

Google

Information und digitale Intelligenz

Google-Suche

Auf gut Glück!

Information und digitale Intelligenz

Masterthesis zur Erlangung des akademischen Grades
„Master of Arts in Arts and Design“

Verfasser

Florian Jindra, BA

Vorgelegt am FH-Studiengang MultiMediaArt, Fachhochschule Salzburg



Begutachtet durch:

Dr. Felix Kramer (Inhaltlicher Gutachter 1)

Dipl. Designer Matthias Edler-Golla (Inhaltlicher Gutachter 2)

Salzburg, 22. August 2012

Eidesstattliche Erklärung

Hiermit versichere ich, Florian Jindra, geboren am 13. Mai 1987 in Oberndorf bei Salzburg, dass ich die Grundsätze wissenschaftlichen Arbeitens nach bestem Wissen und Gewissen eingehalten habe und die vorliegende Masterthesis von mir selbstständig verfasst wurde. Zur Erstellung wurden von mir keine anderen als die angegebenen Quellen und Hilfsmittel verwendet.

Ich versichere, dass ich die Masterthesis weder im In- noch Ausland bisher in irgendeiner Form als Prüfungsarbeit vorgelegt habe und dass diese Arbeit mit der den BegutachterInnen vorgelegten Arbeit übereinstimmt.

Salzburg, am 22. August 2012

1010627017

Florian Jindra, BA

Matrikelnummer

Kurzfassung der Arbeit

<i>Verfasser:</i>	Florian Jindra, BA
<i>Institution:</i>	Fachhochschule Salzburg
<i>Studiengang:</i>	MultiMediaArt
<i>Titel:</i>	Information und digitale Intelligenz
<i>Begutachter 1:</i>	Dr. Felix Kramer
<i>Begutachter 2:</i>	Dipl. Designer Matthias Edler-Golla
<i>Schlagwörter:</i>	1. Information 2. Digitale Intelligenz 3. Google

Die Faszination der Verbindung von Natur und Technik zeigt sich nicht nur in vielen Epochen der Geschichte, die Natur kann darüber hinaus mit ihren Lösungsansätzen für Probleme in vielen Fällen der Forschung und der Wissenschaft entweder als Ausgangspunkt, oder aber als Zielsetzung erkannt werden. Die Verbindung von Natur und Technik, sowie ein damit zusammenhängender gesellschaftlicher Nutzen stellen also auch das Forschungsinteresse der vorliegenden Masterthesis dar.

Die Entwicklung des Computers und in weiterer Folge des Internet hatten die ausgedehnte und umfangreiche Digitalisierung von Wissen zur Folge. Diese immer größer werdende Menge an Wissen resultierte in der Tatsache, dass Wissen einerseits nicht mehr absolut, andererseits nicht mehr zur Gänze erfassbar und überblickbar war. Es fand und findet also ein Wandel der Notwendigkeit von vorhandenem, statischen Wissen zu Information, die bei Bedarf eingeholt und beschafft werden kann, statt. Um dieses Wissen filtern zu können und so ein Informieren, also das Abrufen von Information, zu ermöglichen, sind intelligente Verfahren wie Suchmaschinen notwendig.

Da allerdings die Konzepte aus dem Forschungsfeld der künstlichen Intelligenz nicht für dieses erweiterte Anwendungsgebiet der Suchmaschinen passen, wird der Begriff der *digitalen Intelligenz* als Lösungsansatz eingeführt und anhand des Beispiels Google auf ein mögliches Potenzial hinsichtlich einer Lösung gesellschaftlich relevanter Problemstellungen in der Informationsgesellschaft untersucht.

Abstract

The fascination of combining nature and technology is evident not only in many periods of history, nature's approaches to solving problems can in many cases also be viewed as a role model for research and science. The combination of nature and technology furthermore represents the research interests of this masterthesis, as well as a related set of social benefits for the information society.

The development of the computer and subsequently the Internet led to the digitalization of vast and extensive amounts of knowledge thus becoming available for everyone on the Internet. This growing body of knowledge has resulted in the fact that knowledge is no longer absolute but unseizable in it's entirety for the individual. The need for a change from existing, static knowledge to information that can be obtained when required became obligatory.

To filter this knowledge intelligent methods such as search engines are necessary. However, since the concepts derived from the research field of artificial intelligence do not cover this extended application of search engines, the concept of digital intelligence is being introduced and examined, using the example of Google to review and discuss it's potential in terms of a solution for socially relevant problems within the information society.

Inhaltsverzeichnis

Einleitung.....	8
1 Begriffsklärungen.....	11
1.1 Sammeln und Archivieren von Wissen in Schriftform.....	11
1.2 Das Mundaneum als Vorläufer moderner Dokumentation von Wissen.....	13
1.3 Dokumentation von Wissen im digitalen Zeitalter und ihre Grenzen.....	14
1.4 Die Bedeutungsebenen des Begriffes der Information.....	17
1.5 Informationsverarbeitung und -austausch.....	19
1.6 Digitalisierung von Information.....	21
Résumé.....	22
2 Informationsgesellschaft.....	23
2.1 Die sechs Merkmale der Informationsgesellschaft nach Lehmann.....	26
2.2 Folgen und Probleme der Informationsgesellschaft.....	30
2.3 Theoretische Positionen zur Informationsgesellschaft.....	32
2.3.a Anthony Giddens.....	33
2.3.b Vilem Flusser.....	34
2.3.c Weitere Befunde zur Informationsgesellschaft.....	36
2.4 Praktische Problemlösungsansätze.....	39
2.5 Internet als wesentliche Ausprägung der Informationsgesellschaft.....	43
2.6 Suchmaschinen im Internet als Tool der Informationsgesellschaft.....	46
2.7 Intelligenz und künstliche Intelligenz.....	49
2.8 Von Datenverwaltung zu künstlicher Intelligenz.....	51
2.9 Künstliche neurale Netzwerke „lernen“.....	54
2.9.a Grenzen der Entwicklung.....	56
2.9.b Überprüfung des Vorliegens von Intelligenz.....	58
Résumé.....	61
3 Google.....	63
3.1 Was sammelt Google?.....	66
3.2 Was will Google?.....	67
3.3 Wie lernt Google?.....	71
3.4 Google und das Problem der Menge.....	77
3.5 Google als digitale Intelligenz.....	79
3.6 Google und das Problem der Komplexität.....	87
Erkenntnisse, Résumé und Ausblick.....	90
Quellen.....	95
Bücher.....	95
Beiträge in Sammelbänden.....	96
Zeitschriftenaufsätze.....	97
Onlinequellen.....	97
Audiovisuelle Medien.....	99

Abkürzungsverzeichnis

<i>bzw.</i>	Beziehungsweise
<i>ebd.</i>	ebendort
<i>et al.</i>	et alii
<i>f</i>	und folgende Seite
<i>ff</i>	und folgende Seiten
<i>URL</i>	Uniform Resource Locator
<i>vgl.</i>	vergleiche
<i>z.B.</i>	zum Beispiel
<i>zit. n.</i>	zitiert nach

Einleitung

Die Faszination einer Verbindung von Natur und Technik zeigt sich nicht nur in vielen Epochen der Geschichte, die Natur kann darüber hinaus mit ihren Lösungsansätzen für Probleme in vielen Fällen der Forschung und der Wissenschaft entweder als Ausgangspunkt, oder aber als Zielsetzung erkannt werden. Die Verbindung von Natur und Technik stellt damit auch das Forschungsinteresse und damit den Ausgangspunkt für sowohl die vorliegende Masterthesis, als auch das Werk, das während der konzeptionellen Phase dieser Arbeit entstand, dar.

Das Werk in Form des entwickelten Computerspiels *Of Light & Shadow* setzt sich thematisch mit der Verbindung bzw. Symbiose von Mensch und Maschine auseinander und zeichnet sich besonders im visuellen Bereich durch das Verbinden von technischen und organischen Elementen in einer einzigartigen Umwelt aus. Im Rahmen der vorliegenden Masterthesis soll nun die Fortführung der im Spiel behandelten Thematik durch eine theoretische Auseinandersetzung mit dem Gedanken einer symbiotischen und ergänzenden Herangehensweise an Probleme, die gesellschaftliche Relevanz haben, stattfinden.

Als ein wichtiger Faktor in diesem Zusammenhang kann die Entstehung des Internet gelten, da im Windschatten dieser Entwicklung nicht nur neue multimediale Technologien entstanden, sondern sich auch der Zugang zu Information, Wissen und Kommunikation, sowie der Umgang damit veränderten.

Die durch die Digitalisierung von Information und das Internet weit höheren Geschwindigkeiten von Informationsaustausch und Kommunikation verändern dabei nicht nur das Leben von uns Menschen in vielerlei Hinsicht, im Gegenzug ist damit auch ein gesellschaftspolitisch relevanter Wandel verbunden, der nicht nur die Einzelne und den Einzelnen, sondern die Gesellschaft als Ganzes vor anspruchsvolle Herausforderungen stellt.

Die Wissenschaft hat in den unterschiedlichsten Disziplinen ebenso nach Erklärungs- und Lösungsmodellen gesucht, wie sich auch in der Praxis

Mechanismen herausgebildet haben, die den Umgang mit den für die Gesellschaft relevanten Problemstellungen erleichtern sollen.

Im Internet, als Träger und Flaggschiff der sich in den vergangenen Jahrzehnten in unglaublicher Geschwindigkeit entwickelnden Kommunikationswelten, wurde mithilfe von Suchmaschinen der Schritt von einer anfänglich relativ einfachen, profitorientierten Verwaltung der im Internet verfügbaren Datenmenge zu einer tiefgehenden Strukturierung von digitaler Information getan.

Dieser noch nicht in allen Aspekten abschätzbare Wandel des Umgangs mit Information in verschiedensten Formen führt zur hier zu behandelnden Fragestellung, ob am Beispiel der wirtschaftlich derzeit erfolgreichsten, vor allem jedoch in der Forschung engagiertesten Suchmaschine Google ein Wandel von künstlicher zu digitaler Intelligenz festgestellt werden kann. Daran anknüpfend ist die Frage zu beantworten, welche Relevanz die Feststellung, Google sei bereits als digitale Intelligenz zu bezeichnen, für eine Informationsgesellschaft haben kann und wird.

Forschungsfrage

Kann ein Wandel von künstlicher zu digitaler Intelligenz am Beispiel von Google als Konglomerat verschiedenster multimedialer Angebote festgestellt werden und wenn ja, welche Relevanz hätte diese digitale Intelligenz für eine Informationsgesellschaft?

Um diese Darstellung zu ermöglichen, werden im folgenden ersten Kapitel die für den Einstieg in die Thematik wesentlichen Begriffe Wissen, Information, Informationsgesellschaft und Digitalisierung eingegrenzt bzw. erläutert und die für die vorliegende Arbeit relevanten Unterschiede zwischen diesen Begriffen markiert.

Im zweiten Kapitel werden theoretische und praktische Lösungsmöglichkeiten für die im Zusammenhang mit der Informationsgesellschaft vorgefundenen Problemstellungen dargestellt. Einem Abriss der wesentlichen theoretischen Positionen zu diesen Problemstellungen folgt eine Darstellung praktischer Bewältigungsstrategien, die jedoch für die zuvor im Rahmen des zweiten Kapitels definierten Probleme keine zufriedenstellende Lösung erbringen.

Das Konzept einer *digitalen Intelligenz* wird im dritten Kapitel als möglicher Lösungsansatz eingeführt, wobei sowohl die Merkmale von Intelligenz definiert, als auch weitere relevante Begriffe wie künstliche Intelligenz und neurale Netzwerke erläutert werden. Google als Konglomerat verschiedenster multimedialer Dienste baut dabei auf diesen Konzepten auf, weshalb eine eingehende Beschreibung von Google, soweit für die Arbeit relevant, erfolgt. Im Anschluss daran wird unter Bezugnahme auf verschiedene theoretische und wissenschaftliche Positionen diskutiert, ob im Fall von Google bereits vom Vorliegen einer digitalen Intelligenz gesprochen werden kann. Schlussendlich wird die Relevanz für die in Zusammenhang mit der Informationsgesellschaft erkannten Problemstellungen dargelegt.

1 Begriffsklärungen

Die Grundlage, sowohl für die Untersuchung einer digitalen Intelligenz anhand des Beispiels Google in Kapitel 3, als auch der Betrachtung einer Informationsgesellschaft und ihrer Problemstellungen in Kapitel 2, stellen der Begriff Information und der eng damit verbundene Begriff Wissen dar. Im Laufe dieses Kapitels sollen daher ihre begrifflichen Unterschiede und Überschneidungen, sowie die Bedeutung der beiden Begriffe in einem digitalen Zeitalter beleuchtet werden.

1.1 Sammeln und Archivieren von Wissen in Schriftform

Zu Beginn dieser Untersuchung ist eine Unterscheidung zwischen Wissen und Information zu treffen, da sowohl die für diese Arbeit relevanten Begriffe der Wissensgesellschaft, als auch der Informationsgesellschaft seit Jahrzehnten im Raum stehen und immer wieder aufgegriffen, selten jedoch ihrer Bedeutung gemäß verwendet werden.

„Wissen grenzt sich zu Information dadurch ab, dass ersteres – für das jeweilige Individuum – als vorhanden, als „gewusst“ gelten kann. Informationen dagegen sind akut benötigt und werden dementsprechend beschafft.“ (Lehmann 2005, 34)

Während Wissen also eine passive Ressource ist, die gehortet und, bei Bedarf, abgerufen und benutzt werden kann, ist Information einerseits ständig vorläufig, andererseits auch keine vorhandene Ressource per se, kann aber gewonnen werden und, unter Umständen, zu Wissen in weiterer Folge transformiert werden. (vgl. Lehmann 2005, 34)

Information ist in diesem Zusammenhang daher grundsätzlich nötig, um Wissen an sich überhaupt zu ermöglichen.

Der Versuch des Sammelns und Archivierens von Wissen kann als eines der zentralen Anliegen von Menschen in der Geschichte immer wieder vorgefunden werden. Die Erfindung von Sprache und Schrift haben und hatten nicht nur großen Einfluss darauf, wie wir selbst mit Wissen umgehen

und so etwa lernen, sondern auch darauf, dass und wie wir Erfahrung und Expertise an andere weitergeben können. Diese Formen der Weitergabe oder auch Medien, wie beispielsweise Buch oder Internet, sind in ihren Eigenschaften unterschiedlich, haben und hatten jedoch, offensichtlich, immer eine vermittelnde Position und Funktion. Während nun beispielsweise ein Teil des Internet und Bücher gleichermaßen auf Schrift basieren, unterscheiden sich diese unter anderem in Art und Weise der Möglichkeiten, wie eine Reaktion der beiden Instanzen, zwischen denen vermittelt wird, zustande kommen kann, also im weitesten Sinne eine Kommunikation und damit Wissensaustausch stattfindet. Als Instanz kann in diesem Zusammenhang alles angesehen werden, das miteinander kommuniziert und interagiert, also etwa Menschen miteinander, genauso jedoch etwa Mensch und Computer.

Der Fokus auf Wissen in Schriftform ist historisch gewachsen, meint Frank Hartmann, ein österreichischer Medienphilosoph, sowie Medien- und Kommunikationstheoretiker. „Das Buch ist zu einer universalen Kulturtechnik geworden, zum Instrument der Klassifizierung, Vervielfältigung und Vermittlung des Wissens.“ (2006, 218)

Durch das bisher beinahe ausschließliche Sammeln von Wissen in Schrift- bzw. Buchform ist aber der Fokus für die Archivierung von Wissen im Medium Buch auf ein für die Vielfältigkeit von Information eher ungeeignetes Medium gelegt worden, dem nun möglicherweise das Internet Abhilfe schaffen kann. Ungeeignet deswegen, da das Buch den vielen Verknüpfungen und der vernetzten, verweisenden Natur von Wissen nur schwer entsprechen kann. Das Internet wiederum ermöglicht es, vielleicht nicht optimal, aber momentan zumindest in umfassenderer Art und Weise als Bücher dies könnten, nicht nur Text, also somit Sprache - respektive Schrift - , sondern darüber hinaus auch Bild, Ton, bewegtes Bild und diverse andere Informationsformen zu speichern und in eben dieser verknüpfenden und verweisenden Art zugänglich zu machen. Der Bezug zu Schrift und vor allem dem Buch bleibt dabei, beispielsweise in der Terminologie, in mancher Hinsicht erhalten.

Jeder Diskurs über neue Medien und Medienkultur bleibt direkt oder indirekt bezogen auf das Buch und die Buchkultur. Dies zeigt sich schon auf terminologischer Ebene, wenn vom Hypertext (Theodor Nelson) die Rede ist oder vom Dynabook (Alan Kay), von Software als

Textverarbeitung und von Browserfunktionen als Bookmarks oder vom Synonym kulturellen Wissens als Bibliothek. (Hartmann 2006, 218)

Dabei wurde die Schrift bzw. das Buch als traditionelle Form, Wissen zu speichern, bereits vor dem Internet herausgefordert. Zum einen durch neu entwickelte Technologien, etwa Radio und Fernsehen, aber auch zum anderen durch die moderne Massenproduktion und Druckindustrie, die eine Transformation vom Buch als einem exklusiven Medium der Wissensvermittlung hin zu Massenware bewirkte. Das Aufkommen neuer Medien stellte für Bibliotheken als traditionelle Wissensspeicher also insofern eine Herausforderung dar, als dass die bisherige Form des Archivierens nicht mehr für all die Informationen passend und anwendbar, weil nicht in Schriftform, war. Nicht nur war ungeklärt, was zu archivieren war, sondern auch wie und nach welchem System dies geschehen sollte.

1.2 Das Mundaneum als Vorläufer moderner Dokumentation von Wissen

Als einer der Pioniere auf dem Gebiet der Dokumentation gilt Paul Otlet, der belgische Vater des Informationsmanagements und spätere Gründer der, wenn man so will, ersten analogen Suchmaschine, dem „Mundaneum“. Otlet, der sich besonders für das Dokumentieren, also Archivieren von nicht nur Büchern, sondern Wissen generell, interessierte, publizierte um 1934 die Vision eines multimedialen Buches. Schon mehrere Jahrzehnte früher hatte er „mit dem Aufbau eines bibliographischen Universal-Repertoriums begonnen, einer bibliographischen Datenbank, die nach Otlets Vorstellungen einmal Informationen zu sämtlichen Publikationen, die es gibt und je gab, verfügbar machen würde.“ (Hartmann 2005, 221)

Gemeinsam mit Henri La Fontaine hatte er das Institut Internationale Bibliographie 1895 gegründet und mit dem Aufbau des Universal-Repertorium begonnen, in dem auf Karteikarten standardisierte Informationen dokumentiert werden sollten. Dieses Repertorium wuchs insgesamt auf über elf Millionen Einträge an und wurde per Post als internationaler Such- und Informationsdienst geführt. Es kann also in der Tat als frühe Suchmaschine angesehen werden. Um 1912 wurden beispielsweise bereits mehr als 1500 Anfragen jährlich bearbeitet, darüber

hinaus wurde auch ein Ableger des Repertoriums für Bilder entwickelt, der das ikonographische Gegenstück bilden sollte. (vgl. Rayward 1994, 290)

Für Otlet waren diese beiden Seiten des Repertoriums Teil des größeren Plans, jegliches Wissen, unabhängig davon in welcher Form es zur Verfügung stehen möge, zu dokumentieren und zu sammeln. Gelingen sollte dies durch eine Hyper-Dokumentation, die ein Navigieren durch Informationen und deren Verästelungen ermöglichen sollte. Dies kann also als Vorläufer der Struktur des Hypertextes - respektive des Hyperlinks - gesehen werden. Nach dem selben Prinzip funktioniert heute beispielsweise die freie Wissensplattform Wikipedia, auch dort werden in Texten und Beschreibungen Begriffe, zu denen Artikel bestehen, automatisch als Links zu genau diesen Artikeln angezeigt. Die Herausforderungen, etwa Standardisierung von Information und Datenaustausch, vor denen Otlet dabei bereits vor einem Jahrhundert stand, mussten und müssen auch heute noch von Suchmaschinen wie Google, davor aber natürlich auch um überhaupt erst die Grundstruktur des Internet zu ermöglichen, bewältigt werden. (vgl. Hartmann 2005, 224f)

Das Mundaneum stellt also mit seinem zentralisierten, aber immer für neue Beiträge offenen System der Dokumentation einen Vorläufer von sowohl Wissensarchiven und Online-Enzyklopädien wie Wikipedia, als auch einen Vorläufer von Suchmaschinen wie Google für den Umgang mit einer Vielzahl von Informationen verschiedenster Art, dar. Die Idee eines Systems, das bereits vor der dafür nötigen Technologie vorhanden war.

1.3 Dokumentation von Wissen im digitalen Zeitalter und ihre Grenzen

Die weltweite Vernetzung durch Computer und Internet, wie sie heute für so viele Menschen zur Selbstverständlichkeit geworden ist, hat in den letzten Jahrzehnten ganze Bereiche unseres Lebens, beispielsweise die Kommunikation miteinander oder den Wissenserwerb, wie von Otlet visioniert, revolutioniert und, wenn nicht den bisherigen Usus substituiert, so doch massiv erweitert. Während nicht nur Menschen untereinander vernetzt sind und sowohl grenzübergreifend Kontakt halten, als auch neue Bekanntschaften machen können, so spielt darüber hinaus vor allem die Vernetzung von Wissen immer noch und wieder eine große Rolle.

Gesellschaft und Ökonomie weltweit machen zunehmend Gebrauch von den Möglichkeiten, die diese Vernetzung gerade für die Wissensdokumentation mit sich bringt. Wissen und Expertise sind nicht mehr, wie noch zu Zeiten des Mundaneum, ortsgebunden und noch weniger zeitgebunden. Als Beispiel könnte hier das Google Books Library Project fungieren, welches das gesammelte und geballte Wissen vieler Bibliotheken weltweit zu digitalisieren versucht und so, hoffentlich, ebenfalls konservieren und archivieren, vor allem aber für ein breites Publikum zugänglich machen kann.

Genau dieses Konservieren und Archivieren gestaltet sich allerdings aus verschiedenen Gründen problematisch. Während einerseits die Menge an Informationen und Daten, die es zu bewahren gilt, kaum in Zahlen zu fassen ist, sind auch die Speichermedien, mit deren Hilfe archiviert werden soll, an sich ein Unsicherheitsfaktor. Nicht nur, dass die Lebensdauer der Speichermedien der letzten Jahre, etwa der optischen Datenspeicher CD-ROM, DVD oder Blue-Ray Disc, bei maximal 5 bis 200 Jahren, je nach zu Rate gezogener Institution, liegt, darüber hinaus gelten diese Speichermedien bereits wenige Jahre nach ihrer Einführung schon wieder als veraltet. Dies stellt ihre Vertrauenswürdigkeit und Eignung an sich massiv in Frage. Texte, die beispielsweise vor 15 oder 20 Jahren geschrieben und auf Diskette gespeichert wurden, sind heute nur mehr mit beträchtlichem Aufwand wieder abrufbar zu machen, da weder Hardware noch Software, also Laufwerk, Computer und Programm zur Textverarbeitung, für diese Zwecke noch allgemein zur Verfügung steht. Je weiter diesbezüglich in der Zeit zurück gegangen wird, desto schwieriger und umständlicher gestaltet sich der digitale Zugriff auf Information und Daten.

Doch neben diesen technischen Problemen bereitet auch der exorbitante und stete Anstieg der Daten, die es zu archivieren gilt, Schwierigkeiten, wie Christine Plaß, eine deutsche Sozialwissenschaftlerin, zu bedenken gibt.

Aus Sicht von Bibliothekaren und Archivaren, die vor der Herausforderung stehen, dieses Wissen aufzubewahren, könnte man es auch so formulieren: Während Informationen billig und schnell erzeugt werden können, kostet ihre dauerhafte Speicherung viel Zeit und Geld. (2005, 43)

Zwar könnte diesem Anstieg entgegen gehalten werden, dass viele dieser Daten und Informationen beispielsweise bloß Kopien bereits vorhandener oder ohnehin nicht ausreichend wertvoll für die Allgemeinheit sind, um archiviert zu werden. Dann stellt sich allerdings wiederum die Frage, welche Personen oder Institutionen über Relevanz beispielsweise tatsächlich entscheiden könnten und sollten.

Mit der Menge nehmen darüber hinaus auch die Vielzahl und Verschiedenartigkeit von Information und Daten - respektive deren Formate - zu. Während die Anfänge der Archivierung Text und in Ausnahmefällen auch Bild umfassten, wurden übrige Formen wie Ton, Video und speziell interaktive Medien noch bis vor wenige Jahre aus der digitalen Dokumentation weitgehend ausgeschlossen. So inkludierte beispielsweise die nicht profitorientierte Organisation *The Internet Archive* ab etwa dem Jahr 2000 diverse Bild-, Ton- und Videoformate, während das staatliche *National Digital Information Infrastructure and Preservation Program* der Vereinten Staaten von Amerika erst ab etwa 2003 deren Archivierung mit aufnahm, was aus medienhistorischer Sicht die herausragend revolutionäre Eigenschaft des Internet, multimediale Inhalte zu ermöglichen, egalisiert. (vgl. *The Internet Archive 2012*¹; *Library of Congress 2012*²)

Dabei reicht die Archivierung von Websites und deren Inhalt allein womöglich nicht aus, ist doch gerade die vernetzte Struktur und Natur des Internet eines der Kernelemente digitaler Information. Es müssten also auch alle Links und Verweise, sowie zusätzliche Kriterien, um in Zukunft Relevanz, Zusammenhänge zwischen Websites und deren Herkunft allgemein einschätzen und einordnen zu können, dokumentiert werden. Dies kennzeichnet damit ein Vorgehen, das Google beispielsweise schon durch das Speichern diverser Inhalte inklusive Metainformationen, also Information über Information, einer Website im eigenen Index in gewisser Art und Weise praktiziert. Allerdings muss diese Archivierung ohnehin verhältnismäßig schnell von Statten gehen, wie Brewster Kahle, Gründer von *The Internet Archive* in einem Interview zu bedenken gibt.

1 <http://archive.org/about/>, aufgerufen am 06.08.2012.

2 <http://digitalpreservation.gov/about/background.html>, aufgerufen am 06.08.2012.

Wie sichert man solch einen gigantischen Ozean digitaler Daten für die Nachwelt? Die durchschnittliche Lebensdauer einer Webseite beträgt etwa 44 Tage, die von offiziellen Seiten ungefähr vier Monate. Automatische Archivierungs-Software macht nur sporadisch die Runde durchs Web und bringt eine Momentaufnahme zurück, die bereits veraltet ist, wenn sie online verfügbar ist. (Heuer/Kern 2005, 94)

Wissen und Information müssen also nicht nur gespeichert und archiviert werden, sondern dabei auch ihre multimediale und vernetzte bzw. verweisende Struktur, wie beispielsweise im Internet gegeben, beibehalten, um dessen Möglichkeiten der nicht nur text- und schriftbasierten Wissens- und Informationsvermittlung tatsächlich, wie von Paul Otlet beispielsweise visioniert, ausschöpfen zu können und so eine umfassende Dokumentation zuzulassen. Während Archivierung und Dokumentation essenziell sind, um das Beschaffen von Information für Nutzerinnen und Nutzer in einer Informationsgesellschaft, wie sie in Kapitel 2 noch genauer untersucht werden soll, überhaupt zu ermöglichen, so soll im nachfolgenden Kapitel der Begriff der Information als Grundbestandteil und Voraussetzung einer solchen Informationsgesellschaft erläutert und verschiedene Bedeutungsebenen dieses Begriffes der Information aufgezeigt werden.

