank zutreffend:

*„Eine Webseite hat dann einen hohen Rang, wenn die Summe der Ränge ihrer Backlinks hoch ist“ [Page1998, S. 3].*

Dass dieses Verfahren die Geschichte der Suchmaschinen revolutionierte, steht in dem Erfolg von Google selbst geschrieben. Zirka die Hälfte aller Suchanfragen im World Wide Web wird heute über Google abgewickelt [vgl. Cons2007 und Wiki2007]. Alle anderen angebotenen Suchmaschinen teilen die weiteren 50 Prozent unter sich auf. Als Indiz des Erfolgs ist das Verb „googeln“ bereits in die deutsche Rechtschreibung aufgenommen worden. Im Sprachgebrauch ist damit jegliche Suche im Internet gemeint. Google selbst besteht darauf, dass mit „googeln“ ausschließlich die Suche mittels Google zu verstehen ist. Die Verwendung des Eigennamens im Sprachgebrauch untermauert den Erfolg von Google.

## 2.5 Einführung des ExpertFinding-Systems

Die Größen *Soziales* und *Humanes Kapital* (siehe oben) sind für Unternehmen von großer Bedeutung. Unternehmen können die Produktivität und ihre Innovationskraft durch diese Faktoren steigern [vgl. Reich2005b, S. 325]. Um diese sozialen Netzwerke, die Expertisen, die Fähigkeiten und das Wissen zu identifizieren, zu fördern und sichtbar zu machen, wurde in den letzten Jahren das ExpertFinding-System (EF-System) entwickelt, welches „Experten“ anhand zahlreicher Möglichkeiten herausfiltern soll:

*„Das Expert Finding Framework<sup>6</sup> kann als eine zentrale Datenbank angesehen werden, in der Benutzerprofile generiert, gespeichert, aktualisiert und miteinander verglichen werden. Es versteht sich dabei als kooperations-unterstützendes bzw. –anbahnendes System.“ [Reic2006a, S. 23]*

Mit diesen Worten beschreibt Tim Richling sein entwickeltes System, das zunächst zur Unterstützung der Zusammenarbeit in einer E-Learning Plattform gedacht war. Hier sollte

<sup>6</sup> Im Laufe der Entwicklung hat das EF-System sich nicht nur in technischen Aspekten verändert, sondern auch seinen Namen mehrmals gewechselt.

Lernern die Möglichkeit gegeben werden, Kontakt zu anderen Lernern aufzubauen, die ein bestimmtes Interessen- oder Kenntnisprofil haben [vgl. Reic2004, S. 179]. Somit versteht sich das System von der ersten Entwicklung bis zur heutigen Version als kooperationsunterstützende Software, die ihren Schwerpunkt in der Identifikation von Interessen und Wissen findet. Das System ist von Beginn an in einer offenen Architektur modelliert, die ein schnelles und flexibles Reagieren auf neue gegebene Situationen ermöglicht. Diese Flexibilität stellt die Grundlage für die Nutzung als Unternehmenssuchmaschine dar und wird in Kapitel 5.2 ausführlich beschrieben.

Das EF-System ist von der Lernplattform zu einem allgemeinen Expertisen-Recommendensystem weiterentwickelt worden. Entsprechend wurden die Funktionalitäten des EF-Systems im Laufe der Zeit erweitert und haben Veränderungen durchlaufen. Im Jahr 2004 diente die Software zur Integration in einer Lernplattform [vgl. Beck2004, S. 339]. Der technische Schwerpunkt lag dabei auf den Recommendereigenschaften und Matching-Algorithmen über den statischen Profildaten. Ein Jahr später berichten Reichling, Schubert und Wulf [Reic2005a, S.61] vom EF-System mit der Absicht, benutzerspezifische Daten aus verfassten und gelesenen Texten softwaretechnisch zu extrahieren. Sie legen dabei die Annahme zu Grunde, dass gelesene und eigen verfasste Texte Aussagen über die Expertisen der Benutzer treffen. Die extrahierten Daten werden mit Hilfe der bestehenden Matching-Algorithmen ausgewertet. Diese Weise der Informationsgewinnung hat sich bis heute im System etabliert.

In der Zwischenzeit ist das EF-System eine eigenständige Anwendung, die nach dem Client-Server-Prinzip arbeitet und bereits in einem großen deutschen Industrieverband erfolgreich getestet wurde. Der Benutzerclient dient neben der Kommunikation mit dem Anwender zur Erzeugung der Benutzerprofile und stellt Funktionalitäten zur Kontakthanbahnung bereit. Zur Profilerzeugung werden u.a. die lokal auf der Festplatte abgelegten Textdokumente herangezogen. Die Kontrolle über die bereitzustellenden Dokumente und dem daraus generierten Schlüsselwortprofil bleibt dem Nutzer überlassen. Die Serverseite übernimmt die zentrale Verwaltung der Benutzerprofile sowie das Suchen in den abgelegten Informationen [Reic2006b, S. 4]. Das System kann als eine Erweiterung des Yellow-Page-Ansatzes gesehen werden. Neben den durch die Benutzer eingetragenen personenbezogenen Angaben (wie Kontaktdaten, Name, Anschrift usw.) wird versucht, mittels Indexierungsverfahren Informationen aus Veröffentlichungen und bearbeiteten Dokumenten abzuleiten.

Das EF-System erledigt somit zwei zentrale Aufgaben [Reic2005a, S.65]:

- Erzeugen von Profilen aus verschiedenen Datenquellen,
- Matching (Vergleichen) der Profile miteinander.

## 2.6 Anwendung des ExpertFinding-Systems auf Unternehmen

In einer durch die Marktforschungsgesellschaft GfK durchgeführten Studie zeigt sich, dass 76 Prozent aller „Einkaufsentscheider“ das Internet zur Recherche verwenden. 93 Prozent nutzen dabei Internetsuchmaschinen wie Google oder spezielle Suchmaschinen wie „Wer liefert was?“ [vgl. Werb2005].

Kapitel 2.1 hat gezeigt, dass es für Unternehmen vorteilhaft ist, andere Unternehmen der Region zu finden und zu kennen. Regionale Branchenbücher können dabei nur bedingt Abhilfe schaffen. Suchen Kunden bestimmte Produkte oder Dienstleistungen, bieten Branchenbücher selten die richtigen Informationen und Suchmöglichkeiten. Eine Kategorisierung nach Branchen ist dabei selten hilfreich. Der Suchende ist genötigt, sich selbst ein Urteil über das Dargebotene zu bilden. Dies kostet Recherchearbeit und Zeit. Eine Suchmaschine bietet hingegen die Möglichkeit, eine Empfehlung bezüglich einer Suchanfrage zu geben, hat aber den Nachteil der Informationsflut. Insbesondere im Business-to-Business-Bereich (B2B) werden nicht ausschließlich Informationen über andere Unternehmen angeboten. Private Seiten, Werbung oder Foren stellen nur eine kleine Auswahl dar, die ein Suchender von den gewünschten Informationen filtern muss.

Mit der Entwicklung eines Recommender-Systems für Unternehmen sollen die Vorteile beider Suchsysteme vereint werden. In Kapitel 2.5 wurde das EF-System vorgestellt. Hauptbestandteile des Recommender-Systems sind die leichte Anpassbarkeit an neue Gegebenheiten und das parallele Verwenden von Suchen auf unterschiedlichen Daten. Zum einen werden Expertisen aus den Stammdaten der Benutzer gefiltert. Diese Daten liegen in strukturierter Form (tabellarisch) vor und sind leicht handhabbar. Sie können mit denen eines Branchenbuchs verglichen werden. Zum anderen gewinnt das EF-System durch die Indexierung der unstrukturierten Daten (Textdokumente) Informationen über einzelne Experten. Das Verfahren ist mit dem einer Internetsuchmaschine zu vergleichen. Infolgedessen kann das EF-System als eine Verbindung von Branchenbuch und Suchmaschine verstanden werden, in der einzelne Personen verwaltet werden.

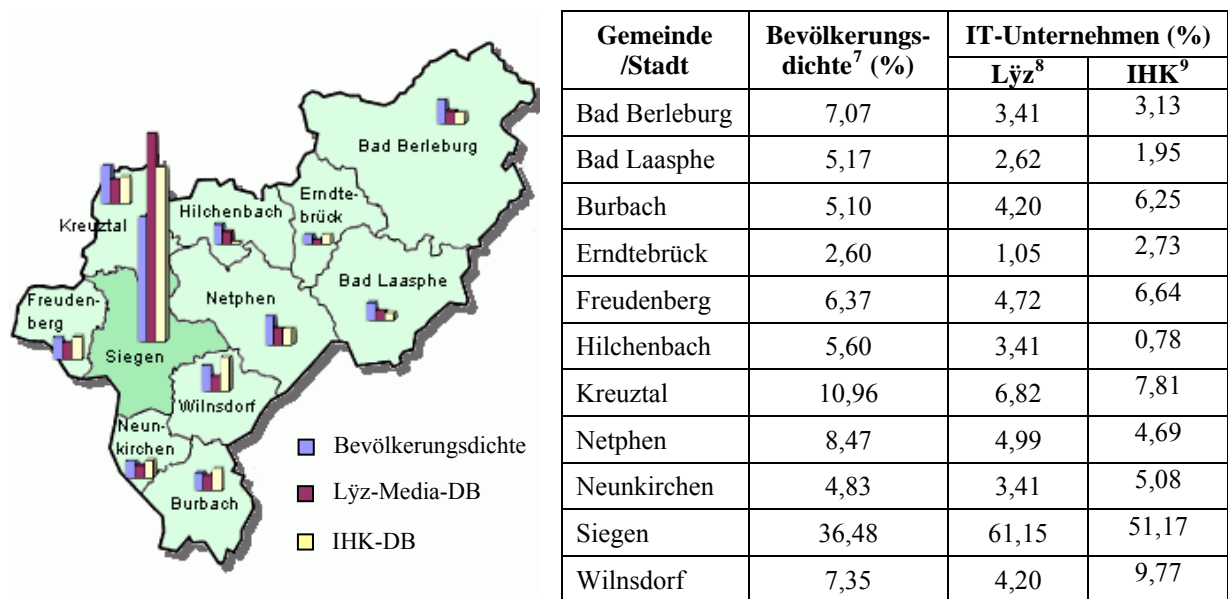
Im Rahmen dieser Arbeit wird das EF-System dahingehend angepasst, dass anstelle von personenspezifischen Daten Unternehmensdaten verwaltet werden. Dabei wird zum einen auf strukturierte Daten der Unternehmen zurückgegriffen. Zum anderen werden Informationen aus unstrukturierten Daten, die in Unternehmen anfallen, extrahiert. Durch das Bereitstellen der Informationen und der Suchmöglichkeiten auf einer Internetplattform soll die Wahrnehmung für regionale Unternehmen gesteigert werden. Damit wird insbesondere eine Steigerung von Kooperation zwischen den ansässigen Unternehmen der Region bezweckt. Nebeneffekt des Systems ist die Wahrnehmung möglicher Konkurrenten und deren technischen Fortschritte. Damit wird ein Anreiz gegeben, das eigene Unternehmen weiter zu entwickeln.

### 3 Forschungsfeld und Forschungsmethode

#### 3.1 Das Forschungsfeld

Im Fokus dieser Arbeit stehen u.a. IT- und Medienunternehmen der Region Siegen. Der Kreis Siegen-Wittgenstein untergliedert sich in elf Kommunen mit ca. 290.000 Einwohner. Mit rund 112.000 Einwohnern ist Siegen das wirtschaftliche und kulturelle Oberzentrum der Region Siegerland-Wittgenstein. Im Dreiländereck Nordrhein-Westfalen, Hessen und Rheinland-Pfalz gelegen, ist Siegen Handels- und Einkaufsstadt, Universitäts- und Kongressstadt im Grünen, in deren Einzugsgebiet rund 600.000 Menschen leben [Sieg2007a].

Tabelle 1 zeigt die Verteilung der Einwohner auf die einzelnen Kommunen.



**Tabelle 1: Verteilung von Medien-Unternehmen und Demografie**

Um einen Bezug zwischen der Bevölkerungsverteilung und den Standorten der einzelnen IT-Unternehmen der Region aufzuzeigen, ist die Lyz-Medien-Datenbank<sup>10</sup> der

<sup>7</sup> Statistisches Bundesamt (31.12.2005) in Prozent zur Gesamtbevölkerung des Kreises Siegen-Wittgenstein: [http://www.destatis.de/gv/suche\\_gv2000.htm](http://www.destatis.de/gv/suche_gv2000.htm)

<sup>8</sup> Lyz-Media-Datenbank der Wirtschaftsförderung des Kreises Siegen-Wittgenstein (stand 27.06.2007): <http://www.lyz-media.de/datenbanken/index.htm>

<sup>9</sup> Daten der IHK-Unternehmensdatenbank. Begrenzt auf Rundfunk- & Nachrichtentechnik, Nachrichtenübermittlung und Datenverarbeitung und Datenbanken (27.06.2007): <http://imteam-suedwestfalen.ihk.de/>



Wirtschaftsförderung des Kreises Siegen-Wittgenstein und die Datenbank der Industrie- und Handelskammer (IHK) in die Auswertung einbezogen worden. Für die IT- und Medienbranche wurden in der Lÿz-Datenbank 381 Unternehmen und im IHK-Branchenbuch 256 Unternehmen für den Kreis Siegen-Wittgenstein ermittelt. Danach sind mehr als 50 Prozent aller IT- und Medienunternehmen in der Stadt Siegen selbst angesiedelt. Das übrige Kreisgebiet lässt sich einteilen in Kommunen mit unmittelbarem Anschluss zur Stadt Siegen und in weiter entfernt liegende Kommunen. Die Ansiedlung von Unternehmen nimmt mit zunehmender Entfernung zur Stadt Siegen ab, wobei die Anzahl der Standorte in keiner Beziehung zur Bevölkerungsdichte der einzelnen Gemeinden stehen.

In einer für die IHK erstellten Studie der Prognos AG [Prog2004] über die „Zukunftsfähigkeit“ von Städten findet sich Siegen im direkten Vergleich zu insgesamt 439 nationalen Städten auf einem mittleren Platz 210 wieder. Die Studie beruht auf 29 Indikatoren, die sich in die Bereiche Demografie, soziale Lage, Arbeitsmarkt, Wettbewerb und Erreichbarkeit aufteilen. Als Ergebnis der Studie stellt die Prognos AG negative und positive Merkmale für den Kreis Siegen-Wittgenstein auf [Prog2004, S. 28]. Eine geringe Arbeitslosenquote und eine geringe Schuldenlast der Kommunen sprechen danach für den Standort Siegen. Negative Merkmale sind insbesondere die Forschungs- und Entwicklungsarbeit (FuE) der Region sowie die Anzahl der Unternehmensgründungen. Der geringe Beschäftigungsanteil in den so genannten „Zukunftsbranchen“, wozu u.a. auch die IT- und Medienbranche gehört, lässt sich durch die hohe Branchenkonzentration auf das metallverarbeitende Gewerbe erklären. Die politischen Maßnahmen der Region sind nach den Worten des früheren Bürgermeisters Ulf Stötzel entsprechend auf diesen Wirtschaftszweig ausgerichtet. Mit rund 50 Prozent aller versicherungspflichtigen Beschäftigten bietet die Metallbranche den Hauptbestandteil lokaler Kompetenzen in Siegen. Diese Kompetenzen sollen durch die kommunalen Wirtschaftsförderungen unterstützt werden [vgl. Sieg2007b, S. 1]. Auch Franz Mockenhaupt, Vorsitzender der IHK-Siegen, hat die Vorteile regionalpolitischer Maßnahmen erkannt und kritisiert in einem Interview mit dem Siegerlandkurier [Sieg2007c, S. 3] die schlechte „Publicity“ der Stadt. Eine Region muss nicht nur attraktiv und leistungsfähig sein, sondern sich ebenfalls am globalen Markt etablieren. Weiter stellt Mockenhaupt heraus, dass die Region Siegen mehr vorzuzeigen hat als ländlichen Raum und traditionell ausgerichtete Industrie. Es existieren erfolgreiche

---

<sup>10</sup> Im Folgenden wird diese Datenbank mit Lÿz-Datenbank abgekürzt.

mittelständische, international tätige Unternehmen, die durchaus Spitzentechnologie anbieten. Mit dieser Ansicht untermauert er die Aussagen von Ulf Stötzel über die Regionalpolitik, kritisiert aber auch gleichzeitig dessen enge Sichtweise. Er stellt heraus, dass die Region Siegen mehr als nur die Metallbranche bietet. Auch andere Branchen werden seiner Meinung nach in und durch die Stadt und Region repräsentiert.

Die Aussage des ehemaligen Bürgermeisters Stötzel zur Konzentration auf die bestehenden Kompetenzen hat durchaus ihre Berechtigung und wird durch die Meinung von Michael E. Porter bestätigt [Port1998a, S. 89] Diese enge Sichtweise erscheint mir für die Region Siegen nicht zeitgemäß. Überzeugender wirkt hier die Aussage des IHK-Vorsitzenden, der in der Region weitere Kompetenzen angesiedelt sieht. Dies belegt z.B. das von der Wirtschaftsförderung des Kreises Siegen regelmäßig veranstaltete *Ljz-Media-Breakfast* für die IT- und Medienbranche. Unternehmen der Region werden zu einer Diskussionsrunde zu ausgewählten IT-Themen eingeladen. Hauptanliegen dabei ist, die heimischen Unternehmen untereinander bekannt zu machen.

### **3.2 Methodisches Vorgehen**

Der gesamte Entwicklungsprozess ist in sechs Phasen gegliedert (vgl. Abbildung 9). Die Abbildung zeigt darüber hinaus, wie das Vorgehen in den AR-Zyklus einzuordnen ist.

In Phase zwei wurde das Grundkonzept des zu entwickelnden Systems festgelegt. Ziel war es, das in der Fachgruppe Wirtschaftsinformatik und Neue Medien entwickelte EF-System auf Unternehmen anzuwenden. Dabei sollte eine Internetplattform zur Verfügung gestellt werden, in der sich Unternehmen finden und eintragen können.

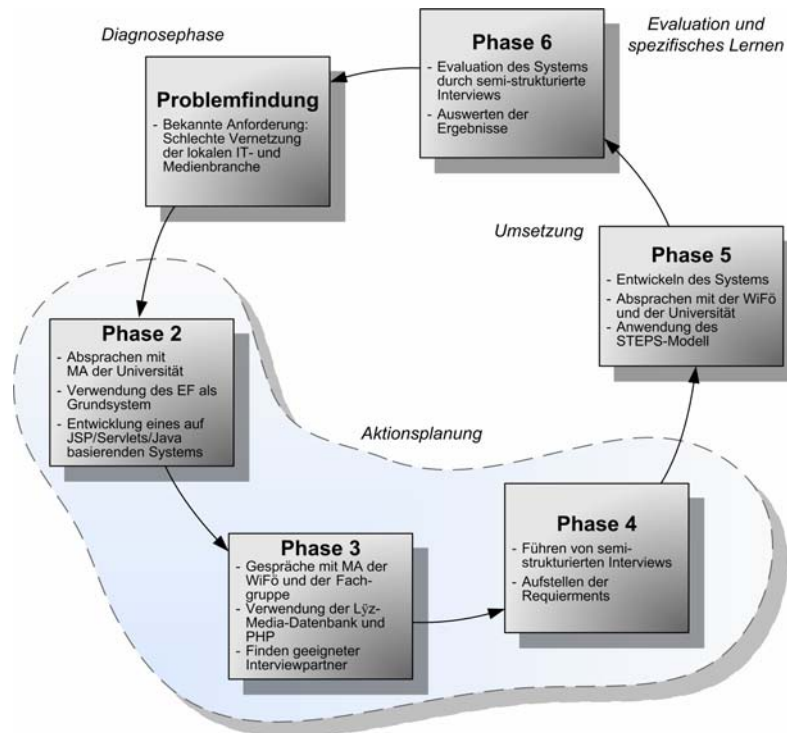


Abbildung 9: Entwicklungsprozess des Business Finder

Um eine rasche Anbindung eines webbasierten Systems an das EF-System zu ermöglichen, wurde zunächst eine auf JSP und Servlets aufgebaute Weboberfläche angedacht, die direkt auf die Komponenten des EF-Systems zugreifen konnte. Zeitgleich wurde das EF-System um eine standardisierte Schnittstelle mittels Webservices erweitert. Diese Schnittstelle wird für die Anbindung an die Weboberfläche genutzt. Die Arbeit wurde durch die Nutzung der Lütz-Media-Datenbank der Wirtschaftsförderung (WiFö) des Kreises Siegen-Wittgenstein unterstützt. Hierbei sollte das zu entwickelnde System in einer Umgebung des bestehenden Systems der WiFö integriert werden. Dieses Branchenbuch sollte Unternehmen der Region in geeigneter Weise finden und präsentieren. Der Wechsel von JSP auf PHP für die Oberflächengenerierung war empfehlenswert, weil das bestehende System der WiFö die dynamischen Oberflächen mittels PHP generiert. In dieser Phase sind geeignete Interviewpartner ausgewählt worden. Die nachstehende Tabelle zeigt die einzelnen Unternehmen und deren Branchen. Als Anbieter werden die Unternehmen bezeichnet, die in der IT-Branche angesiedelt sind. Die Anbieter können darüber hinaus auch als Nachfrager fungieren, indem sie evtl. Zukaufprodukte oder Dienstleistungen von Fremdanbietern in Anspruch nehmen. Die Nachfragerunternehmen produzieren in anderen Branchen und beziehen ausschließlich IT-Komponenten.