1.4 Die Bedeutungsebenen des Begriffes der Information

Wie bereits zum Thema Wissen ausgeführt, unterscheidet sich Information von Wissen dadurch, dass erstere gewonnen werden muss und unter Umständen erst in einem folgenden Schritt zu Wissen transformiert werden kann. Information ist also die Basis einer Informationsgesellschaft, wie sie in weiterer Folge zur Untersuchung steht. Doch hat der Begriff Information, sowohl historisch als auch wissenschaftlich bedingt, verschiedene Bedeutungen und Zuständigkeitsbereiche. Der in Kapitel 1.1 beschriebene Ansatz, Information in Verbindung mit Wissen gesetzt zu sehen, ist dabei einer von mehreren, die nachfolgend erwähnt werden sollen. Generell avancierte Information besonders in jüngster Zeit durch Massenmedien und Internet zur wichtigsten Ressource und Grundlage, die uns zur Verfügung steht und mit der wir das tägliche Leben bestreiten.

Zum einen bedeutet „informare“ das handwerkliche Formen (etwa eines „gewaltigen Rundschilds“ durch den Schmied in der Dichtung des Vergil);

und fast gleichzeitig wird Information als Inhalt eines vollständigen Satzes (etwa bei Cicero) verstanden. (Janich 2006, 20)

Informare als Formen von etwas nimmt hierbei ein Merkmal der Digitalisierung vorweg, nämlich dass die digitale Form von Daten und Information immer auch deren Manipulierbarkeit bedingt, da, einmal digitalisiert, jegliche Information lediglich aus Computer-Codes besteht. Diese tatsächliche Unfassbarkeit und Divergenz von Information unterstreicht dabei deren Eigenständigkeit, die auch Norbert Wiener, ein amerikanischer Professor für Mathematik am *Massachusetts Institute of Technology* und Vorreiter der Kybernetik, attestiert. „Information is information not matter or energy.“ (Wiener 1961, 155)

Dieser Satz, mit dem Wiener etwas scheinbar Triviales und Redundantes formuliert, bringt auch Peter Janich, einen deutschen Philosoph und Professor für Wissenschaftstheorie, einerseits zu der Annahme, dass, wie auch Materie und Energie, Information in den Bereich der Physik zu gehören scheint, andererseits Information aber eben auch etwas Eigenständiges sei, das nicht mit Materie oder Energie verglichen werden könne, sondern möglicherweise mit deren Struktur. Genauer: „einer Struktur, die sich nur an Vorgängen oder Wirkungen“ zeige. (2006, 14)

Information ist darüber hinaus aber sowohl für Lebewesen als auch für Computer essenziell. Information an sich ist hier nicht statisch zu verstehen, vielmehr entstehen Informationen in Lebewesen und Maschinen ständig, werden transformiert und verarbeitet und bilden so die Grundlage für beispielsweise Kommunikation, wie schon im direkten Vergleich mit Wissen Information stets das Vorläufige und Temporäre darstellte. (vgl. Janich 2006, 16)

Im alltäglichen Gebrauch hat Information Geltung und Bedeutung für jeden, der diese sucht, ob auf Google, Wikipedia oder einem anderen analogen oder digitalen Speicherort von Information. (vgl. Janich 2006, 19)

Frank Hartmann leitet den Begriff Information, wie dieser heute dominant ist und verwendet wird, von einer neuen Auffassung von Kommunikation ab. Deren zentrale Rolle und Funktion drehe sich um „zirkuläre Kausalität,

Feedback-Mechanismen und Steuerungsprozesse in Lebewesen und in technischen Systemen". (2008a, 45)

Diese Auffassung von Information sei operational und solle eine Maßeinheit für die „Veränderung im Zustand eines Systems ausdrücken“, erläutert Hartmann weiter. (ebd.)

In diesem Zusammenhang und dank der militärischen Herkunft, aus deren Forschungsprojekten dieser neue Ansatz wohl vorrangig entstand, überschneiden sich in diesem Konglomerat, das der Begriff der Information darstellt, auch Kommunikation, Kontrolle und Steuerung, also Prozesse, die für die Entwicklung und den Fortbestand eines Systemes oder einer Intelligenz von höchster Priorität sein müssen. (vgl. Hartmann 2008a, 45)

Zwar unterstützt diese Auffassung des Informationsbegriffes die Vermutung, dass Information und deren Verarbeitung, Weiterleitung oder Speicherung essenziell für das Überleben eines Organismus seien, Hartmann klärt allerdings nicht, inwiefern sich diese Veränderung innerhalb eines Systemes dann tatsächlich manifestieren könnte, ohne die Auswirkung auf das Verhalten als einzig mess- und erkennbare Referenz zu nehmen. Das Verhalten und dessen tatsächliche Ausdruckskraft im Bezug auf Intelligenz sollen im Kapitel 2.9.b näher untersucht werden.

1.5 Informationsverarbeitung und -austausch

Information als etwas Flüchtliges, das, sobald eingeholt, schon wieder aktualisiert und erneut beschafft werden muss, ist nicht nur in Beruf und Freizeit vieler Menschen fester Bestandteil des Alltags. In diesem Zusammenhang nehmen wir durch das Einholen von Information die virtuelle Welt wahr und informieren uns, wenn wir beispielsweise Google mit einer Suche beauftragen und so anhand der Ergebnisse, die Google liefert, ein Bild eines Begriffs konstruieren.

Peter Bienert, Physiker und Professor für Medien- und Informationstechnik, geht in Anlehnung daran davon aus, dass Menschen und Umwelt nicht nur in der Virtualität offensichtlich in einem ständigen Kontakt zueinander stehen. Das bedeutet, dass wir Einflüsse aus unserer Umgebung einerseits

stets vorhersagen, andererseits darauf reagieren und unsere Vorhersagen somit rekursiv anpassen. "Selbst das 'Funktionieren' des einfachsten Organismus ist von der Fähigkeit abhängig, die Außenwelt zu registrieren und die benötigten Reaktionen darauf im Organismus selbst abzustimmen". (1998, 7)

Es geht Bienert also vorrangig um eine stete Kommunikation, die einen Informationsaustausch bereits auf der elementarsten Ebene des Lebens bedingt, eine Eigenschaft, die für eine digitale Intelligenz also nicht nur von Vorteil, sondern essenziell wäre.

Auch Nicholas Carr, ein amerikanischer Autor, der in den Bereichen Technologie, Wirtschaft und Kultur publiziert und sich in seinem Buch *The Big Switch* an einem Bild von Google als künstlicher Intelligenz versucht, sieht als Gemeinsamkeit aller lebender Systeme, von Amöben bis zu ganzen Nationen oder Staaten, das Verarbeiten von Materie, Energie und Information. Sie alle verwenden Materialien aus ihrer Umgebung, um die Energie dieser Materialien in verschiedene nützliche Substanzen zu transformieren und etwaigen Müll auszufiltern.

This continuous turning of inputs into outputs is controlled through the collection, interpretation, and manipulation of information. The process of control itself has two thrusts. It involves measurement – the comparison of the current state of a system to its desired state. And it involves two-way communication – the transmission of instructions and the collection of feedback on results. The processing of information for the purpose of control may result in the release of a hormone into the bloodstream, the expansion of a factory's production capacity, or the launch of a missile from a warship, but it works in essentially the same way in any living system. (2009, 192)

Wichtig für ein System sind also sowohl das Vergleichen des momentanen mit dem gewünschten Zustand, als auch die Fähigkeit, Feedback, beispielsweise von der Umgebung, einholen zu können, um so mittels Kommunikation im weitesten Sinn die Auswirkungen des eigenen Handelns zu überprüfen. Das ständige und stete Einholen und Beschaffen von Information ist also für jegliche Systeme essenziell. Carr fasst hier also zwei der bisher markierten, benötigten Eigenschaften einer digitalen Intelligenz zusammen. Einerseits das Sammeln und Verarbeiten von Information, wie Bienert und Marshall McLuhan, ein kanadischer Philosoph

und Kommunikationstheoretiker, beispielsweise formulierten, andererseits die rekursive Kommunikation, um auf Veränderungen der Umwelt reagieren zu können, wie sie etwa bei Flusser oder Hartmann, jeweils in anderen Zusammenhängen, dargestellt und unter anderen auch noch in den Abschnitten zu neuronalen Netzwerken und Intelligenz, Kapitel 2.7 bzw. 2.9, essenziell sein wird.

Information am Computer und damit auch im Internet sind nun in besonderem Ausmaß speziell, da sie sich in einem digitalen Umfeld befinden. Im Weiteren soll also kurz beleuchtet werden, welche Bedeutung dem Wort „digital“ in diesem Zusammenhang zukommt.

1.6 Digitalisierung von Information

Damit Daten in elektronischen Computern und verschiedenen Digitalmedien verarbeitet und gespeichert werden können, müssen sie in einem bestimmten Format vorliegen. So gibt es unterschiedliche Datenformate für Text, Bild, Video und Audio. Sie repräsentieren Signale in digitaler Form; wie sie das tun, das hängt von den unterschiedlichen Entscheidungen im Lauf der Computerentwicklung sowie den technischen Notwendigkeiten der Codierung ab. (Hartmann 2008b, 59)

Wenn von Daten, Information, Dokumenten oder Dateien auf einem Computer die Rede ist, kann damit entweder originär digital entstandener Inhalt, oder aber die digitale Repräsentation eines analog bereits vorhandenen Inhaltes, beispielsweise eines Buches, gemeint sein. Während etwa Text in einem Textverarbeitungsprogramm, beispielsweise Google Docs, bereits digital entsteht bei der Eingabe durch Nutzerinnen oder Nutzer – respektive durch diese Eingabe –, so kann ein Buch, wie etwa im Rahmen des Google Books Library Projects, digitalisiert werden, es wird dabei also zu einer digitalen Repräsentation dessen transformiert. Diese Transformation erfolgt zumeist, unabhängig davon, ob es sich um visuelle, auditive oder andere Information handelt, durch eine Abtastung des Objekts, das digitalisiert werden soll.

Ein Bild wird während des Scan-Vorgangs abgetastet und in einzelne Bildpunkte zerlegt, die diverse Informationen beinhalten können, beispielsweise Farbe und Helligkeit. Diese Informationen werden quantisiert, also in Zahlenwerten festgehalten, die als Raster abgespeichert

werden können, also etwa als Bild, das später am Computer geöffnet werden kann. Wichtig dabei ist, dass dieses Zerlegen und die Quantisierung zumindest bisher unabhängig vom Vorgehen der Digitalisierung immer mit einem Verlust an Information verbunden ist, „da es sich um eine Reduktion der vorhandenen Informationen und eine Übersetzung in technische Messwerte handelt“. (Hartmann 2008b, 60)

Digitalisierung bedeutet dabei, dass die Information von ihrer physischen Gebundenheit gelöst wird. Diese Information wird in weiterer Folge zumeist, so erfahren es Nutzerinnen und Nutzer, vom Internet absorbiert. Um wieder Bezug auf das Google Books Library Projects zu nehmen, so wird beispielsweise ein Buch, das von Google ausgewählt wurde, in einem solchen Prozess der Digitalisierung in Code, also eine binäre Kodierung, umgewandelt und in einem passenden Format online, im Internet, verfügbar gemacht. Spannend ist dabei die Tatsache, dass jegliche Information im Internet, ob originär oder digitalisiert, lediglich aus Code besteht. „Mit der Konsequenz, dass die Erreichbarkeit der Informationen immer schon ihre Manipulationsmöglichkeit enthält, ihr Zustand ist damit immer nur ein vorläufiger. (Krug et al. 2005, 20)

Résumé

Im Laufe dieses Kapitels konnte nun gezeigt werden, dass die Begriffe Information und Wissen zwar eng miteinander verbunden sind, sich aber wesentlich in der Dauer ihrer Gültigkeit unterscheiden. Da gerade durch die Vernetzung von Menschen und Computern im Internet die Menge des verfügbaren Wissens dramatisch ansteigt, wird das Informieren, also das Einholen und Beschaffen von Information in kleinen Stücken bei Bedarf, immer wichtiger. Die Archivierung von Wissen stellt eine Herangehensweise dar, mit dieser Menge von Wissen umzugehen. Während die Archivierung allerdings ausschließlich auf Text und Schrift beschränkt war, so kann die Dokumentation nach Paul Otlet mit ihrem Ansatz, Wissen unabhängig von dessen Form zu bewahren und verfügbar zu machen, als Vorgänger von Suchmaschinen wie Google gesehen werden. Welche Auswirkungen die Digitalisierung von Information und Wissen, sowie die Verarbeitung und der Austausch von Information auf die Gesellschaft haben und welche Problemstellungen sich daraus ergeben, soll im nächsten Kapitel genauer betrachtet werden.

2 Informationsgesellschaft

Wie bereits am Ende des ersten Kapitels kurz erwähnt wird ein Wandel von einer auf Wissen basierenden Gesellschaft hin zu einer auf ständig neu zu beschaffender Information aufbauenden Gesellschaft vermutet. In diesem Kontext ist also der Ausdruck Informationsgesellschaft passender als der oft synonym verwendete Begriff der Wissensgesellschaft.

Die Informationsgesellschaft kennzeichnet beispielsweise, dass gerade im vergangenen Jahrhundert ein struktureller Wandel der Ökonomie von Industrie zu Wissensprodukten, also von industrieller Fertigung hin zu Dienstleistungen im Informationsbereich, bei denen die Weitergabe und der Austausch von Wissen einen essenziellen Bestandteil darstellen, erkannt werden kann.

Wissen wird hier nicht gehortet, sondern nur bei Bedarf durch das Beschaffen von Information generiert. Dabei geht es regelmäßig nicht um Erfahrungswissen, also Wissen, das durch eigene Erfahrungen erworben und geprägt wird, sondern eher um wissenschaftliches Wissen im weitesten Sinne, also Wissen, das auf Erkenntnissen aus analytischen Prozessen hervorgeht. (vgl. Lehmann 2005, 34)

Dieser Umgang mit Wissen als essenzieller Ressource ist dabei nicht unbedingt neu, bereits seit jeher hat und hatte Wissen für gesellschaftliche Veränderungen natürlich eine enorme Bedeutung. Entscheidend ist allerdings, dass Wissen nun nicht mehr absolut ist, es wird in Frage gestellt und thematisiert sich selbst.

Als Beispiele dafür, dass die Debatte um Wissen und Information tatsächlich auch auf einer gesellschafts- und weltpolitisch mehr oder weniger relevanten Ebene geführt wird, können die beiden 2003 und 2005 gehaltenen Weltgipfel der Vereinten Nationen mit dem Titel „Weltgipfel zur Informationsgesellschaft“ gelten. In zwei Etappen wurde dabei als deutschsprachiger Beitrag der Entwurf zu einer „Charta der Bürgerrechte für eine nachhaltige Wissensgesellschaft“ erarbeitet, der in verschiedenen Punkten eine Einigung auf bestimmte einzuhaltende Grundsätze und technische Standards in Bezug auf freie Verfügbarkeit von und Umgang mit

Wissen, auch in Zeiten immer häufiger diskutierter Probleme im Namen des Datenschutzes, zu erreichen versuchte. (vgl. Bendrath 2003, 2)

Die „Charta der Bürgerrechte für eine nachhaltige Wissensgesellschaft“ fordert einen an Nachhaltigkeitsprinzipien orientierten freizügigen und inklusiven Umgang mit Wissen und Information. Die Herausforderung der Wissensgesellschaft besteht darin, den Menschen das Wissen anderer über den Zugang zu Information offen zu halten und sie so auf einer sicheren Grundlage handlungsfähig zu machen. Die Charta setzt einen Akzent gegen die zunehmende Privatisierung und Kommerzialisierung von Wissen und Information. Denn eine Gesellschaft, in der der Schutz von geistigem Eigentum das Wissen zunehmend zum knappen Gut macht, ist nicht nachhaltig. (Bendrath 2003, 2)

Obwohl nur Aufruf zu einem freiwilligen und mündigen Umgang mit den Themen Information, Wissen und Kommunikation, so sind diese Charta und der Weltgipfel insgesamt doch Anzeichen dafür, dass Information und Wissen, vor allem der möglichst uneingeschränkte und stete Zugang dazu, eine große Rolle nicht nur für die Zivilgesellschaft, sondern auch die Ökonomie spielen. Ein Augenmerk wird dabei auch auf das Forcieren von kooperativen Formen von Informationsbeschaffung, Wissensproduktion und deren Verbreitung gelegt, was beispielsweise Wikipedia und Google in Teilen betrifft, wie später noch näher erläutert wird.

Doch Wissensproduktion und Informationsverbreitung, hervorgerufen von und zugleich kombiniert mit verschiedensten technischen Entwicklungen, bedingten in den vergangenen Jahren und Jahrzehnten auch einen Anstieg der Geschwindigkeit, mit der Information verfügbar, also abrufbar, ist. Darüber hinaus erhöhte sich auch die Geschwindigkeit von Kommunikation als Informationsaustausch.

Diese zunehmende Geschwindigkeit von Information und Kommunikation ist nach Peter Bienert insofern einer der Schlüsselpunkte unserer heutigen Informationsgesellschaft, als dass wir mit dieser Geschwindigkeit interdisziplinär im Verständnis nicht Schritt halten können. Bienert erkennt die Herausforderung des Informationszeitalters also im „Umgang mit Technologien, deren Folgen schneller eintreten, als dies intellektuell auf Basis vorhandener Erkenntnismodelle verarbeitet werden kann.“ (1998, 2)

**Beschaffung, Verarbeitung
und Vermittlung
als Basis einer
Informationsgesellschaft**

Dabei ist diese Geschwindigkeit mit Sicherheit zu einem großen Teil auf verschiedenste Technologien, die in den vergangenen beiden Jahrhunderten entwickelt und entdeckt wurden, zurückzuführen. Die Idee einer Informationsgesellschaft, in deren Grundfesten das Beschaffen, Verarbeiten und Vermitteln von Information verankert sind, ist also nicht nur, aber ebenfalls auch an Technologie gebunden.

Als wichtige Einschnitte, die in weiterer Folge zur Entwicklung der Informationsgesellschaft führten, nennt Frank Hartmann zum Einen die Erfindung von nichtsprachlichen Speichertechnologien, zum Anderen die Elektrizität bzw. deren Beherrschung. Die Kultur- und Mediengesellschaft, in der wir heute leben, fuße dabei auf den technischen Innovationen, besonders des 19. Jahrhunderts, so Hartmann. Es entstanden vor Allem Technologien zur Telekommunikation, deren Aufzeichnung und Reproduktion, Telegraf, Telefon, Fotografie sowie Film und Druck werden hier genannt. Diese wurden nur wenig später von elektronischen Übertragungstechniken, etwa Funk und Fernsehen, ergänzt und erweitert, sowie durch das momentan vorherrschende Medium zur Datenverarbeitung, den Computer, überflügelt. Besonders die Vernetzung von Information in Form des Internets hebt Hartmann hier hervor, diese bildet natürlich auch jene Grundlage, ohne welche in weiterer Folge das Suchen von Informationen in einem solchen Netzwerk durch ein System wie Google nicht möglich wäre. (vgl. Hartmann 2008a, 7)

Durch das Einbringen des Begriffes der Informationsgesellschaft ergibt sich also ein zumindest vager zeitlicher Rahmen für die zu untersuchende Entwicklung einer digitalen Intelligenz, der nicht allzu weit in das 20. Jahrhundert zurückreichen kann. Im Weiteren soll genauer untersucht und diskutiert werden, welche Indizien tatsächlich auf den Wandel zu einer Informationsgesellschaft hindeuten.

2.1 Die sechs Merkmale der Informationsgesellschaft nach Lehmann

Erstes Merkmal: Wirtschaftliche Bedeutung

Der Soziologe Kai Lehmann nennt sechs Merkmale als Hinweise darauf, dass tatsächlich ein Wandel der Gesellschaft erfolgt. Zu Beginn identifiziert Lehmann, dass „die wirtschaftliche Bedeutung des Informationssektors wächst.“ (2005, 37)

Während Information und Wissen ihren exklusiven Charakter, wie bereits umrissen, nicht verloren, aber doch ein wenig eingebüßt haben, so kann nicht zuletzt Google, eines der größten Unternehmen weltweit, als Paradebeispiel für diesen Trend zu ständig neu zu beschaffender Information gelten. Eric Schmidt, Informatiker und langjähriger Chief Executive Officer des Unternehmens, sieht die Möglichkeiten und Chancen, die das Internet und dessen viele Nutzerinnen und Nutzer für Google beispielsweise bereit hält, ganz pragmatisch. "You get a billion people doing something, there's lots of ways to make money. Absolutely, trust me. We'll get lots of money for it." (Jenkins 2010)

Das Internet als Platz der Zusammenkunft und Ansammlung vieler Leute bietet also gerade für den Werbemarkt, in dem auch Google den größten Teil seines Umsatzes generiert, neue Wege, um Konsumentinnen und Konsumenten zielgruppenspezifischer erreichen zu können.

Zweites Merkmal: Forschung und Wissenschaft

Als zweites Merkmal erkennt Lehmann ein Aufleben von Forschung und Wissenschaft respektive deren Einfluss auf Ökonomie und Gesellschaft. „Naturwissenschaftliche Erkenntnisse nehmen exponentiell zu.“ (2005, 37)

Der im vergangenen Jahrhundert aufgekommene Trend der Analyse und Rationalisierung von Prozessen in verschiedensten Bereichen des Lebens, um sowohl ökonomische als auch soziale Vorgänge planbarer, kontrollierbarer und in weiterer Folge effizienter zu machen, kann in diese Richtung interpretiert werden. Auch in der Produktion nehmen Forschung

und Entwicklung immer öfter einen großen Teil des gesamten Herstellungsprozesses ein, beispielsweise in Bezug auf Mikrochips oder auch technische Geräte im allgemeinen.

**Drittes Merkmal:
Informationsmenge
steigt an**

„Die via Datennetze verfügbare Informationsmenge steigt exponentiell an.“ (Lehmann 2005, 37) Diese Beobachtung kann ebenfalls wohl kaum bestritten werden. Während auf Grund einer fehlenden Kontrollinstanz keine konkreten Zahlen über verfügbare Daten im Internet, in welcher Form auch immer, vorhanden sind und auch Google beispielsweise keine oder nur teilweise Angaben über Speichervolumen oder Seitenaufrufe publiziert, so kann etwa auf die bereits erwähnte Institution *The Internet Archive* verwiesen werden.

Nach eigenen Angaben konnten 2006 bereits mehrere Petabyte an Daten gespeichert werden, also mehrere tausend Terabyte, eine Zahl, die in Anbetracht der noch vor einigen Jahren benützten diversen Diskettenformate mit einer maximalen Größe von wenigen Megabyte absurd erscheint und diese enorme Datenexplosion nur zu deutlich unterstreicht.

**Viertes Merkmal:
Universelle
Multimedia-Dienste**

Lehmann identifiziert die Ausbreitung und Nutzung diverser multimedialer Anwendungen und Dienste im Internet und breite Akzeptanz dieser als ein weiteres Indiz für den Wandel zu einer wissens-, respektive informationsorientierten Gesellschaft. „Universelle Multimedia-Dienste finden weite Verbreitung in der Bevölkerung.“ (Lehmann 2005, 37)

Als eines der bekanntesten Beispiele hierfür kann das soziale Netzwerk Facebook gelten, das seit dem offiziellen Start bzw. Launch im Jahr 2004 nun nach eigenen Angaben bereits mehr als 550 Millionen aktive Nutzerinnen und Nutzer täglich und fast eine Milliarde innerhalb eines

Monats verzeichnet. Die Liste weiterer Dienste, Websites und Anwendungen, die ebenfalls multimediale Angebote online oder offline bereithalten, ist lang, wäre doch eine Netzwelt ohne YouTube, Twitter, Flickr, Skype, diverse Apple-Angebote oder viele mehr nur schwer denkbar. Ob zum Zweck der Kommunikation, Organisation, Kreation oder auch Unterhaltung, die Verbreitung dieser Dienste und Angebote kann anhand vieler Beispiele beobachtet werden. (vgl. Zandy 2012, 1ff)

**Fünftes Merkmal:
Politische
Anerkennung**

Die beiden bereits erwähnten Weltgipfel der Vereinten Nationen zum Thema bestätigen zumindest teilweise auch das Eintreten des nächsten Merkmals, das Kai Lehmann als fünftes von sechs nennt. „Die Notwendigkeit einer informationstechnischen Basisqualifikation für alle wird politisch anerkannt.“ (2005, 37)

Um eine notwendige informationstechnische Basisqualifikation zu ermöglichen, müssen beispielsweise Wissen und Information, aber auch der Zugriff darauf, frei sein, wie es der bereits erwähnte deutsche Beitrag zur Charta fordert. Damit verbunden - respektive daran gebunden - sind also auch Informations- und Kommunikationsfreiheit als Grundrecht, vor allem aber auch eine Ausbildung, die den Umgang mit und Zugang zu Information und Wissen im weiteren Leben ermöglicht.

Es gilt also nicht nur die Rahmenbedingungen anzupassen, sondern auch die individuellen Voraussetzungen, also beispielsweise Fähigkeit im Umgang mit Technik und Internet, zu schaffen. Denn während bis vor wenige Jahrzehnte das Lernen von Schrift und Sprache für Menschen ausreichend war, um Zugang zu Wissen und Information zu erlangen, so müssen bzw. sollten heute zusätzlich dazu der Umgang mit und die Benutzung von Computern respektive Maschinen als Medien, die immer öfter zwischen Menschen und dem gespeicherten Wissen und Informationen stehen, gelernt und beherrscht werden. (vgl. Bendrath 2003, 3f; Plaß 2005, 45)

**Sechstes Merkmal:
Technologie
im Berufsalltag**

Auch das letzte von Lehmann identifizierte Merkmal eines Wandels hin zu einer Informationsgesellschaft kann ohne große Einwände übernommen werden. „Informationstechnologien bestimmen den Berufsalltag auch traditionell technikferner Arbeitsplätze.“ (2005, 37)

Während für Großunternehmen ohnehin bereits seit Jahrzehnten ein Bestehen ohne elektronische Datenverarbeitung, also den Einsatz von Computern in unterschiedlicher Art und Weise, schwer vorstellbar ist, so bieten Computer und Internet auch und gerade für kleine und Mittelbetriebe Vorteile. Ob Organisation, Buchhaltung, Werbung - respektive Akquise - oder Kommunikation, keiner dieser Bereiche kann momentan wohl effizienter abgedeckt werden als mithilfe der nun nicht mehr allzu neuen Medien.

Da das Vorliegen sämtlicher von Lehmann formulierter Merkmale einer Informationsgesellschaft diskutiert und anhand von Beispielen überprüft wurde, kann festgestellt werden, dass der zu Beginn dieses Kapitels nur vermutete Wandel der Gesellschaft hin zu einer Informationsgesellschaft bereits erfolgt ist. Die Informationsgesellschaft nach der Theorie Lehmanns ist also nun belegt. Aus dem Wandel zu dieser auch auf ihre Legitimität untersuchten Informationsgesellschaft ergeben sich Konsequenzen verschiedener Größenordnungen und Natur. Diese sollen im nachfolgenden Kapitel exemplarisch in zwei Problemstellungen formuliert werden.

2.2 Folgen und Probleme der Informationsgesellschaft

Das Problem der Menge an Information

Das am leichtesten zu identifizierende Problem in der Informationsgesellschaft stellt der stete Anstieg von Informations- und Kommunikationsaustausch speziell dahingehend dar, dass die Menge an Information für das Individuum, etwa den mündigen Internetnutzer oder die mündige Internetnutzerin, kaum mehr filterbar ist und überblickt werden kann.