Unternehmen	Branche	IT-Nachfrager/Anbieter
A	Handel	Anbieter
B, K	Webdesign, Werbung	Anbieter
C	Archivierung	Anbieter
D, E, H, J, L, M, O, P	Softwareentwicklung	Anbieter
F	Schulungen	Nachfrager
I	Hardwareaufbau und Handel	Anbieter
G, N	Branchenfremde Großunternehmen	Nachfrager

Tabelle 2: Auflistung der Interviewpartner

Phase vier beschreibt die Interviews mit den Unternehmen, die mit Unterstützung von Mitarbeitern der Fachgruppe Wirtschaftsinformatik und Neue Medien durchgeführt wurden. Die Interviewpartner wurden zu Praxisfällen und zum Bedarf an kooperativen Maßnahmen befragt. Die aufgezeichneten Interviews ermöglichten eine präzise Auswertung. Eine ausführliche Beschreibung der Ergebnisse erfolgt in Kapitel 4. Aus den Ergebnissen der Interviews entstand ein Anforderungskatalog, der die Notwendigkeit der Unterstützung aber auch geäußerte Bedenken gegen die Nutzung eines solchen Systems aufzeigt.

In Phase fünf liegt der Themenschwerpunkt in der Systementwicklung. Dabei ist das STEPS-Modell berücksichtigt worden. Zunächst wurde mittels „rapid prototyping“ eine Oberfläche entwickelt, um eine plastische Darstellung über die geplanten Funktionalitäten des Systems zu gewinnen. Die Mitarbeiter der WiFö waren in den Entwicklungsprozess einbezogen. Zum Abschluss der Phase wurde im Rahmen eines *Ljz-Media-Breakfasts* das entwickelte System präsentiert. Anpassungen an PHP-Versionen und Systemeinstellungen verzögerten die Integration des Systems in die Umgebung der WiFö.

Nachdem das System zu Testzwecken auf einem Universitätsserver installiert war, konnte mit Phase sechs begonnen werden. Eine Evaluation des Systems ist Gegenstand dieser Phase. Weitere Interviews mit Unternehmen der IT- und Medienbranche sowie mit international branchenfremden Mittel- und Großunternehmen wurden durchgeführt. Den Gesprächspartnern wurde das System zu Testzwecken vorgestellt.

## 4 Bedarfsanalyse

Zur Feststellung des Bedarfs an einer technischen Unterstützung der Zusammenarbeit der regionalen Unternehmen ist eine empirische Studie durchgeführt worden. Diese Umfrage erfolgte bei einer repräsentativen Auswahl kooperativer klein- und mittelständischer Unternehmen (KMU) der IT- und Medienbranche der Region Siegen. Die Gespräche wurden in überwiegenden Fällen mit den Entscheidungsträgern der Firmen geführt.

Der vollständige Wortlaut der Interviews ist auf der beigelegten CD-Rom (Anhang D) dokumentiert und wird hier wie folgt referenziert, z.B. (Anhang D, Unternehmer B, Datum, S. 3).

### 4.1 Regionalität, Kooperation und Soziales Kapital

Die Aussagen der befragten KMUs lassen den Schluss zu, dass die Bereitschaft für eine regionale Zusammenarbeit grundsätzlich vorhanden ist. Die überwiegend überregionalen Geschäftsbeziehungen belegen diese Bereitschaft zur Zusammenarbeit allerdings nicht. Konkrete Ursachen für die mangelnden regionalen Beziehungen sehen die Befragten in der IT-Branche und deren Produkten selbst. Digitale Güter sind mit wenigem logistischem Aufwand und über weite Entfernungen wirtschaftlich zu handeln. Daraus folgern insbesondere die Unternehmen der IT-Branche, dass regionale Nähe für eine rentable Geschäftsverbindung nicht erforderlich ist. Dies mag in anderen Industriezweigen abweichend sein: „Ich denke mal im Industriebereich ist das sinnvoller bei großen, schweren Dingen [...]“ (Anhang D, Unternehmer A, 28.06.2006, S. 5).

Die mangelnde regionale Zusammenarbeit wird durch die zwei nachstehenden Kernaussagen untermauert:

- Die Unternehmen der Region sind „zu klein und unbedeutend“ um „größere“ Aufträge auszuführen
- Kein Unternehmen der Region ist in der Lage „größere“ Aufträge anzubieten.

Nach der ersten Aussage werden in der Region „größere“ Aufträge angeboten. Diese werden allerdings an überregionale Unternehmen vergeben in der Annahme, dass regionale Betriebe nicht in der Lage sind, sie auszuführen. Andererseits existieren in der Region lokale Kompetenzen, die „größere“ Aufträge suchen, denen die angebotenen Aufträge aus der Region aber nicht wirtschaftlich genug erscheinen.

Die Aussagen lassen außerdem den Schluss zu, dass mögliche Kooperationen nicht in ausreichendem Maße wahrgenommen werden. So könnte z.B. bei einer entsprechenden Informationsverteilung eine verstärkte gegenseitige Unterstützung bei der Annahme oder der Vergabe von Aufträgen möglich sein.

Folgende Aspekte stehen nach Aussagen der Interviewpartner für eine erfolgreiche Kooperation:

- Geschäftsbeziehungen basieren auf gegenseitigem Vertrauen.
- Der Kooperationspartner muss den Qualitätsansprüchen des eigenen Unternehmens genügen.
- Geschäftsbeziehungen entstehen durch Empfehlungsmarketing.

Grundvoraussetzungen für erfolgreiche Geschäftsbeziehung sind nach Meinung der Unternehmensvertreter gegenseitiges Vertrauen auf der Basis von persönlichen Kontakten und positiven Geschäftserfahrungen. Zur Wahrung der Kundenzufriedenheit werden die eigenen Qualitätsansprüche auf den Kooperationspartner übertragen. Mündliche Empfehlungen und Referenzen von Geschäftspartnern und Kunden bauen neue Geschäftsbeziehungen auf: „Also es läuft praktisch nur über [...] persönlichen Kontakt oder über Mundpropaganda. Ist [...] ein ganz großes Stück Vertrauenssache“ (Anhang D, Unternehmer L, 19.06.2006, S. 2).

## **4.2 Erwartungen und Bedenken**

Alle befragten Unternehmen sind nach ihren Aussagen in zahlreichen Branchenbüchern eingetragen. Während der durchgeführten Interviews war dennoch eine kritische Einstellung gegenüber Branchenbüchern im Allgemeinen unverkennbar. Diese Bedenken resultierten aus Erfahrungswerten der Unternehmen mit solchen Systemen.

Der Nutzen eines Branchenbuches wird von den Unternehmen allgemein als gering betrachtet. Der Aufwand für das Instandhalten korrekter Daten steht nach Auffassung der Unternehmensvertreter in keinem Verhältnis zum wirtschaftlichen Erfolg. Insbesondere ist die Pflege zahlreicher Branchenbücher unrentabel. Lediglich zwei Unternehmen konnten konkrete Geschäftsabschlüsse nachweisen, die auf Eintragungen in einem Branchenbuch zurückzuführen waren. Deshalb beschränken sich die Unternehmen auf die Pflege in wenigen ausgesuchten Branchenbüchern. Ausschlaggebend für eine gezielte Auswahl der Branchenbücher ist Qualität, Aktualität und die Pflege durch den Branchenbuchbetreiber.

Auch das Vertrauen und der Kontakt spielen hierbei eine nicht unwesentliche Rolle. So ist z.B. das Vertrauen eines Interviewpartners durch bewusst oder versehentlich falsche Angaben eines Branchenbetreibers verletzt worden: „Da sind massive Rechtschreibfehler drin, im Namen oder in der Adresse [...] wo man einfach das Gefühl hat [...] die zwingen einen schon indirekt dazu, dass man die Daten aktualisiert“ (Anhang D, Unternehmer M, 23.08.2006, S. 10).

Hinsichtlich der Eintragungen in Branchenbüchern bestehen unterschiedliche Meinungen bei den Unternehmen. Einmal wird nur die Eintragung branchenspezifischer Unternehmen gefordert andererseits besteht auch der Wunsch nach einer umfassenden Datenbank. Dabei wird die Vollständigkeit eines Branchenbuches vorausgesetzt. Wichtig ist den Firmen auch das Erreichen einer kritischen Nutzerzahl, damit solch ein System zum „Selbstläufer“ wird. Einer großen Nutzerzahl steht allerdings auch die Besorgnis gegenüber, mit Anschreiben (Spam) „überflutet“ zu werden.

Entgegen der allgemeinen Aussagen der befragten Unternehmen, dass Branchenbücher keinen wirtschaftlichen Vorteil bringen, zeigt sich aber doch das vorhandene Interesse in der geäußerten Bereitschaft aller Teilnehmer, den Business Finder nach seiner Fertigstellung zu testen. Weiter bestätigt die Kritik an den „herkömmlichen“ Branchenbüchern die Notwendigkeit eines neuen Ansatzes.

### **4.3 Anforderungen**

Die in 4.2 genannten Erwartungen der Firmen beziehen sich insbesondere auf die Aktualität und auf einen möglichst geringen Pflegeaufwand. Das System sollte Daten automatisiert und ohne Mehraufwand für das Unternehmen aktualisieren können. Da der Business Finder das EF-System erweitert, lag es nahe, aktuelle Informationen aus unternehmensspezifischen Textdokumenten zu extrahieren. Neben Werbematerial wie Flyer und Produktbeschreibungen sind die Webseite sowie der regelmäßig versendete Newsletter zur aktuellen Informationsgewinnung genannt worden. Während Webseiten von einigen Unternehmen nur sporadisch aktualisiert werden (in manchen Fällen alle drei Jahre), bietet der Newsletter aktuelle Informationen zum Unternehmensstatus. Dies verschafft eine bisher ungenutzte Möglichkeit, Informationen über ein Unternehmen zu sammeln.

Bei diesen „öffentlichen“ Datenquellen bestanden keine Sicherheitsbedenken der Unternehmer. Unternehmensinterne Dokumente, die nicht für die Öffentlichkeit bestimmt

sind, sollten allerdings nicht für das System freigegeben werden. Selbst dann nicht, wenn lediglich die daraus extrahierten Stichworte gespeichert würden.

Außerdem bemängelten die Unternehmen, dass das Auswählen von Dokumenten keine Aufwandsersparnis der Pflege sei, sondern lediglich eine Aufwandsverschiebung. Webseiten werden gewartet und bieten ständig ausreichend Informationen über das Unternehmen. Befürworter der Dokumentenanalyse empfehlen die Berücksichtigung des Lebenszyklus der durchsuchten Dokumente. Dies müsste bei der Bewertung entsprechend berücksichtigt werden.

Ebenfalls angesprochen wurde die Frage, wie Dokumente dem System bereitgestellt werden sollen. Die eine Möglichkeit, für das Durchsuchen der Dokumente ein lokales Programm (einen Client) zu installieren, wurde von den meisten Befragten abgelehnt. Bereitgestellte Dokumente sollen mittels „Upload“, also über eine Webschnittstelle dem System zugeführt werden. Daraus resultierend ergeben sich allerdings die oben angesprochenen Sicherheitsbedenken. Die Bedenken gegenüber einem System, das die lokale Festplatte durchsucht, sind dennoch größer, so dass der Webschnittstelle der Vorrang eingeräumt wird.

Als abschließender Punkt wurde allgemein der Wunsch nach einem benutzerfreundlichen System geäußert.

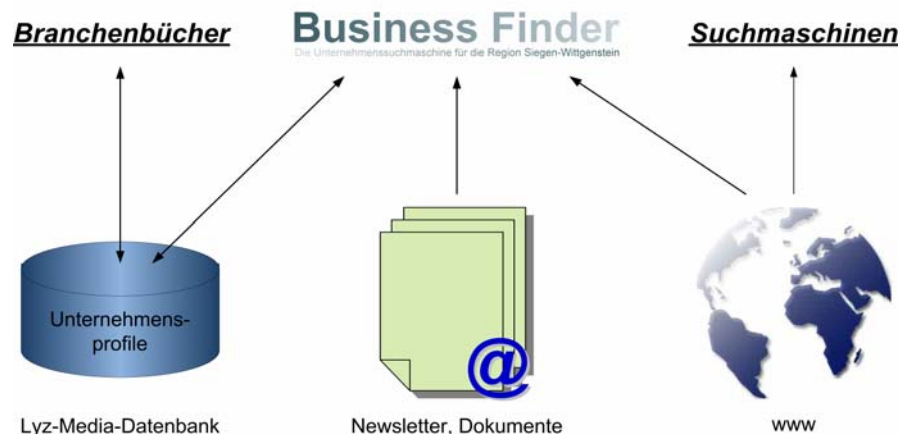


## 5 Der Business Finder

### 5.1 Konzept

Die sich aus den Interviews ergebene Anforderung, die Kooperation zwischen Unternehmen der Region weiter auszubauen, führte zur Entwicklung und Einführung des Business Finders.

Der Business Finder (BF) stellt eine zentrale Datenbank dar, in der Unternehmensprofile verwaltet, aktualisiert, miteinander verglichen und bewertet werden. Er soll für Internetnutzer eine transparente und übersichtliche Plattform zur Unternehmenssuche schaffen. Dabei bedient sich der Business Finder verschiedener technischer Verfahren. Zum einen werden Stammdaten von Unternehmen wie Name, Anschrift, Ansprechpartner oder der Weblink zur Unternehmens-Webseite verwendet, die durch die Wirtschaftsförderung des Kreises gepflegt werden. Dieses Verfahren ist an die bekannten Online-Branchenbücher angelehnt.



**Abbildung 10: Informationsgewinnung des Business Finder**

Zum anderen - und im Gegensatz zu den üblichen Branchenbüchern - wertet der Business Finder Webseiten der eingetragenen Unternehmen aus und erstellt daraus ein Schlüsselwortprofil, ähnlich dem Vorgehen von Internetsuchmaschinen wie Google.

Zu den gesammelten Daten aus der Datenbank und dem Internet ist es möglich, mit dem BF unternehmensspezifische Dokumente über das Portal hochzuladen. Dies können Produktkataloge, Werbeflyer, Preislisten oder sonstige Dokumente sein, die zumeist bereits

existieren. Die Dokumente werden nach Schlüsselworten durchsucht, die dann in das Unternehmensprofil eingepflegt werden.

Eine zusätzliche Funktion, die weder Branchenbücher noch Suchmaschinen unterstützen, ist das Durchsuchen von Newslettern. Eingetragene Unternehmen haben die Möglichkeit, ihren Kundennewsletter an das System zu senden. Auch dieser wird in das Unternehmensprofil eingepflegt.

## **5.2 Technische Grundlage: Das EF-System**

### **5.2.1 Modularität**

Bei der Entwicklung des EF-Systems spielten seinerzeit verschiedene Aspekte eine Rolle. Ein wichtiger Punkt war die Anpassbarkeit an neue Gegebenheiten. Wie sich im Entwicklungsverlauf des Systems gezeigt hat, hat von der ersten Umsetzung bis zum derzeitigen Stand eine umfangreiche Veränderung stattgefunden. Der modulare Aufbau des Systems hat sich bewährt. Bei der Entwicklung einer neuen Datenquelle zur Expertenbestimmung, ist es ohne größeren Aufwand möglich, ein entsprechendes Modul in das System einzubinden. Dies ist für die Verwendung des EF-Systems im Business Finder wichtig.

Derzeit werden Daten für die Bewertung (Matching) aus zwei Quellen gewonnen. Für jedes Matching existiert jeweils ein eigenes Modul. Das sind einerseits die eingetragenen Daten der Benutzer und andererseits der aus den Dokumenten gewonnene Index.

Beim Business Finder werden lediglich neue Datenquellen eingebunden. Er ist ein EF-System, das nicht auf Personendaten, sondern auf Unternehmensdaten zugreift.

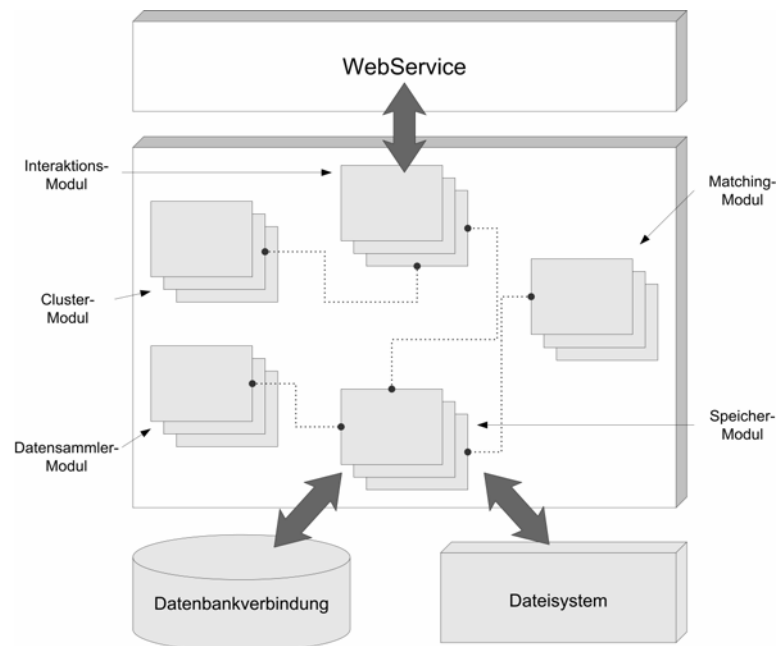


Abbildung 11: Architektur des EF-Systems [nach Reic2005a, S. 65]

Die einzelnen Module des EF-Systems werden in Kategorien zusammengefasst:

- *Interaktionsmodule* stellen die Verbindung zu anderen Systemen her, hier beispielsweise zu den Webservices
- *Clustermodule* realisieren die Clusterfunktionalitäten<sup>11</sup>
- *Datensammler-Module* extrahieren aus unterschiedlichen Personenquellen Daten
- *Speichermodule* legen in der Datenbank oder im Dateisystem die Daten ab
- *Matching-Module* beinhalten die Möglichkeit von Bewertungsverfahren.

Über die Webservice-Schnittstelle bietet das EF-System eine standardisierte Möglichkeit, mit anderen Softwareprodukten zu kommunizieren. Der Business Finder nutzt diese Schnittstelle für die Anbindung der Weboberfläche. Die hohe Flexibilität des EF-Systems ermöglicht eine schnelle und effiziente Integration neuer Module jeglicher Art.

### 5.2.2 Profilerzeugung und Matching

Für die Profilerzeugung werden unterschiedliche Arten von Daten herangezogen (vgl. Kapitel 2.5). Einmal erfolgt die Datengewinnung aus den Kontaktdaten (Name, Vorname, Schulbildung, Weiterbildung, etc.), die von den Benutzern einzutragen sind. Weiter werden

<sup>11</sup> Ein Relikt aus der Lernplattform. Diese Module erstellen „Kompetenzlandschaften“ zur Analyse der Community-Struktur [vgl. Reic2004, S. 181].

Informationen aus den lokal abgelegten Dokumenten der Benutzer gewonnen [vgl. Reic2005a, S. 63]. Über das Client-System des EF-Systems werden die Dokumente von der lokalen Festplatte des Nutzers nur mit dessen Einverständnis ausgewählt und dem EF-System zur Verfügung gestellt. Dem Nutzer bleibt dabei die Kontrolle über die Auswahl der Dokumente vorbehalten.

Bei der im EF-System verwendeten Textanalyse von digitalen Dokumenten ist das Erkennen des Dateiformates und das Extrahieren des Klartextes aus diesen Dateien erforderlich [vgl. Reic2005a, S. 65]. Grund hierfür ist die unterschiedliche Speicherung der jeweiligen Dateiformate. Reine Textdokumente sind einfach zu interpretieren, da sie nur *reinen* Text speichern. Microsoft Word-Dateien, PDF-Dokumente oder andere Dateien enthalten neben dem Klartext Metainformationen zur Textformatierung und -strukturierung. Die Gewinnung des Klartextes erfolgt in der Programmiersprache Java durch bestehende Bibliotheken.

Das System unterstützt die folgenden Dateiformate, die einen Großteil der in Unternehmen verwendeten Textformate bilden:

- txt (einfache Textdateien),
- rtf (Rich-Text-Format-Dateien),
- pdf (Adobe Acrobat Reader Dateien),
- doc (Microsoft Word-Dokumente),
- xls (Microsoft Excel-Arbeitsmappen),
- pps, ppt (Microsoft Powerpoint-Präsentationen) und
- html, htm etc. (Standard-Webseitenformat).

Das Erkennen der Sprache ist notwendig für die Stoppwortfilterung. Ist die Sprache nicht erkannt, können Stoppworte nicht entfernt werden. So ist z.B. dem Wort „die“ in der deutschen Sprache eine andere Bedeutung zuzumessen, als im Englischen. Als bestimmter Artikel der deutschen Sprache kann er problemlos bei der Indexierung ignoriert werden. In der englischen Sprache hat das Wort jedoch eine Bedeutung. Sind alle Stoppworte entfernt, liegt eine Wortfolge vor, in der jedem Wort eine Bedeutung zuzumessen ist. Aus dieser Wortfolge werden sowohl Teilfolgen als auch einzelne Terme extrahiert, die dann in einem Finalschrift zum Benutzerprofil zusammengefügt werden. Dabei werden die Kombinationen und einzelnen Deskriptoren mit der Häufigkeit ihres Auftretens gespeichert. Beim Matching

wird schließlich eine Bewertung vorgenommen, die die Dokumentenhäufigkeit einzelner Deskriptoren einbezieht.

Die Antwort auf die möglicherweise fiktiv aufkommende Frage, warum der BF nicht das in Google erfolgreich angewandte Verfahren PageRank verwendet, das die beste Möglichkeit bietet, Informationen aus Texten zu analysieren, liegt in dem Bezug zur Aufgabe der Systeme. Google versucht eine umfassende Datensammlung von Dokumenten zu erhalten und stellt diese Dokumente in Beziehungen zueinander. Der BF soll Unternehmen identifizieren und eine Kontaktabahnung mit diesen ermöglichen. Die Dokumente, die die Unternehmen bereitstellen, können unterschiedlichster Art sein. Hier sind nicht ausschließlich wissenschaftliche Arbeiten propagiert. In Organisationen fallen häufig eine Reihe spezifischer Dokumente an, die die Unternehmen widerspiegeln und dennoch keinen wissenschaftlichen Charakter (Verweise auf weitere Dokumente) aufweisen. Vor diesem Hintergrund hätte ein Verfahren wie PageRank ohne jegliche Zitate und Verweise wenig Erfolgsaussichten.

### **5.3 Die Architektur des Business Finder**

Der Business Finder ist eine webbasierte Anwendung aus vier Komponenten (vgl. Abbildung 12). Er ist eine Client-Serverarchitektur, bei der ein internetfähiger Browser das Client-System darstellt. Die Abbildung zeigt die vier verwendeten Komponenten und den Informationsfluss bei einer Suchanfrage. Ausgehend vom lokalen Rechner (dem Client) wird die Suchanfrage an das System über das Internet (1 und 2) weitergeleitet. Die Anfrage wird zuerst an einen PHP-fähigen Webserver gesendet, der u.a. die Benutzerschnittstelle zur Verfügung stellt. PHP ist speziell für die Kommunikation mit Datenbanken vorgesehen. Einfache Auflistungen und Manipulationen an den Stammdaten von Unternehmen sind zum Teil von der WiFö bereits vorhanden und werden deshalb von PHP direkt bearbeitet (9 und 10). Bei Stichwortsuchanfragen wird die Anfrage über Webservices und das EF-System geleitet.

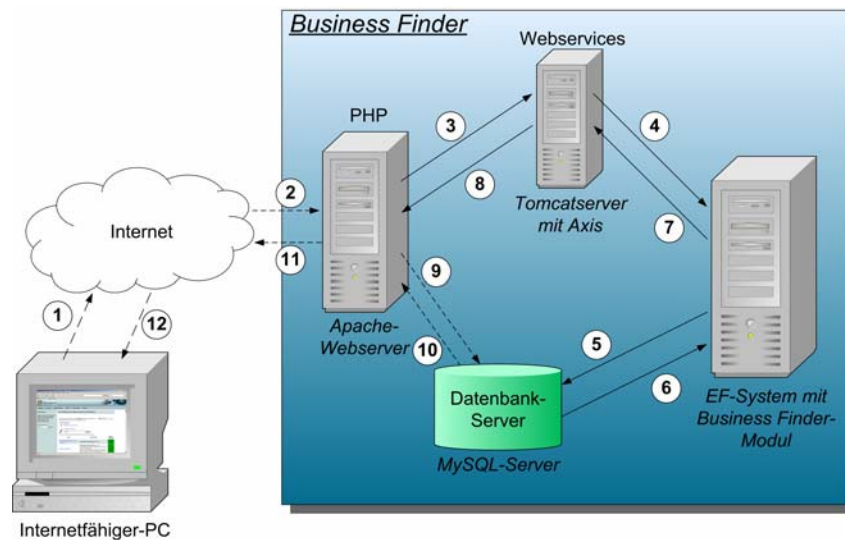


Abbildung 12: Die vier Komponenten des Business Finder

Die Webservices des EF-Systems bilden die Verbindung zwischen dem EF-System und der „Außenwelt“. Hiermit werden verschiedene Funktionalitäten des EF-Systems für fremde Softwaresysteme in standardisierter Form bereitgestellt. Es ist somit möglich, auf einfache Weise das EF-System zu nutzen und in andere Systeme einzubinden. Suchbegriffe an den BF werden aus der Eingabemaske (also der PHP-Seite) an die Webservices übermittelt (3). Dieser leitet die Anfrage an das EF-System weiter (4) und wartet auf das Ergebnis. Für die Webservices wird ein zusätzlicher Container benötigt, der sie verwaltet und bereitstellt.

Das EF-System benötigt keine weiteren Softwaresysteme, da es ein selbstständig laufendes Programm ist. Dabei erledigt es zwei zentrale Aufgaben:

- Erzeugen von Schlüsselwortprofilen aus verschiedenen Datenquellen und
- Berechnen der Ergebnisse bei Suchanfragen.

Da die Struktur und die Aufgabe des Business Finder von der des EF-Systems abweichen, sind Anpassungen notwendig. Es existiert kein eigenständiges Clientsystem. Deshalb sind die Module, die im EF-Client die Profilerzeugung generieren (vgl. Kapitel 5.2.2), auf den Server zu portieren. Diese Module müssen entsprechend angepasst werden, damit sie die Daten sowohl aus den hochgeladenen Dateien, aus den bereitgestellten Newslettern als auch aus den Internetauftritten der Unternehmen extrahieren. Internetseiten liegen permanent vor und die daraus extrahierten Daten müssen in bestimmten Abständen aktualisiert werden. Für einen möglichst geringen Arbeitsaufwand des Nutzers ist ein Verfahren notwendig, welches in

bestimmten Zeitabständen automatisch die Profilvergenerierung initialisiert und somit die verschiedenen Dokumente analysiert.

Um an die Daten der Internetseiten zu gelangen, ist das Modul `BFWebConnection` entwickelt worden. Es verbindet sich mit einer Unternehmens-Webseite und stellt den Inhalt zur Verfügung. Dabei durchläuft das System rekursiv die Seiten und sammelt auf diese Weise möglichst viele Informationen über das Unternehmen. Empfangene Newsletter werden automatisch nach Stichworten durchsucht. Dafür ist das EF-System um einen Emailclient erweitert worden. Außerdem waren Anpassungen zur Datenbankverbindung nötig.

Bei Suchanfragen wird auf die dort abgelegten Daten zurückgegriffen (6) und eine Ergebnisliste berechnet. Diese Trefferliste wird an die Webservices zurück geliefert (7). Er generiert daraus eine Antwort für die PHP-Seite. Die zurückgegebene Tabelle beinhaltet pro Zeile (Datensatz) die Identifizierungsnummer der Unternehmen, die berechnete prozentuale Übereinstimmung der verschiedenen Matchingmodule sowie einen Kommentar, wie dieser Wert berechnet wurde. Zurzeit finden zwei Matchingmodule ihren Platz im System. Ein Modul berechnet die Übereinstimmung mit den Stammdaten der Unternehmen, das andere vergleicht das Schlüsselwortprofil mit der Suchanfrage. Hierzu sind auf Basis vorhandener Module des EF-Systems neue Module implementiert worden.

Mit den aus der Berechnung des EF-Systems gewonnenen Daten generiert der BF eine nach prozentualer Übereinstimmung absteigende Tabelle, um so eine Empfehlung bezüglich der Suchanfrage zu geben. Die dynamisch generierte Webseite wird über das Internet dem verbundenen Browser übermittelt (11 und 12).

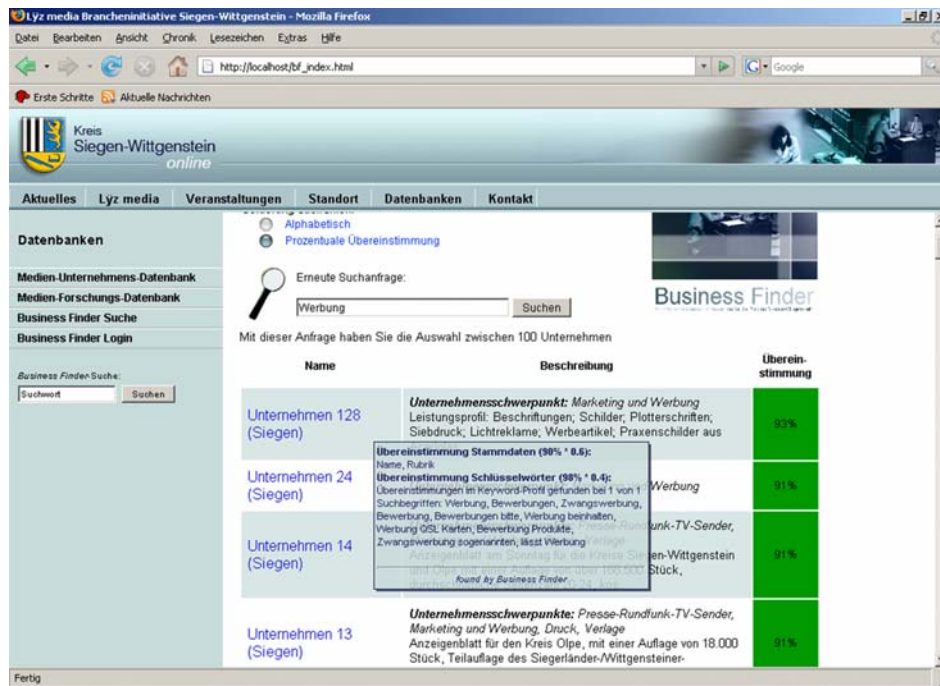


Abbildung 13: Matchingergebnis des Business Finder

Das Ergebnis für eine Suchanfrage z.B. nach dem Wort „Werbung“ zeigt Abbildung 13. Neben dem Unternehmensnamen (hier anonymisiert) wird eine kurze Beschreibung mit den Unternehmensschwerpunkten aufgeführt. In der letzten Spalte wird die gesamte prozentuale Übereinstimmung des Unternehmensprofils mit der Suchanfrage dargestellt. Ebenfalls ist ein so genanntes Tooltip-Fenster zu sehen welches beim Überfahren eines Eintrags mit der Maus erscheint. Es enthält detaillierte Informationen über die Bewertung des Unternehmens. Oberhalb der Tabelle besteht die Möglichkeit eine neue Suchanfrage zu generieren oder die Tabelle alphabetisch zu sortieren. Darüber hinaus ist die Einbettung des BF in die Umgebung der Wirtschaftsförderung des Kreises Siegen-Wittgenstein erkennbar.

Abbildung 14 zeigt die erweiterten Optionen für eingetragene Unternehmen. Sie haben jetzt die Möglichkeiten, die Webseite, Newsletter und Dokumente für die Stichwortsuche bereitzustellen. Im Bereich *Hochladen und Scannen von Dokumenten* können Dokumente von der lokalen Festplatte auf den Server geladen werden. In dem Listenfeld werden alle für das Unternehmen aktuell zu scannende Dokumente aufgeführt und die seit dem letzten Login durchgeführten Scans dargestellt. Die Webseite und die Absenderadresse werden ebenfalls auf dieser Seite eingetragen. Darüber hinaus können die Unternehmen selbstständig entscheiden, was mit dem Schlüsselwortprofil geschieht. Der Standardfall ist, dass die bereitgestellten Dokumente mit dem bestehenden Profil gemischt werden. Darüber hinaus besteht die Möglichkeit, ein neues Profil zu erzeugen oder das Bestehende beizubehalten.



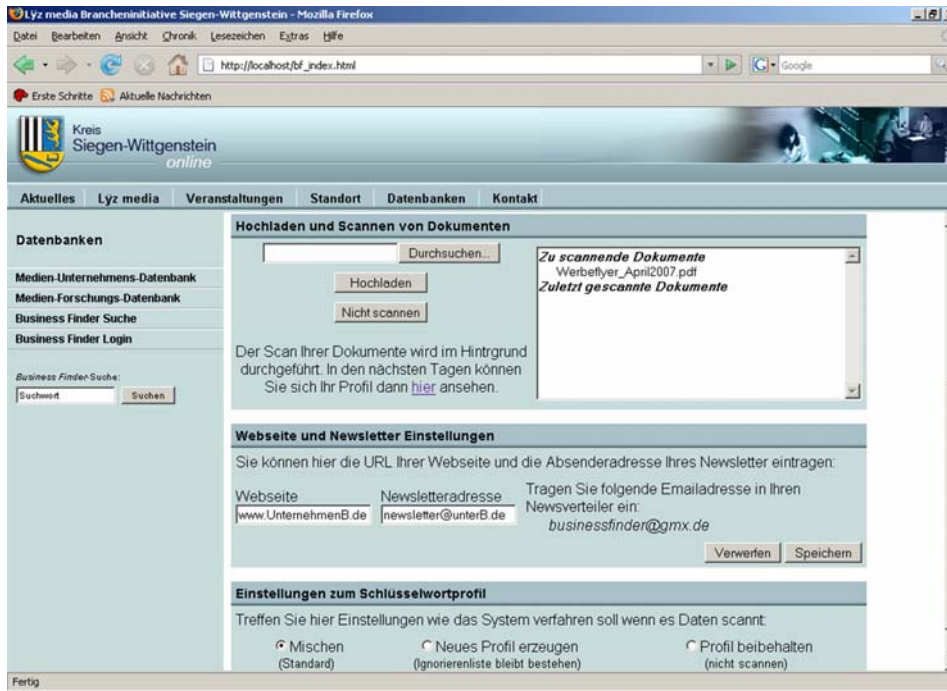


Abbildung 14: Spezifische Einstellungen zur Profilgenerierung

Den Benutzern soll eine möglichst große Kontrolle über ihr Profil im System gegeben werden. Dennoch sind Einschränkungen notwendig, um das System vor ungewollten Manipulationen zu schützen. Es ist möglich, gezielt Begriffe aus dem Profil zu entfernen, um sie auf eine „Ignorierenliste“ zu setzen (Abbildung 15). Begriffe dieser Liste werden bei Suchanfragen und Stichwortanalysen nicht berücksichtigt, sie werden „ignoriert“.

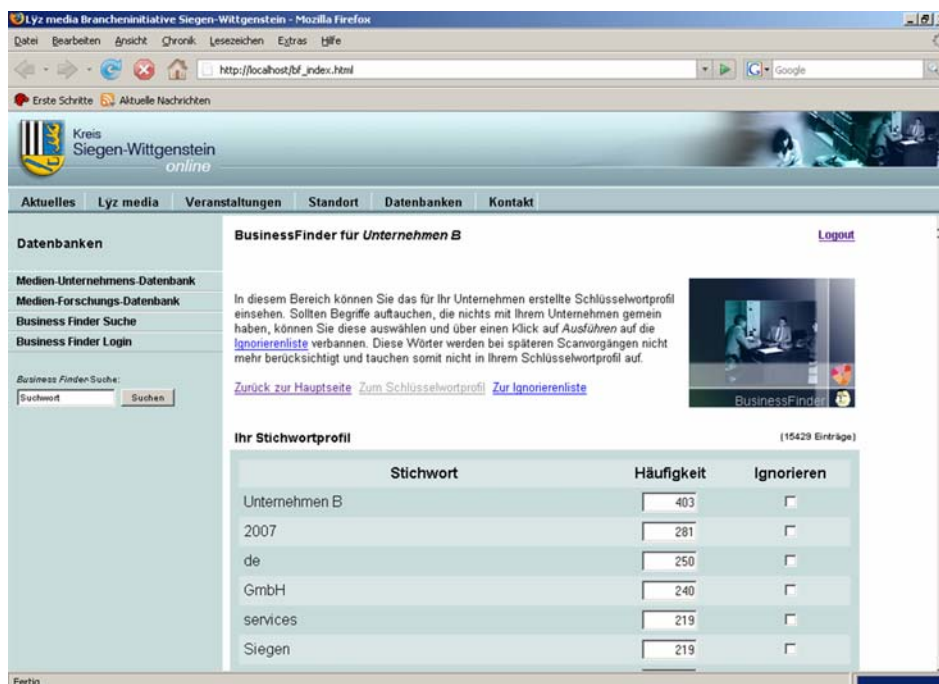


Abbildung 15: Schlüsselwortprofil von Unternehmen B

## 5.4 Implementierung

Die Architektur des Business Finders gibt zwei Programmiersprachen vor: *Java* für die Module des EF-Systems und *PHP* für das Branchenbuch der Wirtschaftsförderung des Kreises Siegen-Wittgenstein. Für eine ausführliche Dokumentation der implementierten Klassen befinden sich im Anhang C die entsprechenden Klassendiagramme sowie auf der beigelegten CD-Rom (Anhang D) der vollständige Quellcode mit Dokumentation.

### 5.4.1 PHP-Funktionalitäten

Zur Verdeutlichung der Funktionalitäten des Systems ist zunächst im „rapid prototyping“-Verfahren eine Weboberfläche mittels PHP erstellt worden. Dabei war das bereits bestehende Branchenbuch der WiFö zu berücksichtigen. Farben, Funktionalitäten und Prozessabläufe sind entsprechend in das Umfeld des bestehenden Systems integriert.