Dies stellt vermutlich einen der offenkundigsten Gründe für die Entstehung von Suchmaschinen, allen voran Google, dar. Der stete und exorbitante Anstieg der Menge verfügbarer Information kann von uns Menschen ohne technische Hilfe weder momentan noch in Zukunft bewältigt werden.

Doch nicht nur die Menge an Information, die für uns verfügbar ist, sondern auch die Art und Weise, wie Informationen und Daten verknüpft sind - respektive von uns vernetzt betrachtet werden müssen - , hat ein Ausmaß erreicht, das vermutlich nicht nur bloß fragmentarisch durchschaubar, sondern kaum überhaupt zur Gänze erfahrbar ist.

Vernetzte, komplexe Systeme, wie etwa „Wirtschaftssysteme, Entwicklungen in der Gesellschaft, Beziehungen zwischen Organisationen und Staaten, die miteinander verkoppelten Ökokreisläufe, die Vielfalt sozialer Beziehungen zwischen Menschen, Informationsflüsse im Internet und vieles mehr“ werden durch das Einwirken einer Vielzahl an Akteuren, ob Mensch oder Computer, und deren Wechselwirkungen auf sowohl das System als auch einander für uns zunehmend unfassbar und dementsprechend auch unvorhersehbar oder zumindest nicht präzise planbar. (Ritter 2008, 130)

**Das Problem
der Komplexität
von Information**

Als zweites Problem in Zusammenhang mit Information kann daher festgestellt werden, dass in Verbindung mit der Menge an verfügbarer Information auch die Komplexität von Systemen - respektive auch Informationssystemen mit komplexen Zusammenhängen - steigt und auch aus diesem Grund für mündige Nutzer und Nutzerinnen ohne technische Hilfe weder momentan noch in Zukunft bewältigt werden kann.

Praktisch veranschaulicht wird beispielsweise das zuerst genannte Problem der Menge, wenn nach einer beliebigen Information, etwa einem Begriff, gesucht wird. Da Google nicht nur im amerikanischen Raum, sondern weltweit bereits nahezu seit dessen Bestehen weit mehr als die Hälfte des Marktes der Suchmaschinen, neben beispielsweise Bing, Yahoo oder ähnlichen, für sich beansprucht, soll an dieser Stelle das bereits seit 2004 erstmals im deutschsprachigen Duden geführte Verb „googeln“ verwendet werden. (vgl. Duden 2004³)

Wenn also eine Nutzerin oder ein Nutzer nach einer Information sucht und einen beliebigen Begriff, beispielsweise Intelligenz, googelt, werden von Google auf Grund von vielen Faktoren, die in Kapitel 3.1 noch detailliert untersucht werden sollen, innerhalb des Bruchteils einer Sekunde zum momentanen Stand 9.520.000 Ergebnisse geliefert, also Webseiten und Links, die in einem mehr oder weniger relevanten Zusammenhang mit diesem Begriff Intelligenz stehen. Diese Zahl ist für uns als mündige Internetnutzerinnen und mündige Internetnutzer mit großer Sicherheit völlig absurd.

Während Google natürlich, wie bereits erwähnt, nach einer Reihe von Faktoren die Relevanz des gesuchten Begriffes auf einer Website im Vergleich zu anderen Websites gewichtet, also beispielsweise, ob der Begriff bereits im Titel des Dokumentes oder der Website, möglicherweise aber auch in einer der Überschriften eines Absatzes oder Kapitels dort, platziert ist, so versucht die Suchmaschine darüber hinaus uns in Relation mit dem gesuchten Begriff zu setzen und so Suchergebnisse zu bevorzugen, die nicht nur auf Grund des gesuchten Begriffes, sondern auch

3 <http://www.duden.de/rechtschreibung/googeln>, aufgerufen am 06.08.2012.

diverser anderer Informationen oder Annahmen, die Google von uns hat, für uns passen könnten.

So beschreibt etwa Eric Schmidt in einem Interview, welches Bild von einer Nutzerin oder einem Nutzer bestehen und somit Faktoren bei einer Suche darstellen kann. "We know roughly who you are, roughly what you care about, roughly who your friends are." (Jenkins 2010⁴)

Natürlich stehen diese Informationen auch uns zu einem Teil zur Verfügung, allerdings werden bei einer Suche faktisch kaum mehr als die ersten zehn Ergebnisse bzw. die ersten drei Seiten an Ergebnissen von Nutzerinnen und Nutzern angesehen. (vgl. Sumako 2009⁵)

Was Google also besonders interessant macht und von anderen Technologien abhebt, ist die Vermutung, dass für beide Probleme, das der Menge an Informationen und das der Komplexität dieser Informationen in Systemen, schon heute Lösungen angeboten werden, die diesen zumindest teilweise beikommen können. Diese Vermutung wird in Kapitel 3.4 noch genauer auf ihre Richtigkeit untersucht werden.

Diese beiden formulierten Probleme der Informationsgesellschaft, nämlich Menge und Komplexität von Information, wurden und werden dabei, da interdisziplinär relevant, von verschiedenen Bereichen der Wissenschaft thematisiert und untersucht. Einige für diese Untersuchung relevanten theoretischen Positionen dazu sollen im Weiteren etwas genauer betrachtet werden.

2.3 Theoretische Positionen zur Informationsgesellschaft

Mit den Entwicklungen und zunehmend neuen, weil veränderten, Umständen in der Gesellschaft haben sich in den vergangenen Jahrzehnten verschiedene Denkschulen beschäftigt. Ob Soziologie, Philosophie, Medien- oder Kommunikationswissenschaften, in diesem Kapitel sollen einige interessante und markante Positionen, soweit im Zusammenhang

4 <http://online.wsj.com/article/SB10001424052748704901104575423294099527212.html>, aufgerufen am 06.08.2012.

5 <http://www.21v.de/>, aufgerufen am 06.08.2012.

von Informationsgesellschaft und den beiden Problemen der Menge und Komplexität nötig, dargestellt und diskutiert werden.

2.3.a Anthony Giddens

Vor allem die Fülle an Information, die verfügbar ist und stetig ansteigt, veranlasste Anthony Giddens, britischer Professor für Soziologie, von einer Orientierungslosigkeit zu sprechen, die das Gefühl nährt, „systematisches Wissen über die Organisation der Gesellschaft sei nicht zu haben“. (1996, 10)

Giddens zufolge genüge die Einführung eines Begriffes wie „Postmoderne“ nicht für die Rechtfertigung dieser Sachlage, sondern es erfordere ein erneutes Befassen mit der Moderne, vor Allem in den Sozialwissenschaften. Nach Giddens haben die Lebensformen, also sozial und gesellschaftlich geprägte Arten des Zusammenlebens, die während und durch die Moderne seit dem 19. Jahrhundert entstanden, uns von althergebrachten traditionellen Typen einer sozialen Ordnung losgelöst. Sowohl auf extensionaler als auch auf intensionaler Ebene, also nicht nur die Gesellschaft, sondern auch jede und jeden einzeln betreffend, wären die Auswirkungen und damit einhergehenden Umgestaltungen der Moderne tiefgreifender, als das bei früheren Perioden und deren charakteristischer Art des Wandels der Fall gewesen sei. (vgl. 1996, 10ff)

Einerseits hätten in extensionaler Hinsicht die Umstellungen vor allem dem Entstehen einer den Globus umspannenden Form der sozialen Verbindung gedient, andererseits würden sie aus intensionaler Sicht auch intimste und persönlichste Elemente und Merkmale unseres Alltages und unserer täglichen Existenz ändern. Die Auswirkungen, seien sie auch unterschiedlich in Art und Umfang, betreffen also sowohl das gesamte gesellschaftliche Gefüge, als auch den Einzelnen und die Einzelne in persönlicher Denk- und Handlungsweise, was logisch erscheint. (vgl. Giddens 1996, 13)

2.3.b Vilem Flusser

Vilem Flusser, ein mittlerweile verstorbener Medienphilosoph und Kommunikationswissenschaftler, dessen Idee der telematischen Gesellschaft später noch erwähnt werden soll, sieht die Veränderungen im gesellschaftlichen Gefüge dahingehend von Bedeutung für den Einzelnen und die Einzelne, als vor Allem die Kommunikation, die nach Flusser als Infrastruktur der Gesellschaft dient und durch die eine Gesellschaft näher zusammenrücken kann als es ihr eigentlich möglich wäre, einem Wandel unterzogen ist. (vgl. 1997, 145)

Das Befassen mit der Moderne und den gesellschaftlichen Konsequenzen, die neue Technologien mit sich führen, bringt Flusser unter anderem zum Postulieren einer Annahme über die Informationsgesellschaft als eine soziale Struktur, in der Information bzw. deren Verteilung und Verarbeitung, also auch Kommunikation, eine wichtige Rolle einnimmt. Flusser attestiert dieser Annahme einer Informationsgesellschaft das Entstehen aus der Industriegesellschaft heraus und erläutert weiter, dass die sich nun seit mehreren Jahrzehnten entwickelnde Informationsgesellschaft die Industriegesellschaft schrittweise immer mehr verdrängen werde. (vgl. 1997, 143)

Diesen Zugang zu weit Entferntem bezeichnet bei Flusser die Telematik, „eine Technik zum selbstbewegten Näherrücken von Entferntem“. (1997, 145)

Die Telematik setzt Apparate, in diesem Fall Computer, voraus, die der Informationsgesellschaft in verschiedener Hinsicht dienen, besonders aber in Bezug auf die Verwirklichung des einen im anderen, das „Abschaffen des Selbst zugunsten der intersubjektiven Verwirklichung“. (Flusser 1997, 146)

Computer sind Apparate zum Verwirklichen von innermenschlichen, zwischenmenschlichen und außermenschlichen Möglichkeiten dank des exakten kalkulatorischen Denkens. Diese Formulierung kann als eine mögliche Definition von „Computer“ verstanden werden. Wir sind nicht mehr Subjekte einer gegebenen objektiven Welt, sondern Projekte von alternativen Welten. Aus der unterwürfigen subjektiven Stellung haben wir uns ins Projizieren aufgerichtet. Wir werden erwachsen. (Flusser 1997, 213)

Computer geben uns in der Interpretation Flussers also nicht nur, wie bisher großteils im Umgang mit anderen Medien, die Möglichkeit, passiver Teil einer statischen Welt zu sein, sondern diese, wenngleich oft virtuelle Welt, zu beeinflussen, zu gestalten und darüber hinaus eigene Welten zu erschaffen, mit denen bzw. innerhalb derer beispielsweise wiederum andere Menschen interagieren können. Während etwa Spiele, auf verschiedensten Plattformen wie Computer, diversen Konsolen oder vielen anderen Geräten, als mittlerweile sehr großer Teil der Unterhaltungsbranche sicherlich zumindest teilweise als Beleg für diese Sichtweise zulässig sind, nimmt Flusser damit auch in gewisser Weise die eingangs in Kapitel 2 formulierte Beobachtung, dass Wissen sich selbst thematisiert und kritisiert, vorweg bzw. unterstreicht diese.

Als Beispiel für die intersubjektive Verwirklichung nach Flusser könnten wohl durchaus soziale Netzwerke wie beispielsweise Facebook oder Google Plus genommen werden, da dort, vermutlich noch mehr als sonst in multimedialen Anwendungen und Diensten im Internet üblich, vor allem das Erschaffen des digitalen Selbst nicht nur dem Zweck der Kommunikation dient, sondern auch als Teil des Prozesses der Selbstfindung und Persönlichkeitsbildung in einer, nun immer mehr virtuellen Gesellschaft angesehen werden könnte. Auch das Hinterlassen digitaler Spuren, die etwa beim Googeln des eigenen Namens zumindest teilweise zu Tage befördert werden können, werden immer mehr Teil dieses Prozesses, da sie nicht nur dokumentieren und erinnern, sondern in Zukunft auch als kaum zu beherrschender Zeuge unseres virtuellen Daseins Einblick in die abgebildete Persönlichkeit und Teile der Vergangenheit eines Menschen geben werden.

Generell sieht Flusser dieser von ihm als „telematische Gesellschaft“ bezeichneten Veränderung, sowohl auf persönlicher, als auch gesellschaftlicher Ebene, mithilfe von Technik und somit der Informationsgesellschaft positiv und optimistisch entgegen. Mit dieser weltweiten Vernetzung wird nicht nur das gegenseitige Näherücken, sondern auch das aufeinander Zugehen ermöglicht bzw. stark vereinfacht, ja fast automatisiert. Voraussetzung dafür sind Apparate in Form von beispielsweise Computern, welche diese durch zu- und miteinander geprägte Informationsgesellschaft ermöglichen. (vgl. 1997, 145f)

Der altruistische Gedanke Flussers erfordere ein Umdenken der Informationsgesellschaft, da, je mehr wir miteinander vernetzt sind und dieses Netzwerk von Kommunikation uns definiert, wir umso mehr eine Gesellschaftsform erstrebenswert finden müssen, in der jeder Mensch seine Verwirklichung im Informationsaustausch findet. (vgl. 1997, 144)

Flussers Vision der telematischen Informationsgesellschaft gleicht also einer möglichen Utopie, die hauptsächlich durch Apparate und deren Vernetzung die Menschen nach einer Zeit der großen Konflikte im frühen 20. Jahrhundert und der eher durch einseitige Kommunikationsmöglichkeiten geprägten Medien Radio und Fernsehen wieder zueinander finden lassen soll.

2.3.c Weitere Befunde zur Informationsgesellschaft

Auch Frank Hartmann spricht von einer zu Beginn unscheinbaren, aber doch tiefgreifenden kulturellen Veränderung durch die Digitalisierung von Kommunikation. Durch die digitalisierte Information bestimmen Medien einerseits stärker den Alltag, da sie auch immer mehr in die Privatsphäre vordringen, andererseits wird so eine globalisierte Kommunikation vorangetrieben. Doch es sind nicht die einzelnen Computer, die uns von einem neuen Medium sprechen lassen, sondern erst die in den letzten beiden Jahrzehnten des 20. Jahrhunderts entstandenen Netzwerke von Computern, die den Menschen Zugang zu weit Entferntem, nach Flusser, und dadurch einen vergrößerten Rahmen für Kommunikation ermöglichten. Eine Eigenschaft, die vorangegangene Kulturtechniken in diesem Umfang wohl nicht besaßen. (vgl. 1999, 69)

Hartmann formuliert die Tragweite dieser technologischen Entwicklungen und deren Auswirkungen auf uns Menschen auch etwas konkreter bezogen auf die Denkweise:

Die philosophische Dimension dieser Entwicklung liegt auf der Hand: Mit einer Transformation, die Information und Kommunikation in ein neues Verhältnis zueinander stellt, verändert sich das gesellschaftliche Wissensgefüge und damit unsere Art zu denken ebenso, wie die mediale Entwicklung zugleich Ausdruck solcher Veränderungen ist. (Hartmann 1999, 70)

Hinsichtlich der nach Peter Bienert nötigen Anpassung unserer Erkenntnismodelle, wie eingangs in Kapitel 2 erörtert, kann auch bei Michel Foucault, einem französischen Philosophen, Historiker, Psychologen und Soziologen, die Notwendigkeit für eine Kultur herausgelesen werden, von Zeit zu Zeit die althergebrachter Art und Weise zu Denken umzuwerfen, um anders denken zu können. (vgl. 1974, 83)

Marshall McLuhan wiederum bezeichnet dieses Umdenken auch als eine neue Art der Teilnahme an und der Zugehörigkeit zu einer Gesellschaft. Menschen seien nun nomadische Sammler von Information und Wissen, ausgestattet mit Information und Wissen wie nie zuvor und nicht mehr dazu genötigt, Spezialisten zu sein. Auch sieht McLuhan die Menschen in den totalen sozialen Prozess wie nie zuvor involviert, da wir mit der Elektrizität und später wohl mit den Computernetzwerken unser zentrales Nervensystem global ausdehnen würden, um uns in einem Augenblick mit allen Menschen und deren Erfahrungen verbinden zu können. Sowohl Flusser als auch McLuhan sehen diese Entwicklung der Vernetztheit also eher positiv, um nicht zu sagen utopisch. (vgl. 1964, 311)

Auch Nicholas Negroponte, Gründer des *Media Lab* am *Massachusetts Institute of Technology*, geht davon aus, dass der wahre Wert eines Netzwerkes nicht so sehr in der Zurverfügungstellung von Information, sondern vor allem im Schaffen eines neuen sozialen Gefüges bestehe und die dadurch mögliche verstärkte Kommunikation. (vgl. 1996, 183)

Dieses anders Denken, die neue Art der Informationsbeschaffung und Wissensproduktion, ist wie in Kapitel 2.1 beschrieben, einer der Aspekte, der sich durch das Aufkommen des Internet und dem damit verbundenen Informationsüberfluss verändert hat. Wenn Foucault in Anlehnung an Rene Descartes davon ausgeht, dass „überhaupt jede Erkenntnis [...] durch die Vergleichung zweier oder mehrerer Dinge miteinander erworben wird“, nimmt dieser die Notwendigkeit der Filterfunktion, die unser Hirn im täglichen Online-Alltag nunmehr inne haben muss und die nun auch multimediale Dienste wie beispielsweise Google versuchen zu optimieren, vorweg. (Descartes 1906 zit. n. Foucault 1974, 85)

Foucault beschreibt dieses Filtern als einen Akt der Intelligenz, der es uns ermöglicht, durch den Vergleich verschiedener Sachverhalte und eine im Anschluss daran stattfindende Deduktion Offensichtliches erkennen und

verbinden zu können. Wenn Foucault von einer Deduktion von Offensichtlichem durch Intelligenz ausgeht, nimmt er damit in gewisser Weise die Musterbildung und -erkennung vorweg, durch die sich Intelligenz nach Jeff Hawkins, dem amerikanischen Erfinder des *Palm Pilot* und Gründer eines Instituts für theoretische Neurowissenschaften, in dessen Buch „On Intelligence“ unter anderem definiert und die in Kapitel 2.9.b noch genauer erörtert werden soll. (vgl. Foucault 1974, 83)

Klaus Mainzer, ein deutscher Philosoph und Professor für Wissenschaftstheorie, sieht für uns Menschen in naher Zukunft die Notwendigkeit, aus der Technologie des Internet und dessen vernetzter Natur Kapital zu schlagen.

Im Fall des Internets entsteht verteilte Künstliche Intelligenz durch Kooperation mit im Netz verteilten Informationssystemen. Die Interaktion mit dem Internet zeigt ferner, dass die alte Schnittstellendiskussion Mensch-Maschine überholt ist. Menschen stehen als Nutzer nicht mehr einzelnen Maschinen gegenüber [...] Vielmehr agieren sie in Infrastrukturen des Internets mit verteilten Dienstleistungssystemen. Situations- und personenbezogene Interaktionsmodelle rücken daher in das Zentrum der Computerphilosophie. (Mainzer 2003, 164)

Dieser gedankliche Ansatz stimmt auf die in Kapitel 3 stattfindende Diskussion um Google ein, da Suchmaschinen immer mehr diese Funktion der Schnittstelle einnehmen können. Daraus ergibt sich die Notwendigkeit einer Filterfunktion für Wissen und Information, was dazu führt, dass Suchmaschinen wie etwa Google in Vergangenheit und Gegenwart versuchen, immer intelligenter zu werden, um uns schrittweise als intelligente menschliche Filter zu entlasten. In Anlehnung an Mainzer gibt Sven Türpe, ein deutscher Informatiker und Mitgründer des Labors für Sicherheitsanalyse von Software am Fraunhofer Institut, ebenfalls einen Ausblick darauf, was im Laufe dieses und der nächsten Kapitel erörtert und erweitert werden soll.

Im Grunde genommen tut Google aber nichts anderes als unser Gehirn, nur ohne den Filter unserer Sinnesorgane, ohne Abgleich mit Lehrbuchwissen und mit viel mehr Daten und Aspekten dieser Daten als uns normalerweise bewusst werden. Google lernt Sprachen - oder Expertenintuition - wie wir, nur schneller und ohne den Umweg über

Übungen und explizite Regeln direkt aus Beispielen und Feedback. Und wir sind die Lehrer. (Türpe 2012⁶)

Dem Begriff der Intelligenz ist vor allem das Kapitel 2.9.a gewidmet, da, um die Idee einer digitalen Intelligenz untersuchen zu können, erst eine Hinführung zu diesen Begriffen erfolgen soll, um Bedeutung und Zusammenhang erkennbar zu machen. Während in diesem Kapitel nun die theoretischen Erklärungsmodelle verschiedener Denkschulen dargelegt wurden, so soll im Folgenden ein kurzer Einblick in praktische Ansätze zur Bewältigung der beiden Probleme der Menge und Komplexität gegeben werden.

2.4 Praktische Problemlösungsansätze

Wie können die Menge an Information und deren komplexe Verweise und Zusammenhänge im Internet bewältigt werden? Durch das Öffnen von Plattformen, um über Mitwirkung und Hilfe einer hohen Anzahl an Nutzerinnen und Nutzern deren Potenzial, Fähigkeiten und Kenntnisse auf verschiedensten Gebieten einsetzen zu können. Um also einer Menge an Information und dem hohen Komplexitätsgrad des zu erstellenden Inhaltes beikommen zu können, setzt Wikipedia als freie Wissensplattform und Enzyklopädie beispielsweise auf eine möglichst hohe Anzahl von Autorinnen und Autoren.

Spannend in diesem Zusammenhang ist die Hinwendung vieler Systeme, Organisationen oder Unternehmen aus unterschiedlichsten Sparten und Interessengebieten zum so genannten *Crowdsourcing*, also dem eben beschriebenen zu Nutze machen der Menge. Während dieser Begriff besonders im Bezug auf Google noch etwas später essenziell sein wird, so kann etwa am Beispiel der Wissensplattform Wikipedia das Prinzip des *Open Source*, einem mit dem *Crowdsourcing* verwandten, auf nicht-kommerzielle Nutzung ausgelegten Begriff, klar identifiziert werden.

„Wikipedia is written collaboratively by largely anonymous Internet volunteers who write without pay. Anyone with Internet access can write and make changes to Wikipedia articles, except in limited cases where

⁶ <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241-5.html>, aufgerufen am 06.08.2012.

editing is restricted to prevent disruption or vandalism. Users can contribute anonymously, under a pseudonym, or, if they choose to, with their real identity.“ (Wikipedia 2012⁷)

Open Source zeichnet dabei aus, dass die Quellen des Schaffens offengelegt werden und zugänglich sein müssen. Wikipedia stellt hierbei also beispielsweise eine Plattform für Wissen zur Verfügung und vertraut im Weiteren darauf, dass ein Artikel von jeder Nutzerin und jedem Nutzer selbst geschrieben, darüber hinaus auch jeder andere Artikel von allen Nutzerinnen und Nutzern zu jeder Zeit bearbeitet und somit das Wissen der Menge angezapft werden kann. Die Quellen des Wissens sind also offen in der Hinsicht, dass alle Nutzerinnen und Nutzer selbst der Ursprung des Wissens sind, welches auf der Plattform verfügbar ist.

Im Fall von Wikipedia muss die Motivation, wie bei anderen *Open Source* Projekten, intrinsisch vorhanden und vorherrschend sein, also der Wille und Drang, einen Artikel zu schreiben oder zu ändern, da keine extrinsische Motivation, wie etwa eine Entlohnung für jeden geschriebenen oder geänderten Artikel vorhanden ist und sein wird. Dieses Prinzip des Wandels von „Eine/Einer für Viele“ zu „Alle für Alle“ ist aber nicht nur wie im Fall Wikipedia bei unkommerziellen Systemen und Organisationen, wie eben auch vielen *Open-Source* Projekten, mittlerweile weit verbreitet, auch viele profitorientierte Unternehmen versuchen seit jeher, die Menge, „crowd“, bzw. deren Zeit, Kraft und Können zu mobilisieren und einen Nutzen daraus zu ziehen.

In Falle einer kommerziellen Nutzung, wenn also nicht nur intrinsische, sondern auch extrinsische Motivation in Form einer Entlohnung, beispielsweise finanziell, eine Rolle spielt, wird aus *Open Source* das bereits erwähnte *Crowdsourcing*, das in Hinblick auf die Betrachtung von Google noch relevant sein soll.

Crowdsourcing ist eine interaktive Form der Leistungserbringung, die kollaborativ oder wettbewerbsorientiert organisiert ist und eine große Anzahl extrinsisch oder intrinsisch motivierter Akteure unterschiedlichen Wissensstands unter Verwendung moderner IuK-Systeme auf Basis von Web 2.0 einbezieht. Leistungsobjekt sind Produkte oder Dienstleistungen

7 <http://en.wikipedia.org/wiki/Wikipedia:About>, aufgerufen am 06.08.2012.

unterschiedlichen Innovationsgrades, welche durch das Netzwerk der Partizipierenden reaktiv aufgrund externer Anstöße oder proaktiv durch selbsttätiges Identifizieren von Bedarfslücken bzw. Opportunitäten entwickelt werden. (Martin/Lessmann 2007⁸)

Obwohl Wikipedia gerade hinsichtlich Open Source einen derzeitigen Endpunkt markiert, ist zur Form des verwalteten Wissens zu bemerken, dass sich Wikipedia hauptsächlich auf verschriftlichtes Wissen beschränkt, hier daher bei weitem noch nicht alle Möglichkeiten der digitalen Informationsspeicherung ausgeschöpft werden. Die Intertextualität wiederum zeugt von moderner Herangehensweise an die Problemstellung.

Inhaltliche Gegenpole hierzu bilden beispielsweise Flickr und YouTube, da diese jeweils die Plattform für eine weitere Medienform, nämlich Bild und Video, darstellen. Die auf Flickr und YouTube verfügbaren, von Nutzerinnen und Nutzern generierten Daten, Videos und Photos gehen dabei in die Milliarden, wobei hier das Prinzip des Andere-arbeiten-Lassens praktiziert wird. Sowohl YouTube als auch Flickr bieten lediglich eine Plattform an, auf die eigene Daten, beispielsweise Videos oder Fotos, hochgeladen und von anderen Nutzerinnen und Nutzern angesehen werden können. Während bei Flickr beispielsweise sowohl das Hochladen, als auch das Ansehen von Bildern gratis ist, können Nutzerinnen und Nutzer für ihre Bilder eine Lizenzierung aktivieren, diese also für eine mögliche kommerzielle Nutzung vormerken, etwa durch die Partnerseite GettyImages, oder aber etwaige sonstige Interessenten. (vgl. Flickr 2012⁹)

Es können also zwei Arten von relevanten Diensten, die auf dem Prinzip des Crowdsourcing beruhen, unterschieden werden. Einerseits solche, die lediglich eine Plattform bieten, auf die möglichst viele Nutzerinnen und Nutzer aufmerksam gemacht werden sollen, um in weiterer Folge mehr Einnahmen aus Werbeschaltungen lukrieren zu können. Dies wird oft mittels der Möglichkeit für Nutzerinnen und Nutzer, eigene Inhalte zu kreieren oder diese präsentieren zu können, versucht zu erreichen. Andererseits gibt es Dienste, die tatsächlich auf dem Prinzip des Crowdsourcing basieren und die Menge an Personen zu unterschiedlichen Zwecken nutzt.

8 http://ibis.in.tum.de/mkwi08/18_Kooperationssysteme/05_Martin.pdf, aufgerufen am 06.08.2012.

9 <http://www.flickr.com/help/gettyimages/>, aufgerufen am 06.08.2012.

Google Image Labeler - respektive dessen Vorgänger, das ESP-Game erfunden von Luis von Ahn - legen dieses Prinzip des *Crowdsourcing* unter dem Titel *Human Computing* auf Prozesse und Arbeitsschritte um, die für Computer allein momentan noch nicht ausreichend oder zufriedenstellend lösbar sind, für Menschen aber keinerlei Hindernis bzw. kaum Aufwand bedeuten, allerdings in der Masse ebenfalls nicht lösbar wären. Von Ahn formuliert dabei eine wichtige Facette von *Crowdsourcing*. „I build systems that combine humans and computers to solve large-scale problems that neither can solve alone. I call this Human Computation, but others sometimes call it Crowdsourcing.“ (2006¹⁰)

Konkret werden also Aufgaben von Computern auf Menschen ausgelagert, da so eine effizientere Bewerkstelligung dieser Aufgaben möglich ist. Anders gesagt ermöglicht beispielsweise das ESP-Game einem Computer das Lernen von menschlicher Interaktion und stellt somit eine weitere Informationsquelle neben vielen anderen für Google dar. Im Falle des *Google Image Labeler* respektive dessen von Google erstandenen Vorgängers, dem ESP-Game, ist das Erkennen und Beschreiben von Bildinhalten, da für Computer bzw. Programme schwer möglich, in ein Spiel für Nutzerinnen und Nutzer verpackt. (vgl. Von Ahn 2006¹¹)

Das Spiel funktioniert nach einem einfachen Prinzip, zwei unbekannte und anonyme Nutzerinnen oder Nutzer werden einander zufällig, ohne jegliche Möglichkeit der Kommunikation, zugewiesen. Im Verlauf des Spiels werden beiden Bilder gezeigt, zu denen sie Schlagwörter, Tags, eingeben können, die eingegebenen Begriffe des virtuellen Partners können dabei allerdings erst nach Ablauf einer Runde eingesehen werden. Eine Runde endet jeweils dann, wenn von den eingegebenen Begriffen und Schlagwörtern der beiden Spielerinnen oder Spieler zumindest eine Übereinstimmung gegeben ist. Während jedes abgeschlossene Bild eine gewisse Punktzahl auf das eigene Konto bringt, so wurde ein Highscore-System als Motivator verwendet. (vgl. Schwartz 2012¹²)

Die von Google intendierte Funktion ist dabei, wie bereits erwähnt, das Kreieren von Metadaten. Diese aus dem *Crowdsourcing* gewonnenen

10 <http://www.cs.cmu.edu/~biglou/>, aufgerufen am 06.08.2012.

11 Ebd.

12 <http://www.seroundtable.com/google-image-labeler-dead-14663.html>, aufgerufen am 06.08.2012.

Metadaten können wiederum für verschiedene Zwecke nützlich sein und eingesetzt werden. Einerseits beispielsweise als Unterstützung und Präzisierung der Suche, im Fall des *Google Image Labeler* nach Bildern, oder andererseits auch um Websites mit eingearbeiteten Bildern so gestalten und anzeigen zu können, dass diese etwa von Personen mit visueller Beeinträchtigung, die sich Texte von Websites per Programm vorlesen lassen, ebenfalls sinnvoll wahrgenommen werden können. Google macht sich also, ganz nach dem Prinzip des *Crowdsourcing* - respektive *Human Computing* wie von Luis von Ahn beschrieben - Wissen und Fähigkeiten von Vielen zunutze, um den beiden Problemen der Menge und der Komplexität beikommen zu können.

Als der relevanteste praktische Lösungsansatz kann daher die Kombination von Computersystemen und Menschen im Prozess der Lösung komplexer Aufgaben oder aber der Bewältigung großer Mengen an Aufgaben festgestellt werden. Dieser Aspekt soll in Kapitel 3.4 noch unter anderen Gesichtspunkten untersucht werden. Teil von Problem und Lösung zugleich ist hierbei also das Internet, das sowohl den Überfluss an Information und Daten ermöglicht, gleichzeitig aber auch die Infrastruktur für das Bekämpfen des Problems darstellt.

2.5 Internet als wesentliche Ausprägung der Informationsgesellschaft

Das Internet ist nicht nur ein weltweites Netzwerk, genauer ist es ein Netzwerk von Netzwerken. Wie beispielsweise heute ein lokales Netzwerk, sprich Local Area Network oder auch LAN, über das Internet mit vielen anderen lokalen Netzwerken, etwa in der Nachbarschaft, Universitäten oder an einer Vielzahl anderer Standorte, verbunden ist, so liegen auch die Anfänge des Internet im Verbinden einzelner Forschungseinrichtungen und deren Computer-Netzwerke.

Dabei zu unterscheiden sind heute zumindest zwei Arten des Internet, einerseits das *surface web* und andererseits das darunter liegende *deep web*. Während das *surface web* für alle Nutzerinnen und Nutzer zugänglich ist und an der Oberfläche, also beispielsweise ohne Notwendigkeit eines Accounts, liegt, so ist das *deep web* für Suchmaschinen wie Google nicht indizierbar und gilt somit als unsichtbar oder versteckt, könnte aber

trotzdem kostenfrei erreicht werden. Die Gründe für diese mangelnde Indexfähigkeit sind vielfältig, so können etwa dynamische Inhalte, private Websites, fehlende Links von und zu anderen Websites oder ganz einfach von Suchmaschinen nicht unterstützte Daten-Formate ausschlaggebend sein. (vgl. Bergman 2001, 47)

Zwar fehlen aktuelle Angaben, aber bereits 2001 umfasste das *deep web* mehr als sieben Petabyte, verglichen mit 19 Terabyte an Daten und Informationen, die im *surface web* verfügbar waren. Generell wurde bereits damals die Menge an Information, die öffentlich im *deep web* zugänglich war, auf etwa 400 bis 550 mal größer als die des *surface web* und die Reichweite diverser Suchmaschinen wie Google auf etwa 0,03 Prozent des gesamten verfügbaren Inhalts des Internet eingeschätzt. Diese Zahlen haben sich in den letzten zehn Jahren sicherlich drastisch verändert, geben aber zumindest einen vagen Ausgangspunkt vor. (vgl. Bergman 2001, 45)

Umfang, Größe und die rasante Entwicklung des Internet kommen jedoch noch deutlicher zum Vorschein, wenn sie in Relation zu den Anfängen dieser heutzutage weltweiten Vernetzung gesetzt werden. Die Grundidee etwa des ARPAnet, der Verknüpfung von vier Universitäten und deren Computer durch die Advanced Research Projects Agency, eine von der US-Regierung nach dem so genannten Sputnik-Schock gegründete Forschungsbehörde, in der frühen zweiten Hälfte des 20. Jahrhunderts, war es, Rechenkapazitäten der jeweiligen Computer in den Universitäten sowohl anderen zur Verfügung stellen, als auch selbst die eigenen Möglichkeiten erweitern zu können. Eine Grundintention also, die in gewisser Art und Weise nicht nur den Einzelnen und die Einzelne mit der Menge zu verknüpfen suchte, sondern darüber hinaus auch die Möglichkeiten zur Verfügung stellte, Probleme von einzelnen Individuen mit Hilfe einer großen Anzahl an anderen Individuen möglicherweise erfolgreich lösen zu können. (vgl. Hartmann 2008b, 82)

Diese Ideen waren für einige der federführenden Gründungsväter des ARPAnet, das als Vorgänger des Internet gilt, Motivation und Ziel, das diese mit ihrer Forschung verfolgten. Robert Taylor, der erst für die NASA im Bereich der Forschung gearbeitet hatte und 1966 als Direktor des vom Pentagon initiierten Advanced Research Projects Agency Information Processing Techniques Office eingesetzt wurde, entwickelte unter anderen die Idee des ARPAnet auf Basis von Forschung und Konzepten Joseph Carl

Robnett Lickliders, ebenfalls für die ARPA tätig. Taylor, in dessen Büro sich zu dieser Zeit drei Computer-Terminals befanden, die jeweils eine Verbindung zur Systems Development Corporation in Santa Monica, der Universität von Berkeley und dem Massachusetts Institute of Technology darstellten. Die Idee, wie Taylor sie beschreibt, war so simpel wie revolutionär.

We had in my office three terminals to three different programs that ARPA was supporting. One was to the Systems Development Corporation in Santa Monica. There was another terminal to the Genie Project at U.C. Berkeley. The third terminal was to the C.T.S.S. project that later became the Multics project at M.I.T. [...]

For each of these three terminals, I had three different sets of user commands. So if I was talking online with someone at S.D.C. and I wanted to talk to someone I knew at Berkeley or M.I.T. about this, I had to get up from the S.D.C. terminal, go over and log into the other terminal and get in touch with them.

I said, oh, man, it's obvious what to do: If you have these three terminals, there ought to be one terminal that goes anywhere you want to go where you have interactive computing. That idea is the ARPAnet. (Taylor 1999¹³)

Diese Vernetzung von Ressourcen, Systemen und Individuen hatte, obwohl beabsichtigt, Folgen, die nicht abzuschätzen waren. Taylor beschreibt etwa, dass sich innerhalb der ersten Monate des ARPAnet bereits Gruppen und Gemeinschaften rund um die vernetzten Institute gebildet hatten, die gemeinsame Interessen ausloteten und beispielsweise Daten austauschten oder per E-Mail miteinander verkehrten, obwohl gegenseitig völlig unbekannt. Taylor war von dieser völlig neuen Art von Kommunikation und Beziehung begeistert, während Joseph Carl Robnett Licklider, dessen Serie von Memos und Schriften aus dem Jahre 1962 rund um das Konzept eines „galactic network“ den Grundstein für die Idee Taylors gelegt hatten, nicht nur die Kommunikation der Menschen untereinander, sondern speziell auch zwischen Mensch und Maschine als Ziel verfolgte. (vgl. Leiner et al. 1997¹⁴)

13 <http://web.archive.org/web/20080922095019/http://partners.nytimes.com/library/tech/99/12/biztech/articles/122099outlook-bobb.html>, aufgerufen am 06.08.2012.

14 <http://www.internetsociety.org/internet/internet-51/history-internet/brief-history-internet>, aufgerufen am 06.08.2012.

Besonders in Lickliders Manuskript „Man-Computer Symbiosis“, welches dieser noch vor seiner Zeit bei ARPA verfasste, galt für Robert Taylor und viele andere als Stein des Anstoßes, in diese Richtung zu forschen. Darin gibt Licklider nicht nur einen Ausblick darauf, wie in Zukunft eine Symbiose zwischen Mensch und Maschine funktionieren könnte, sondern auch, warum die bisherigen Konzepte der künstlichen Intelligenz und des mechanisch erweiterten Menschen nicht in der momentanen Form funktionieren würden.

„The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today.“ (Licklider 1960¹⁵)

Dieser Ansatz ist sowohl dem in Kapitel 2.4 bereits erwähnten Human Computing nach Luis von Ahn, als auch der Ansicht von Sergey Brin, amerikanischer Informatiker und Google Mitgründer, sehr ähnlich, obwohl schon mehr als 40 Jahre früher ausgesprochen, doch dazu in Kapitel 3.4 mehr. Während das Internet also sowohl Teil der Probleme ist, die damit einhergehen, so sind Suchmaschinen wie etwa Google ein praktischer Ansatz, um diese beiden Probleme der Menge und der Komplexität zu lösen. Im nächsten Kapitel sollen daher Funktion und Bedeutung von Suchmaschinen kurz erläutert werden.

2.6 Suchmaschinen im Internet als Tool der Informationsgesellschaft

Der Wert von Information als Ressource wurde bereits in Kapitel 1.4 umrissen, dementsprechend offensichtlich ist die Existenzgrundlage für Suchmaschinen. Die Menge an Information und Daten, die momentan und in Zukunft im Internet quasi frei verfügbar sind und sein werden, ist für sich gesehen schon revolutionär. Darüber hinaus spannend ist allerdings die Kombination von dieser Menge an im Internet gespeichertem Wissen mit einer Möglichkeit, dieses Wissen gezielt zu durchforsten und als Information abzurufen. Suchmaschinen bieten daher einen idealen,

15 <http://groups.csail.mit.edu/medg/people/psz/Licklider.html>, aufgerufen am 06.08.2012.

zentralen Einstiegspunkt in dieses Wissen - respektive das Internet - , um von diesem Punkt aus in die große vernetzte Wissenswelt aufzubrechen.

Suchmaschinen wie Google können daher momentan als eine Art „Universalschnittstelle zwischen Mensch und Information“ angesehen werden, Google kann etwa nicht nur Text, sondern auch Bilder, Videos, Dokumente verschiedenster Art, Weltkarten und vieles mehr nicht nur suchen, sondern vor allem finden. (Krug et al. 2005, 20) Für Suchmaschinen ist damit das Wissen fragmentarisch, Google beispielsweise interessiert, zumindest im Moment, noch nicht, in welchem Absatz oder auf welcher Seite, noch in welchem Kontext die gewünschte Information zu finden ist. Die Folge ist eine Patchwork-Information, die einer ständigen Vorläufigkeit unterworfen ist.

Anders betrachtet bedeutet das wiederum, dass Suchmaschinen einen Teil unserer Realität konstruieren. Was beispielsweise von Google nicht gefunden werden kann, existiert de facto für viele Menschen nicht. Bei der Suche nach einem Begriff werden vielleicht ein oder zwei Suchmaschinen, vielleicht sogar noch ein paar mehr in Ausnahmefällen, zu Rate gezogen. Sollten diese aber geschlossen kein Ergebnis auf die Anfrage liefern können oder wollen, so wäre wohl, möglicherweise durch andere Medien zwar größtenteils noch überwindbarer, aber doch merklich ein blinder Fleck erkennbar. Was nicht gesucht - respektive gefunden - wird, existiert also zumindest für Nutzerinnen und Nutzer von Google nicht. „Dies macht Suchmaschinen zu Realitäts- und Machtmaschinen: das Internet, das Wissen und die Realität werden von ihnen kolonisiert und strukturiert.“ (Krug et al. 2005, 21)

Es ist daher legitim, von einer medialen Erschaffung von Realität respektive Wirklichkeit zu sprechen. Dieses Phänomen kann zwar schon zeitlich vor dem Internet an anderen Massenmedien identifiziert werden, erreicht nun aber durch die noch größere Ansammlung und in weiterer Folge Filterung von Information eine neue Dimension, vor allem auch, da Google beispielsweise für Nutzerinnen und Nutzer bei einer Suche als nur im Hintergrund agierend wahrgenommen wird. Dabei ist der Einfluss von Medien oder Suchmaschinen beispielsweise „umso bedeutsamer, je größer die Bereiche des menschlichen Wissens sind, die von ihnen erfasst werden, und je mehr diese Medien genutzt werden.“ (ebd., 21f)

Spannend ist die Erschaffung der virtuellen Wirklichkeit in Anbetracht der Tatsache, dass Google beispielsweise, wie auch jede andere Suchmaschine, nicht mit Information an sich, sondern Meta-Information, also Information über Information, arbeitet und diese an Nutzerinnen und Nutzer weitergibt. Mit Hilfe eines zuvor erstellten Index, in dem die verfügbaren Informationen aus dem Internet abgespeichert und mit Meta-Daten versehen werden, kann Google später für Nutzerinnen und Nutzer Suchergebnisse ermitteln. Das bedeutet, dass, um eine Auswahl zu treffen, Informationen manipuliert werden, zwar nicht im Einzelnen, jedoch statistisch gesehen. (vgl. Krug et al. 2005, 28)

Womit Suchmaschinen generell und damit auch Google aber eigentlich handeln, ist die Aufmerksamkeit von Nutzerinnen und Nutzern bzw. „der Aufmerksamkeitsvorteil, der entsteht, wenn Informations- und andere Angebote selektiv zu Gunsten bestimmter Anbieter zur Verfügung gestellt werden.“ (ebd., 29)

Damit dieser Aufmerksamkeitsvorteil optimal erreicht werden kann, ist eben diese Manipulation von Information und das damit verbundene Erstellen eines Index von Nöten, um gebündelt und selektiert genau das bieten zu können, was Nutzerinnen und Nutzer suchen beziehungsweise in weiterer Folge finden. Daraus ergibt sich, dass Suchmaschinen wie Google nicht mit, sondern mittels Information, handeln, sich dieser also lediglich bedienen, um beispielsweise Werbungen platzieren oder Inserate schalten zu können, was den eigentlichen wirtschaftlichen Mechanismus darstellt. Es darf hier nicht vergessen werden, dass die am häufigsten benutzten Suchmaschinen, allen voran Google, aus ökonomischen Gründen und nicht etwa vor dem Hintergrund altruistischer Motive entwickelt wurden.

Trotzdem kann hinsichtlich Information als Ressource ein Wandel festgestellt werden. Während noch vor und während des 20. Jahrhunderts großteils für Information gezahlt wurde, da diese ein exklusives Gut darstellte, ist ein großer Teil der im Internet verfügbaren Information heutzutage, sobald einmal veröffentlicht, durch die stete Spiegelung und somit Wiederveröffentlichung quasi nicht mehr exklusiv zu halten. Der Tausch und Handel von Information gegen Geld wurde und wird also zu Gunsten eines Handels von Meta-Informationen gegen Geld ausgetauscht, mit dem Unterschied, dass Google als Suchmaschine nicht Nutzerinnen

und Nutzer selbst zur Kasse bittet, sondern stattdessen mit Werbeeinnahmen dieses Geld generiert und damit den Handel vorantreibt.

Damit Google jedoch bei einer Suche die richtigen - respektive am ehesten relevanten - Inhalte von über das Internet verstreuten Websites findet, muss eine Filterung und in weiterer Folge eine Vorhersage in Form einer Liste mit vorgeschlagenen Ergebnissen stattfinden. Nach Foucault ist dieses Filtern und die Deduktion von Offensichtlichem, also in diesem Zusammenhang das Erarbeiten einer Ergebnisliste, ein Akt von Intelligenz. Bevor nun aber die Intelligenz von Google an Hand dieser Prozesse zur Diskussion stehen kann, sind die Begriffe der Intelligenz und der künstlichen Intelligenz im nachfolgenden Kapitel kurz zu umreißen und zu diskutieren.

2.7 Intelligenz und künstliche Intelligenz

Die Intelligenz eines Menschen wird nach dem amerikanischen Psychologen Philip Zimbardo in der Weise definiert, als damit die allgemeine Fähigkeit des Individuums, die Welt, in der es lebt, zu verstehen und sich in ihr zurecht zu finden, bezeichnet. Daraus abgeleitet setzt sich Intelligenz aus der Anpassung an neue Situationen und sich verändernde Anforderungen, dem Lernen oder der optimalen Nutzung von Erfahrung oder Übung und abstraktem Denken, sowie dem Gebrauch von Symbolen und Begriffen, zusammen. (vgl. 1988, 543)

Während diese allgemein anerkannte Herangehensweise an den Begriff Intelligenz aus dem Bereich der Psychologie stammt, vereint das Forschungsfeld der künstlichen Intelligenz in seiner Geschichte seit etwa 1950 Beiträge aus originär anderen Fach- und Forschungsbereichen, etwa der Kybernetik, Biologie, Kommunikationswissenschaften und natürlich auch der Mathematik. Ob Norbert Wiener, John McCarthy, Claude Shannon, Walther Pitt, John von Neumann oder Noam Chomsky, künstliche Intelligenz in Form eines Computers verbindet seit jeher Ideen, die interdisziplinär entstehen oder fortgeführt werden.

AI in its formative years was influenced by ideas from many disciplines. These came from people working in engineering (such as Norbert Wiener's work on cybernetics, which includes feedback and control), biology (for

example, W. Ross Ashby and Warren McCulloch and Walter Pitts's work on neural networks in simple organisms), experimental psychology (see Newell and Simon [1972]), communication theory (for example, Claude Shannon's theoretical work), game theory (notably by John Von Neumann and Oskar Morgenstern), mathematics and statistics (for example, Irving J. Good), logic and philosophy (for example, Alan Turing, Alonzo Church, and Carl Hempel), and linguistics (such as Noam Chomsky's work on grammar). (Buchanan 2006, 56)

Während die Technologie der Computer seit etwa 1940 vorhanden war, so wurde erst etwa zehn Jahre später der Gedanke der Intelligenz einer Maschine bzw. eines Computers konkret verfolgt. Norbert Wiener war einer der ersten Wissenschaftler, die das Prinzip des Feedbacks und der damit verbundenen Feedback-Theorie in Bezug auf Computer - respektive Maschinen - beobachteten und erforschten.

Als bekanntes Beispiel kann ein Thermostat fungieren. Dieser kontrolliert die Temperatur der Umgebung, etwa per Messung in einem Haus, und vergleicht die gesammelten Daten mit der gewünschten Temperatur. Als Ergebnis wird die Heizung schlussendlich entweder an oder abgestellt. Die spannende und für den weiteren Verlauf einer Geschichte der künstlichen Intelligenz wichtige Erkenntnis Wieners dabei war, dass er die Theorie aufstellte, alles intelligente Verhalten wäre das Ergebnis von Feedback-Mechanismen, die möglicherweise von Maschinen oder Computern simuliert werden könnten. Auch Jeff Hawkins wird später diese wichtige Beobachtung bzw. Annahme im Zusammenhang mit Intelligenz wieder aufgreifen. (vgl. Dyess et al. 1996¹⁶)

Generell können nach Reiner Ruffing, einem deutschen Philosophen, im Forschungsfeld der künstlichen Intelligenz der symbolische und der konnektionistische Ansatz unterschieden werden. Während die symbolische künstliche Intelligenz von einem möglichst umfassenden Regelwerk für Maschinen ausgeht, mit dessen Hilfe in weiterer Folge Antworten auf Fragen in spezifischen Fachbereichen erreicht werden können, so sollen im Rahmen einer konnektionistischen künstlichen Intelligenz vor allem die Lernvorgänge des Gehirns simuliert und erforscht werden. Zu diesem Zweck wird bei den konnektionistischen Ansätzen auch stärker auf die Umwelt bzw. Feedback und ständig neue Daten

16 <http://library.thinkquest.org/2705/>, aufgerufen am 06.08.2012.

eingegangen, also ein Faktor, der schon für Norbert Wiener nicht zu vernachlässigen war. (vgl. Ruffing 2005, 268)

Die beiden Ansätze stehen sich insoweit gegenüber, als sie von zwei entgegengesetzten Punkten ausgehen. Der symbolischen künstlichen Intelligenz stehen von Beginn an alle Daten und Funktionen zur Verfügung, mit deren Hilfe ein Lösungsweg gefunden werden soll. Es wird also von oben nach unten operiert. Die konnektionistische künstliche Intelligenz hingegen baut auf dem Gedanken des Lernens auf, es werden etwa Lernfähigkeit, Situationserkennung und Fehlertoleranz propagiert und somit das System von unten nach oben, also von einer Anpassungsfähigkeit ausgehend zu einem erfolgreichen Lösungsansatz eines Problems als Ziel, gestaltet. Der klassische Ansatz der künstlichen Intelligenz als von Regeln bestimmtes Handeln steht hier der Idee von lernfähigen, neuronalen Netzwerken gegenüber, die ihre Inspiration klar von der Architektur des Gehirns nehmen. (vgl. Ruffing 2005, 269)

Ausgehend von dieser Definition ergeben sich, nach Ruffing, die jeweiligen Einsatzgebiete und Stärken dieser Herangehensweisen. „Konnektionistische Systeme haben gegenüber herkömmlichen Modellen v. a. in Bereichen der Mustererkennung besondere Vorteile, während digitale Systeme besser zur Erforschung argumentativer Prozesse geeignet sind.“ (2005, 268)

Dieser Fokus auf Muster und Mustererkennung soll speziell im Bezug auf neurale Netzwerke, Intelligenz allgemein, als auch auf Google später noch relevant sein.

2.8 Von Datenverwaltung zu künstlicher Intelligenz

Die Geschichte von perfekt auf eine Funktionalität ausgerichtete Computern, im Sinne der künstlichen Intelligenz also von einem symbolischen Ansatz ausgehend, ist, entgegen der von funktionierenden Betriebssystemen, verhältnismäßig lang. Bereits relativ früh wurde beispielsweise versucht, das Regelwerk von Schach auf Computer zu übertragen und so einen vermeintlich intelligenten Schachcomputer zu erschaffen. Als einer der Vorreiter künstlicher Intelligenz im allgemeinen widmete sich unter anderen der Mathematiker Alan Turing schon ab etwa 1940 der Herausforderung, Regelwerk bzw. Spielzüge bei einem

Schachspiel in ein Programm, vorerst allerdings auf Papier, zu übersetzen. Dieses bestand wie auch viele heutzutage verwendete Programme „aus einer Reihe von Befehlen, die nacheinander abgearbeitet wurden, um zu entscheiden, welcher Antwortzug bei welcher Stellung der Schachfiguren ausgeführt werden sollte.“ (Standage 2002, 193)

Während Turing bei einem ersten Testspiel 1952 mit einem Kollegen so handelte und zog, wie das Programm mit seinen Befehlen es ihm auftrug und damit vorgab, eine Maschine zu sein, so wurde bereits wenige Jahre später, 1958, ein erstes tatsächliches Schachprogramm von dem Amerikaner Alex Bernstein erstmals auf einem IBM Computer mit einem menschlichen Gegner konfrontiert. Als denkwürdiges Ereignis, sowohl im Bereich der Schachcomputer, als auch der künstlichen Intelligenz, ging allerdings die Serie von Spielen Garry Kasparovs gegen den Computer Deep Blue in den Jahren 1996 und 1997 ein. Kasparov, langjähriger Schachweltmeister und oft bester Schachspieler aller Zeiten genannt, trat wie schon sechs Jahre zuvor gegen Deep Blue, einen von der Carnegie-Mellon-Universität über viele Jahre entwickelten, spezialisierten Schachcomputer, in mehreren Spielen an. Während Kasparov 1996 die Konfrontation für sich entscheiden konnte und in sechs Spielen nur eine Niederlage hinnehmen musste, so konnte er nur ein Jahr später einen Sieg von Deep Blue insgesamt nicht verhindern. (vgl. Standage 2002, 198ff)

Dieser Sieg eines Computers über einen Menschen hatte für viele eine historische Bedeutung, jedoch stecken, außer einem ausgeklügelten und über viele Jahre iterierten Programm, Prozessoren, bzw. im Falle von Deep Blue viele Prozessoren dahinter, die es dem Computer ermöglichen, 200 Millionen Stellungen pro Sekunde und somit sehr viele verschiedene Züge zu berechnen, etwas, das Kasparov wohl in die Niederlage trieb. Im Vergleich dazu wirken die etwa 3000 Berechnungen innerhalb mehrerer Minuten, die auch schon zu Alan Turings Lebzeiten möglich waren, unscheinbar, allerdings veranlassten sie Turing bereits damals zu einer Annahme, die nicht nur das Forschungsfeld der künstlichen Intelligenz, in abgewandelter Form, als Turing Test kennt. „Eine intelligente Maschine sollte also nach Turing zwei Dinge können: Schach spielen und ein normales Gespräch führen.“ (Standage 2002, 194)

Doch während die Kritik an der vermeintlichen Intelligenz von Computern, die Schach spielen, schon in deren frühen Jahren laut wurde, so gab es

auch im Team von Deep Blue und dessen Vorgänger Deep Thought klare Aussagen gegen einen intelligenten Schachcomputer. So fand beispielsweise Murray Campbell, kanadischer Informatiker und einer der Ingenieure, die an Deep Blue gearbeitet hatten, dass dieser ganz klar nur dazu gebaut worden war, um einen einzigen Zweck zu erfüllen, einen winzigen Teilbereich des Anforderungsspektrum an menschliche Intelligenz zu imitieren. „I never considered Deep Blue intelligent in any way. It's just an excellent problem solver in this very specific domain.“ (1997, 92)

Konkret bedeutet das also, dass, um die Intelligenz eines Computers zu beweisen, dieser einen Menschen in irgendeiner Art und Weise täuschen muss. Anders gesagt und auch schon von Alan Turing bemerkt, gibt es keinen objektiven Test für Intelligenz, der außerhalb der menschlichen Wahrnehmung liegt, ein Problem, das besonders Jeff Hawkins ausführlich untersucht hat und das in Kapitel 2.9.a genauer behandelt werden soll.