Es waren weitere Funktionalitäten in PHP zu realisieren:

- Kommunikation mit dem Nutzer (GUI<sup>12</sup>),
- Kommunikation der Oberfläche mit dem EF-System über Webservices,
- Sortierung und Darstellung der Suchergebnisse,
- Anzeigen und manipulieren des Schlüsselwortprofils,
- Anzeigen und manipulieren der Ignorierliste,
- Dateien „upload“ und
- Authentifizierungsmechanismen.

Bei der Gestaltung der GUI werden dynamische PHP-Kontexte in einfache HTML-Seiten eingebunden. Für diese Dynamik sind Datenbankroutinen hinterlegt. PHP ist in seiner Konzeption darauf ausgerichtet, dynamische Webseiten mittels einer Anbindung an eine Datenbank zu gestalten. Entsprechende Bibliotheken sind in der Sprache selbst enthalten. Eine effiziente und einfache Anpassung an die vorhandene Lÿz-Datenbank ist somit möglich. Komplexer gestaltete sich die Implementierung der Suchprozesse. Bei einer Suchanfrage übermitteln die Webservices der Klasse *Matching* die Ergebnisse als eine Zeichenfolge in HTML-Tabellenform. Versuche, komplexere Datenstrukturen als eindimensionale Arrays über die Webservices zu senden, konnten nicht umgesetzt werden [vgl. Brei2007, S. 58]. Es

---

<sup>12</sup> Graphical User Interface – Grafische Benutzerschnittstelle

ist möglich, diese Tabelle direkt in eine PHP-Seite einzubinden. Da bei der Suchanfrage wegen des Webservicesproblems ausschließlich die Identifikationsnummer sowie Informationen über die Suchergebnisse zurückgegeben werden, ist es notwendig, diese Tabelle in einzelne Datensätze und Datenfelder zu zerlegen. Dies übernimmt in der Klasse `Matching` die Methode `splitResult()`, die ein zweidimensionales Array aus den Datensätzen erzeugt. Dieses Array wird dann an die Methode `buildTable()` weitergegeben, die mit Hilfe von Datenbankroutinen und der Identifikationsnummer weitere Daten aus der Datenbank anfordert. Aus den gesamten Daten generiert sie die komplette Ergebnistabelle, die in der PHP-Seite `findings.php` angezeigt wird. Damit das Bild dieser Tabelle mit den Ergebnistabellen der WiFö übereinstimmt, sind detaillierte Informationen zu den Suchergebnissen in kleine Fenster ausgelagert (vgl. Abbildung 13). Diese so genannten Tooltips erscheinen beim „Überfahren“ der Ergebnisse mit der Maus. Da diese Funktion von einem Ereignis abhängig ist, das lokal in dem Browser des Benutzers ausgelöst wird, ist das Abfangen dieses Ereignisses nicht mittels PHP möglich. PHP ist eine serverseitige Skriptsprache, die nicht auf clientseitige Ereignisse reagieren kann. Eine weitere Programmiersprache ist notwendig. *Javascript* bietet die benötigten Funktionalitäten. Für die Tooltips ist eine fertige Javascript-Datei von Walter Zorn<sup>13</sup> verwendet worden. Die Datei bietet umfangreiche Einstellungsmöglichkeiten und kann mittels Javascript angesprochen werden.

Die alphabetische Sortierung ist durch die Erweiterung der Methode `buildTable()` um einen Sortieralgorithmus möglich.

Abbildung 16 zeigt den Vorgang von dem Erhalten der Ergebniszeichenkette bis zur fertigen Tabelle.

---

<sup>13</sup> Siehe <http://www.walterzorn.de/>

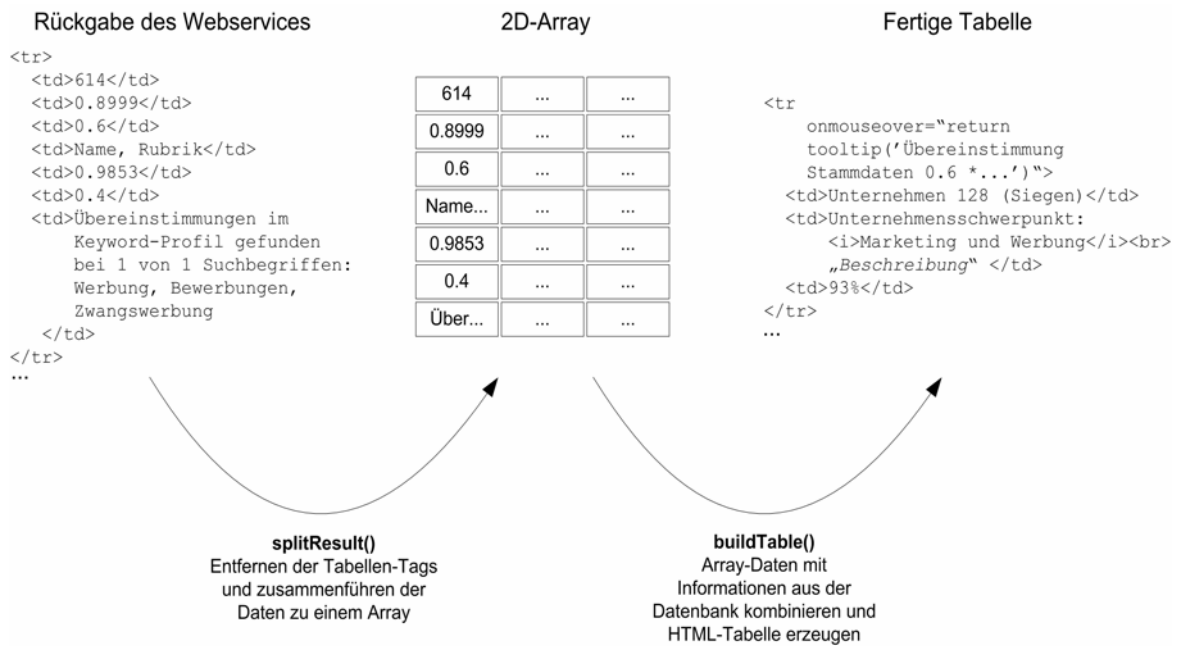


Abbildung 16: Verarbeitung der Webservice-Rückgabe

Die bisher beschriebenen Funktionalitäten beziehen sich auf das Matchingverfahren. Eingetragene Unternehmen haben die Möglichkeit, ihr Profil zu bearbeiten. Dies umfasst die Manipulation des Schlüsselwortprofils und der Ignorierliste sowie die Option, Dokumente auf den Server zu transferieren.

Die Manipulation des Schlüsselwortprofils und der Ignorierliste ist durch standardisierte Datenbankabfragen realisiert. Für den Dateienupload ist die zusätzliche Klasse `FileSharing` implementiert worden. Damit werden die hochgeladenen Dateien auf dem Server gespeichert und verwaltet.

Der Algorithmus besteht aus fünf Schritten:

1. Temporären Dateinamen erzeugen.
2. Überprüfen, ob die Datei größer als 0 Byte ist, also einen Inhalt besitzt.
3. Feststellen der Dateierweiterung und Überprüfen auf Zulässigkeit.
4. Bei Gültigkeit von 2. und 3. speichern der Datei in einem Verzeichnis mit der Identifikationsnummer des Unternehmens und unter einem generierten Schlüssel mit Datum als Dateinamen.
5. Datenbankeintrag durchführen.

Schritt 5 des Algorithmus ist notwendig, damit die BF-Module im EF-System die Dateien den richtigen Unternehmen zuordnen können. Grund ist die zeitliche Verzögerung, die zwischen dem Hochladen einer Datei und ihrem Scanvorgang liegen kann.

Als weitere Funktion ist ein Authentifizierungsmechanismus implementiert, der die Unternehmensprofile vor unberechtigtem Zugriff schützt. PHP liefert für solche Vorgänge die entsprechenden Bibliotheken. Die Klasse `SessionTracking` erzeugt nach der Authentifizierung mit Unternehmensnamen und Passwort eine neue *Session (Sitzung)*. Der Benutzer meldet sich einmalig an und kann sich dann frei zwischen den kritischen Seiten bewegen, bis er die Sitzung beendet. Als zusätzlicher Sicherheitsmechanismus wird die Sitzung bei Inaktivität nach einem vorgegebenen Zeitfenster automatisch durch das System beendet.

#### **5.4.2 Erweiterungen des EF-Systems durch BF-Module**

Die im Zuge dieser Arbeit für das EF-System implementierten Module sind in zwei Gruppen zu unterscheiden. Eine Gruppe ist für die Datenbeschaffung und Verwaltung zuständig, während die Matchingmodule die Bewertung der vorhandenen Daten bei Suchanfragen vornehmen. Zu den Verwaltungsmodulen gehören das übergeordnete Modul `Businessfinder` sowie die Klasse `BFTimer`. Diese beiden regeln gemeinsam den technischen und den zeitlichen Ablauf der Datenbeschaffung. Module werden im EF-System durch zwei Klassen repräsentiert<sup>14</sup>. Eine Klasse dient der eigentlichen Ausführung der Aufgabe. Zu dieser Klasse wird eine zweite Klasse implementiert, die zur Erzeugung eines oder mehrerer Objekte der ersten Klasse dient. Üblicherweise tragen diese Klassen den ursprünglichen Klassennamen mit dem zusätzlichen Postfix „Factory“. Im Falle des Businessfinder-Moduls besteht dieses Modul also aus den Klassen `Businessfinder` und `BusinessfinderFactory`. Die Klasse `BFTimer` ist eine „normale“ Javaklasse, die einen Zeitstempel generiert. Dieser Zeitstempel löst alle fünfzehn Minuten die Suche nach hochgeladenen Dokumenten und eingegangenen Newslettern auf dem Server aus. Er wird außerdem verwendet, um in einem einstellbaren Zeitraum das Durchsuchen von Webseiten zu initialisieren. Dazu erzeugt die Klasse `BFTimer` ein Objekt der Klasse `BFScanning`. Abhängig vom Aufruf dieser Klasse wird entweder ein kompletter Scanvorgang (Newsletter,

---

<sup>14</sup> Dieses Verfahren ist an das allgemein bekannte Entwurfsmuster „Factory Method“ angelehnt [vgl. Gamm2004, S. 131]

Dokumente und Webseiten) gestartet oder eine Analyse ohne die Berücksichtigung der Webseiten ausgeführt. Diese Unterscheidung geschieht aus Performancegründen, da das Durchsuchen sämtlicher Webseiten einen enormen Aufwand darstellt.

Für das Analysieren von Dokumenten und Newslettern sind jeweils nur wenige Zugriffe notwendig. Ergänzend zu dem in Kapitel 5.2.2 beschriebenen Vorgehen sind aufgrund des Zeitunterschiedes zwischen dem Hochladen einer Datei und der Textanalyse zwei zusätzliche Schritte notwendig. Der Ablauf des Dokumentenscans für ein Unternehmen lässt sich wie folgt beschreiben:

1. Information (Speicherort, Name der Datei) über hochgeladene Dokumente aus der Datenbank laden.
2. Den Text wie in Kapitel 5.2.2 beschrieben nach Schlagworten durchsuchen und die Häufigkeit des Auftretens bestimmen.
3. Speichern der gewonnenen Informationen in der Datenbank.

Der Algorithmus für das Durchsuchen von eingegangenen Newslettern gestaltet sich ähnlich. Zu Beginn werden alle eingegangenen Mails mittels des Moduls `BFPop3Connection` von einem Mailserver in das EF-System heruntergeladen. Danach findet eine Zuordnung der Absenderadresse der Mail zu einem Unternehmen statt. Die weitere Verfahrensweise ist ähnlich dem genannten Ablauf des Dokumentenscans: Bestimmen des Dateityps der Dateianhänge, Parsen, Schlagworte extrahieren und speichern in der Datenbank.

Die Analyse einer Webseite ist eine komplexe Aufgabe. Webseiten sind in einem hohen Maße heterogen aufgebaut und weisen meist eine geschachtelte Struktur auf. Informationen stehen nicht ausschließlich auf der ersten Seite einer Internetpräsenz. Zunächst wird die URL des Unternehmens aus der Datenbank geladen und eine Verbindung zur Webseite aufgebaut. Sie wird nach Schlagworten durchsucht und die Schritte drei bis fünf werden wie oben beschrieben durchgeführt.

Um nicht nur die Informationen der ersten Seite zu ermitteln, ist ein Algorithmus notwendig, der die Seiten zunächst nach weiterführenden Weblinks durchsucht. Diese können allerdings auf Internetseiten von Fremdanbietern verweisen und Informationen beinhalten, die nichts mit dem aktuell ausgewählten Unternehmen gemein haben. Folglich ist eine Differenzierung zwischen externen und internen Links notwendig. Der Algorithmus, der die Identifizierung

der Links vornimmt, findet sich in der Klasse `BFWebSearch` wieder. Unter Zuhilfenahme des Moduls `BFWebConnection` verbindet sich die Klasse mit der übergebenen Webseite und durchsucht sie nach internen Links.

Für die Suche nach internen Links wird davon ausgegangen, dass die Maindomain, die übergeordnete URL, bekannt ist. Weiter liegt die Annahme zugrunde, dass Weblinks im HTML-Code in einer der beiden folgenden Formen vorliegt:

- `<a ... href="LINK">` oder
- `<frameset ... src="LINK">`

Die Punkte symbolisieren eine beliebige Anzahl von Zeichen mit Ausnahme des Kleinerzeichens (`>`), da dies eine Anweisung in HTML schließt. *LINK* steht für die gesuchte URL. Das Auffinden aller URLs ist durch eine einfache lexikalische Analyse möglich: Zunächst werden alle nicht druckbaren Zeichen (Leerzeichen, Zeilenumbruch etc.) entfernt. Die Zeichenkette wird dann nach „`<a`“ bzw. „`<frameset`“ durchsucht. Anschließend sind die Attribute „`href`“ bzw. „`src`“ zu finden und die Zeichenkette zwischen den folgenden beiden Anführungszeichen zu extrahieren. Es ist zu beachten, dass diese Links vor dem nächsten Kleinerzeichen stehen um somit dem „Tag“ zuzuordnen sind.

Internetseiten können als gerichtete Graphen gesehen werden, bei denen die Knoten Webseiten repräsentieren und die Kanten als Weblinks der Seite. Um einen Graphen zu durchsuchen, werden im Allgemeinen zwei Arten unterschieden: Die *Breitensuche* (*breadth-first-traversal*) und die *Tiefensuche* (*depth-first-traversal*). Bei der Breitensuche werden ausgehend von einem Knoten zunächst alle direkten Nachfolger untersucht. Danach wird ein Kindsknoten nach dem anderen auf die gleiche Art betrachtet. Es werden also immer zuerst alle Knoten einer Ebene betrachtet. Bei der Tiefensuche wird dagegen eine vollständige Kantenfolge behandelt. Von einem Knoten wird ein direkter Nachfolger ausgewählt und von diesem wiederum ein Kindsknoten. Dies wird solange durchgeführt, bis ein Knoten keine weitere Kante mehr besitzt oder nur noch auf bereits besuchte Knoten verlinkt. Dann wird eine Ebene zurückgegangen und der Vorgang rekursiv mit der nächsten Kante als Startpunkt durchgeführt.

Bei der Suche nach internen Weblinks auf einer Seite findet das Tiefensuchverfahren Anwendung. Die Suche erfolgt in sieben Schritten:

1. Suche alle Links auf der angegebenen Seite (wie oben beschrieben) und speichere sie in einer Liste.
2. Füge alle Links, die die URL der Unternehmenswebseite enthalten, zu einer neuen Liste „gültige Links“ hinzu und entferne sie aus der Liste der zu untersuchenden Links.
3. Alle Links die mit „/“ oder „../“ beginnen, sind relative Links<sup>15</sup> innerhalb der Webseite.
  - a. Füge allen Links die mit „/“ anfangen, die aktuell beobachteten URL als Präfix hinzu.
  - b. Falls der Link mit „../“ beginnt, entferne den letzten Abschnitt der aktuell beobachteten URL, ab dem vorletzten „/“. Entferne „../“ des Links. Sind die ersten Zeichen des Links wieder „../“, wiederhole diesen Vorgang, bis der Link nicht mehr mit „../“ beginnt. Füge dann die beobachtete URL und den Link aneinander.
  - c. Speichere die auf diese Weise erzeugten Links in der Liste der gültigen Links und entferne sie aus der Liste der zu untersuchenden Links.
4. Entferne alle Links, deren Anfang „www.“ oder „http://“ ist aus der Liste der zu untersuchenden Links.
5. Für alle übrigen Links füge das Präfix der aktuell beobachteten URL hinzu und überprüfe, ob das Ergebnis eine gültige Adresse ist. Wenn ja, füge es der Liste der gültigen Links hinzu, ansonsten lösche den Eintrag aus der Liste der zu untersuchenden Links.
6. Füge die aktuell beobachtete URL zu der Liste der gültigen Links hinzu.
7. Entferne alle doppelten Einträge in der Liste der gültigen Links.

---

<sup>15</sup> Das Zeichen „/“ verweist auf ein Unterverzeichnis der aktuellen Webseite. Mit „../“ wird von dem Ort der aktuellen Webseite eine Verzeichnisebene aufwärts beschrieben. Ausgehend vom aktuellen „Ort“ wird also *relativ* auf das Ziel verlinkt.



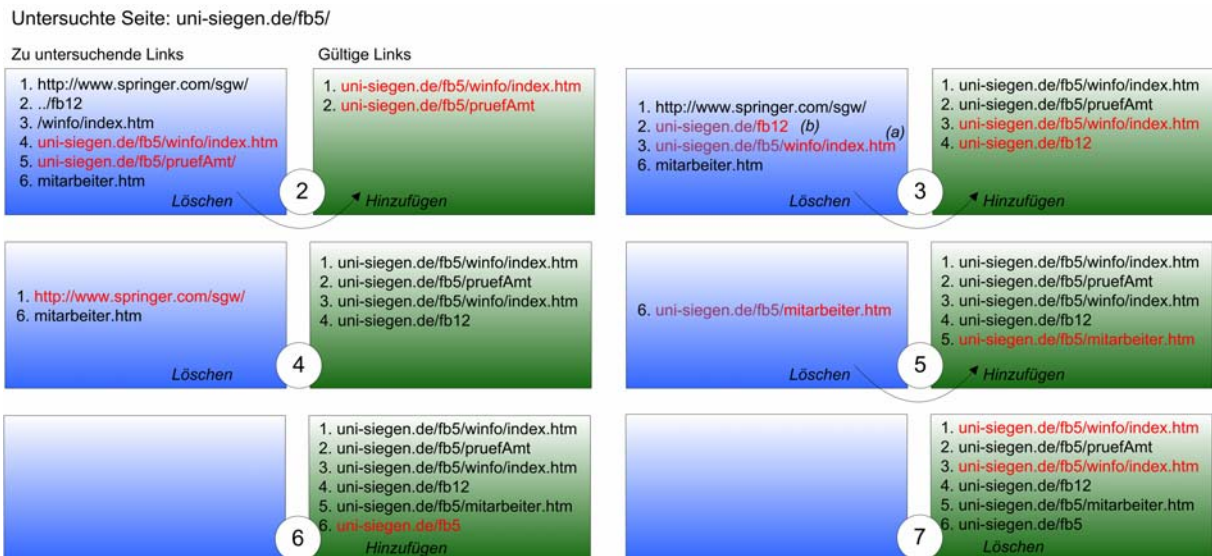


Abbildung 17: Beispiel des Algorithmus zum Filtern interner Links

Die Liste der „internen Links“ enthält alle internen Links auf der aktuell beobachteten Seite. Durch einen rekursiven Aufruf kann so der gesamte Inhalt einer Webseite durchsucht werden. Im aktuellen Status durchsucht der Business Finder drei Ebenen einer Webseite. Nachdem alle Links gefunden sind, wird sukzessiv eine Seite nach der anderen nach Schlagworten auf die oben beschriebene Weise durchsucht.