Auch John Rogers Searle, ein amerikanischer Philosoph, der sich bereits relativ früh mit der Thematik einer künstlichen Intelligenz auseinandersetzte, formulierte ein Gedankenexperiment mit dem Namen „The Chinese Room“. Searle geht darin von einem Raum aus, in dem eine beispielsweise englischsprachige Person sitzt. In dem Raum befinden sich außer der Person auch ein englischsprachiges Regelwerk für die Art und Weise, wie chinesische Schriftzeichen aneinandergereiht und angeordnet werden sollen. Wenn nun jemand beispielsweise ein Blatt Papier mit einem chinesischen Text durch einen Schlitz in der Tür schiebt, kann die Person in dem Raum an Hand des Regelwerks und dem bekommen Blatt Papier auf einem anderen Blatt chinesische Schriftzeichen anordnen und zurück durch den Schlitz geben. Die Person wird allerdings, unabhängig davon, wie oft oder wie lange dies passiert, nichts über chinesische Sprache, über Semantik, lernen, noch wissen, ob das erhaltene Blatt eine Geschichte, eine Frage oder eine Antwort thematisiert. Die Person wird also Operationen durchführen, in diesem Fall die Aneinanderreihung von Schriftzeichen, aber keine Zusammenhänge respektive ein größeres Ganzes erkennen können. (vgl. 1986, 30)

Die künstliche Intelligenz eines Computers beschränkt sich also zu einem Teil auf das Ausführen von Operationen nach einem Regelwerk, eine Situation, die beispielsweise Grund für das Entstehen des Forschungsfeldes

der neuronalen Netzwerke war. Auch George Dyson, amerikanischer Wissenschafts- und Technikhistoriker, formuliert einen möglichen Stillstand in dieser Forschungsthematik der künstlichen Intelligenz ohne ein Umdenken. „Computers have been getting better and better at providing answers – but only to questions that programmers are able to ask.“ (2005¹⁷)

Carr führt diesen Gedanken, wie Campbell und Turing, fort und fügt hinzu, dass dieser Ansatz entgegen dem eines lebenden Systems steht, insofern unser Gehirn beispielsweise einen großen Teil des Lebens Antworten auf Fragen geben muss, die nicht gefragt wurden oder zumindest nicht in präziser Art und Weise. (vgl. Carr 2009, 224)

2.9 Künstliche neuronale Netzwerke „lernen“

Das Forschungsfeld der künstlichen neuronalen Netzwerke, oder auch generell neuronale Netzwerke, wählt im Gegensatz zur symbolischen den Weg der konnektionistischen künstlichen Intelligenz und bezieht seine Motivation aus der Erkenntnis, dass das menschliche Hirn in einer gänzlich unterschiedlichen Art und Weise funktioniert als ein herkömmlicher Computer. Neuronale Netzwerke gehen nicht nur davon aus, dass die Art und Weise der Vernetzung und Verknüpfung von beispielsweise Neuronen im Gehirn ausschlaggebend für Intelligenz ist, sondern darüber hinaus auch die Beschaffenheit der Objekte, die miteinander verbunden sind, keine maßgebliche Rolle spielt, seien es Neuronen oder auch Chips.

Simon Haykin, ein kanadischer Professor für Elektrotechnik und Spezialist auf dem Gebiet der neuronalen Netzwerke, nennt in Anlehnung an eine mögliche Definition neuronaler Netzwerke das Gehirn einen hochgradig komplexen, nonlinearen und parallelen Computer oder auch Informations-Verarbeitungs-System. Als das wichtigste Merkmal des Gehirns wird hier die Fähigkeit erkannt, die strukturellen Bestandteile, auch Neuronen genannt, zu organisieren und in weiterer Folge unterschiedlichste Aktionen oder Berechnungen im weitesten Sinne durchführen zu können. Beispiele für diese Berechnungen wären etwa Mustererkennung, Wahrnehmung oder die Motorik. Haykin ist hier auch der Meinung, dass das menschliche Hirn

17 http://www.edge.org/3rd_culture/dyson05/dyson05_index.html, aufgerufen am 06.08.2012.

diese Aufgaben signifikant schneller abwickeln könne als das zum heutigen Tage die schnellsten Computer könnten. (vgl. 2009, 31)

Neurale Netzwerke sind also der Versuch, ein künstliches Modell einer Maschine zu entwerfen, welches dem menschlichen Gehirn und dessen Art und Weise, Aufgaben oder Funktionen zu bewerkstelligen, nachgebildet ist. Während es innerhalb des Forschungsfeldes verschiedene Herangehensweisen gibt, beschränkt sich Haykin auf lernende neurale Netzwerke und definiert diese als adaptive Maschinen.

A neural network is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

1. *Knowledge is acquired by the network from its environment through a learning process.*
2. *Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. (2009, 32)*

Der Prozess des Lernens wird hierbei durch einen so genannten *learning algorithm* zu Stande gebracht, wobei zwei große Arten des Lernens unterschieden werden, das *supervised learning* und das *unsupervised learning*. Während beim supervised learning eine Instanz, etwa eine Lehrperson, auf eine Frage eine gewünschte Antwort vorgibt und so das lernende System so lange in einer Feedback-Schleife verhaftet, bis dessen Antwort mit der vorgegebenen, gewünschten Antwort übereinstimmt, bleiben beim unsupervised learning zwei Unterarten offen. Das *reinforcement learning* geht davon aus, dass, nachdem das lernende System eine Handlung getätigt hat, die Umgebung dieses Systems reagiert und das lernende System so überprüfen kann, ob die gewählte Handlung in allen ausgeführten Schritten und Teilen erfolgreich war bzw. funktioniert, oder aber angepasst werden muss. Das reine unsupervised learning an sich wiederum geht davon aus, dass es keine Feedback-Schleife gibt, also das lernende System die Reaktionen der Umgebung auf die getätigte Handlung nicht abwartet, sondern durch fortlaufende Interaktion die eigene Performanz optimiert. (vgl. Haykin 2009, 64f)

Das reinforcement learning steht also genau genommen zwischen den beiden Sparten supervised learning und unsupervised learning, da hier die Reaktion und Antwort der Umgebung eine maßgebliche Rolle spielen. Während beim supervised learning also vereinfacht gesagt alle Gegenstände, die das lernende System erfassen soll, bekannt und von einer Lehrinstanz ausgewiesen sind, fehlen diese Informationen beim unsupervised learning. Als eine bestehende Schwäche der neuronalen Netzwerke sieht Haykin das Problem, das ein lernendes System lösen muss, wenn es mit einer sehr großen Anzahl an Mustern konfrontiert wird, die es zu erkennen und zu klassifizieren gilt. Hier könnte das *semisupervised learning* Abhilfe schaffen, in dem einem lernenden System sowohl bekannte bzw. ausgewiesene als auch unbekannte bzw. nicht ausgewiesene Lerngegenstände während der selben Lerneinheit zum erfassen gegeben werden. (vgl. 2009, 75)

Neurale Netzwerke verwenden, so Haykin, ein vereinfachtes Modell des Nervensystems. Zwischen Reiz und Reizantwort liegen demnach die Rezeptoren, das neurale Netz und die Effektoren. (vgl. 2009, 39)

2.9.a Grenzen der Entwicklung

Diese Übersimplifizierung des Gehirnes ist einer der Kritikpunkte, die Jeff Hawkins im Bezug auf neurale Netzwerke nennt, da der überaus komplexen Struktur und Architektur des Gehirnes und des Neokortex, der für diesen Prozess der Erkennung, Verarbeitung und Adaption von Mustern zentral sei, nicht Rechnung getragen werde. Darüber hinaus werde der Faktor Zeit kaum in Betracht gezogen, wenn es um das Verarbeiten von Information im Gehirn gehe, es gebe keinen statischen Fluss an Information sondern viele verschiedene, die sich wiederum ständig verändern würden. Auch die Feedback-Schleife zwischen lernendem System und Umgebung wird für Hawkins nicht ausreichend in die Untersuchung mit einbezogen, obwohl die Bandbreite des Rückflusskanals zu den Rezeptoren den des eigentlichen Input-Signales um den Faktor zehn übersteigt, es wird also stets eine große Menge mehr Information zum Auge geschickt, als dieses tatsächlich liefert. Das Feedback dominiert also, nach Hawkins, die Verbindungen des Gehirnes mit den Rezeptoren bzw. Sinnesorganen. (vgl. 2004, 25)

Zur originär aus der Gehirnforschung stammenden Definition des Begriffes der Intelligenz nun ist besonders Hawkins Position interessant, die dieser in seinem Buch „On Intelligence“ extensiv darlegt und schildert. Hawkins befasst sich dabei nicht nur mit einer möglichen Funktionsweise und umfassenden Theorie des Gehirns, als Basis für menschliche oder organische Intelligenz, darüber hinaus werden auch jene Arten künstlicher Intelligenz untersucht, die in Zusammenhang mit der technologischen Entwicklung des Computers und dessen vernetzter Form stehen und daher für diese Arbeit besonders interessant sind. Besonders die Gedanken zu künstlicher Intelligenz und neuronalen Netzwerken bzw. deren Grenzen sind, wie teilweise bereits erwähnt, relevant und sollen zur weiteren Abgrenzung und Eingrenzung der Idee einer digitalen Intelligenz in Kapitel 3.4 aufgegriffen werden.

Der Ausgangspunkt für Hawkins Arbeit ist der Neokortex. Dieser ist der Teil des Gehirns, in dem Wahrnehmung, Sprache, logisches Denken, und Vorstellungskraft passieren, also ein großer Teil dessen, was allgemein unter Intelligenz zusammengefasst wird. Dieser Fokus auf den Neokortex und dessen Arbeitsweise in Bezug auf Erinnerung und Vorhersage als Schlüssel echter Intelligenz basiert auf der Forschungsarbeit von Vernon Mountcastle. Mountcastle beobachtete, dass im wesentlichen alle Hirnregionen, unabhängig davon ob für Sehen, Hören, Motorik oder Verstehen zuständig, nicht nur erstaunlich ähnlich aussähen, sondern darüber hinaus auch in derselben Art und Weise arbeiteten. Wenn also Hirnregionen theoretisch gleiche Funktionen ausführten, könne der Grund dafür, warum das Sehzentrum visuell und der Frontallappen motorisch vorherrschend aktiv seien, nicht Unterschiede in Funktion und Arbeitsweise, sondern lediglich die Verknüpfung miteinander und den restlichen Teilen des zentralen Nervensystems sein. (vgl. Johnston 2008, 396)

Unterstützt wird diese Annahme, nach Hawkins, durch die extreme Flexibilität der Verknüpfung und die Uniformität aller Input-Signale in den Neokortex, die ausnahmslos aus elektrochemischen Aktionspotenzialen oder Spitzen bestehen, unabhängig davon, ob nun ein Signal ursprünglich von Ohr, Auge, Nase, Haut oder etwa Muskeln kommt. Die Anpassungsfähigkeit des Neokortex bedeutet nun, dass die Verknüpfungen variabel sind und je nach Typus des Input-Signals verändert und adaptiert

werden können. Veranschaulicht wird diese enorme interne Adaptionsfähigkeit beispielsweise bei dem Verlust eines Sinnesorgans.

Adults who are born deaf process visual information in areas that normally become auditory regions. And congenitally blind adults use the rearmost portion of their cortex, which ordinarily becomes dedicated to vision, to read braille. Since braille involves touch, you might think it would primarily activate touch regions – but apparently no area of cortex is content to represent nothing. The visual cortex, not receiving information from the eyes like it is „supposed“ to, casts around for other input patterns to sift through – in this case, from other cortical regions. (2004, 54)

Diese ungemeine Flexibilität ist essenziell für die Funktionsweise des Gehirns und damit unserer Intelligenz. Als ein zentrales Problem beim Erschaffen eines intelligenten Computers, unabhängig davon, ob künstliche Intelligenz, neurales Netzwerk oder etwas Ähnliches, sieht Hawkins also den Ansatz, einen Computer zu programmieren, wo doch unser Gehirn nur zu einem vermutlich sehr begrenzten Teil programmiert, sondern zu einem großen Teil selbstlernend ist und, wie eben aufgezeigt, sich durch Flexibilität auszeichnet. Ein Computer soll seine Arbeit perfekt in jeder Situation und Umgebung verrichten, das Gehirn jedoch ist natürlich flexibel und toleriert Fehler, es soll mit diesen ohne eine vorgegebene Routine umgehen können. Es kann sich der Mensch daher im Sinne der eingangs erwähnten Definition Wechslers durch diese Fähigkeit an neue Situationen und verändernde Anforderungen leicht anpassen. Computer haben einen zentralen Prozessor, der für alle Handlungen und Abläufe zuständig ist, das Gehirn jedoch verfügt über keine zentrale Kontrolleinheit, ein Umstand, den Hawkins im Verlauf seiner Arbeit diskutiert und versucht zu belegen. (vgl. 2004, 12)

Die Weiterentwicklung von künstlicher zu digitaler Intelligenz soll dabei aber gesondert in Kapitel 3.4 untersucht werden.

2.9.b Überprüfung des Vorliegens von Intelligenz

Hawkins stellt darüber hinaus die wohl bekannteste, aber auch am häufigsten kritisierte Art, um künstliche Intelligenz im weitesten Sinn zu überprüfen, in Frage. Der schon erwähnte Turing-Test, erfunden und

eingeführt von Alan Turing um 1950, misst vermeintlich die Intelligenz eines Testsubjektes oder -objektes. Konkret wird hierbei eine menschliche Testperson mit zwei ihm oder ihr unbekanntem Testpersonen über einen Bildschirm ohne Sicht- oder Hörkontakt, sondern lediglich per Text verbunden. Kann während der folgenden Unterhaltung der oder die Fragenstellende nicht zwischen Mensch oder Maschine unterscheiden, so fällt der Test für die Maschine positiv aus. Seit Einführung dieser Testpraxis wurden einige strittige Punkte gefunden, Hawkins Kritik fußt hierbei auf die Tatsache, dass der Turing-Test Intelligenz mit Verhalten gleichsetzt. (vgl. 2004, 33)

Verhalten kann aber dann nur eine Manifestation von Intelligenz sein, wenn davon ausgegangen werden muss, dass ein System oder Lebewesen auch ohne Sprechen, Schreiben oder Handeln intelligent sein kann. Hawkins beschreibt etwa, dass im Laufe der Tests beispielsweise ein der Behandlung durch einen Psychotherapeuten nachempfundenes System, das auf Fragen mit einer Gegenfrage antwortete, den Test bestanden hätte, obwohl nach anderen Maßstäben wohl kaum von Intelligenz gesprochen hätte werden können. Diesen Fokus auf Verhalten bezeichnet er weiter als vorrangigen Kritikpunkt sowohl an der Theorie der neuronalen Netzwerke, als auch der Erforschung der künstlichen Intelligenz. Unabhängig davon, ob diese Verhalten als „Antworten“, „Muster“ oder „Output“ bezeichnen, gehen beide Theorien von Intelligenz aus, wenn nach einem gegebenen „Input“ von einem untersuchten System erwartetes Verhalten produziert wird. (vgl. Hawkins 2004, 29)

Im Sinne des Funktionalismus, so Hawkins, wäre Intelligenz etwas, das nicht so sehr abhängig ist von den einzelnen Bestandteilen eines Systems oder eines Organismus, sondern vielmehr wie diese organisiert sind. Von dieser Annahme ausgehend würde so etwas wie ein Verstand oder eine Intelligenz also in jedem System auftreten, dessen zusammenhängende Bestandteile in der richtigen Beziehung und Anordnung zueinander stünden, unabhängig davon, ob diese einzelnen Bestandteile nun Neuronen, Chips oder etwas völlig anderes seien. (vgl. 2004, 36)

Das würde also bedeuten, dass, wenn ein künstliches System den selben funktionalen Aufbau und die selbe Architektur wie auch ein funktionierendes und intelligentes Gehirn hätte, dieses unweigerlich ebenfalls intelligent sein müsste. Genau dieser Aufbau aber ist es, der, so

Hawkins, den Forschungsfeldern Probleme bereite und zwar in der Hinsicht, dass für das Funktionieren des Gehirns und somit unserer Intelligenz keine umfassende, übergeordnete Theorie bestünde, die eine groß angelegte Erklärung für das System, die Zusammengehörigkeiten, Verbindungen und Relationen innerhalb des Gehirns plausibel erklären könnte.

Doch auch neurale Netzwerke können der revolutionären Rolle, die angestrebt war, nicht ganz gerecht werden. Obwohl die Herangehensweise neuraler Netzwerke dem Aufbau des Gehirns nachempfunden ist, so ergeben sich bei ihrer Funktion zwei Probleme. Einerseits das Problem der Größe bzw. des Maßstabs, mit dem die Wissenschaft im Vergleich zum Gehirn arbeitet. Der Komplexität der vielen Millionen von miteinander verbundenen Neuronen können die verhältnismäßig simpel gestalteten künstlichen neuronalen Netzwerke bei weitem nicht erreichen. Andererseits fehlt darüber hinaus aber auch das Verständnis, welche Funktionen verschiedenste Teile des Gehirns erfüllen bzw. eine Gesamtheorie, die den Zusammenhang zwischen dem Verstehen von neuronalen Mechanismen und intelligentem Verhalten umfasst. (vgl. Johnston 2008, 389)

Genau diese Reduktion und Gleichsetzung von Intelligenz mit Verhalten sieht Hawkins als das fundamentalste Problem hinsichtlich der Herangehensweise von neuronalen Netzwerken an Intelligenz, wie schon bei künstlicher Intelligenz. Unabhängig davon, wovon Intelligenz in den Forschungsfeldern abgeleitet würde, ob von Antworten, Mustern oder generellem Output, so werde immer das Verhalten, das ein Programm, ein neurales Netzwerk oder eine künstliche Intelligenz nach dem Verarbeiten eines gegebenen Inputs produziert und aufweist, als Anhaltspunkt für diese Intelligenz genommen. Verhalten, wie bereits erwähnt, könne nur eine Manifestation von Intelligenz sein, so Hawkins. Auch wenn man im Dunkeln liegt und nachdenkt ist man intelligent, weist aber keinerlei auffälliges Verhalten auf. Es gilt also, nicht das Verhalten, sondern die Vorgänge im Kopf und im Gehirn zu verstehen, um Intelligenz und somit in weiterer Folge intelligente Maschinen herstellen zu können. Eine legitime Metrik wäre hingegen wie das Hirn etwas zum einen speichert und in weiterer Folge diese Erinnerung für Vorhersagen aller Art verwendet wird. (vgl. 2004, 29)

Dieses Treffen von Vorhersagen über die Zukunft ist nun der essenzielle Punkt in Zusammenhang mit Intelligenz und auch für diese Arbeit in weiterer Folge von großer Relevanz.

Prediction, not behavior, is the proof of intelligence. We are now ready to delve into the details of this new idea of the memory-prediction framework of the brain. To make predictions of future events, your neocortex has to store sequences of patterns. To recall the appropriate memories, it has to retrieve patterns by their similarity to past patterns (auto-associative recall). And, finally, memories have to be stored in an invariant form so that the knowledge of past events can be applied to new situations that are similar but not identical to the past. (Hawkins 2004, 105)

Es muss also einem intelligenten System möglich sein, mithilfe von Sinnen, in welcher Art und Weise auch immer, die Umwelt wahrzunehmen, Muster aus dieser zu extrahieren, diese in einem Erinnerungssystem so abzuspeichern, dass Ähnlichkeiten zwischen Mustern erkannt und diese adaptiert, also auf neue Situationen angepasst werden können. Diese Herangehensweise ist dann essenziell, wenn das Wissen um erinnerte Situationen und Ereignisse zur Lösung und Bewältigung von Probleme und Herausforderungen in der Gegenwart herangezogen werden kann.

Résumé

In diesem Kapitel wurde anhand von sechs Merkmalen nach Lehmann ein gesellschaftlicher Wandel zu einer Informationsgesellschaft festgestellt und diskutiert. Im Rahmen dieser Diskussion konnten zwei Problemstellungen im Zusammenhang mit einer Informationsgesellschaft identifiziert werden, das Problem der Menge an verfügbarer Information im Internet, sowie das Problem der Komplexität dieser Menge an Information. Daran anschließend wurden ausgewählte theoretische Positionen einerseits zur Informationsgesellschaft selbst, andererseits zu den damit verbundenen Problemstellungen der Menge und Komplexität von Information erläutert. Da sich in der Praxis bereits verschiedene Lösungsansätze für diese Problemstellungen herausgebildet haben, wurden einige Herangehensweisen eingeführt und die Bedeutung von Suchmaschinen, speziell Google, in diesem Zusammenhang als universelle Schnittstellen zwischen Mensch und Computer erläutert.

Der Prozess einer Suche nach Information ist als ein Akt von Intelligenz einzustufen, weshalb sowohl Konzepte aus dem weiten Forschungsfeld der künstlichen Intelligenz, als auch deren Grenzen aufgezeigt wurden. Darüber hinaus konnte mit den beiden Ansätzen Crowdsourcing und Human Computing ein Ausblick darauf gegeben werden, wie eine effektive Funktionsweise einer digitalen Intelligenz als Kombination der Fähigkeiten von Mensch und Computer aussehen könnte. Da nun sowohl die Problemstellungen, als auch verschiedene Konzepte zur Bewältigung dieser vorgestellt wurden, kann im letzten Kapitel mit der Untersuchung von Google als Konglomerat verschiedenster multimedialer Dienste begonnen werden. In weiterer Folge soll Googles Potenzial als digitale Intelligenz und somit Lösungsansatz für die beiden Probleme der Menge und Komplexität von Information ausgelotet werden.

3 Google

Google ist allgemein bekannt als eine Suchmaschine. Während die zentrale Funktion der Suche nach Text, die den ursprünglichen Ausgangspunkt darstellt, im Kern immer noch erhalten ist, so wurde aus der Suchmaschine über die Jahre ein Konglomerat aus verschiedensten multimedialen Funktionen und Diensten, deren Verbindung und Zusammenwirken Google einzigartig macht. Zu Beginn dieses Kapitels soll daher ein kurzer Abriss der Geschichte von Google erfolgen, um dessen Dimension und Potenzial, aber auch Funktionsweisen und Absichten besser einordnen zu können. Die bisherigen Beobachtungen, Vermutungen und Erkenntnisse sollen in der anschließenden Untersuchung von Google als mögliche digitale Intelligenz zusammengeführt werden.

Doch zu Beginn gibt sich Google unscheinbar, wie so oft, unscheinbar. Die kurze Zeit, zumeist weniger als eine Sekunde, zwischen einer Sucheingabe und den vielen, oft Millionen, Suchergebnissen auf www.google.com oder einer von über 180 weiteren Domains des Unternehmens Google geben kaum einen Hinweis darauf, was diese Suche ermöglicht, nach welchen Kriterien bei dieser vorgegangen wird und was Google darüber hinaus alles kann, macht und in Entwicklung hat. Google ist dabei weder die einzige, noch die am frühesten entwickelte Suchmaschine weltweit. *Archie*, *The Wanderer*, *AltaVista* und *Yahoo!* sind die bekanntesten Dienste im Internet der Geschichte vor Google. Aufgrund der weltweit faktisch kaum bestrittenen Dominanz und Marktvorherrschaft Googles, die bereits nahezu seit der Gründung vor mehr als zehn Jahren besteht, soll im weiteren vornehmlich Google betrachtet werden, auch um den Rahmen der Arbeit nicht zu sprengen. (vgl. Patzwaldt 2005, 75ff)

Die Anfänge dieser Suchmaschine finden sich im Jahre 1996 an der Universität von Stanford in den Vereinigten Staaten von Amerika, genauer Kalifornien. Dort entwickelten die beiden Informatik-Studenten Larry Page und Sergey Brin eine „Suchmaschine mit dem Namen *BackRub*, die mithilfe von Links die Relevanz einzelner Webseiten ermittelte.“ (Google 2012b¹⁸) *BackRub* wurde an der Universität Stanford als Suchmaschine eingesetzt, allerdings überstieg diese bald die Kapazitäten der Server und wurde wieder deaktiviert. Bereits ein Jahr später kam es einerseits zu einem neuen

18 <http://www.google.com/intl/en/about/company/history/>, aufgerufen am 06.08.2012.

Namen der Suchmaschine, Google, einem Wortspiel auf *googol*, das den mathematischen Terminus für eine 1 gefolgt von 100 Nullen und damit eine Reflexion auf die schier unfassbaren Mengen an Daten und Informationen, die das Internet füllen, darstellte. (vgl. Google 2012b,¹⁹)

Bereits 1998 übersiedelt Google vorerst in eine Garage, Brin und Page stellen erstmals einen befreundeten Studenten an, das Unternehmen Google wird gegründet. Schon im Jahr 2000 umfasst der Such-Index mehr als eine Milliarde *URLs* und ist damit zu diesem Zeitpunkt die weltweit größte Suchmaschine, darüber hinaus wird die bisher verfügbare Sprache Englisch um 14 zusätzliche erweitert, sowohl aus dem europäischen, als auch dem asiatischen Sprachraum. Nur zwei Jahre später ist Google in 72 Sprachen verfügbar und die neuen Bereiche *Google Image Search*, *Google News* und *Google Books* werden offiziell eingeführt. 2004 erreicht der Index eine Größe von 8 Milliarden Einträgen und die Grundsteine für spätere Dienste wie *Google Earth*, *Google Maps*, *Google Scholar* und sogar *Google Plus* sind gelegt. Auch geht das Unternehmen in diesem Jahr Partnerschaften mit Bibliotheken verschiedener Universitäten wie etwa Harvard, Oxford, der New York Public Library oder auch etwa der Österreichischen Nationalbibliothek, ein und beginnt damit, systematisch deren gesammelte Bücher und Werke zu digitalisieren. Auf dieses *Google Books Library Project* soll etwas später noch genauer eingegangen werden. (vgl. Google 2012b²⁰)

Während also die Quellen, die Google erfasst und indiziert, immer vielfältiger und zahlloser werden, bildet die Suche nach Text, wie eingangs erwähnt, den ursprünglichen Ausgangspunkt. Herzstück der Suche bildet der von Page und Brin während ihrer gemeinsamen Studienzeit entwickelte Algorithmus *PageRank*, der erstmals im Rahmen der Suchmaschine BackRub zum Einsatz kam. Dieser Algorithmus zur Analyse und Priorisierung von Websites, auf die am häufigsten von anderen Websites verlinkt wird, ist immer noch unter dem Namen PageRank in den Prozess der Google-Suche eingebunden. Bereits zwei Jahre später, 1998, gründen Page und Brin das Unternehmen Google. Im Laufe der Jahre wurden über die reine Text-Suche hinaus eine Vielzahl weiterer Produkte und Dienste entwickelt, von denen ein Teil ebenfalls mit einer Suche, etwa nach Bildern,

19 <http://www.google.com/intl/en/about/company/history/>, aufgerufen am 06.08.2012.

20 Ebd.

Videos oder auch Straßenkarten, betraut und beauftragt werden kann. (vgl. Google 2012d²¹)

Als Ausgangspunkt der weiteren Überlegungen soll vorerst wieder die Text-Suche dienen. Bei einer Suche wird nicht das gesamte Internet durchsucht, sondern ein Index, den Google anlegt. Dieser Index wird durch den Einsatz von *Spiders* gewonnen, automatisierten Programmen, auch *Bots* genannt, die systematisch etwa bei einer Website beginnen, diese durchsuchen bzw. indizieren, um in weiterer Folge ebenfalls alle Links von dieser Website weg zu untersuchen und diesen auf die verweisenden Websites zu folgen. Es entsteht also von einem kleinen Ausgangspunkt aus jede Sekunde ein immer vollständigeres Bild jeglichen Inhaltes, der auf einer Website, auf die eine andere Website verlinkt, verfügbar ist. Diese Spiders also durchforsten das Internet und legen alles Gefundene in einem Index ab, den Google für die Nutzerin oder den Nutzer durchsuchen kann. (vgl. Google 2012d²²)

Bereits beim Eintippen eines Suchbegriffes beginnt nun Google damit, eine Vielzahl möglicherweise beabsichtigter Variationen des Begriffes vorzuschlagen, diese Funktion ist auch als *Google Instant* bekannt. Sobald nun die Suche gestartet wird, beginnt Google damit, den angelegten Index nach einem speziellen Auswahlverfahren zu durchsuchen. Google stellt im Rahmen dieses Auswahlverfahrens viele Fragen, genau genommen über 200. Ein Teil dieser Fragen ist bekannt, etwa, wie oft eine Website die eingegebenen Suchbegriffe, also *keywords*, enthält, ob diese möglicherweise im Titel der Website aufscheinen oder vielleicht in der Web-Adresse der Website, der URL. Auch wird die direkte Nachbarschaft einer Website mit in Betracht gezogen, enthalten eventuell verlinkte Websites einen oder mehrere Suchbegriffe oder vielleicht ein Synonym für einen dieser, ist die Website qualitativ hochwertig oder eher minderwertig, versucht diese möglicherweise via Suchmaschinen-*Spamming* bzw. Index-*Spamming*, also durch Kumulieren einer Vielzahl an unzusammenhängenden und für Nutzer irrelevanter Suchbegriffe auf einer ansonsten nutzlosen Website, das Suchsystem zu überlisten und so das eigene Ranking zu manipulieren? (vgl. Google 2012d²³)

21 <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=70897&topic=2370570&ctx=topic>, aufgerufen am 06.08.2012.

22 Ebd.

23 Ebd.

Natürlich ebenfalls in diesen Prozess des Filterns inbegriffen und eine der Grundlagen für die Bewertung einer Website ist der schon im Rahmen von BackRub entstandene PageRank-Algorithmus zur Gewichtung der Popularität und Relevanz eines Links bzw. der wiederum darauf verweisenden Links. Hypothetisches Ziel dieses Algorithmus ist es, eine Nutzerin oder einen Nutzer nachzubilden, der zufällig durch das Internet surft und in weiterer Folge die Wahrscheinlichkeit zu errechnen, mit der diese Nutzerin oder dieser Nutzer im Laufe der virtuellen Reise auf die zur Debatte stehende Website stoßen würde. Alle diese Faktoren werden schlussendlich zusammengerechnet und ergeben eine Wertung, die schließlich dem Nutzer oder der Nutzerin als Liste von Ergebnissen, beispielsweise in Text- oder Bildform, präsentiert wird. (vgl. Google 2012b²⁴)

Nach Googles eigenen Angaben gibt es auch keinen direkten finanziellen Weg, um die Wertung einer Website, die Frequenz in der diese einem update unterzogen wird oder überhaupt in den Index aufgenommen wird, zu manipulieren. Eine erfolgreiche Suche bringt als Ergebnis also einen Titel, die „URL“ der Website und einen Teil des textlichen Inhaltes dieser mit den hervorgehobenen, eingegebenen Suchbegriffen. (vgl. Google 2012c²⁵)

3.1 Was sammelt Google?

Google sammelt, etwa wie schon erwähnt durch den Einsatz von Spiders, unglaubliche Mengen an Daten zumindest im surface web, möglicherweise aber auch im deep web. Doch nicht nur entgeht der Technik Googles vermutlich nur ein Bruchteil des publizierten Inhaltes im surface web wie in Kapitel 2.5 beschrieben, darüber hinaus können jede Sekunde aus einer Vielzahl von Nutzerhandlungen zusätzliche Mengen an Daten und Information gesammelt werden, also nicht nur statische Daten einer durchsuchten Website, sondern auch Information aus der Interaktion mit Nutzerinnen und Nutzern. Das Unternehmen sitzt damit in der Mitte des Internet und streift irgendwann nahezu jede veröffentlichte Information, vorausgesetzt diese ist frei zugänglich und entweder bei der Suchmaschine registriert, oder aber von oder zu dieser Website wird verlinkt. Zusätzlich dazu werden auch Informationen außerhalb des Internet gesammelt, etwa

24 <https://www.google.at/intl/de/about/company/history/>, aufgerufen am 06.08.2012.

25 <http://www.google.de/intl/de/about/company/philosophy/>, aufgerufen am 06.08.2012.

im Google Books Library Project, Google Art Project, bei der Erforschung sauberer Energie oder vielen anderen Initiativen, die oft auch von Google.org, einem der nicht profitorientierten Ableger des Unternehmens, vorangetrieben werden.

Dabei werden sowohl originär digitale, als auch digitalisierte Inhalte und Informationen nach analoger Vorlage bei Google zusammengeführt. Ein Blick in die Optionen der Suche zeigt die vielen verschiedenen Inhalte, die gefunden werden können. Ob Inhalte von Websites, Bilder, Videos, Bücher, Blogs, geografische Karten, wissenschaftliche Dokumente, interaktive Inhalte oder vieles mehr, der enorme Index umfasst nicht nur diese Daten, sondern auch Nutzerdaten, also bei jeder Nutzung eines Dienstes von Google automatisch generierte und gespeicherte Informationen zu Interaktion, Kommunikation und Verhalten von Nutzerinnen und Nutzern, die quasi Verhaltensforschung im Industriemaßstab ermöglichen. Während in Kapitel 3.3 dieser Umstand noch etwas genauer betrachtet werden soll, so gibt Eric Schmidt, langjähriger Vorstandsvorsitzender Googles, einen Einblick in die Dimension der Datenmengen, mit denen Google täglich konfrontiert ist. „There were five exabytes of information created by the entire world between the dawn of civilization and 2003. Now that same amount is created every two days.“ (Schmidt 2010²⁶)

Google als Unternehmen und Suchmaschine in vermittelnder Instanz zwischen Nutzerinnen bzw. Nutzern und dem Internet profitiert dabei von diesen Mengen, wie im folgenden Kapitel kurz angeschnitten werden soll.

3.2 Was will Google?

Google ist ein Wirtschaftsunternehmen. Die Firma erzielt Umsätze, indem sie anderen Unternehmen Suchtechnologien anbietet und Anzeigen verkauft, die auf der Website von Google und anderen Websites im Internet geschaltet werden. Hunderttausende von Kunden werben weltweit mit Google AdWords für ihre Produkte und Hunderttausende von Webseitenbetreibern nutzen unser Programm Google AdSense zur Bereitstellung relevanter Anzeigen für ihre Webinhalte. (Google 2012c²⁷)

26 <http://www.youtube.com/watch?v=UAcClSrAq70>, aufgerufen am 06.08.2012.

27 <http://www.google.com/about/company/philosophy/>, aufgerufen am 06.08.2012.

Die Ziele des Unternehmens Google sind, trotz einer Vielzahl von verschiedenen Produkten, simpel gehalten, wenn man dessen eigene Angaben als Referenz nimmt. Im Laufe der nun schon über ein Jahrzehnt als Unternehmen währenden Schaffensperiode hat Google eine öffentlich zugängliche Liste von Dingen, die das Unternehmen für wichtig befindet, erstellt. Der Tenor dieser Zielübersicht ist so bestechend klar wie nichtssagend. Egal welches Produkt oder welcher Dienst, essenziell sind Fokus auf Nutzerinnen und Nutzer, Perfektion in der Ausführung, die Geschwindigkeit des Service und die Überzeugung, dass Geld auch verdient werden kann, ohne dabei böse zu sein. Während dieses Mantra wohl auf den ersten Blick wie ein Best-of verschiedenster Marketing-Credos wirkt, so können daraus durchaus spannende Erkenntnisse gewonnen werden, wenn in die richtige Relation gesetzt. (vgl. Google 2012c²⁸)

Im Grunde genommen können tatsächlich die meisten Produkte Googles auf diese zehn Leitideen zurückgeführt werden, nämlich die möglichst schnelle und akkurate Versorgung mit möglichst relevanten Informationen für Nutzerinnen und Nutzer zu erreichen, wo oder wie auch immer diese Informationen benötigen. Wenn nun einzelne Dienste beispielhaft in Relation dazu gesetzt werden, etwa der E-Mail Dienst *Google Mail*, das Office Paket *Google Docs*, das zum Organisieren und Präsentieren von Bildern nützliche *Picasa*, ja selbst das soziale Netzwerk *Google Plus* und das mobile Betriebssystem *Android*, so kann die Zutraglichkeit zu diesem Ansatz, Zugang zu Information und Vermittlung derselben, zwischen welchen Instanzen auch immer, klar erkannt werden.

Die Möglichkeit für Google, daraus Kapital zu schlagen, besteht in der Verwendung und Einbindung der hauseigenen Werbesysteme *AdWords*, *AdSense* und *AdMob* in diese Prozesse. Der Anteil dieser Werbesysteme am Gesamtumsatz des Unternehmens beträgt dabei knapp 96%, Google hat also auch diesen Faktor sehr stark optimiert in doppelter Hinsicht, wie gleich noch im nächsten Kapitel genauer erwähnt werden soll. (vgl. Google 2012a²⁹)

Google sammelt also eine unvorstellbare Menge an Daten zu jeder Zeit an vielen Orten dieser Welt. Durch den Zugriff auf diese Vielzahl an

28 <http://www.google.com/about/company/philosophy/>, aufgerufen am 06.08.2012.

29 <http://investor.google.com/financial/tables.html>, aufgerufen am 06.08.2012.

Informationen können die Computer und Server, die das technische Rückgrat der Suchmaschine bilden, lernen und aus einem faktisch nie versiegenden Quell an Input schöpfen. Damit ergeben sich Anwendungsbereiche, die in ihrem Umfang wohl kaum abzusehen sind, die Google aber sowohl in den Forschungseinrichtungen unter dem Namen *Google Labs*, als auch an vielen anderen Standorten versucht zu erweitern. Alfred Spector, *Vice President of Research and Special Initiatives* bei Google, fasst das größere Ziel, das Google nicht nur als Unternehmen, sondern vor allem im Rahmen der Forschung verfolgt, zusammen.

Wir arbeiten zwar auch an Sachen, die nützlich für Google und seine Nutzer sind. Aber unser Zeithorizont ist weiter. Wir versuchen, den Stand der Technik voranzutreiben, auf Gebieten wie Maschinenlernen, Verarbeitung natürlicher Sprache, Übersetzung oder Bilderkennung. Diese Themen werden traditionell zum Gebiet der Künstlichen Intelligenz gezählt. Unser entscheidender Vorteil ist, dass wir gewissermaßen „in vitro“ arbeiten können: Dank der riesigen Google-Systeme können wir auf große Datenmengen und eine große Anzahl von Nutzern zurückgreifen. (Simonite 2011³⁰)

Während diese Absicht auf den ersten Blick möglicherweise nicht allzu naheliegend erscheint, ergibt das vermutlich überaus hoch gesteckte Ziel nicht nur im Bereich der Forschung, sondern auch für Google als Unternehmen Sinn. Im momentanen Zustand kann Google nach der Eingabe eines Suchbegriffes durch die Nutzerin oder den Nutzer eine oder mehrere Antworten suchen, und zumeist in vielen verschiedenen medialen Ausprägungen liefern, etwa Texte, Bilder oder Videos. Ein kurzer Blick in die bereits erwähnte Liste der „Zehn Grundsätze“ auf der Website des Unternehmens macht klar, dass praktisch alle Faktoren, die in den Prozess einer Suchanfrage involviert sind, einer ständigen Optimierung unterzogen werden, etwa die Geschwindigkeit, in der diese abgewickelt werden kann, die Eingabemöglichkeiten durch Nutzerinnen und Nutzer oder auch wie viel Information zum Durchsuchen zur Verfügung steht. (Google 2012c³¹)

Es liegt also in der Natur dieses Optimierungsprozesses, bereits ultimativ die Antwort zu liefern, bevor die Frage überhaupt gestellt wurde. In einem

30 <http://www.heise.de/tr/artikel/Wir-koennen-in-vitro-arbeiten-1350475.html>, aufgerufen am 06.08.2012.

31 <http://www.google.com/about/company/philosophy/>, aufgerufen am 06.08.2012.

Interview im Jahre 2004 mit dem *Playboy Magazine* geben Sergey Brin und Larry Page einen kleinen Einblick in ihre Vorstellung der Zukunft im Bezug auf Google. Beide erklären, dass die Art, wie Information für uns zugänglich und nützlich wird, sich immer mehr von der reinen Anwendung auf einem Computer lösen wird. „You want access to as much [information] as possible so you can discern what is most relevant and correct. The solution isn't to limit the information you receive. Ultimately you want to have the entire world's knowledge connected directly to your mind.“ (Brin 2004³²)

Page und Brin geben dabei einen Einblick in die Überlegung und Planung Googles für die Zeit nach dem reinen Suchfenster, denn Forschung und Entwicklung bleiben nicht bei dem kleinen Eingabefeld auf der Startseite stehen. In einem Interview mit dem *Wall Street Journal* aus dem Jahr 2010 legt Eric Schmidt ebenfalls etwas genauer dar, welche Sicht Google auf die Zukunft hat und welche Rolle Individualisierung von Angeboten spielt.

"We're trying to figure out what the future of search is. I mean that in a positive way. We're still happy to be in search, believe me. But one idea is that more and more searches are done on your behalf without you needing to type. I actually think most people don't want Google to answer their questions, they want Google to tell them what they should be doing next. [...] The power of individual targeting - the technology will be so good it will be very hard for people to watch or consume something that has not in some sense been tailored for them." (Jenkins 2010)

Diese Personalisierung von Angeboten macht für Google durchaus Sinn, erwirtschaftet das Unternehmen doch, wie auch beispielsweise Facebook, einen großen Teil des Umsatzes ausschließlich durch Werbung in verschiedensten Formen. AdSense, AdWords und Admob, die drei Dienste, die sowohl für die Anwendung auf Google.com, als auch für das Einbinden auf verschiedensten Websites und Geräten optimiert sind, profitieren von allen Mechanismen, Algorithmen und Möglichkeiten, um Werbung noch interessanter und relevanter für Nutzerinnen und Nutzer des Internet zu gestalten. Bei einem Einkauf über Amazon werden beispielsweise ebenfalls Vorschläge erteilt, üblicherweise basierend darauf, was andere Kundinnen

32 http://www.sec.gov/Archives/edgar/data/1288776/000119312504139655/ds1a.htm#toc59330_25b, aufgerufen am 06.08.2012.

und Kunden davor, danach oder zusammen mit dem jeweiligen Produkt kauften, bzw. in Anlehnung an die vergangenen persönlichen Einkäufe.

Während diese Praxis bereits eine mehr oder weniger fundierte Vorhersage interessanter Produkte basierend auf verhältnismäßig wenig Information darstellt, so ist das Potenzial von Google in diesem Zusammenhang offensichtlich um ein Vielfaches höher. Nicht nur, da Google persönliche Daten speichert, beispielsweise bei jeder Suche, jedem Klick auf einer Website mit *Google Analytics*, jeder Nutzung eines Google Mail-, YouTube- oder Google Plus-Accounts. Vor allem die Tatsache, dass neben der Verknüpfung der Daten dieser vielen einzelnen Services im Konglomerat Google über eine Milliarde monatliche Nutzerinnen und Nutzer mit ihrer Nutzung von Google als Suchmaschine Verhaltens- und Gesellschaftsforschung auf einer mehr oder weniger globalen Ebene ermöglichen. Das wiederum lässt Google Vorhersagen treffen, die eine weitaus höhere Chance haben, tatsächlich interessant zu sein. (vgl. Efrati 2011³³)

Da Google als Unternehmen, wie nun umrissen wurde, stets daran interessiert ist, sowohl von Nutzerinnen und Nutzern für die Suche, als auch die eigenen Werbesysteme zu lernen, soll im nachfolgenden Kapitel eingehender beleuchtet werden, wie dieser Prozess des Lernens aussehen könnte.

3.3 Wie lernt Google?

Bei einer Unmenge an Daten und Information, wie sie das Internet einerseits, aber auch andererseits die Nutzerinnen und Nutzer, die dieses bedienen, ständig zur Verfügung stellen, scheinen beste Voraussetzungen gegeben zu sein, um einen Computer lernen lassen zu können, wenn man sich, wie Google, im Zentrum des Geschehens befindet.

Als ein Beispiel dafür, wie Google auch das Geschäftsmodell so ausgelegt hat, dass zwischen Google und dessen Nutzerinnen und Nutzern ein rekursiver Fluss an Daten und Verknüpfungen nicht nur bestehen bleibt,

³³ <http://blogs.wsj.com/digits/2011/06/21/google-notches-one-billion-unique-visitors-per-month/?mod=e2tw>, aufgerufen am 06.08.2012.

sondern sich auch stetig ausbreitet und erweitert, kann das bereits erwähnte Google AdSense angesehen werden.

The Google AdSense box on the home page of my blog [...] makes me part of Google's empire. Google sends me money for those ads. Google sends me readers via search. Google benefits by showing those readers more of its ads, which it can make more relevant, effective, and profitable because it knows what my site is about. I invited Google in because Google helps me do what I want to do.

I, in turn, help spread Google by putting its ads on my page and by embedding its YouTube videos, Google Maps, and Google search box on my blog. When I link to a page on the internet, I help Google understand what that page is about and how popular it is. I make Google smarter. With our clicks and links, we all do. Google is clever enough to organize that knowledge and take advantage of it. (Jarvis 2009, 5f)

Google durchsucht also nicht nur selbständig das Internet nach Information, sondern hat darüber hinaus auch Möglichkeiten, um mit Nutzerinnen und Nutzern eine Art von Informations- und Datenaustausch aufrecht zu erhalten. Gerade aber, wenn es um das Nutzungsverhalten und somit nicht nur um das Sammeln von mehr oder weniger statischen Daten auf Websites in Form von Text, Bild oder anderen Medien geht, sondern um Daten über die Nutzung der Dienste selbst und somit viele Nutzerinnen und Nutzer analysiert werden müssen, ergeben sich diverse Probleme.

Sven Türpe beschreibt, dass der klassische Ansatz, Daten zu sammeln und zu verwalten, etwa in einer Datenbank, wohl eine Möglichkeit, aber im Fall von Google gleichzeitig eine Unmöglichkeit ist. Türpe zieht die Analogie zu einer fiktiven Welttankstelle, die jegliche Sorte Treibstoff, die jemals irgendwo angeboten, benannt und entwickelt wurde an einer eigenen Zapfsäule vertreibt. Diese Welttankstelle wäre riesig und unübersichtlich, sie würde aus einer Vielzahl von Zapfsäulen bestehen, von denen aber wohl die Meisten nur selten zum Einsatz kämen. Wie sollte also nun jede Fahrerin und jeder Fahrer an die für sie oder ihn richtige Zapfsäule weitergeleitet werden? Man müsste vermutlich zuerst alle Sprachen sprechen, um sich mit den einzelnen Kundinnen und Kunden über ihre Wünsche unterhalten zu können, manche wüssten aber womöglich nicht genau, welche Zapfsäule wohl die richtige für sie wäre, etwa aus purer Vergesslichkeit, oder aber

auch, weil das derzeitige Fahrzeug geliehen wäre. Sie hätten darüber hinaus auch vermutlich unterschiedliche Präferenzen, die vielleicht, vielleicht aber auch nicht mit ihrem momentanen Fahrzeug zusammenhängen, bei manchen würde vielleicht sogar die gewählte Zapfsäule von Mal zu Mal variieren. Und darüber hinaus könnten Fahrerinnen und Fahrer ihr Fahrzeug jederzeit wechseln und so beim nächsten Tankbesuch nicht wiedererkannt werden. (Türpe 2012³⁴)

Alfred Spector bestätigt die Existenz des angesprochenen Problems und gibt einen Ausblick auf Googles Ansatz, dieses zu lösen und damit einer Maschine Lernen zu lernen.

Wir nutzen eine gemischte oder „hybride KI“. Das heißt, wir lernen von unserer Nutzergemeinde. Wenn die einen Sachverhalt mit einer bestimmten Bedeutung versieht, machen wir uns diese zunutze. Haben wir beispielsweise bei der Sprachsuche eine Äußerung richtig verstanden, folgt darauf das Anklicken eines Treffers. Das System ist so angelegt, dass es sich mit solchen Vorgängen selbst trainiert, also besser wird, je häufiger es genutzt wird. (Simonite 2012³⁵)

Der Vergleich mit einer Welttankstelle erscheint also in gewisser Weise sinnvoll. Zum einen müsste Google wohl jede neue Nutzerin und jeden neuen Nutzer über Präferenzen befragen, dazu deren Sprache beherrschen und trotz allem in Kauf nehmen, dass entweder diese ihre Wünsche gar nicht kennen, oder aber von einer anderen Schnittstelle aus auf Dienste zugreifen, etwa Netbook, Smartphone oder Desktop-PC und somit jedes Mal möglicherweise als neue Nutzerin oder neuer Nutzer erkannt werden. Die sprachliche Barriere und damit der Prozess, wie Google eine Sprache lernt, soll an anderer Stelle noch genauer untersucht werden.

Eine Datenbank würde also im Fall der Welttankstelle zu einem sehr großen Teil aus Spezialfällen bestehen und jedes Mal scheitern, wenn eine dem System unbekannte Nutzerin oder Nutzer eine Anfrage stellt. Für die fiktive Welttankstelle schlägt Türpe nun den Bau eines Klassifikators vor, der lernfähig sein soll und somit selbständig Kundinnen und Kunden an die

34 <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241-4.html>, aufgerufen am 06.08.2012.

35 <http://www.heise.de/tr/artikel/Wir-koennen-in-vitro-arbeiten-1350475.html>, aufgerufen am 06.08.2012.

jeweils für sie richtige Zapfsäule weiterleitet, darüber hinaus vorhandene Informationen generalisiert, um auch für neue Nutzerinnen und Nutzer mit großer Wahrscheinlichkeit die richtige Wahl treffen zu können. Auch sollte dieser sich selbständig aktuell halten können und auf Veränderungen in der Umgebung ohne ein nötiges Update reagieren und eingehen können ohne äußeres Zutun. Für den Bau dieses Klassifikators sollen zu Beginn alle Merkmale, die beobachtet werden können an Kundinnen und Kunden der Welttankstelle bzw. Nutzerinnen und Nutzern von Google, aufgezeichnet werden. Welche davon schlussendlich verwendet werden, ist noch nicht abzusehen. (vgl. Türpe 2012³⁶)

Damit unser Klassifikator gut funktioniert, benötigt er Beschreibungen der einzelnen Cluster, ihrer Clustergrenzen und gegebenenfalls der Überschneidungen mehrerer Cluster. Um einen Datenpunkt zu klassifizieren, also eine Entscheidung zu treffen, müssen wir ihn dem passendsten Cluster zuordnen. Erweist sich eine Klassifikation als falsch, so wollen wir außerdem die betroffenen Clusterbeschreibungen anpassen, ohne uns jedoch von einzelnen Ausreißern unsere Statistik kaputtmachen zu lassen. Anstelle der Rohdaten im Datenbankformat verwendet man dafür Repräsentanten: Für jeden Cluster bestimmt man einen Satz von Punkten, der diesen Cluster gut repräsentiert. Diese Repräsentanten liegen irgendwo zwischen den Rohdatenpunkten; ihre Anzahl ist in der Regel geringer. (Türpe 2012³⁷)

Wenn nun eine neue Nutzerin oder ein neuer Nutzer eine Anfrage startet, wird diese oder dieser zu einem neuen Datenpunkt und das System versucht mit Hilfe eines generalisierenden Klassifikators etwa, eine Annahme darüber zu treffen, zu welchem Cluster dieser neue Datenpunkt zugehörig sein könnte. Bei einer Suchanfrage über Google also etwa zu dem Zweck, die richtige Werbung auf die Seite der Suchergebnisse zu schalten. Jedes mögliche Ziel, bei Google Ads also beispielsweise Werbung für verschiedenste Firmen, kann als Cluster verstanden werden und bekommt dabei einen oder mehrere Repräsentanten zugeteilt.

Diese Repräsentanten bestehen aus gemittelten Werten der Merkmale, die von der anzusprechenden Gruppe an Nutzerinnen und Nutzern abgeleitet

36 <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241-4.html>, aufgerufen am 06.08.2012.

37 Ebd.

werden können. Um also nun einen neuen Datenpunkt, eine neue Nutzerin oder einen neuen Nutzer beispielsweise, klassifizieren zu können, wird diesem ein Repräsentant zugeteilt. Durch die Zuteilung dieses Repräsentanten ergibt sich eine Wahrscheinlichkeit für eine oder mehrere Clusterzugehörigkeiten. Durch jedes Feedback des Nutzers kann nun dessen Verhalten beobachtet und daraus gelernt werden, ob die getätigte Annahme richtig war oder nicht. War die Entscheidung falsch, wurde also etwa auf eine für diese Nutzerin oder diesen Nutzer falsche Werbung verwiesen, wird der Fehler analysiert und der betroffene Repräsentant wird marginal angepasst, also etwa näher in Richtung eines anderen Clusters. Das soll das System und den Klassifikator schützen und minimal träge machen. Da das System also aus Fehlern besonders gut lernen kann, können zu Beginn die Repräsentanten zufällig verteilt werden, um dies zu gewährleisten. „Erst wenn systematische Fehler auftreten, akkumulieren sich viele gleichartige Korrekturen zu einer nennenswerten Verschiebung der Repräsentanten und Clustergrenzen.“ (Türpe 2012³⁸)

Als eines der prominentesten, wenn vielleicht auch unauffälligsten Beispiele für diesen nun vereinfacht dargelegten Prozess des Lernens könnte etwa Google Instant gesehen werden. Dieser Service, der seit 2012 bei jeder Suchanfrage über Google.com zum Einsatz kommt, schlägt bereits während der Eingabe eines Suchbegriffes diejenigen Begriffe vor, deren Wahrscheinlichkeit am größten ist, dass die Nutzerin oder der Nutzer diese eingeben wollen wird oder wollte.

Die Korrektur der Schreibweise in der Websuche baut auf demselben Ansatz auf. Als Barack Obama sich um das Präsidentenamt bewarb, gab es noch viele Leute, die nicht die korrekte Schreibweise seines Namens wussten und es mit unterschiedlichen Varianten versuchten. War schließlich ein Ergebnis dabei, das ihnen sinnvoll vorkam, klickten sie darauf. Unser System registrierte, welche Schreibweise zu diesem Ergebnis führte. Anhand solcher Vorgänge konnten wir Eingaben automatisch korrigieren. (Simonite 2011³⁹)

Durch diesen Mechanismus ist es dem System also möglich, selbständig, sprich ohne äußeres Zutun durch eine lehrende Instanz im Sinne der

38 <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241-4.html>, aufgerufen am 06.08.2012.