Zu den Datenbeschaffungsmodulen des EF-Systems besteht ein weiterer Unterschied. Der BF besitzt nicht die Möglichkeit, hochgeladene Dokumente beliebig lange auf dem Server vorzuhalten. Während der Client des EF-Systems jederzeit auf die Dokumente der lokalen Festplatte zugreifen kann, werden die hochgeladenen Dokumente beim BF aus Platz- und Datenschutzgründen nach einem Scanvorgang vom Server gelöscht. Damit die Informationen aus den vergangenen Dokumentenscans nicht verloren gehen, wird das bestehende Stichwortprofil mit den jeweils neu erzeugten Daten gemischt. Das EF-System dagegen erzeugt bei jedem Scannen ein komplett neues Profil.

Zum Mischen der Profile wird für jeden Term geprüft, ob er bereits im eigentlichen Schlüsselwortprofil vorhanden ist. Für diesen Fall wird aus der aktuell ermittelten Dokumentenhäufigkeit und dem alten Wert aus dem Schlüsselwortprofil das arithmetische Mittel berechnet und gespeichert. Existiert der Term noch nicht im Schlüsselwortprofil, so wird die ermittelte Häufigkeit durch zwei dividiert und in das Profil eingefügt. Dadurch werden die Werte um den Faktor zwei normalisiert und so ein ausgeglicheneres Matching ermöglicht.

Neben der Datenerzeugung existieren Module der Datenverarbeitung. Informationen, die mit Hilfe der beschriebenen Verfahren erzeugt werden, müssen entsprechend verwertet werden. Der BF beinhaltet dafür zwei Matchingmodule. Die Module `BFDBMasterData` und `BFDBProfileStorage` übernehmen das Generieren der Matchingergebnisse über die Stammdaten bzw. das Schlüsselwortprofil. Das `BFDBProfileStorage`-Modul ist eine Abwandlung des im EF-System verwendeten `DBProfileStorage`-Moduls. Das Modul ist an die Gegebenheiten der Lÿz-Datenbank angepasst. Der Ablauf beim Matching über Schlüsselworte entspricht dem Vorgehen des EF-Systems (vgl. Kapitel 5.2.2).

Das Modul `BFDBMasterData` bewertet die Trefferquote nach dem Auffinden des Suchwortes innerhalb eines Datensatzes der Stammdaten. Eine Übereinstimmung des Suchbegriffs mit dem Unternehmensnamen ist höher zu bewerten als z.B. die Übereinstimmung mit dem Straßennamen. Eine Abstufung der Werte vom Unternehmensnamen mit der höchsten Priorität über Rubrik, Beschreibung, Internetadresse, Geschäftsführer, Ansprechpartner bis zur Adresse des Unternehmens sorgt für eine „gerechte“ Bewertung der Suchanfrage.

Beide Module berechnen unabhängig voneinander die Übereinstimmungen der Daten mit dem Suchbegriff. Das Modul `ExpertSearcher` des EF-Systems stellt die Ergebnisse zusammen. Hier wird der modulare Aufbau des EF-Systems deutlich. Das `ExpertSearcher`-Modul übernimmt die Verwaltung der einzelnen Matchingmodule. So kann durch einfaches Hinzufügen neuer Module das EF-System und somit auch der BF angepasst werden. Die Konfiguration der zu verwendenden Module wird bereits bei der Anfrage durch PHP an die Webservices mitgeteilt. PHP findet die zu verwendenden Module in einer frei konfigurierbaren XML-Datei, um eine möglichst flexible Gestaltung fortzusetzen. Diese XML-Datei beinhaltet nicht nur die Konfigurationsmöglichkeiten der zu verwendenden Module, sondern darüber hinaus Einstellungen, die das Mailkonto, den Datenbankpfad, die Webseitenanalyse, gültige Dateierweiterungen und die Sicherheitsmechanismen betreffen.

## 6 Evaluation

Der BF wurde im März 2007 anlässlich eines Lÿz-Media-Breakfasts erstmals der Öffentlichkeit vorgestellt (Anhang B). Bei der zeitnahen Integration der Software in der Kommunalen Datenzentrale des Kreises Siegen-Wittgenstein (KDZ), bei der die Seiten der WiFö verwaltet werden, traten technische und organisatorische Probleme auf. Deshalb wurde der BF zunächst auf einem Testserver der Universität Siegen installiert. Ein optimaler Betrieb war mangels ausreichender Kapazitäten des Testservers nicht möglich. Beim Vorstellen des Systems waren verlängerte Wartezeiten bei Suchanfragen hinderlich.

Der BF ist während der Evaluation auf zwei unterschiedliche Arten betrachtet und bewertet worden. In einer zweiten Interviewphase mit vier Unternehmen wurden Themen der regionalen Kooperation und der BF als unterstützendes System untersucht. Die vier Interviewpartner, die an der zweiten Interviewphase teilnahmen, unterscheiden sich nach IT-Anbietern und IT-Nachfragern (vgl. Kapitel 3.2). Zu den IT-Nachfragern gehören zwei in der Region angesiedelte, international tätige Großunternehmen. Die beiden IT-Anbieter waren bereits an der ersten Interviewphase beteiligt. Anschließend wurde der BF mit Suchmaschinen und Branchenbüchern der Region verglichen und bewertet.

Beide Nachfrager-Unternehmen berichten unabhängig voneinander, dass in bestimmten Bereichen und Größenordnungen darauf geachtet wird, regionale Anbieter zu berücksichtigen. Begründet wird dies aus logistischen Gründen und dem Erfordernis, flexibel reagieren zu können. Dies gilt insbesondere bei Sonderanfertigungen und bei der Inanspruchnahme unterstützender Dienstleistungen in IT-Bereichen. Hierbei wird im Allgemeinen auf bestehende soziale Beziehungen zu regionalen Geschäftspartnern zurückgegriffen.

Der Business Finder kann nach Aussagen der Unternehmen in verschiedenen Abteilungen unterschiedlich genutzt werden. Nach den Interviews kann z.B. seitens der Abteilung Einkauf ein größeres Interesse am System bestehen als in der IT-Abteilung oder der Abteilung Qualitätsmanagement. Auch im strategischen Management bieten sich Möglichkeiten für einen sinnvollen Einsatz. Die IT-Anbieter-Firmen wurden zu den Themen Regionalität, und Kooperation in der ersten Interviewphase befragt.

Das System wurde allen vier Unternehmen vorgestellt. Dabei sollten Funktionalitäten und Nutzbarkeit aber auch eventuelle Verbesserungsmöglichkeiten untersucht werden. Hierbei ist nach den Anbietern, die im System eingetragen sind und den Nachfragern, die ausschließlich die Suche nutzen, zu unterscheiden. Die IT-Anbieter haben die Funktionalitäten der Profilbearbeitung analysiert.

Während der zweiten Interviewphase war durchweg eine positive Resonanz festzustellen. Hervorgehoben wurde dabei insbesondere die Benutzerfreundlichkeit und die automatische Datenerfassung aus den Webseiten. Bedenken wurden hinsichtlich der Möglichkeit der Manipulation der Suchergebnisse geäußert. Dies erklärt sich beim Erreichen einer kritischen Zahl an Nutzern darin, dass sich die einzelnen Akteure um eine bestmögliche Platzierung bemühen werden. Beliebige Textdokumente hochzuladen, fördert diese Manipulationsmöglichkeit. Nutzer könnten Dokumente zusammenstellen in denen ausgewählte Begriffe häufig auftauchen. Auf den Server hochgeladen und gescannt könnte auf diese Weise das Schlüsselwortprofil bewusst manipuliert werden, um zu einem bestimmten Begriff besser positioniert zu sein.

Den Teilnehmern wurden anonymisierte Suchergebnisse verschiedener Suchsysteme vorgelegt, die sie bewerten sollten. Dazu wurde die bestehende Lÿz-Datenbank, Google und der BF mit der gleichen Suchanfrage<sup>16</sup> beauftragt und die Ergebnisse in einer neutralen Darstellung den Beteiligten vorgelegt. Google wurde vor dem Hintergrund als Vertreter sämtlicher Suchmaschinen ausgewählt und die Lÿz-Datenbank als Vertreter für herkömmliche Branchenbücher. Weiter kann durch die Wahl der Lÿz-Datenbank ein direkter Vergleich zum Suchverfahren des BF hergestellt werden.

Alle Wertungen fielen zu Gunsten entweder der Lÿz-Datenbank oder dem BF aus. Google enthielt natürlich auch Angaben, die nicht zu Unternehmensseiten<sup>17</sup> führten. Bemerkenswert ist, dass die Angabe des Rankings des BF keinen entscheidenden Unterschied zu der alphabetischen Sortierung der Lÿz-Datenbank bei den Befragten darstellt. Bekannte Unternehmen in den Suchergebnissen beeinflussten die Bewertung durch die Interviewpartner.

---

<sup>16</sup> Die Suchanfrage der Suchmaschine Google wurde durch die ergänzende Eingabe des Wortes „Siegen“ auf die Region beschränkt.

<sup>17</sup> Die Unternehmer wurden vorher darauf hingewiesen, dass es sich dabei um eine gezielte Unternehmenssuche handelt.

Kritisiert wurden an dieser Stelle die Ergebnisse der Bewertung. Speziell ein Unternehmer (L) hat auf die diesbezüglich unklare Zusammenstellung der Prozentangaben hingewiesen. Im anschließenden Praxistest der Business-Finder-Suche konnte dann dieses Problem ausführlich besprochen und klargestellt werden. Kritik war also nur an der Darstellung von Informationen gegeben, nicht jedoch am eigentlichen Konzept des BF.

Beiden IT-Anbietern wurde jeweils eine Liste mit der Platzierung ihres Unternehmens bei verschiedenen Suchmaschinen vorgelegt. Diese wurden zuvor mit einer Auswahl von Suchbegriffen ermittelt. (vgl. z.B. Tabelle 3).

Suchbegriff	Business Finder	Google	Lyz-Datenbank
Internet	24	-	69
Datenbanken	3	179	4
PHP	1	-	3
Server	21	-	-
Programmierung	6	139	14
Java	1	364	2

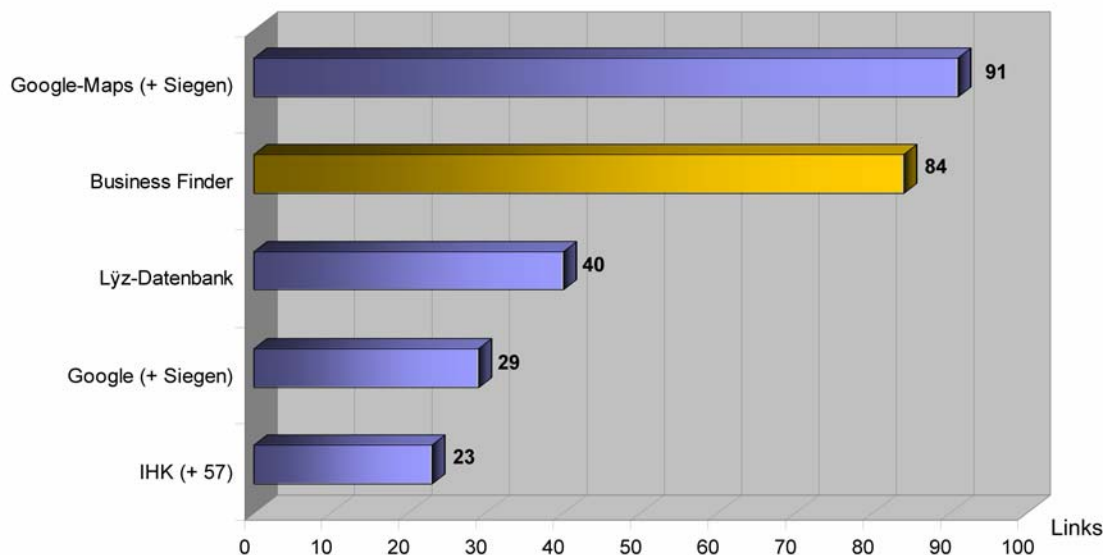
**Tabelle 3: Platzierung bei verschiedenen Suchanfragen von Unternehmen L**

Nach Auffassung der Interviewpartner war die Auswahl der Suchbegriffe zu allgemein, um ihre Unternehmen zu beschreiben. Daraus entstand dann die Bereitschaft, den BF mit den von den Unternehmen ausgewählten Suchbegriffen zu testen. Dieser Praxistest wurde positiv aufgenommen. Ein Unternehmen wurde stets auf eine der oberen fünf Platzierungen gelistet. Der Unternehmer L betonte, dass er der Lyz-Datenbank bisher keine Beachtung geschenkt hat, die Suchergebnisse aber für einen Einsatz des BF sprechen würden. Das Unternehmen wird gefunden, obwohl kein Aufwand in das System gesteckt wurde: „Weil man sieht, ich wurde sogar gefunden wenn ich meine Daten nicht gepflegt habe“ (Anhang D, Unternehmer L, 14.07.2007, S. 20).

Hier wird sehr deutlich, welchen Vorteil die Schlüsselwortergänzung zu den eigentlichen Stammdaten hervorbringt. Die Stärke dieser Option zeigt sich bei der Suche nach ausgefallenen und seltenen Begriffen.

Darüber hinaus zeigt sich, dass die Unternehmen im BF auf höheren Positionen gelistet werden als bei Google. Dabei kommt der Vorteil des BF zum Tragen, dass er auf die Region Siegen bezogen ist und eine gezielte regionale Auswahl getroffen wird.

Um einen Vergleich der Suchergebnisse zu ermöglichen, sind die nachstehenden Suchmaschinen bzw. Branchenbüchern mit zehn verschiedenen Suchbegriffen<sup>18</sup> überprüft worden. Als repräsentative regionale Branchenbücher wurden das IHK-Branchenbuch und die Lüz-Datenbank getestet. Daneben wird Google und Google-Maps<sup>19</sup> mit dem zusätzlichen Schlüsselwort „Siegen“ in den Vergleich einbezogen. Jeweils die ersten 20 Suchergebnisse sind repräsentativ. Abbildung 18 zeigt die Anzahl der durch die verschiedenen Suchmaschinen „korrekt“ gefunden Unternehmen. „Korrekt“ bedeutet in diesem Zusammenhang, dass das gefundene Unternehmen zu dem Suchbegriff passt und in der Region ansässig ist. Google-Maps geht z.B. bei einer niedrigen Trefferquote dazu über, die Region weiter zu fassen und auch Unternehmen entfernterer Regionen zu listen.



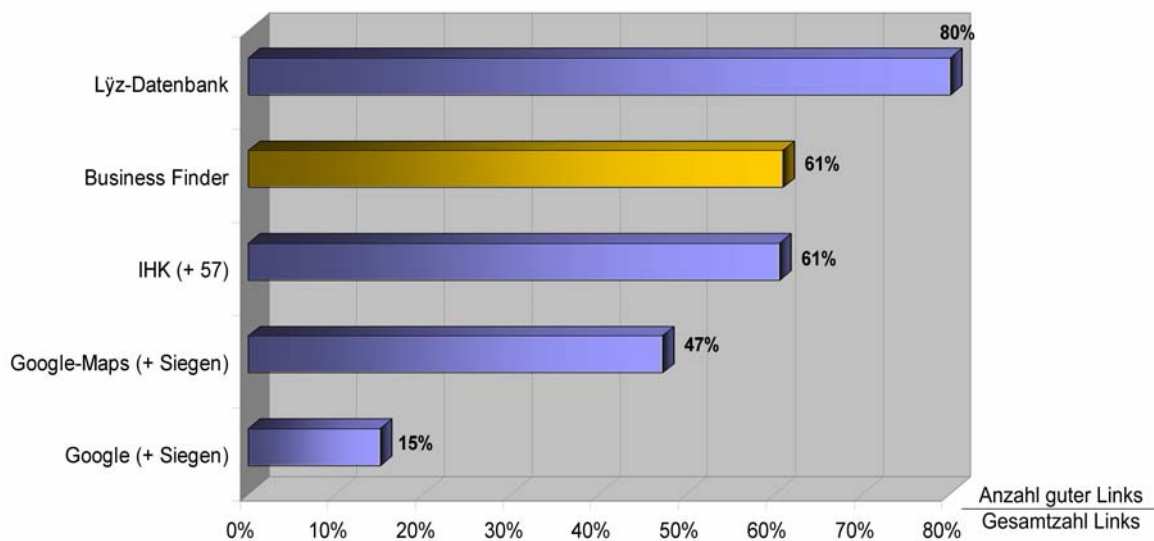
**Abbildung 18: Gültige Links zu Unternehmensseiten der Region**

Der maximalen gültigen Trefferzahl steht das Risiko der Informationsüberflutung gegenüber. Im Verhältnis zu der Gesamtzahl der jeweils betrachteten Links ist Google-Maps von der anfänglichen ersten Position auf die fünfte Position abgefallen. Nur 47 Prozent der aufgeführten Links sind entsprechende, lokale Unternehmen (Abbildung 19). Dass die Lüz-Datenbank in diesem Vergleich das beste Ergebnis erzielt, lässt sich an der Gesamtzahl der

<sup>18</sup> Die zehn Suchbegriffe lauten: Archivierung, CAD, CMOS, Cobol, Dokumentenmanagement, Druckerei, ERP-Systeme, Offset Print, Werbung und Workflow Management.

<sup>19</sup> Google-Maps ist ein neuer Dienst von Google der Kartentechnologie und Informationen zu Unternehmen wie Standort und Kontaktinformationen liefert.

gefundenen Links erklären. Vor der IHK lieferte sie die zweitwenigsten Treffer insgesamt und bestätigt damit das Problem von Branchenbüchern, eine zu geringe Auswahl an Unternehmen zu bieten. Die Lÿz-Datenbank bewertet ausschließlich die durch die Unternehmen eingefügten Stammdaten. Durch die niedrige Trefferzahl bietet die Datenbank allerdings trotz der hohen Korrektheit der Ergebnisse nur einen geringen Nutzen bei der Suche nach speziellen Unternehmen.



**Abbildung 19: Verhältnis von gültigen Links zu der Gesamtzahl der Links**

Der Business Finder findet sich dagegen in beiden Auswertungen auf Platz zwei wieder. Er bildet somit ein ausgewogenes Verhältnis zwischen der Menge an Informationen und deren Gültigkeit. Dem Nutzer wird eine ausreichend große Zahl an potentiellen Unternehmen geliefert, die dann in der Lage sind, die gewünschte Leistung zu erbringen.

## 7 Zusammenfassung und Diskussion

Mit der Verbreitung der Internettechnologie Ende des 20. Jahrhunderts hat sich die Meinung gebildet, dass „herkömmliche“ Standortfaktoren wie Nähe zu Kunden und Lieferanten nebensächlich sind. Ein „Zerfall von Ballungsräumen“ wurde postuliert. Heute gewinnen regionale Netzwerke zunehmend an Bedeutung und es zeigt sich, dass die Digitalisierung dazu beiträgt. Die hohe Konzentration auf digitale Prozessabwicklungen und Kommunikationswege führt dazu, dass Unterschiede in konkurrierenden Unternehmen überwiegend in den nicht-digitalen Bereichen zu suchen sind. Diese nicht-digitalen Bereiche sind die Sektoren, die Clusterunternehmen von Inselunternehmen unterscheiden. Hauptaspekte sind das Soziale Kapital und der persönliche Kontakt im Cluster. „Mund-zu-Mund-Propaganda“ und „Flurgespräche“ führen zu Erfahrungsaustausch und einer schnellen Informationsverbreitung. Dabei wird neben den Informationen über Bedarf an neuen Produkten und Innovationsschritten konkurrierender Unternehmen auch Wissen konstruiert. Dies führt zu kompetenteren Arbeitskräften. Die Unternehmen können spezialisierte Fachkräfte einstellen, die zur Steigerung von Innovation führen.

Aufgrund dessen werden politische Maßnahmen ergriffen, die zur Förderung von Clustern beitragen. Diese beschränken sich vorwiegend auf Veranstaltungen, in denen der persönliche Kontakt zu weiteren Unternehmen geknüpft werden soll. Auch in der Region Siegen-Wittgenstein ist bis zur Einführung des Business Finders die Förderung des regionalen Netzwerkes durch Veranstaltungen wie z.B. das Lÿz-Media-Breakfast erfolgt. Abgesehen von Clusterwebseiten, in denen ein Cluster mit seinen teilnehmenden Unternehmen dargestellt wird, findet sich keine aktive technische Unterstützung, die speziell auf die Förderung eines agglomerierten Wirtschaftsraums abzielt.

Der Business Finder und die durchgeführte Studie sind ein erster Ansatz, mit Hilfe eines Recommender-Systems die Unternehmen der Region Siegen-Wittgenstein besser zu vernetzen und die Clusterbildung zu unterstützen. Die Aussagen der befragten Unternehmen zeigen, dass in der Region Siegen noch kein vollständiges IT-Cluster vorhanden ist. Die ansässigen Unternehmen suchen ihre Abnehmer außerhalb der Region, da sie davon ausgehen, dass keine regionalen Geschäftspartner existieren. Andererseits wissen lokale Konsumenten nicht einzuschätzen, welches Anbieterpotential die Region bietet und suchen



IT-Dienstleister im nationalen und/oder internationalen Raum. Ein Grund hierfür ist in der mangelnden Transparenz der Region zu sehen.

Der Business Finder bietet mit seiner Zusammensetzung ein überschaubares Softwaresystem, das zur Steigerung dieser Transparenz beiträgt. Die kritische Einstellung der Interviewpartner gegenüber dem BF änderte sich erkennbar nach dessen Einführung. Dies zeigte sich u.a. in der Bereitschaft zum Testen des Systems, nachdem man sich anfänglich skeptisch geäußert hatte.

Die Verstärkung einer regionalen Vernetzung wird alleine durch den Business Finder dennoch nicht zu erwarten sein. Er kann vielmehr als zusätzliche Unterstützung neben anderen Förderungsmaßnahmen gesehen werden. Eine konkrete Aussage über den Erfolg ist noch nicht möglich. Eine längerfristige Beobachtung des Ballungsraums und der Zugriffszahlen auf das System können Aufschluss darüber geben.

Der Erfolg von allgemeinen Kontaktbörsen für Personen wie „Xing“ demonstriert darüber hinaus die Bedeutung von Vernetzung im Allgemeinen. Die Leistungsfähigkeit des Systems hat sich in einem Fall gegen Ende der hier vorgestellten Studie bereits bestätigt. Das (mit der Universität Siegen kooperierende) Fraunhofer-Institut für Angewandte Informationstechnik (FIT) suchte einen kompetenten Industriepartner aus dem Bereich „Radio Frequency Identification“ (RFID), um ein darauf basierendes Navigationssystem für Feuerwehrleute zu entwickeln. In diesem Zusammenhang wurde zunächst ein nationales Unternehmen gesucht, das in dem Bereich der RFID spezialisiert ist. Nachdem eine Internetrecherche (mit Google) kein relevantes Ergebnis lieferte, wurde auch der Business Finder in die Suche eingebunden. Der Business Finder lieferte bei einer Suche zum Begriff „RFID“ ein Unternehmen aus der Region. Es zeigte sich, dass das Unternehmen tatsächlich deutliche Kompetenzen in diesem Bereich aufwies. Das Unternehmen wurde nicht aufgrund der Stammdaten durch den Business Finder gefunden, sondern durch die indexierte Webseite (auf die auch Google zugreift). Das Unternehmen wurde schließlich erfolgreich für die Kooperation gewonnen.

Nicht unbeachtet lassen darf man die Sicherheitsbedenken der Unternehmen, betriebsinterne Informationen über das Internet zu versenden. Dies bezieht sich in erster Linie auf Dokumente, die zur Verfügung gestellt werden können. Weitere Bedenken der Unternehmer betreffen das gezielte Hochladen spezifischer Dokumente, die das Stichwortprofil

beeinflussen (vgl. Kapitel 6). Damit verliert das System an Aussagekraft. Dem entgegen zu wirken bedarf es eines hohen technischen Aufwands, dessen Verwirklichung nicht Gegenstand dieser Arbeit ist.

Mit ca. 380 Unternehmen bietet die Lÿz-Datenbank zwar eine solide Grundlage; für den Erfolg ist aber eine kritische Masse an Unternehmen zu überschreiten. Dies kann nur mit der Öffnung der Datenbank für weitere Branchen erzielt werden, um so weiteren Unternehmen die evtl. an der Wertschöpfungskette beteiligt sind Zugang zum System zu gewähren.

Ein Ansatz für die Erweiterung des Business Finders im Kontext regionalpolitischer Maßnahmen der WiFö des Kreises Siegen-Wittgenstein wäre das Einbinden der bereits bestehenden Datenbank kunststoffverarbeitender Unternehmen der Region. Die lokal stark vertretene Metallbranche sollte berücksichtigt werden, um so die vorhandenen lokalen Kompetenzen zu stärken [vgl. Port1998a, S.89]. Das Einbeziehen dieser Branchen kann zum Erfolg des Systems und damit auch zur Stärkung der Region beitragen.

Neben der Möglichkeit, die lokale Wirtschaft Siegens zu fördern, ist der Einsatz des Business Finders in weiteren regionalen Clustern denkbar. In einem bestehenden Cluster ist die Bereitschaft zur Kooperation und die Regionalität gegeben. Es ist möglich, dass der Business Finder die Transparenz weiter steigert und somit die Innovation fördert.

Neben den theoretischen Ansätzen sind auch noch technische Verbesserungen möglich. Die Sicherheitsbedenken der Nutzer können gemindert werden, wenn das System beim Einloggen zu dem jeweiligen Browser eine sichere Verbindung mittels SSL<sup>20</sup> aufbaut.

Ein weiterer Ansatz wäre nach den Ergebnissen der Umfrage das Berücksichtigen von Referenzen im System. Während der Studie bestätigten die Unternehmer, dass Referenzen eine hilfreiche Information über Unternehmen darstellen. Es bestehen allerdings Bedenken über die Validität der Angaben. Der Business Finder könnte das bei Personenkontaktbörsen verwendete Prinzip verwenden, nur auf eingetragene Personen bzw. die im Business Finder eingetragenen Unternehmen zu verweisen. Das referenzierte Unternehmen müsste wie bei gängigen Personenkontaktbörsen eine Bestätigung ggf. mit einer Bewertung zurücksenden,

---

<sup>20</sup> SSL ist ein Verschlüsselungsprotokoll für Datenübertragungen im Internet.

bevor die Referenz Gültigkeit erhält. In Anlehnung an das PageRank-Verfahren könnte damit eine weitere Matchingmöglichkeit geschaffen werden.

Auch im Bereich der Suchverfahren sind weitere Optimierungsmöglichkeiten gegeben. Bei dem System der WiFö ist neben der Stichwortsuche auch die Selektion nach IT-Branche und Ort vorhanden. Eine Kombination der Suchen, d.h. die Auswahl einer Branche und eines Ortes mit der zusätzlichen Einschränkung der Stichwortsuche verbessert die Bedienbarkeit des Systems. Außerdem kann die Eingabe um logische Operatoren wie „und“, „oder“ und „nicht“ erweitert werden. Um dies zu ermöglichen, müssten aus der Suchzeichenkette alle Stoppworte entfernt werden. Als Nebeneffekt wird dadurch die Suche beschleunigt. Das System sucht nicht unnötig nach unbedeutenden Worten.

Mit einer verbesserten semantischen Analyse könnte im Sinne einer Clustersuchmaschine versucht werden, Unternehmen anhand des Schlüsselwortprofils in Branchen einzuordnen. Zusätzlich könnte sie zum Auffinden synonyme Begriffe genutzt werden. Den Unternehmern würde dadurch die Eingabe und die Auswahl erleichtert.

Zusammenfassend ist festzuhalten, dass der Business Finder nicht dazu beitragen wird, neue Cluster zu bilden. Sein Einsatz kann in einer späteren Phase durch die Steigerung der Transparenz dazu beitragen, Kooperationen, Konkurrenzdenken und die Vernetzung von Unternehmen zu fördern. Regionale Netzwerke und Cluster werden auch in Zukunft die Innovationszentren sein. Förderungsmaßnahmen jeder Art sind weiter zu erforschen. IT kann ein einfaches und effizientes Mittel für Clusterinitiativen sein.

## Literaturverzeichnis

[Amen1992]

*Amendola, Giovanni; Guerrieri Paolo; Padoan Pier Carlo*: International Patterns of Technological Accumulation and Trade. In: *Journal of International and Comparative Economics* (1992) 1, S. 173-197.

[Bask1999]

*Baskerville, Richard L.*: Tutorial – Investigating information systems with action research. In: *Communications of the AIS* (1999) 2. URL: <http://cais.aisnet.org/articles/2-19/default.asp?View=pdf&x=56&y=11> (letzter Zugriff: 18.07.2007).

[Batt2005]

*Battenfeld, Jochen*: Benutzer-Matching auf Basis automatischer Textanalyse – Ein Ansatz zur Ähnlichkeitsbestimmung von Benutzern durch Dokumentenanalyse für das ExpertFinder Framework. Diplomarbeit, Universität Siegen, 2005.

[Beck2004]

*Becks, Andreas; Reichling, Tim; Wulf, Volker*: Expertise Finding – Approaches to Foster Social Capital. In: *Huysman, Marleen; Wulf, Volker (Hrsg.): Social Capital and Information Technology*. The MIT Press, Cambridge 2004, S. 333- 254.

[Beka2002]

*Bekavec, Bernard; Lenich, Andreas; Meyer Renate*: Methoden und Verfahren von Suchdiensten im Internet (2002). URL: [http://www.inf-wiss.uni-konstanz.de/suche/tutorial/such\\_tutorial\\_advanced.html](http://www.inf-wiss.uni-konstanz.de/suche/tutorial/such_tutorial_advanced.html) (letzter Zugriff: 23.07.2007).

[Berg01]

*Bergman, Michael K.*: The Deep Web: surfacing Hidden Value. In: *The Journal of Electronic Publishing* (2001) 1, URL: <http://www.press.umich.edu/jep/07-01/bergman.html> (letzter Zugriff: 23.07.2007).

[Bieg2003]

*Bieger, Thomas; Scherer Roland*: Clustering und integratives Standortmanagement: von einem theoretischen Konzept zu konkreten Handlungsstrategien. In: *Bieger, Thomas; Scherer Roland (Hrsg.): Clustering – das Zauberwort der Wirtschaftsförderung*. Paul Haupt, Bern 2003, S. 9-26.

[Boeh1988]

*Boehm Barry W.*: A Spiral Model of Software Development and Enhancement. In: *Computer* (1988) 5, S. 61-72.

[Bour1997]

*Bourdieu, Pierre*: Ökonomisches Kapital – Kulturelles Kapital – Soziales Kapital. In: *Bourdieu, Pierre (Hrsg.): Verborgene Mechanismen der Macht*. VSA-Verlag 1997, Hamburg, S. 49-80.

[Brin1998]

*Brin, Sergey Michailowitsch; Page, Lawrence: The Anatomy of a Large-Scale Hypertextual Web Search Engine (1998). URL: <http://infolab.stanford.edu/~backrub/google.html> (letzter Zugriff: 25.07.2007).*

[Coms2007]

*ComScore Incorporated: comScore Releases April U.S. Search Engine Rankings (2007). URL: <http://www.comscore.com/press/release.asp?press=1447> (letzter Zugriff: 26.07.2007)*

[Deer1990]

*Deerwester, Scott; Dumais, Susan T.; Harshman, Richard: Indexing by Latent Semantic Analysis. In: Journal of the American Society of Information Science (1990) 6, S. 391-407*

[Donh2006]

*Donhauser, Stefan: Aktivierung von Wachstumspotenzialen durch Netzwerke - Clusterbildung in Baden-Württemberg. In: Statistisches Monatsheft Baden-Württemberg (2006) 4, S. 18-23.*

[Down1998]

*Downes, Larry (1998): Beyond Porter – Drawing on research for a fascinating new book – the author shows how three shock waves are overwhelming the famous Five Forces. URL: <http://www.contextmag.com/setFrameRedirect.asp?src=/archives/199712/technosynthesis.asp> (letzter Zugriff: 22.06.2007).*

[Ellw2002]

*Ellwein, Christian: Suche im Internet für Industrie und Wissenschaft. Oldenbourg Industrieverlag GmbH, München 2002.*

[Erlh2005]

*Erlhofer, Sebastian: Suchmaschinenoptimierung – Für Webentwickler – Grundlagen, Funktionsweisen und Ranking-Optimierung. Galileo Press GmbH, Bonn 2005.*

[Eule2006]

*Euler, Mark (Hrsg.): Soziales Kapital – Ein Brückenschlag zwischen Individuum und Gesellschaft. BIS-Verlag, Oldenburg 2006.*

[Euro2002]

*European Comission: Regional Clusters in Europe - Observatory of European SMEs 2002, No. 3. Luxemburg 2002.*

[Floy1989]

*Floyd, Christiane: STEPS to Software Development with Users. In: Ghezzi, Carlo; McDermid, John A.: ESEC' 89: 2<sup>nd</sup> European Software Engineering Conference, University of Warwick, Coventry, UK, Proceedings - Lecture Notes in Computer Science. Springer Verlag, Berlin 1989, S. 48-64.*

[Fran1997]

*Frants, Valery I.; Shapiro, Jacob; Voiskunskii, Vladimir G.*: Automated Information Retrieval – Theory and Methods. Academic Press, San Diego 1997.

[Fran1998]

*Frank, Ulrich; et al.*: Aktionsforschung in der WI – Einsatzpotentiale und –probleme. In: *Schütte, R. (Hrsg.); Siedentopf, J.; Zelewski, S.*: Wirtschaftsinformatik und Wirtschaftstheorie, Grundposition und Theoriekerne, Arbeitsberichte des Institutes für Produktion und Industrielles Informationsmanagement (1998) 4, S. 71-90.

[Gamm2004]

*Gamma, Erich; et. al.*: Entwurfsmuster – Elemente wieder verwendbarer objektorientierter Software. Addison Wesley Verlag, München 2004.

[Glög2003]

*Glögger, Michael*: Suchmaschinen im Internet. Springer Verlag, Berlin 2003.

[Gunz1986]

*Gunz, Josef (Hrsg.)*: Handlungsforschung – Vom Wandel der distanzierten zur engagierten Sozialforschung. Wilhelm Braumüller, Wien 1986.

[Höls2002]

*Hölscher, Christoph*: Die Rolle des Wissens im Internet – Gezielt suchen und kompetent auswählen. Clett-Cotta, Stuttgart 2002.

[Kahl1996]

*Kahler, Helge; Rohde, Markus; Wulf Volker*: Einführung von Groupware im Prozess integrierter Organisations- und Technikentwicklung. In: *Brödner, Peter; Paul, Hansjürgen; Hamburg, Ilena (Hrsg.)*: Kooperative Konstruktion und Entwicklung – die Zukunft von CAD-Systemen. Rainer Hampp Verlag, München 1996, S. 95-115.

[Kahl1997]

*Kahler, Helge; Stiernerling, Oliver*: Einführung eines CSCW-Systems zur Kooperationsunterstützung im Komponentenentwicklungsprozess. In: *Industrie Management* (1997) 6, S. 43-46.

[Kete2006]

*Ketels, Christian H. M.*: Cluster – Kurzfristiger Hype oder nachhaltige Strategie?. In: Grenzüberschreitender Workshop Cluster-Management 2006. URL: <http://www.standort.zh.ch/internet/vd/awa/standort/de/meinzuerich/aktuell.html> (letzter Zugriff: 18.07.2007).

[Kock2004]

*Kock, Ned*: The three threats of action research – a discussion of methodological antidotes in the context of an information systems study. In: *Decision Support Systems* (2004) 2, S. 265-286.

[Kolk2001]

*Kolko Jed David*: Silicon Mountains, Silicon Molehills – Geographic Concentration and Convergence of Internet Industries in the US. UNU World Institute for Development Economics Research, Helsinki 2001.

[Krät2001]

*Krätke, Stefan; Scheuplein, Christoph*: Produktionscluster in Ostdeutschland: Methoden der Identifizierung und Analyse -Kurzfassung- Eine Studie im Auftrag der Otto Brenner Stiftung. Berlin 2001.

[Lawr1999]

*Lawrence, Steve; Giles, C. Lee*: Accessibility of Information on the Web. In: *Nature* (1999) 8, S. 107-109.

[Leun2003]

*Leuninger, Stefan; Held, Holger*: Kommunale Wirtschaftsförderung im Umbruch – Kundenmanagement in Bestandsentwicklung und im Standortmarketing praxisorientiert umsetzen. In: *Standort – Zeitschrift für Angewandte Geographie* (2003) 4, S. 161-166.

[Lewi1953]

*Lewin, Kurt*: Die Lösung sozialer Konflikte – Ausgewählte Abhandlung über Gruppendynamik. 3. Aufl., Christian-Verlag, Bad-Nauheim 1968.

[Maig2003]

*Maignan, Carole* (2003): ICT, clusters and regional cohesion: A summary of theoretical and empirical research. URL: <http://www.feem.it/NR/rdonlyres/EFEEBC72-A6FD-40B7-AC22-C3CFAE26BBFE/990/5805.pdf> (letzter Zugriff; 22.07.2007).

[Male1998]

*Malek, Miroslaw* (1998): Silicon Valley – Das Geheimnis des Erfolgs. URL: <http://www.berlinews.de/wista/archiv/95.shtml> (letzter Zugriff: 22.06.2007).

[Müll2006]

*Müller, Claudia*: Eine Ethnografie der Vernetzung von Akteuren aus Wissenschaft und Industrie in der IT-Branche der Region Bonn/ Rhein-Sieg – Perspektiven und Potentiale des Sozialkapitalbegriffs von Pierre Bourdieu. Magisterarbeit, Rheinische Friedrich-Wilhelms- Universität, Bonn 2006.

[Page1998]

*Lawrence Page et al.*: The PageRank Citation Ranking: Bringing Order to the Web. Manuscript in progress (1998). URL: <http://infolab.stanford.edu/~backrub/pageranksub.ps> (letzter Zugriff: 25.07.2007).

[Pola1946]

*Polanyi, Michael*: Science, Faith and Society – A searching examination of the meaning and nature of scientific inquiry – with a new introduction by the author. Taschenbuchausgabe, The University of Chicago Press, London 1964.

[Port1998a]

*Porter, Michael Eugene*: Clusters and the New Economics of Competition. In: *Edward, Elgar (Hrsg.)*: Systems of Innovation: Growth, Competitiveness and Employment. MPG Books Ltd., Cornwall 2000, S. 309-322.

[Port1998b]

*Porter, Michael Eugene (Hrsg.)*: Clusters and Competition – New Agendas for Companies, Governments, and Institutions. In: On Competition (Harvard Business Review). McGraw-Hill Professional, Boston 1998, S. 197-288.

[Port2000]

*Porter, Michael Eugene*: Locations, Clusters, and Company Strategy. In: *Clark, G. L.; Feldman, M. P.; Gertler, M. S. (Hrsg.)*: Oxford Handbook of Economic Geography. Oxford University Press, New York 2000, S. 253-274.

[Prog2004]

*Prognos AG*: Prognos-Zukunftsatlas 2004 – Deutschlands Region im Zukunftswettbewerb – Ergebnisse der Sonderauswertung für die Landkreise Siegen-Wittgenstein und Olpe und die Region Siegen (2004). URL: [http://www.ihk-siegen.de/fileadmin/Geschaeftsfelder/Standortpolitik/Broschueren\\_und\\_Studien/Prognos\\_Zukunftsatlas\\_2004.pdf](http://www.ihk-siegen.de/fileadmin/Geschaeftsfelder/Standortpolitik/Broschueren_und_Studien/Prognos_Zukunftsatlas_2004.pdf) (letzter Zugriff: 05.08.2007).

[Putn2001]

*Putnam, Robert David (Hrsg.)*: Gesellschaft und Gemeinsinn – Sozialkapital im internationalen Vergleich. Verlag Bertelsmann Stiftung, Gütersloh 2001.