39 <http://www.heise.de/tr/artikel/Wir-koennen-in-vitro-arbeiten-1350475.html>, aufgerufen am 06.08.2012.

neuralen Netzwerke, zu lernen und sich auf dem aktuellen Stand zu halten. Dank der steten Feedback-Ströme etwa der Autokorrektur von Google Instant wird beispielsweise jede Rechtschreibreform automatisch in den Suchalgorithmus eingespielt. Das Korrigieren in der eigenen Sprache ist auch in Hinsicht auf *Google Translate* von großem Nutzen. Dies ist „ein kostenloser Dienst für sofortige Übersetzungen in und aus 64 verschiedenen Sprachen. Google Übersetzer kann Wörter, Sätze oder ganze Webseiten übersetzen - und das zwischen beliebigen Kombinationen unserer unterstützten Sprachen.“ (Google 2012e⁴⁰)

Wenn eine Nutzerin oder ein Nutzer also eine Textstelle beispielsweise übersetzen möchte, greift Google Translate auf eine Vielzahl an frei verfügbaren Dokumenten im Internet zu, die bereits von Menschen übersetzt wurden. Die Europäische Union etwa übersetzt alle veröffentlichten Dokumente in ihre 23 Amtssprachen und bietet damit zweifellos eine gute Basis für viele Übersetzungen. Auch wenn Nutzerinnen und Nutzer die Übersetzungen korrigieren, bewerten oder kommentieren, kann dieses Feedback, wie schon bei Google Instant, wertvoll sein. Sätze können mittlerweile auch grammatikalisch analysiert und zergliedert werden. Durch dieses *Parsing* können Subjekt, Verb und Objekt innerhalb eines Satzes etwa erkannt und dementsprechend besser und korrekter übersetzt werden. (vgl. Simonite 2011⁴¹)

Besondere Relevanz gewinnt in diesem Zusammenhang nun das bereits erwähnte Google Books Library Project. Dabei handelt es sich um ein ambitioniertes Unterfangen mit dem Ziel, Bücher, Schriften und darüber hinausgehende Werke in Obhut von National-, Universitäts- oder auch Privatbibliotheken an Standorten der ganzen Welt zu digitalisieren und so über den Dienst Google Books einer größeren Öffentlichkeit zugänglich machen zu können. Seit der offiziellen Ankündigung im Jahr 2004 konnten mit Hilfe der über 20 Partnerschaften mit Bibliotheken mittlerweile mehr als 12 Millionen Bücher digitalisiert werden. (vgl. Google 2011b⁴²)

Darüber hinaus verbindet Google aber auch die eigenen Services und deren Nutzerdaten untereinander. Dieser Aspekt unterscheidet Google nun

40 http://translate.google.at/about/intl/de_ALL/, aufgerufen am 06.08.2012.

41 <http://www.heise.de/tr/artikel/Wir-koennen-in-vitro-arbeiten-1350475.html>, aufgerufen am 06.08.2012.

42 <http://books.google.at/intl/de/googlebooks/history.html>, aufgerufen am 06.08.2012.

von anderen Unternehmen oder Organisationen, deren Produkte bzw. Dienste in Funktionalität oder Verwendung ähnlich erscheinen oder sich überschneiden. In jüngster Zeit etwa wurden Facebook und Google Plus zumeist als direkte Konkurrenten auf dem Markt der Social Networks beschrieben. Doch Sven Türpe gibt zu bedenken, dass, obwohl oberflächlich ähnlich, unterschiedliche Ziele und Interessen hinter beiden Diensten stehen.

Facebook (der Dienst) ist für Facebook (die Firma) der Kern des Geschäftsmodells. Für Google dagegen ist maschinelles Lernen aus allen Daten dieser Welt der Kern des Geschäftsmodells - und Google+ vor allem eine weitere Quelle interessanter Daten. Wer interagiert wie mit wem? Welche Inhalte verbreiten sich in welchen Kreisen? Wie reagieren Nutzer auf personalisierte Suchergebnisse? Welche Transformationen durchläuft ein Gerücht? (2012⁴³)

Mit diesem Kapitel über Google respektive den vorangegangenen Teilen der Arbeit ist nun die Grundlage dafür geschaffen, das Potenzial von Google als digitale Intelligenz und somit als möglichen Lösungsansatz auf die beiden Probleme der Informationsgesellschaft, nämlich Menge und Komplexität von Information, zu untersuchen.

3.4 Google und das Problem der Menge

Das Forschungsfeld der künstlichen Intelligenz, wie bereits beschrieben, hat eine Zeit des Stillstands hinter sich, die wohl eine Folge der Annahme war, Intelligenz könne durch ein Regelwerk, sei es noch so komplex, simuliert - respektive nachgebildet - werden. Während Computer und deren Programme in der Lösung spezifischer Probleme teilweise schon seit Jahrzehnten unabdingbar einigen menschlichen Fähigkeiten überlegen sind, so bleibt doch selbst die Simulation und Imitation des menschlichen Verhaltens, ob in Sprache, Mimik oder Gestik beispielsweise, zwar nicht unerforscht, jedoch noch unerreicht, obwohl schon Pioniere der künstlichen Intelligenz wie etwa Alan Turing wussten, dass es sich hierbei

⁴³ <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241.html>, aufgerufen am 06.08.2012.

nur um eine möglichst perfekte Täuschung handeln könne, nicht aber um intelligentes Handeln per se.

Doch nicht nur bei der Erforschung künstlicher Intelligenz, sondern auch der Funktionsweise des Gehirns und somit menschlicher Intelligenz können zwar Erfolge im Detail verbucht werden, allerdings scheint das System dahinter bzw. die Frage, was Intelligenz tatsächlich ist und ausmacht, nur in Ansätzen geklärt. Während etwa verschiedene Hirnregionen bzw. deren Funktionen vermeintlich identifiziert werden konnten, stellen nicht nur Vernon Mountcastle und Jeff Hawkins die Frage, ob möglicherweise doch alle Teile des Hirns die selben Voraussetzungen und Funktionen hätten und erst durch die variable Vernetzung verschiedenster Regionen Intelligenz und spezifische Aufgabenbereiche entstünden. Die Fortschritte, nicht nur die natürliche Intelligenz der gesamten Tierwelt inklusive der des Menschen zu ergründen und zu erforschen, sondern diese Erkenntnisse dann auch in einer Maschine oder einem Computer wieder hervorbringen zu können und diesen wahrhaftig intelligent, um nicht zu sagen lernfähig zu machen, scheinen also dem Ziel noch nicht zur Gänze gerecht zu werden.

***Das Problem
der Menge
an Information***

Wie passt nun Google in dieses Bild einer Intelligenz? Google bietet vermeintlich Lösungen für beide der eingangs beschriebenen Probleme durch die Verbindung der Fähigkeiten von Mensch und Maschine im weitesten Sinn. Auf Grund der stetig exorbitant zunehmenden Menge an Information, die für uns zugänglich und verfügbar ist, können Menschen selbst diese nicht mehr ausreichend ohne technische Hilfe filtern. Dieses Problem der Menge wurde primär identifiziert.

Funktionalität und Möglichkeiten von Google in diesem Zusammenhang sind klar, das System wurde explizit zu dem Zweck konstruiert, um diesem Wucher beikommen zu können. Als Suchmaschine streift Google nahezu jegliche Information des surface web, ob Text, Bild, Video oder in welcher Form auch immer, und indiziert diese wie bereits beschreiben. Einmal indiziert, also in einem Index zusammen mit Metadaten abgespeichert, bleibt die Information verfügbar bzw. es kann nach ihr gesucht werden.

An Hand des Prozesses einer solchen Suche wird klar, aus welchen Gründen Google jedem rein menschlichen Filtersystem bereits völlig überlegen ist, obwohl oder gerade weil es mithilfe der menschlichen Nutzerinnen und Nutzer lernt. Mehr als zweihundert Faktoren nehmen innerhalb von Bruchteilen einer Sekunde Einfluss auf eine Auswahl an Suchergebnissen, die durch die hohe Bezugnahme auf Informationen, die Google über uns und viele andere Suchende hat, mit einer sehr hohen Wahrscheinlichkeit die Informationen über den Suchbegriff beinhalten, die gesucht wurden. Google hat dabei zwar teilweise Informationen über uns, beispielsweise wo und wer wir sind, mit wem wir befreundet und welche unsere Vorlieben sind, die viel wichtigere Tatsache scheint hier allerdings zu sein, dass Google dieses Wissen auch über einen Großteil aller anderen Suchenden besitzt und so zu jeder Zeit wie ausgeführt Verhaltensforschung in einem kaum vorstellbaren Ausmaß betreiben kann. Das macht für Google durchaus Sinn, kann es doch genau diese Daten bestens für die eigenen Systeme für Werbung, AdSense und AdWords, mit denen das Unternehmen nahezu den gesamten Umsatz erzielt, gebrauchen. (vgl. Google 2012a⁴⁴)

Google tut dabei also, grob genommen, nichts anderes als unser Gehirn, nämlich eingehende Datenströme filtern, auswerten und Vorhersagen treffen. Während allerdings das Gehirn bei der Aufnahme von Informationen und dem Erfassen der Welt durch Funktionsweise und Grenzen der Sinnesorgane limitiert ist, so kann Google auf eine große Menge von Aspekten und Daten mehr und schneller zugreifen, als es für uns möglich wäre und lernt darüber hinaus von sehr vielen Nutzerinnen und Nutzern des Internet. Die Technik hilft uns also in diesem Fall, mit der Fülle an Informationen aus aller Welt zurecht zu kommen und setzt für uns Prioritäten, die wir momentan noch zu einem gewissen Teil adjustieren können, insofern wie wir die Suchbegriffe genauer spezifizieren. (vgl. Türpe 2012⁴⁵)

3.5 Google als digitale Intelligenz

Alle Eingaben von Nutzerinnen und Nutzern, ob richtig oder falsch mit Korrektur, jegliche Information im Internet, ob Text, Bild, Video oder in anderer Form, digitalisierte Bücher aus Bibliotheken auf der ganzen Welt,

44 <http://investor.google.com/financial/tables.html>, aufgerufen am 06.08.2012.

45 <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241-5.html>, aufgerufen am 06.08.2012.

digitalisierte Bilder und Kunstwerke von Museen und Galerien über den Globus verstreut, Aufnahmen sowohl aus dem Weltall, als auch der Erdumlaufbahn oder direkt von Straßen, Plätzen und Städten dieser Erde sowie die Kommunikation und Interaktion von Menschen im Internet, sind Informationsquellen für Google, aus denen gelernt werden kann. Wie Sven Türpe, der uns bereits in der Rolle einer Lehrperson Feedback an Google geben sieht, definiert auch Jeff Hawkins ein mögliches Modell einer intelligenten Maschine, die von ihrer Umwelt lernt. Die in diesem Zusammenhang formulierten Thesen sind auf digitale Intelligenz zwanglos umzulegen.

The intelligent machine must learn via observation of its world, including input from an instructor when necessary. Once our intelligent machine has created a model of its world, it can then see analogies to past experiences, make predictions of future events, propose solutions to new problems, and make this knowledge available to us.(2004, 209)

Teilweise erfüllt Google diese Voraussetzungen, wie von Hawkins formuliert, bereits. Google observiert auf sehr vielen Ebenen dieser Welt nicht nur uns und unsere Umgebung, sondern zusätzlich dazu auch vieles, wofür wir uns interessieren und worüber wir selbst Daten sammeln. Darauf basierend kann ein immer genaueres Bild der Welt entstehen, das im Gegenzug theoretisch allen anderen Menschen auf der Suche nach Information zur Verfügung steht. Den von Hawkins angesprochenen *instructor*, ähnlich dem reinforcement learning im Zusammenhang mit neuronalen Netzwerken, stellen dabei beispielsweise all jene Nutzerinnen und Nutzer dar, die einen erst falsch geschriebenen Suchbegriff ausbessern, einen von Google übersetzten Text als unpassend oder nicht korrekt markieren, auf eine von Google angezeigte Werbung klicken bzw. diese ignorieren, oder aber auch einfach nur eine Website auf einer anderen verlinken. Alle diese und noch unzählige mehr Handlungsweisen geben Google Feedback auf Annahmen, die von Google selbst getroffen wurden auf Grund von Erinnerungen im weitesten Sinn, also Erfahrungen, die etwa bereits mit anderen Nutzerinnen und Nutzern gemacht wurden, ganz wie beim Human Computing nach Luis von Ahn in Kapitel 2.4 beschrieben.

Dieses Treffen von Vorhersagen nun ist nach Hawkins der Schlüsselpunkt von Intelligenz im allgemeinen und damit in weiterer Folge auch

künstlicher und digitaler Intelligenz. Um Vorhersagen über die Zukunft treffen zu können, muss die Fähigkeit Muster zu erkennen vorhanden sein, etwas, das Google in großem Stil bereits seit mehr als einem Jahrzehnt ständig aufs Neue mit noch nicht dagewesenen Strömen an Input bewerkstelligt. Diese Muster müssen abgespeichert werden, um später wieder zur Verfügung zu stehen, auch diesen Punkt erfüllt Google mit einer schier unendlichen Menge an Daten und Informationen zu jeder Sekunde, ebenfalls bereits seit mehreren Jahren. Um eine passende Erinnerung, eine Information oder etwa einen Suchbegriff im Fall von Google abrufen zu können, müssen Muster auf Grund ihrer Ähnlichkeit zu anderen Mustern in der Vergangenheit erkannt werden, was wohl das hauptsächliche Einsatzgebiet von Google bei Nutzerinnen und Nutzern darstellt. Schlussendlich müssen diese Erinnerungen nach Hawkins in einer invarianten Form gespeichert werden, sodass das Wissen aus der Vergangenheit auch auf neue Situationen angepasst werden kann, die ähnlich, aber nicht ident mit dieser sind. Die konkrete Vorgehensweise, wie eine im definierten Sinne intelligente Maschine nach Hawkins aufgebaut sein muss, wird nachfolgend näher erläutert.

Das Abspeichern eines Musters in invarianter Form bewirkt, dass etwa beim Fangen eines Balls die Erinnerung vom Fangen eines Balls automatisch ins Gedächtnis gerufen wird, das bedeutet autoassoziativ, also ohne dass wir uns bewusst daran erinnern müssen. Diese Erinnerung an das Fangen eines Balls ist also im Gedächtnis, trifft aber möglicherweise auf die momentane Situation nicht zu, sie muss also angepasst werden, damit der Ball gefangen werden kann. Das Abspeichern einer Erinnerung in invarianter Form meint also, dass ein Muster abgespeichert wird, das autoassoziativ ein im Gedächtnis befindliches Muster hervorrufen kann. Dieses muss nicht gleich sein, sondern nur ähnlich oder in einer Art und Weise für die Situation behilflich, um in weiterer Folge abgeändert oder angepasst zu werden. (vgl. Hawkins 2004, 69)

Der Ansatz der künstlichen Intelligenz bisher, wieder bezogen auf das Beispiel des Fangens, war unabhängig von einer Erinnerung. Es wurde also versucht, alle nötigen Berechnungen etwa der Flugbahn, der Reibung des Balls, eines geometrischen Modells des Balls etc. in einer für das Fangen eines Balls notwendigen Geschwindigkeit, in Echtzeit, anzustellen. Doch obwohl die verfügbare Prozessorleistung unser Gehirn an Kapazität bereits seit mehreren Jahren übersteigt, so war die Herangehensweise eine falsche,

wie Helge Ritter, ein deutscher Physiker und Professor für Neuroinformatik, beschreibt.

Die Idee war dabei, vor der tatsächlichen Ausführung einer Handlung ihre möglichen unterschiedlichen Handlungsschritte und die darauf von Seiten der Umgebung zu erwartenden Reaktionen – gewissermaßen wie in einem überdimensionalen Schachspiel – in einem „inneren Modell“ der Situation durchzuspielen und die dabei gefundene beste Möglichkeit zur wirklichen Ausführung zu bringen. (2008, 120)

Wie schon beim Schachspielen sollten alle möglichen Züge, im Fall der Hand also Stellungen, Positionen, Geschwindigkeit und Flugkurve des Balls beim Fangen etc., durchgespielt werden, um den Ball exakt abfangen zu können. Diese Herangehensweise des Ausschlusses aller Zufälligkeiten kann aber wohl kaum als intelligent angesehen werden, sondern vielmehr als eine nach einem auf die Welt aufgesetzten Regelwerk berechnende, automatisierte Herangehensweise, die darüber hinaus ob der rechnerischen Überlegenheit von Computern ineffizient mit den zur Verfügung stehenden Ressourcen umgeht.

Ein Umgehen mit den oft vagen, vielfältigen und vor allem neuartigen Herausforderungen benötigt eine hohe Flexibilität in Bezug auf Abstimmungen des eigenen Handelns auf neue Situationen, etwas, das Google teilweise schon beherrscht, darüber hinaus auch akribisch und in verschiedensten Projekten erforscht. Auch Eric Schmidt erläutert in einem Interview mit dem Wall Street Journal aus dem Jahr 2010, welche Ziele das Unternehmen vorwiegend verfolgt.

"As you go from the search box to the next phase of Google, you really want to go from syntax to semantics, from what you typed to what you meant. And that's basically the role of Artificial Intelligence. I think we will be the world leader in that for a long time." (Jenkins 2010⁴⁶)

Dieser Sprung vom Suchbegriff zum tatsächlich Gemeinten würde wohl einer intelligenten Handlung schon sehr nahe kommen und im Sinne der

46 <http://online.wsj.com/article/SB10001424052748704901104575423294099527212.html>, aufgerufen am 06.08.2012.

Vorhersage nach Hawkins auch das Erkennen von Mustern in invarianter Form bedingen. Spannend an dieser Aussage Schmidts ist auch der Begriff Semantik in diesem Zusammenhang. So verwendet beispielsweise John Rogers Searle, ein amerikanischer Philosoph, von dem auch das bereits erwähnte Gedankenexperiment des Chinese Room stammt, das Wort semantisch im Zusammenhang mit Geist oder Bewusstsein. „Der Geist ist semantisch – semantisch in dem Sinne, dass er mehr hat als eine formale Struktur: er hat einen Gehalt.“ (1986, 30)

Semantisch würde in diesem Kontext bezogen auf Google etwa also bedeuten, Zusammenhänge verstehend und auf die Welt oder Umgebung beziehend, also bewusst, sein im weitesten Sinn. Während das zu diesem Zeitpunkt nicht Gegenstand der Untersuchung sein soll, so scheint das durchaus eines der Ziele für Google in Zukunft zu sein, auch die Anhäufung und das Sammeln von Erinnerungen, oder im Fall Google Daten und Informationen, würden durchaus in diese Richtung gehen, sind doch Erinnerungen und das Gedächtnis Bestandteile einer Identitäts- respektive Persönlichkeitsbildung.

In Bezug auf Hawkins Definition von Intelligenz, basierend auf Erinnerungen, um Voraussagen treffen zu können, ist Google schon heute im Besitz der Grundvoraussetzungen dafür. Hier kann die Frage, ob bzw. nach welchen Prinzipien eine künstliche Intelligenz geschaffen und konstruiert werden kann, wie folgt beantwortet werden.

The first step [...] is to start with a set of senses that extract patterns from the world, and they don't have to be equivalent to human senses. The second is to attach to the senses a hierarchical memory system that works like the Neocortex. And the third is to train the memory system so that the machine can begin to build a model of the world as seen through its senses. (Johnston 2008, 399f)

In weiterer Folge könnte die künstliche Intelligenz Analogien zu vergangenen Erfahrungen und Erinnerungen erkennen, Voraussagen über Ereignisse in der Zukunft treffen und dieses Wissen für uns in weiterer Folge verfügbar machen. Als Basis sind also zuallererst das Schaffen und verfügbar - respektive benutzbar - Machen einer enormen, Erinnerungen speichernden Instanz oder Kapazität notwendig. Diese stellt in weiterer

Folge ein Gedächtnis für die assoziativen Erinnerungen von invarianten Repräsentationen verschiedenster Muster dar. Google oder möglicherweise auch Facebook besitzen wie bereits beschrieben diese verhältnismäßig riesigen Speicherkapazitäten, teilweise gefüllt mit einerseits gesammelten und digitalisierten Informationen, sowie originär digital erstellten Inhalten, andererseits mit dem umso spannenderen Wissen um Handlungen, Verhalten, Kommunikation und Interaktion von Millionen, wenn nicht Milliarden von Menschen. Der Ausgangspunkt, ein großes Gedächtnis, kann also als gegeben betrachtet werden.

Das Hinzufügen bzw. Anknüpfen von Sinnen, die Muster in der Umgebung erkennen und extrahieren können, kann als erster Schritt im Sinne der zitierten Ausführungen von Johnston identifiziert werden. Diese Sinne nun müssen nicht mit unseren menschlichen Sinnen vergleichbar sein, sondern lediglich dazu befähigen, Information welcher Art auch immer wahrzunehmen, damit diese in weiterer Folge verarbeitet werden kann, also eine Wahrnehmung der Umgebung und der Welt ermöglichen. Ein Beispiel dafür, wie eine solche Wahrnehmung aussehen kann, liefert in gewisser Art und Weise ebenfalls Google.

Als Umwelt ist dabei die Menge an Information anzusehen, die zu jeder Zeit gesammelt wird, ob durch Digitalisierung, beispielsweise im Falle des Google Books Library Project oder bei Google Earth, Durchforsten von digitalen Inhalten, also des Internet mithilfe der Spiders, oder aber durch das Aufzeichnen von Interaktion und Kommunikation, etwa Google Plus, respektive deren Analyse. Wie in Kapitel 3.1 bereits beschrieben, lernt Google anders, als Menschen das beispielsweise tun würden. Die einzelnen gespeicherten Informationen sind dabei nur eine Übergangslösung, bis aus genügend verschiedenen Datenpunkten Muster erkannt und invariante Repräsentationen, dort von Sven Türpe Repräsentanten genannt, gebildet und für weitere Verwendung abgespeichert werden können.

Wie der Prozess des Lernens momentan bei Google in all seinen Einzelschritten aussieht, kann wegen fehlender Angaben nur vermutet werden. Tatsache ist allerdings, dass Google lernt, dies wird sowohl von diversen Funktionären und Repräsentanten des Unternehmens kommuniziert, als es auch von Nutzerinnen und Nutzern bei vielen verschiedenen Handlungen, wie beschrieben, nachvollzogen werden kann. Wie in Kapitel 3.3 mithilfe des Beispiels der Welttankstelle nach Türpe

allerdings erläutert wurde, so ist bzw. war der bereits langjährige Einsatz von Klassifikatoren bei Google ein erster Schritt in diese Richtung des lernenden Computers, wenn auch in eingeschränkter Form.

Ein Klassifikator häuft nicht einfach Daten an. Er besitzt eine Konfiguration und zwei Grundfunktionen: Lernen und Klassifizieren. In der Funktion Klassifizieren erhält er einen Datensatz als Eingabe und gibt eine Entscheidung oder Entscheidungsempfehlung aus. In der Funktion Lernen passt er seine Konfiguration an, um die Rate der Fehlentscheidungen zu reduzieren. (Türpe 2012⁴⁷)

Um wieder zurück auf die Konstruktion einer intelligenten Maschine zu kommen, so soll im nächsten Schritt an die Sinne ein hierarchisches Erinnerungssystem, das ähnlich dem Neokortex des Gehirns funktioniert, angeschlossen werden. Der Neokortex bewältigt dabei mehrere Aufgaben, etwa das Abspeichern von Mustern, das Erkennen von Ähnlichkeiten verschiedener Muster sowohl der Gegenwart als auch der Vergangenheit und der autoassoziative Wiederaufruf dieser, sowie das Anpassen von gespeicherten Mustern auf neue Situationen. (vgl. Hawkins 2005, 105)

Bezogen auf Google gibt es mehrere Beispiele, die darauf hindeuten, dass bereits nach dem Prinzip der Mustererkennung gearbeitet wird. Eines davon ist das bereits erwähnte Google Instant, eine Funktion der Suchmaschine, die bereits während dem Vorgang des Eintippens eines Suchbegriffes Ergebnisse liefert und so einerseits Tippfehler korrigieren bzw. ignorieren kann, als auch die Geschwindigkeit und Effizienz der Suche insgesamt erhöht, sowie sofortiges und dynamisches Feedback ermöglicht. (vgl. Google 2010⁴⁸)

Ein weiteres Beispiel stellt eine Funktion der Suche nach Bildern dar. Während Bilder einerseits nach wie vor per Eingabe eines Suchbegriffes gesucht werden können, so ist es nun andererseits auch möglich, ähnliche Bilder, verglichen mit entweder einem Bild aus dem Internet, oder aber einem eigenen, selbst angegebenen Bild, beispielsweise einem Foto, zu finden. Google sucht dann nach optisch ähnlichen Bildern im Internet und kann darüber hinaus auch nach verschiedenen im Internet vorhandenen

47 <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241.html>, aufgerufen am 06.08.2012.

48 <http://www.google.com/insidesearch/features/instant/about.html>, aufgerufen am 06.08.2012.

Größen eines Bildes suchen. Diese beiden Beispiele zeigen also, dass Google teilweise bereits sehr wohl imstande ist, Muster zu erkennen, diese zu extrahieren und auf Ähnlichkeiten zu anderen Mustern zu überprüfen. (vgl. Google 2011a⁴⁹)

Als dritter und letzter Schritt ist für die Konstruktion einer intelligenten Maschine das Trainieren des Systems respektive der Erinnerungen speichernden Instanz notwendig, damit die Sinne, in welcher Art und Weise auch immer, rezipieren und ein eigenes Bild der Welt erstellen können. (vgl. Johnston 2008, 399f)

Während andere Suchmaschinen ebenfalls das Internet durchforsten, unabhängig davon ob deep web oder surface web, so hat Google ein bzw. viele Alleinstellungsmerkmale, was die Anzahl der Datenquellen betrifft. Google vereint nicht nur viele verschiedene Dienste als Unternehmen, sondern führt darüber hinaus seit dem Jahr 2012 auch die Nutzerdaten, die bei der Benutzung von fast jedem dieser Dienste durch Nutzerinnen und Nutzer gespeichert werden, zusammen.

Mit der Zusammenführung erlaubt sich Google nämlich auch, die Daten der Nutzer aus allen Services zusammenzulegen, sobald ein Nutzer einen Service benutzt. Über diese Bündelung der Nutzer-Informationen wird zum einen die besagte engere Verzahnung und die Bereitstellung von spezifischeren, für den Nutzer relevanten Informationen erreicht; zum anderen erlaubt es Google aber auch, seine Kunden noch genauer zu verstehen und so beispielsweise noch personalisiertere Anzeigen zu präsentieren. (Abdi 2012⁵⁰)

Diese gesammelten und zusammengeführten Nutzerdaten ermöglichen dem System Google also das Lernen und Trainieren in einem Ausmaß wie nie zuvor. Mit diesem letzten Schritt nach Johnston bzw. Hawkins kann also von Google als einer intelligenten Maschine, wenn auch im Anfangsstadium, gesprochen werden und so eine Zusammenführung der in den vorangegangenen Kapiteln behandelten Konzepte im Begriff der digitalen Intelligenz stattfinden.

49 <http://support.google.com/images/bin/answer.py?hl=de&answer=1325808>, aufgerufen am 06.08.2012.

50 <http://www.computerbase.de/news/2012-01/google-fuehrt-daten-aus-allen-diensten-zusammen/>, aufgerufen am 06.08.2012.