[Rapo1970]

*Rapoport, Robert Norman*: Three dilemmas in action research – With special reference to the Tavistock Experience. In: Human Relations (1970) 6, S. 499-513.

[Reck2001]

*Recklies, Dagmar* (2001): Beyond Porter – Strategie in der Internet-Ökonomie. URL: <http://www.themanagement.de/pdf/Beyond%20Porter.PDF> (letzter Zugriff: 22.06.2007).

[Reic2004]

*Reichling, Tim et al.*: Kontakthanbahnung in Lernplattformen. In: *Keil-Slawik, Reinhard; Selke, Harald; Szwillus, Gerd (Hrsg.)*: Mensch & Computer 2004 – Allgegenwärtige Interaktion. Oldenbourg Verlag, München 2004, S. 179-188.

[Reic2005a]

*Reichling, Tim; Schubert, Kai; Wulf, Volker*: Matching Human Actors based on their Texts - Design and Evaluation of an Instance of the ExpertFinding Framework. In: Conference on Supporting Group Work - Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work. ACM Press, New York 2005, S. 61-70.

[Reic2005b]

*Reichling, Tim; Veith Michael*: Expertise Sharing in a Heterogeneous Organizational Environment. In: Proceedings of the Ninth European Conference on Computer Supported Cooperative Work. Springer, Dordrecht 2005, S. 325-345.



[Reic2006a]

*Reichling, Tim; Wirth, Matthias*: Yellow Pages in einem Verband – Expertin finden und zusammenführen im VDMA. In: *Journal Arbeit* (2006) 1, S. 22-23.

[Reic2006b]

*Reichling, Tim*: ExpertFinding – Auffinden von Experten in großen Organisationen. In: *Heinecke, Andreas M.; Paul, Hansjürgen (Hrsg.): Mensch & Computer 2006 – Mensch und Computer im StrukturWandel*. Oldenbourg Verlag, München 2006.

[Resn1997]

*Resnick Paul; Varian, Hal R.*: Recommender systems. In: *Communications of the ACM* (1997) 40, S. 56-58.

[Royc1970]

*Royce, Winston W.*: Managing the development of large software systems – Concepts and techniques. In: *WESCON Technical Papers*. WESCON, Los Angeles 1970, S. 1-9. Nachdruck in: *Proceedings of the Ninth International Conference on Software Engineering*, 1987, S. 328-338.

[Sand1998]

*Sander-Beuermann, Wolfgang; Schomberg, Mario*: Internet Information Retrieval – The Further Development of Meta-Searchengine Technology. In: *Internet Society: Proceedings of the Internet Summit*. Genf 1998. URL: [http://www.isoc.org/inet98/proceedings/1c/1c\\_2.htm](http://www.isoc.org/inet98/proceedings/1c/1c_2.htm) (letzter Zugriff: 23.07.2007).

[Salt1983]

*Salton, Gerard; McGill Michael J.*: Information Retrieval - Grundlegendes für Informationswissenschaftler. McGraw-Hill, Hamburg 1987.

[Saut2004]

*Sautter, Björn*: Regionale Cluster – Konzept, Analyse und Strategien zur Wirtschaftsförderung. In: *Standort - Zeitschrift für Angewandte Geographie* (2004) 2, S. 66-72.

[Schi2003]

*Schiele, Holger (Hrsg.)*: Der Standort-Faktor. Wie Unternehmen durch regionale Cluster ihre Produktivität und Innovationskraft steigern. 1. Aufl., Wiley-VCH Verlag GmbH & Co KGaA, Weinheim 2003.

[Sieg2007a]

*Stadt Siegen, Medien- und Öffentlichkeitsarbeit*: Webseite der Stadt Siegen – Stadtinfo (2007). URL: <http://www.siegen.de/standard/page.sys/183.htm> (letzter Zugriff: 02.08.2007).

[Sieg2007b]

*Stadt Siegen, Medien- und Öffentlichkeitsarbeit*: Mit Siegen gewinnen – Universitätsstadt Siegen- Die Wirtschaftsmetropole Südwestfalens. URL: <http://www.siegen.de/doc.cfm?seite=27&urlDoc=pfaddownloads/27downloads/broschuere.pdf> (letzter Zugriff: 05.08.2007).

- [Sieg2007c]  
*Siegerlandkurier*: Mit den Pfunden wuchern – Franz-J. Mockenhaupt (IHK) über unsere Region und wie sie gesehen wird. In: *Siegerlandkurier* (Lokalzeitung vom 11.03.2007), S. 3.
- [Sölv2003]  
*Sölvell, Örjan; Lindqvist, Göran; Ketels, Christian (Hrsg.): The Cluster Initiative Greenbook*. Bromma tryck AB, Stockholm 2003.
- [Susm1978]  
*Susman, Gerald I.; Evered Roger D.: An Assessment of the Scientific Merits of Action Research*. In: *Administrative Science Quarterly* (1978) 23, S. 582-603.
- [Süsk1994]  
*Süskind, Patrick: Das Parfum – Die Geschichte eines Mörders*. Taschenbuchausgabe, Diagonales Verlag AG, Zürich 1994.
- [Vena2001]  
*Venables, Anthony J.: Geography and International Inequalities – the impact of new technologies*. In: *Journal of Industry, Competition and Trade* (2001) 2, S. 135-159.
- [Werb2005]  
*Werben und Verkaufen: Studie – Einkäufer suchen gerne im Internet* (2005). URL: <http://www.wuv.de/news/archiv/2005/01/39452/index.html> (letzter Zugriff 31.07.2007).
- [Wiki2007]  
*Wikipedia Die freie Enzyklopädie: Suchmaschine* (2007). URL: <http://de.wikipedia.org/wiki/Suchmaschine> (letzter Zugriff 27.07.2007).
- [Will1998]  
*Willke, Helmut: Systematisches Wissensmanagement*. Lucius & Lucius Verlagsgesellschaft mbH, Stuttgart 1998.
- [Wolf2002]  
*Wolfe, David Alexander: Social Capital and Cluster Development in Learning Regions*. In: *Holbrook, J. Adam; Wolfe, David Alexander (Hrsg.): Knowledge, Clusters and Learning Regions*. Queen's University, Kingston 2002. URL: [http://www.utoronto.ca/progris/pdf\\_files/Wolfe\\_SocialCapital.pdf](http://www.utoronto.ca/progris/pdf_files/Wolfe_SocialCapital.pdf) (letzter Zugriff: 09.07.2007)
- [Wulf1995]  
*Wulf, Volker; Rohde Markus: Towards an Integrated Organization and Technology Development*. In: *Proceedings of the Symposium on Designing Interactive Systems*. ACM-Press, New York 1995, S. 55-64.

## Anhang

### A Interviewleitfäden

#### Bedarfsanalyse

*Anbieter*

##### **Einleitende Fragen:**

1. Welche Größe hat Ihr Unternehmen?
2. Wie lautet Ihre Unternehmensphilosophie?
3. Was produzieren Sie? Welche Dienste bieten Sie an?
4. Welche Aufgaben haben Sie im Unternehmen?

##### **Kundenkontakt:**

5. Haben Sie direkten Kontakt zu Kunden?
6. (Wie) werben Sie um Kunden? Wie werden Kunden auf Sie aufmerksam?
7. Sind Ihre Kundenverhältnisse eher langfristig oder eher kurzfristig?
8. Wie viele Mitbewerber in Ihrer Branche haben Sie? Wie stark sind diese nach Ihrer Einschätzung?
9. Gibt es Kunden, von denen Sie besonders „abhängig“ sind?
10. Wie viel Prozent Ihrer Aufträge erhalten Sie von Unternehmen aus der Region?
11. Sind es oft die gleichen Unternehmen von denen Sie Aufträge bekommen oder wechseln Ihre Auftraggeber?
12. Wie viele dieser Unternehmen kommen aus einer ähnlichen Branche wie Ihrer?
13. Wissen Sie wie diese Unternehmen auf Sie aufmerksam geworden sind?
14. Wie wichtig ist es für Sie, diese Unternehmen zu kennen?
15. Wie wichtig ist es, *neue* Abnehmer zu finden, denen Sie Ihre Dienste anbieten können? Z.B. für gezielte Marketing-Maßnahmen.
16. Fahren Sie gezielt Marketing-Strategien an lokale Unternehmen?
17. Ist es für Sie von Bedeutung andere Unternehmen Ihrer Branche zu kennen und ausfindig zu machen?
18. Wie oft kommt es vor, dass Sie Anfragen von Unternehmen erhalten, mit denen Sie noch nie zusammen gearbeitet haben?

##### **Problemlösung:**

19. Woran arbeiten Sie im Augenblick?  
Sind in die Lösung weitere Unternehmen einbezogen?
20. Lösen Sie im Allgemeinen Ihre eigenen IT-Probleme selbst oder vergeben Sie viele Aufträge an andere Unternehmen?
21. Legen Sie Wert auf eine engere Zusammenarbeit der Unternehmen dieser Region?
22. Wenn Sie sich zwischen zwei Aufträgen entscheiden müssten, könnte dann die lokale Nähe ein Entscheidungskriterium sein?

### **Fragen zur Suche nach Unternehmen:**

23. Ist Ihnen ein konkreter Fall bekannt, bei dem Sie von einem Ihnen unbekanntem Unternehmen angesprochen wurden?  
Hat das Unternehmen gesagt wie es auf Sie aufmerksam wurde?
24. Welche Zeit nimmt die Suche nach möglichen Kunden in Anspruch?
25. Gibt es Informationen innerhalb Ihres Unternehmens, die für andere Unternehmen wichtig sein könnten (um gerade Ihnen einen Auftrag zu erteilen), die Sie aber ungern veröffentlichen wollen? (Produktdaten, Strategiepläne...)?  
Welche dieser Daten würden Sie einem Computersystem anvertrauen?
26. Ist Ihnen die Lÿz-Media-Datenbank ein Begriff?
27. Werden Ihre Daten dort auf dem Laufenden gehalten?
28. Was gefällt Ihnen (nicht) an der Lÿz-Media-Datenbank?
29. Sind Sie mit der Darstellungsart in der Lÿz-Media-Datenbank zufrieden? Was möchten Sie lieber von Ihrem Unternehmen darstellen?
30. Welche Kontaktdaten möchten Sie anderen Unternehmen zur Verfügung stellen?
31. Mit dem ExpertFinding Framework ist es möglich, auch weitere Daten als ausschließlich selbst angelegte Profile auszuwerten. Somit ist eine genauere Bestimmung für ein passendes Unternehmen möglich. Welche Voraussetzungen müssten gegeben sein, dass Sie ein solches System nutzen?
32. Würden Sie lieber „vertrauliche“ Dokumente auf einen Webserver hochladen, so dass Sie dort nach Schlüsselwörtern durchsucht werden können oder wäre es Ihnen lieber, eine Software lokal zu installieren, die dann nur noch die erstellten Keywordprofile auf dem Server ablegt?
33. Anfangs werden in dem ExpertFinding die Internetpräsenzen der Unternehmen eingepflegt, anhand derer die besagten Keywordprofile erstellt werden.  
Ist es für Sie von Interesse, die URLs selbst zu verwalten, ggf. weitere hinzuzufügen oder Subdomains auszuschließen?
34. Welche Erwartungen stellen Sie an ein solches Recherche-Tool?
35. In wie weit ist Ihr Unternehmen dazu bereit, Daten in einem solchen System zu pflegen?
36. Könnten Sie sich Dokumente vorstellen, die für ein solches System interessant wären?
37. Denken Sie, Sie werden diese Daten eher pflegen als in der Lÿz-Media-Datenbank?
38. In welchem Rahmen macht ein solches Netzwerk Sinn? Oder macht es keinen Sinn?

### *Nachfrager*

### **Einleitende Fragen:**

1. Welche Größe hat Ihr Unternehmen?
2. Wie lautet Ihre Unternehmensphilosophie?
3. Was produzieren Sie? Welche Dienste bieten Sie an?

### **Verbindung zu anderen Unternehmen:**

4. Woran arbeiten Sie im Augenblick?  
Sind in die Lösung Fremdunternehmen einbezogen?
5. Lösen Sie im Allgemeinen Ihre IT-Probleme selbst oder vergeben Sie viele Aufträge an andere Unternehmen?
6. Welche Leistungen beziehen Sie von externen Dienstleistern?
7. Wie viel Prozent davon sind ungefähr Unternehmen aus dieser Region?

8. Sind es feste Unternehmen, mit denen Sie zusammen arbeiten oder wechseln die Unternehmen von Auftrag zu Auftrag?
9. Wie wichtig ist es für Sie, diese Unternehmen zu kennen?
10. Welche Informationen sind wichtig für Sie? Nach welchen Kriterien wählen Sie Anbieter aus?
11. Wie oft/genau vergleichen Sie die Angebote verschiedener Unternehmen?
12. Wie wichtig ist es, *neue* Unternehmen zu finden, deren Dienste Sie in Anspruch nehmen können? Wie oft suchen Sie nach neuen Unternehmen?
13. Welche Informationsmaterialien ziehen Sie bei der Suche nach Anbietern in Betracht?
14. Welche Informationen wären sonst noch wünschenswert?
15. Ist es für Sie von Bedeutung, andere Unternehmen Ihrer Branche zu kennen und ausfindig zu machen?
16. Legen Sie Wert auf eine engere Zusammenarbeit der Unternehmen dieser Region?
17. Würden Sie ein Unternehmen aus der Region einem anderen vorziehen?
18. Wie oft kommt es vor, dass Sie nach anderen Unternehmen suchen?

### **Vorgehen bei der Suche:**

19. Wie sind Sie auf Ihre Partnerunternehmen aufmerksam geworden, wie haben Sie diese gefunden?
20. Ist Ihnen ein konkreter Fall bekannt, bei dem Sie ein anderes Unternehmen gesucht haben?
  - a. Wie sind Sie dabei vorgegangen?
  - b. Welche Zeit hat es in Anspruch genommen?
21. Welche Zeit nimmt die Suche durchschnittlich in Anspruch? Wie oft kommen Sie vor?
22. Welche Informationen über andere Unternehmen erscheinen Ihnen wichtig?
23. Haben Sie bei Ihrer Suche die Lÿz-Media-Datenbank einbezogen?
24. Was gefällt Ihnen (nicht) an der Lÿz-Media-Datenbank?
25. Welche Erwartungen stellen Sie an ein Recherche-Tool um andere Unternehmen ausfindig zu machen?
26. Sollten Personen (bspw. aus der Forschung) auch in die Datenbank integriert werden?
27. Was erwarten Sie von den verbesserten Möglichkeiten des Expert-Finding?
28. Denken Sie, dass Sie dieses System eher bei Ihren Recherchen nutzen, als die Lÿz-Media-Datenbank?
29. Was denken Sie, in welchem Rahmen ein solches Netzwerk Sinn machen würde? Oder macht es überhaupt keinen Sinn?
  - Größe der Region,
  - Beteiligte Unternehmen/Personen usw.

## Evaluationsstudie

### Anbieter

1. Sie sind beispielsweise bei verschiedenen Branchenbüchern eingetragen. Wie aktuell halten Sie Ihre Daten dort?
2. Schauen Sie oft dort rein?  
Wie wichtig ist es für Sie dort eingetragen zu sein? Nutzen Sie es oft?
3. Welchem dieser Suchergebnisse (Google, BF, Lÿz) würden Sie den Vorrang geben?
4. Was vermuten Sie, an welcher Stelle Sie bei folgenden Suchbegriffen auftauchen?
5. Was sind für Sie typische Suchbegriffe?
6. *BusinessFinder vorstellen...*
7. Was gefällt Ihnen am Konzept des Business Finders?
8. Könnte man den Business Finder an bestimmten Stellen verbessern?
  - a. Was würden Sie sich von einem solchen System noch wünschen?
  - b. Vermissen Sie etwas?
9. Was halten Sie von der Möglichkeit, Dokumente hochzuladen?
10. Wie finden Sie die Idee, dass Ihr Newsletter nach Stichworten durchsucht wird?
11. Welchem dieser Suchergebnisse (Google, BF, Lÿz) würden Sie den Vorrang geben?
12. Könnten Sie sich vorstellen, aufgrund eines solchen Tools mehr mit Anbietern/Nachfragern aus der Region zusammen zu arbeiten?
13. Haben Sie die Lÿz-Media-Datenbank bisher in Ihre Recherchen einbezogen? Werden Sie das in Zukunft eher tun?
14. Wie wichtig ist Ihnen ein direktes Feedback durch das System (z.B. Beim Hochladen der Dokumente)?
15. Würden Sie das System eher nutzen, wenn es national oder international wäre?

### Nachfrager

1. Woher beziehen Sie Ihre IT-Komponenten?  
Aus der Region? Kann man das an etwas festmachen?
2. Wie sind Sie an die Anbieter gekommen?
3. Arbeiten Sie mit Unternehmen aus der Region zusammen?
4. In welchen Branchen sind diese Unternehmen angesiedelt?
5. Wie gehen Sie vor, wenn Sie neue Komponenten „kaufen/integrieren“ wollen?  
Können Sie das an einem Beispiel verdeutlichen?
6. Gibt es konkrete Beispiele, wo Sie Kontakt zu anderen Unternehmen über das Internet geknüpft haben? Wie werden allgemein Kontakte geknüpft?
7. *Business Finder vorstellen...*
8. Was gefällt Ihnen am Konzept des Business Finder?
9. Wie könnte man den Business Finder an bestimmten Stellen verbessern?
10. Was sind für Sie typische Suchbegriffe?
11. Welchem dieser Suchergebnisse (Google, BF, Lÿz) würden Sie den Vorrang geben?
12. Könnten Sie sich vorstellen, aufgrund eines solchen Tools mehr mit Anbietern/Nachfragern aus der Region zusammen zu arbeiten?
13. Haben Sie die Lÿz-Media-Datenbank bisher in Ihre Recherchen einbezogen? Werden Sie das in Zukunft eher tun?
14. Wie wichtig sind solche Partnerschaften für ein Unternehmen Ihrer Größe?

## B Lyz-Media-Breakfast Werbeflyer

### Anmeldung

Melden Sie sich bitte telefonisch unter 0271 333-2407, per E-Mail unter [info@lyz-media.de](mailto:info@lyz-media.de), online unter [www.lyz-media.de](http://www.lyz-media.de) an oder senden Sie uns diese Anmeldung per Fax unter 0271 333-2419 zu.

Hiermit melde ich

Name: \_\_\_\_\_

Vorname: \_\_\_\_\_

Firma: \_\_\_\_\_

Straße: \_\_\_\_\_

PLZ/Ort: \_\_\_\_\_

Tel./Fax: \_\_\_\_\_

e-mail: \_\_\_\_\_

mich verbindlich zur Lyz media Veranstaltung am 08. März 2007 an.

Datum, Unterschrift

Den Kostenbeitrag für das Frühstück in Höhe von 5,00 Euro bitten wir am Veranstaltungstag zu entrichten. Die Veranstaltung findet im Bistro Lyz statt.



### Veranstalter & Kontakt

Kreis Siegen-Wittgenstein  
Wirtschaftsförderung und Strukturentwicklung  
Koblenzer Straße 73, 57072 Siegen  
Klaus Irie  
Telefon: 0271 333-2407  
Telefax: 0271 333-2419  
E-Mail: [a\\_schmitz-kaempfer@siegen-wittgenstein.de](mailto:a_schmitz-kaempfer@siegen-wittgenstein.de)  
Internet: [www.lyz-media.de](http://www.lyz-media.de)

### WIRTSCHAFTSFÖRDERUNG

Informieren ■ beraten ■ vermitteln ■ koordinieren



Gezielte Internetsuche nach Unternehmen mit dem *BusinessFinder*

### Einladung

Donnerstag, 08. März 2007  
8.15 - 9.30 Uhr

Bistro Lyz  
Eingang B  
Medien- und Kulturhaus Lyz  
St.-Johann-Str. 18  
57074 Siegen



#### Anfahrt:



### Vorwort

Um Kooperationen zwischen den Unternehmen in der Region weiter auszubauen und Medienunternehmungen zu stärken, führt die Wirtschaftsförderung in Zusammenarbeit mit der Fachgruppe „Wirtschaftsinformatik und Neue Medien“ der Universität Siegen den *BusinessFinder* zur Suche von Unternehmen im Internet ein.

Ziel dieser Veranstaltung ist es, Ihnen den *BusinessFinder* vorzustellen und zu zeigen, wie er sich gewinnbringend für Ihr Unternehmen einsetzen lässt. Den *BusinessFinder* können Sie sich wie eine Verknüpfung von einem Branchenbuch und einer Suchmaschine vorstellen. Diese Verbindung hat einen entscheidenden Vorteil: Während ein Branchenbuch allein nur die jeweils eingetragenen Daten verwaltet, können Suchmaschinen aktiv Internetpräsenzen, Newsletter und beliebige Textdokumente (z.B. Werbematerialien, Produktbeschreibungen, etc.) von den am Projekt beteiligten Unternehmen nach bestimmten Informationen durchsuchen. Aus den hier gesammelten Daten bildet das System dann ein individuelles, sehr umfassendes und aussagekräftiges Schlüsselwortprofil für jedes einzelne Unternehmen als sinnvolle Ergänzung unserer regionalen Unternehmensdatenbank.

Im Anschluss an den Vortrag haben Sie die Gelegenheit, in angenehmer Atmosphäre, verbunden mit einem Frühstück, mit interessanten Gesprächspartnern zu diskutieren und neue Kontakte zu knüpfen.

Reinhard Kämpfer  
Referent für Wirtschaftsförderung und Strukturentwicklung des Kreises Siegen-Wittgenstein

■ informieren ■ beraten ■ vermitteln ■ koordinieren

### Programm

**8:15 Uhr**  
**Eröffnung und Einführung in das Thema „Der BusinessFinder - ein neues, komfortables Modul für die regionale Medien-Unternehmensdatenbank“**  
Klaus Irie, Wirtschaftsförderung Kreis Siegen-Wittgenstein

**8:20 - 9:00 Uhr**  
**Benjamin Moos stellt mit Tim Reichling ein neues Programm-Modul vor, die unsere regionale Marktübersicht für Sie gewinnbringend ergänzen soll.**  
Das Programm sucht in von Ihnen bereitgestellten Dokumenten nach häufig vorkommenden, markanten Worten, die dann in ein gesondertes Verzeichnis eingetragen werden. Diese Begriffe sorgen später dafür, dass die Unternehmen, die bereits in der regionalen Medien-Unternehmensdatenbank gespeichert sind, bei der Stichwortsuche eines Internetbesuchers zusätzlich auch unter den ausgewählten Begriffen gefunden werden.

**9:00 - 9:30 Uhr**  
**Gemeinsames Frühstück mit offener Gesprächsrunde**  
**ca. 9:30 Uhr**  
**Ende der Veranstaltung**

■ informieren ■ beraten ■ vermitteln ■ koordinieren

### Referent



**Benjamin Moos**  
Universität Siegen

Diplomand im Fach Wirtschaftsinformatik „Neue Medien“ der Universität Siegen, Prof. Dr. Wulf

Benjamin Moos stellt das Ergebnis seiner Diplomarbeit zum Thema „Unterstützung regionaler Vernetzung durch das webbasierte System *BusinessFinder*“ vor.



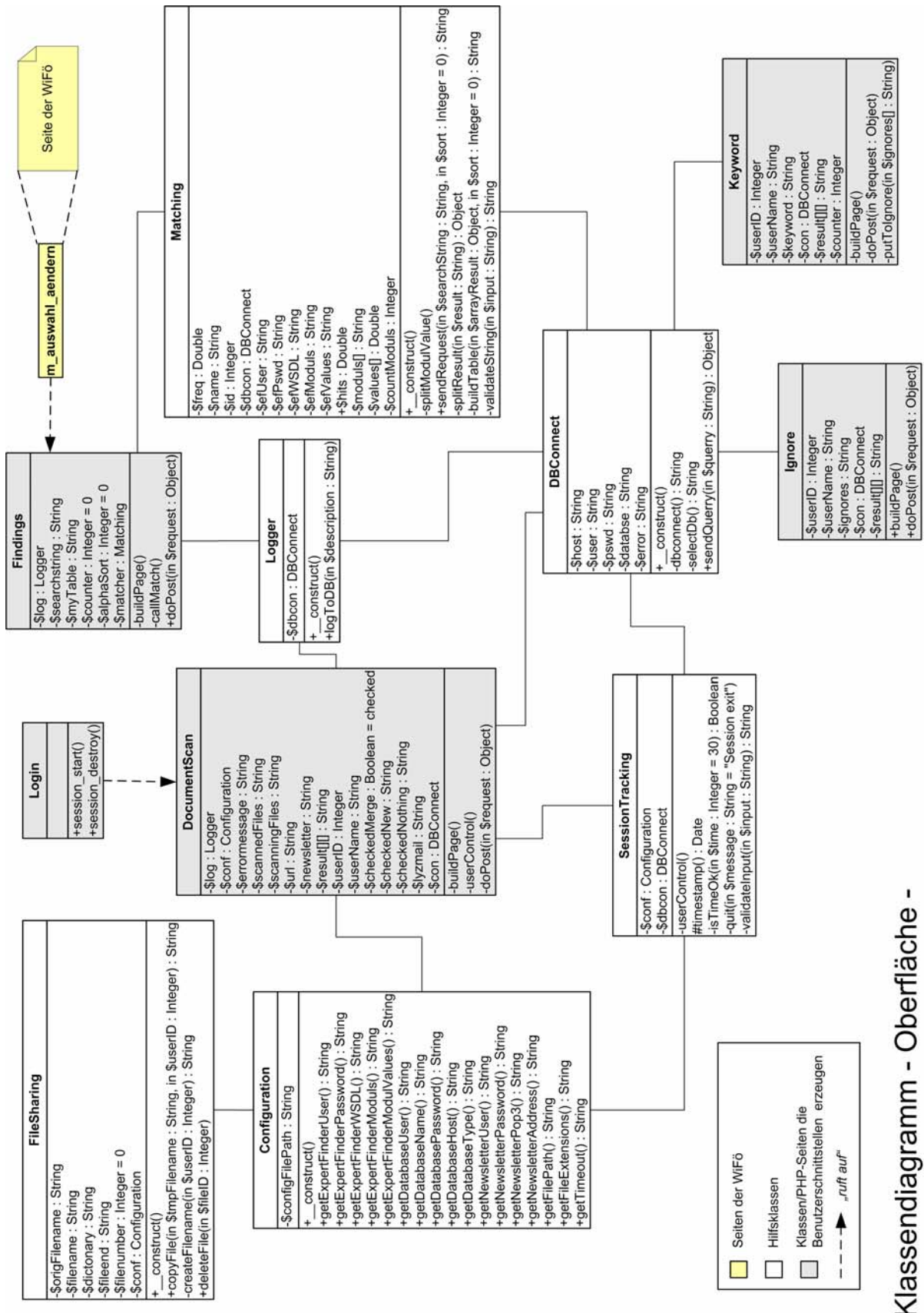
**Tim Reichling, Dipl.-Inf.**  
Universität Siegen

Wissenschaftlicher Mitarbeiter der Universität Siegen im Bereich Wirtschaftsinformatik „Neue Medien“

Tim Reichling erforscht im Forschungsprojekt WIN (Wissensmanagement in Informationsnetzwerken) die Möglichkeiten und Potenziale des Einsatzes von Wissensmanagementsystemen und -strategien in einem bedeutenden europäischen Industrieverband.

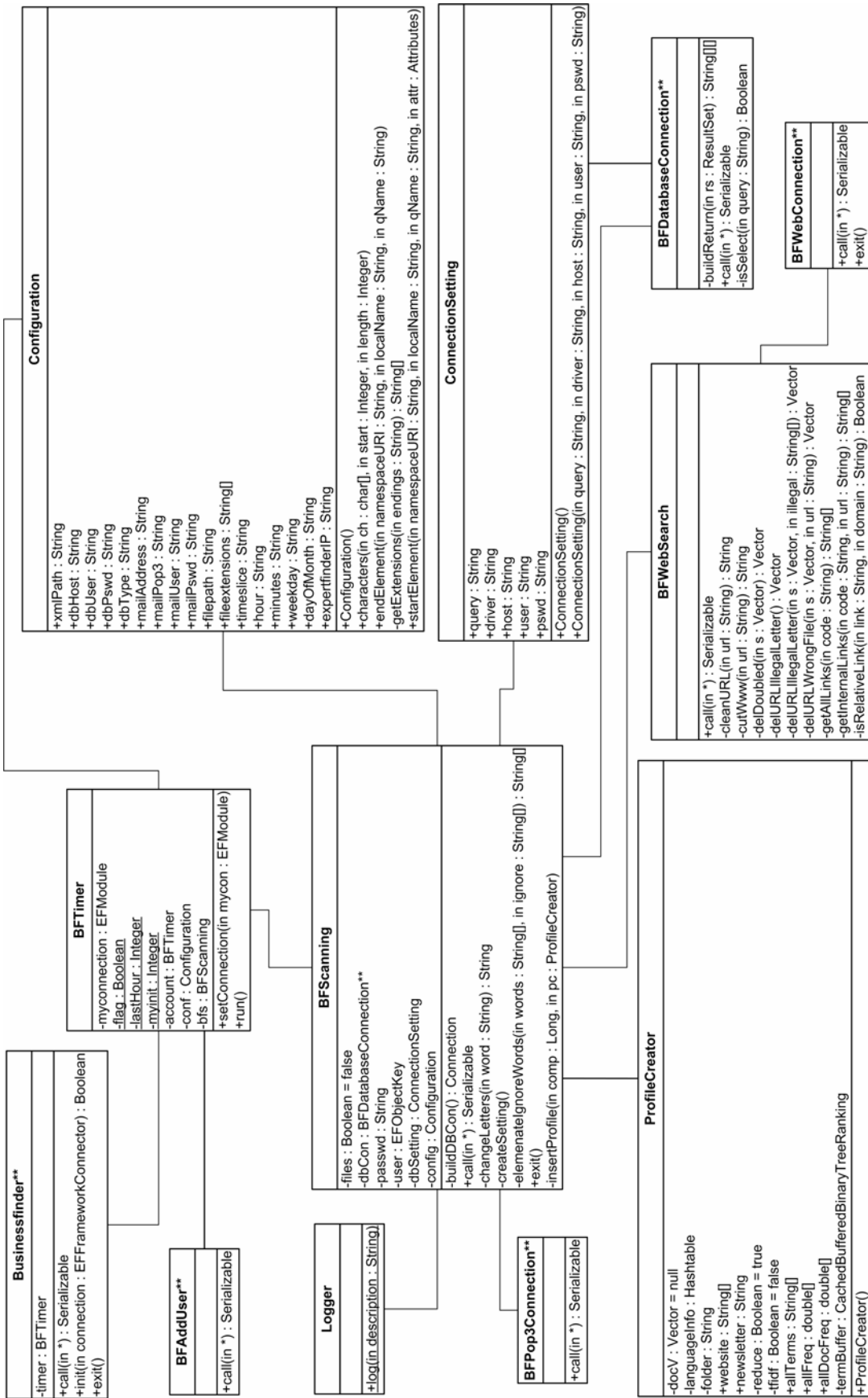
■ informieren ■ beraten ■ vermitteln ■ koordinieren

# C Diagramme



Klassendiagramm - Oberfläche -

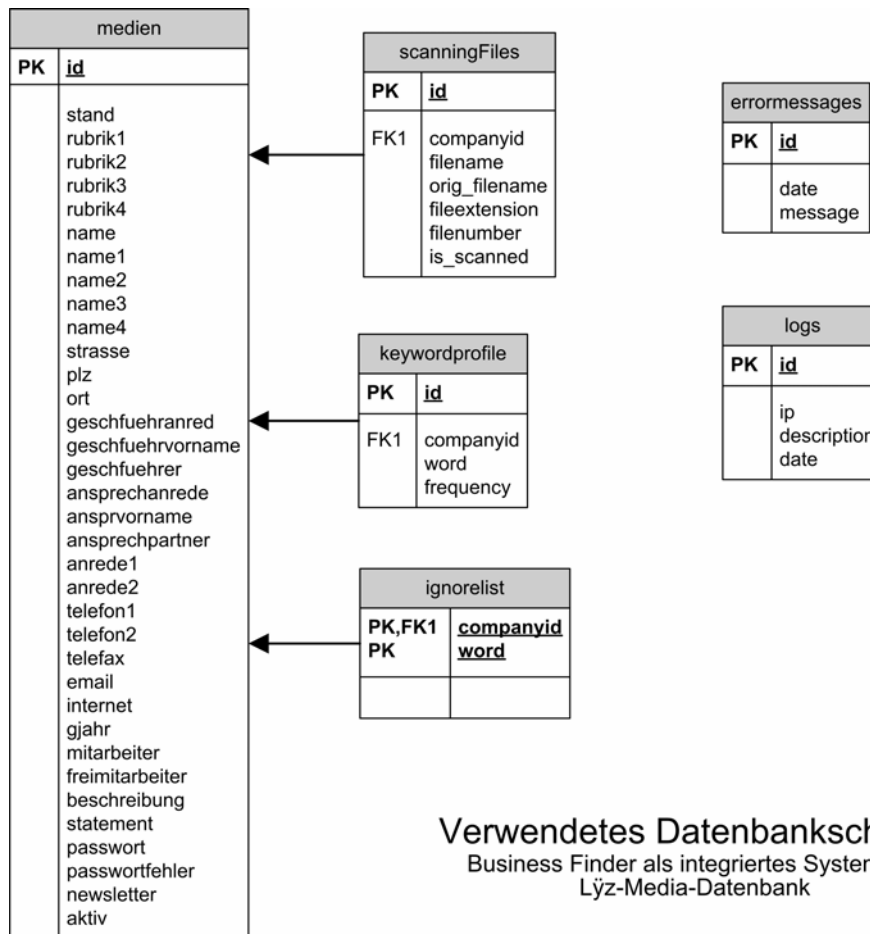
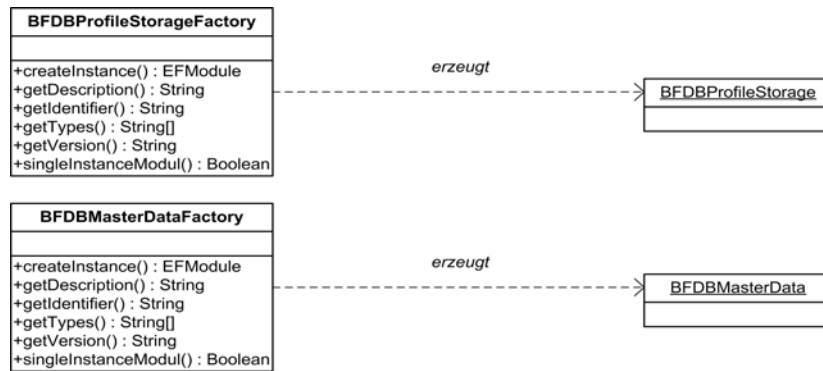


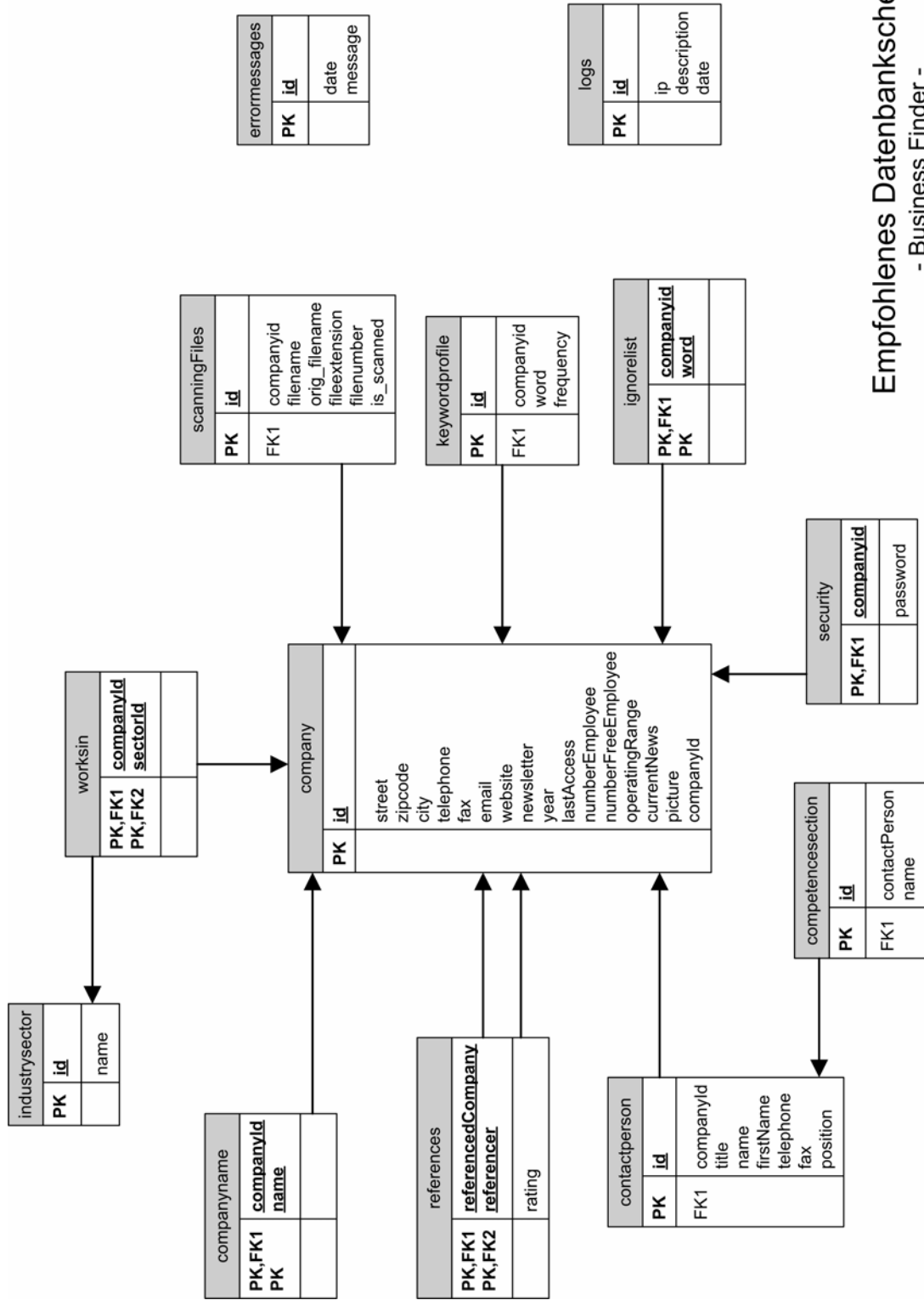


**Klassendiagramm - Business Finder Suchmaschine -**

- \* +call(in attribute : String, in user : EFOBJECTKey, in passwd : String, in param : Serializable) : Serializable
- \*\* Die Klassen besitzen eine Factory-Klasse. Der Aufbau und die Anbindung ist gleich dem der Matching-Module (folgende Seite).

**Klassendiagramm – Business Finder Matching-Module –**





Empfohlenes Datenbankschema  
- Business Finder -

## D CD-Rom

### Inhalt:

- Business Finder
  - Installationsanleitung
  - Quellcode
    - BF-Module für das EF-System
    - PHP-Oberfläche
  - Dokumentation
    - Java-Doc der BF-Module
    - Oberflächenklassendiagramm
    - Datenbankschema
- Beigefügte Dokumente
  - Die Diplomarbeit
  - Lÿz-Media-Breakfast-Werbeflyer
  - Foliensatz des Vortrags beim Lÿz-Media-Breakfast
  - Anonymisierte Interviewtranskriptionen
    - Bedarfsanalyse
    - Evaluation