3.6 Google und das Problem der Komplexität

Google als digitale Intelligenz

Ganz im Sinne Luis von Ahns Human Computing, kombiniert mit dem Konzept des Crowdsourcing, Joseph Lickliders Vision einer Symbiose von Gehirn und Computer, dem Entstehen einer intelligenten Maschine nach Jeff Hawkins und auch John Johnston, Klaus Mainzers Vision der Kooperation mit Informationssystemen, der hybriden Intelligenz nach Alfred Spector und Eric Schmidt entsteht also mit Google eine digitale Intelligenz, die sich durch eine symbiotische Form der Intelligenz auszeichnet.

Symbiotisch meint dabei das Zusammenspiel und die Kombination der Fähigkeiten von sowohl Mensch, als auch Computer, ermöglicht durch die digitale Form von Interaktion im Medium Internet. Während in Kapitel 3.3 einerseits dargelegt wurde, dass Google von uns Menschen, durch unsere erstellten Inhalte, Interaktion und Kommunikation, lernt, so wurde der umgekehrte Sachverhalt, die zunehmende Computerisierung und die sich daraus ergebenden Umstände und Problemstellungen der Informationsgesellschaft, in Kapitel 2.2 eingehend beschrieben.

Im Sinne des deutschen Systemtheoretikers und Soziologen Niklas Luhmann kann hier wohl von einer Interpenetration der beiden Systeme Mensch und Computer, ermöglicht durch das Internet, gesprochen werden.

Von Penetration wollen wir sprechen, wenn ein System die eigene Komplexität [...] zum Aufbau eines anderen Systems zur Verfügung stellt. [...] Interpenetration liegt entsprechend dann vor, wenn dieser Sachverhalt wechselseitig gegeben ist, wenn also beide Systeme sich wechselseitig dadurch ermöglichen, daß das Verhalten des penetrierenden Systems durch das aufnehmende System mitbestimmt wird (und eventuell außerhalb dieses Systems orientierungslos und erratisch abläuft wie das einer Ameise ohne Kontakt zum Ameisenhaufen). (1984, 290)

Während beide Systeme jeweils unterschiedliche Komplexitäten nach Luhmann, beispielsweise Menge und Komplexität von Information auf der einen und schnelle, analytische Datenverarbeitung auf der anderen Seite, zur Verfügung stellen, so zeichnet das Konzept der digitalen Intelligenz für

beide Systeme ein Bild von sowohl Mitbestimmung des und durch den anderen, als auch einen gegenseitigen Nutzen insgesamt.

Denn während Google als Gesamtkomplex von menschlicher Kommunikation respektive Interaktion, sowie generierten, digitalen Inhalten lernen und damit in weiterer Folge eine eigene Wahrnehmung der Datenwelt im Internet entwickeln kann, so bietet die Suchmaschine Google, wie bereits erwähnt, schon jetzt eine funktionierende Lösung für das in Kapitel 2.2 identifizierte Problem der Menge.

Wenn sich sowohl die Zahl der Nutzerinnen und Nutzer wie auch der ausgeführten Nutzeraktionen nun stetig erweitert und damit das Lernpotenzial steigt, kann davon ausgegangen werden, dass bald auch eine immer umfassendere bzw. potentere Lösung für das zweite erkannte Problem der Komplexität von Information mithilfe dieser symbiotischen, digitalen Intelligenz realistisch ist.

Während PageRank, also der Algorithmus, um relevante Websites erkennen und filtern zu können, bereits ein praktischer Lösungsansatz für ein komplexes System von Verweisen im Internet darstellt, so muss in weiterer Folge die Adaption an andere komplexe Systeme, welcher Art auch immer, erfolgen, um tatsächlich die Rolle einer wertvollen Lösung für das Problem der Komplexität bestätigen und umfassender ausfüllen zu können.

Helge Ritter erkennt ebenfalls die Möglichkeit um nicht zu sagen Notwendigkeit, Erkenntnisse aus dem Forschungsfeld der künstlichen Intelligenz und der Hirnforschung für das Verständnis von vernetzten, komplexen Systemen außerhalb des Gehirns einzusetzen und zu nützen. In diesem Zusammenhang könnten die verschiedenen Herangehensweisen von Mensch und Computer in ergänzender Art und Weise von Nutzen sein, da für Menschen gerade in komplexen Systemen selbst simple Zusammenhänge oft unerkannt bleiben. (vgl. 2008, 131)

Der Umgang mit vielen Daten und Informationen könnte so auch „im Fall einer faktischen Nichtkonstruierbarkeit umfassender Modelle und eines Fehlens überschaubarer Gesetzessysteme dennoch intelligentes, auf ein Ziel hin erfolgreich strukturiertes Verhalten organisiert werden“. (Ritter 2008, 131)

Gerade diese momentan noch verhältnismäßig klare Trennung der unterschiedlichen Stärken beider Systeme macht auch für Eric Schmidt das enorme Potenzial für die Zukunft deutlich.

A reasonable model is that ten years from now computers will do what computers are really good at, which is doing deep analysis, remembering things exactly, accurately and so forth. And humans will do, what humans are really good at, which is intuition, feelings and so on. That separation of power is a reasonable one. The fact of the matter is that I think we're not quite ready for the application of modern Artificial Intelligence technology on top of that. (Schmidt 2010⁵¹)

Schmidt geht dabei, im Unterschied zu Ritter, einen Schritt weiter und sieht neben Computer und Mensch die Verbindung von deren Fähigkeiten mit einer künstlichen Intelligenz als eigentlich fortschrittliche Entwicklung, die für viele Menschen allerdings noch nicht vorstellbar ist. Schmidt geht dabei, wie bereits beschrieben, von eben jenem Begriff der digitalen Intelligenz aus, der durch die symbiotische Form der Interaktion miteinander entsteht und sieht Google in genau dieser Rolle.

Auch Klaus Mainzer, der diese Entwicklung einer Intelligenz wie bereits beschrieben ebenfalls als essenziell für die Zukunft der Informationsgesellschaft sieht, stellt fest, dass dieses Potenzial nicht in einem Gerät, „sondern vielmehr in der vernetzten Infrastruktur von verschiedenen Geräten, die eine intelligente Nutzerumgebung schaffen“ steckt. (2003, 175)

Um abschließend Googles Eignung für die Lösung des Problems der Komplexität beurteilen zu können, soll dieses noch einmal in Erinnerung gerufen werden.

51 <http://www.youtube.com/watch?v=UAcClSrAq70>, aufgerufen am 06.08.2012.

**Das Problem
der Komplexität
von Information**

Das in Kapitel 2.2 formulierte identifizierte Problem der Komplexität von Information stellt fest, dass in Verbindung mit der Menge an verfügbarer Information auch die Komplexität von Systemen - respektive auch Informationssystemen mit komplexen Zusammenhängen - steigt und auch aus diesem Grund für mündige Nutzer und Nutzerinnen ohne technische Hilfe weder momentan noch in Zukunft bewältigt werden kann.

Zusammenfassend kann zum Problem der Komplexität von Information daher festgehalten werden, dass Google in der Funktion einer digitalen Intelligenz bereits zu einem gewissen Grad eine Lösung anbieten kann. Dieser Befund muss deshalb eingeschränkt ausfallen, weil noch nicht alle Bereiche von Google abgedeckt werden, also erst eine Ausdehnung auf noch nicht erschlossene Bereiche, etwa Ökonomie, Soziologie oder Ökologie, erfolgen muss. Es liegt nunmehr daran, sowohl Funktionalitäten als auch Möglichkeiten optimal einzusetzen und auszuschöpfen.

Erkenntnisse, Résumé und Ausblick

Am Schluss der Untersuchung kann die eingangs gestellte Frage, die durch das Formulieren einer bzw. mehrerer Problemstellungen den Ausgangspunkt dieser Arbeit darstellt, klar beantwortet werden.

Forschungsfrage

Kann ein Wandel von künstlicher zu digitaler Intelligenz am Beispiel von Google als Konglomerat verschiedenster multimedialer Angebote festgestellt werden und wenn ja, welche Relevanz hätte diese digitale Intelligenz für eine Informationsgesellschaft?

Es kann ein Wandel von künstlicher zu digitaler Intelligenz am Beispiel Google festgestellt werden. Relevant ist diese Entwicklung insofern, als sie einen Lösungsansatz für die Problemstellungen, die sich in einer Informationsgesellschaft ergeben, bietet. Die Beantwortung dieser Frage schließt dabei die Beobachtungen, die aus den einzelnen Kapiteln

resultieren, ein und führt diese zusammen, daher sollen diese noch einmal kurz in Erinnerung gerufen werden.

In Kapitel 1 wurden die Begriffe Information und Wissen, sowie Dokumentation, Digitalisierung und Bedeutung dieser beiden für die Gesellschaft dargelegt, um einerseits die begrifflichen Grundlagen für die weitere Untersuchung zu schaffen, andererseits aber auch auf Information als Basis und Problemstellung für die nachfolgend behandelte Informationsgesellschaft hinzuführen.

Diese Informationsgesellschaft wurde in Kapitel 2 anhand verschiedener Merkmale charakterisiert und verschiedene theoretische Positionen und Herangehensweisen in diesem Zusammenhang erläutert. Im Prozess der Untersuchung dieser Informationsgesellschaft konnten konkret zwei Problemstellungen in Bezug auf Information, nämlich Menge und Komplexität von Information, identifiziert werden. Im weiteren Verlauf des Kapitels wurden sowohl theoretische, als auch praktische Lösungsansätze für diese beiden Probleme, etwa das Internet, Suchmaschinen, Crowdsourcing und künstliche Intelligenz, sowie deren Grenzen aufgezeigt.

In Kapitel 3 wurde der Versuch unternommen, Funktionsweise und Absichten von Google als fortschrittlichster Suchmaschine und Konglomerat verschiedenster medialer Dienste zu erörtern, sowie Googles Potenzial, durch Kombination und Weiterentwicklung der in Kapitel 2 vorgestellten Lösungsansätze im Konzept einer digitalen Intelligenz eine Antwort auf die beiden Problemstellungen der Menge und Komplexität von Information, gezeigt.

Wie in Kapitel 2 beschrieben, können also zunächst ein gesellschaftlicher Wandel hin zu einer Informationsgesellschaft und als Konsequenz zwei Problemstellungen festgestellt werden. Die anfängliche Vermutung, dass Google als Weiterentwicklung bzw. Erweiterung des Forschungsfeldes der künstlichen zu einer digitalen Intelligenz per Definition gelten kann, konnte nicht nur bestätigt, sondern auch die Relevanz dieser Untersuchung von Google als digitaler Intelligenz erfolgreich in Zusammenhang mit der Lösung der Problemstellungen einer Informationsgesellschaft gebracht werden.

Zusammengefasst kann also ein Wandel von künstlicher zu digitaler Intelligenz am Beispiel Google beobachtet werden. Die Relevanz für die Informationsgesellschaft ist primär das enorme Potenzial, das Google als mögliche Lösung in Bezug auf die beiden formulierten Probleme der Menge und Komplexität von Information bietet.

Die Lösung und Bewältigung dieser beiden Probleme der Menge und Komplexität von Information sind essenziell für die Zukunft der Informationsgesellschaft, in der wir uns befinden. Klaus Mainzer fasst diese Notwendigkeit der Entwicklung von Lösungsstrategien ebenfalls zusammen.

Der Benutzer des Internets nimmt den Datenstau der Datenströme nur wahr, wenn sich die Informations- und Kommunikationssysteme verlangsamen oder gar zusammenbrechen. Ähnlich nehmen wir in unserem Körper die komplexen Datenströme im Nervensystem nicht bewusst wahr. Was wir aber in der Informationsgesellschaft beobachten, ist die zunehmende Flut von Informationen analog einer Überreizung des Nervensystems durch Geräusche, Bilder und andere Sinnesdaten. Das Internet als Zentralnervensystem (World Wide Web) der Informationsgesellschaft wird daher intelligente Verfahren zur Informationsfilterung und -bewältigung ähnlich wie ein lebender Organismus benötigen. Um die Schnittstelle mit dem Internet möglichst nutzerfreundlich zu gestalten, sollten sich diese Computerverfahren an der menschlichen Informations- und Kommunikationsbewältigung orientieren. Gemeint sind fehlertolerante, flexible und fuzzy Lern- und Selektionsverfahren, wie sie im Soft Computing angestrebt werden. (2003, 162)

Wir als Nutzerinnen und Nutzer des Internets sind überlastet und benötigen technische Hilfe, um gemeinsam Information filtern und bewältigen zu können. Diese Formulierung kann als Bestätigung einerseits des Bestehens der beiden Probleme der Menge und Komplexität, andererseits auch der notwendigen Entwicklung eines am Menschen orientierten Verfahrens, um diese erfolgreich lösen zu können, aufgefasst werden. Hierbei ist die Lösung in Form von Google also nicht nur am Mensch orientiert, sondern darüber hinaus auch eine lernende, symbiotische Kombination von Mensch und Computer bzw. deren Fähigkeiten.

Eine der größten Herausforderungen der Informationsgesellschaft liegt im „Umgang mit Technologien, deren Folgen schneller eintreten, als dies

intellektuell auf Basis vorhandener Erkenntnismodelle verarbeitet werden kann“, wie bereits eingangs in Kapitel 2 festgehalten wurde. (Bienert 1998, 2)

Auf Google trifft dies nun in besonderem Maße und in unterschiedlichen Zusammenhängen zu, nicht nur als Suchmaschine oder Unternehmen, sondern als Sinnbild für eine Gesellschaft, die einen immer größer werdenden Teil des Lebens im Umgang mit dem Internet in all seinen Facetten verbringt.

Interessant sind beispielsweise mögliche Auswirkungen auf soziale Formen des Umgangs miteinander, etwa Freundschaften und Beziehungen. Google oder auch Facebook schaffen die Möglichkeit, Freunde über ein Leben hinweg nicht aus den Augen zu verlieren, vernetzt zu bleiben und mit all den verfügbaren Medienformen eine zumindest virtuelle Nähe zu erhalten. Dieser Umstand bedeutet aber auch, dass speziell Google jeder Nutzerin und jedem Nutzer Einblicke in eben diese digitalen Spuren, die im Lauf dieses virtuellen Daseins hinterlassen wurden, gewähren kann. Die virtuelle Vergangenheit resultierend aus der Menge an Information und Inhalten, die zu einem großen Teil von Nutzerinnen und Nutzern selbst erstellt wurden, ob Fotos, Videos, Text oder vieles mehr, kann damit reale Personen einholen. (vgl. Jarvis 2009, 231)

In diesem Zusammenhang soll ein möglichst neutrales Bild von Google, weder als an die Wand gemalter Teufel, noch übermächtiges Allheilmittel, vermittelt werden. Es muss dabei aber auf die Tatsache hingewiesen werden, dass in einer Gesellschaft, die Information als fundamentale Ressource ansieht, Google als Instanz zwischen Mensch und Information eine kritisch zu betrachtende Rolle und Position einnimmt, die nicht unterschätzt werden sollte.

Erwähnt soll hierbei auch werden, dass Google in verschiedenen Projekten Möglichkeiten erforscht, um nicht nur zwischen Mensch und Information, sondern auch zwischen Mensch und Umwelt stehen bzw. vermitteln zu können, also digitale Information von der virtuellen Welt des Computers zu lösen und für die physische Welt des Alltags nutzbar zu machen. An dieser

Stelle kann weiterführend auf Googles Smartphones, vor allem aber auf das ambitionierte Project Glass verwiesen werden. (vgl. Manjoo 2012⁵²)

Diese Verbindung von Natur und Technik ist darüber hinaus auch der thematische Kern des Masterprojekts „Of Light & Shadow“, das parallel zu dieser Arbeit entstand. Dabei handelt es sich um ein Computerspiel, das von einem fachbereichs- und studiengangübergreifenden Team zwischen September 2010 und Juli 2012 entwickelt wurde. Das Medium Spiel zeichnet dabei einerseits die hohe Interaktivität, ja notwendige Interaktion mit anderen Spielern, oder aber dem Computer, sowie die Zusammenführung mehrerer verschiedener audiovisueller Medienformen aus. Ganz im Sinne Flussers, wie in Kapitel 2.3.b dargelegt, können Spiele als die Erschaffung alternativer wenngleich virtueller Welten gesehen werden. Das Internet ermöglichte in diesem Zusammenhang gerade in den vergangenen beiden Jahrzehnten den Fokus vieler Spiele auf die Interaktion von vielen verschiedenen Menschen miteinander und spiegelt so den Zeitgeist der Informationsgesellschaft wider. Sowohl Werk als auch These haben dabei den Wert eines interaktiven Zugangs zu verschiedensten Medienformen gezeigt und den Einblick gegeben, dass nicht nur eingehendes Befassen mit etwas, sondern auch ausgedehnte Interaktion, ob mit einem Menschen oder einem Computer, tiefgründige Gedanken ermöglichen kann. (vgl. Flusser 1997, 213)

Die Folgen und Konsequenzen, die eine digitale Intelligenz wie Google mit sich bringt, sind in Anlehnung an Bienert wohl noch nicht in ihrem vollen Umfang abzusehen. Im Laufe dieser Arbeit wurde aber gezeigt, dass nicht nur die Möglichkeiten und Chancen für die Informationsgesellschaft vorhanden, sondern kritisch und essenziell für eine erfolgreiche Herangehensweise an Problemstellungen, die sich damit verbunden eröffnen, sind. Während im Verlauf der Untersuchung Referenzen und Verweise auf Werke der Science Fiction, obwohl thematisch naheliegend, aus Gründen der Wissenschaftlichkeit vermieden wurden, darf ein abschließendes Zitat verziehen werden.

"When our machines overtook us, too complex and efficient for us to control, they did it so fast and so smoothly and so usefully, only a fool or a prophet would have dared complain." (Ings 1999, 2)

52 <http://www.technologyreview.com/review/428212/you-will-want-google-goggles/>, aufgerufen am 06.08.2012.

Quellen

Bücher

Bienert, Peter (1998): Information & Kommunikation. Technik und Anwendung in Wirtschaft und Medien. Berlin: Springer Verlag.

Carr, Nicholas (2009): The big switch. Rewiring the world, from Edison to Google. New York: Norton & Company, Inc.

Descartes, Rene (1906): Regeln zur Leitung des Geistes. Leipzig: Meiner.

Flusser, Vilem (1997): Medienkultur. Frankfurt am Main: Fischer.

Foucault, Michel (1974): Die Ordnung der Dinge. Frankfurt am Main: Suhrkamp Verlag.

Giddens, Anthony (1996): Konsequenzen der Moderne. Frankfurt am Main: Suhrkamp.

Hartmann, Frank (1999): Cyber.Philosophy. Medientheoretische Auslotungen. Wien: Passagen Verlag.

Hartmann, Frank (2006): Globale Medienkultur. Technik, Geschichte, Theorien. Wien: Facultas Verlags- und Buchhandels AG.

Hartmann, Frank (2008a): Medien und Kommunikation. Wien: Facultas Verlags- und Buchhandels AG.

Hartmann, Frank (2008b): Multimedia. Wien: Facultas Verlags- und Buchhandels AG.

Hawkins, Jeff (2004): On Intelligence. New York: Times Books.

Haykin, Simon (2009): Neural Networks and Learning Machines. New Jersey: Pearson Education, Inc.

Ings, Simon (1999): Headlong. London: HarperCollins Publishers.

Janich, Peter (2006): Was ist Information? Frankfurt am Main: Suhrkamp.

Jeff, Jarvis (2009): What Would Google Do? New York: HarperCollins Publishers.

Johnston, John (2008): The Allure of Machinic Life. Cybernetics, Artificial Life, and the New AI. Cambridge: MIT Press.

Luhmann, Niklas (1984): *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt am Main: Suhrkamp.

Mainzer, Klaus (2003): *Computerphilosophie zur Einführung*. Hamburg: Junius Verlag.

McLuhan, Marshall (1964): *Understanding Media: The Extensions of Man*. Cambridge: MIT Press.

Negroponte, Nicholas (1996): *Being Digital*. New York: Vintage Books.

Ruffing, Rainer (2005): *Einführung in die Philosophie der Gegenwart*. Paderborn: Wilhelm Fink Verlag.

Searle, John Rogers (1986): *Geist, Hirn und Wissenschaft*. Frankfurt am Main: Suhrkamp.

Standage, Tom (2002): *Der Türke. Die Geschichte des ersten Schachautomaten und seiner abenteuerlichen Reise um die Welt*. Frankfurt/Main: Campus Verlag GmbH.

Wiener, Norbert (1961): *Cybernetics, or control and communication in the animal and the machine*. Cambridge: MIT Press.

Zimbardo, Philip G. (1988): *Psychologie*. Berlin: Springer Verlag.

Beiträge in Sammelbänden

Campbell, Murray (1997): *An Enjoyable Game. How HAL Plays Chess*. In: *HAL's Legacy. 2001's Computer as Dream and Reality*. Cambridge: MIT Press.

Krug, Thomas / Lehmann, Kai / Schetsche, Michael (2005) *Zehn Prinzipien der neuen Wissensordnung*. In: *Lehmann, Kai / Schetsche, Michael (2005): Die Google Gesellschaft. Vom digitalen Wandel des Wissens*. Bielefeld: transcript Verlag, 17-32.

Lehmann, Kai (2005) *Der lange Weg zur Wissensgesellschaft*. In: *Lehmann, Kai / Schetsche, Michael (2005): Die Google Gesellschaft. Vom digitalen Wandel des Wissens*. Bielefeld: transcript Verlag, 33-39.

Patzwaldt, Klaus (2005) *Suchmaschinenlandschaften*. In: *Lehmann, Kai / Schetsche, Michael (2005): Die Google Gesellschaft. Vom digitalen Wandel des Wissens*. Bielefeld: transcript Verlag, 75-82.

Plaß, Christine (2005) *Das große Vergessen. Datenschwund im digitalen Zeitalter*. In: *Lehmann, Kai / Schetsche, Michael (2005): Die Google Gesellschaft. Vom digitalen Wandel des Wissens*. Bielefeld: transcript Verlag, 41-46.

Rayward, W. Boyd (1994) *The International Federation for Information and Documentation*. In: Wiegand, Wayne A. / Davis, Don G. (1994): *Encyclopedia of Library History*. New York: Garland Press.

Ritter, Helge (2008): *Die neuen Denkmuster der künstlichen Intelligenz*. In: Mutius, Bernhard von (Hg.): *Die andere Intelligenz. Wie wir morgen denken werden*. Stuttgart: J.G. Cotta'sche Buchhandlung Nachfolger GmbH, 118-131.

Zeitschriftenaufsätze

Bergman, Michael K. (2001): *The Deep Web. Surfacing Hidden Value*. In: *The Journal of Electronic Publishing* 7/2001, 45-62.

Buchanan, Bruce G. (2006): *A (Very) Brief History of Artificial Intelligence*. In: *AI Magazine* 4/2006, 53-60.

Heuer, Steffan / Kern, Thomas (2005): *Das ganze Netz im Online-Archiv*. In: *Technology Review* 1/2005, 92-95.

Onlinequellen

Abdi, Sasan (2012): *Google führt Daten aus allen Diensten zusammen*. In: <http://www.computerbase.de/news/2012-01/google-fuehrt-daten-aus-allen-diensten-zusammen/>, aufgerufen am 06.08.2012.

Bendrath, Ralf (2003): *Charta der Bürgerrechte für eine nachhaltige Wissensgesellschaft*. In: http://www.worldsummit2005.de/download_de/Charta-Flyer-deutsch.pdf, aufgerufen am 06.08.2012.

Duden (2004): *Googeln. Schwaches Verb*. In: <http://www.duden.de/rechtschreibung/googeln>, aufgerufen am 06.08.2012.

Dyess, Ken / Shaw, Susan / Williams, Angela (1996): *The History of Artificial Intelligence*. In: <http://library.thinkquest.org/2705/>, aufgerufen am 06.08.2012.

Dyson, George (2005): *Turing's Cathedral. A visit to Google on the occasion of the 60th anniversary of John von Neumann's proposal for a digital computer*. In: http://www.edge.org/3rd_culture/dyson05/dyson05_index.html, aufgerufen am 06.08.2012.

- Efrati, Amir (2011): Google Notches One Billion Unique Visitors Per Month. In: <http://blogs.wsj.com/digits/2011/06/21/google-notches-one-billion-unique-visitors-per-month/?mod=e2tw>, aufgerufen am 06.08.2012.
- Flickr (2012): FAQ. Getty Images. In: <http://www.flickr.com/help/gettyimages/>, aufgerufen am 06.08.2012.
- Google (2010): Über Google Instant. In: <http://www.google.com/insidesearch/features/instant/about.html>, aufgerufen am 06.08.2012.
- Google (2011a): Suche anhand von Bildern. In: <http://support.google.com/images/bin/answer.py?hl=de&answer=1325808>, aufgerufen am 06.08.2012.
- Google (2011b): Über die Google Buchsuche. In: <http://books.google.at/intl/de/googlebooks/history.html>, aufgerufen am 06.08.2012.
- Google (2012a): 2012 Financial Tables. In: <http://investor.google.com/financial/tables.html>, aufgerufen am 06.08.2012.
- Google (2012b): Unternehmensgeschichte im Detail. In: <https://www.google.at/intl/de/about/company/history/>, aufgerufen am 06.08.2012.
- Google (2012c): Unsere Zehn Grundsätze. In: <http://www.google.com/about/company/philosophy/>, aufgerufen am 06.08.2012.
- Google (2012d): Google Basics. Learn how Google discovers, crawls, and serves web pages. In: <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=70897&topic=2370570&ctx=topic>, aufgerufen am 06.08.2012.
- Google (2012e): Erfahren Sie, wie unsere Übersetzungen zustande kommen. In: http://translate.google.at/about/intl/de_ALL/, aufgerufen am 06.08.2012.
- Jenkins, Holman W. Jr (2010): Google and the Search for the Future. The Web icon's CEO on the mobile computing revolution, the future of newspapers, and privacy in the digital age. In: <http://online.wsj.com/article/SB10001424052748704901104575423294099527212.html>, aufgerufen am 06.08.2012.
- Leiner, Barry M. / Cerf, Vinton G. / Clark, David D. / Kahn, Robert E. / Kleinrock, Leonard / Lynch, Daniel C. / Postel, Jon / Roberts, Larry G. / Wolff, Stephen (1997): Brief History of the Internet. In: <http://www.internetsociety.org/internet/internet-51/history-internet/brief-history-internet#f3>, aufgerufen am 06.08.2012.
- Manjoo, Farhad (2012): You Will Want Google Goggles. In: <http://www.technologyreview.com/review/428212/you-will-want-google-goggles/>,

aufgerufen am 06.08.2012.

Schwartz, Barry (2012): Addictive Google Image Labeler Game Gone. In: <http://www.seroundtable.com/google-image-labeler-dead-14663.html>, aufgerufen am 06.08.2012.

Simonite, Tom (2011): Wir können in vitro arbeiten. In: <http://www.heise.de/tr/artikel/Wir-koennen-in-vitro-arbeiten-1350475.html>, aufgerufen am 06.08.2012.

Sumako (2011b): 21v. Suchmaschinen Studie. In: <http://www.21v.de/>, aufgerufen am 06.08.2012.

The Internet Archive (2012): About The Internet Archive. In: <http://archive.org/about/>, aufgerufen am 06.08.2012.

Türpe, Sven (2012): IMHO: Warum Googles Datensammeln gar nicht so böse ist. In: <http://www.golem.de/news/imho-warum-googles-datensammeln-gar-nicht-so-boese-ist-1203-90241.html>, aufgerufen am 06.08.2012.

Von Ahn, Luis (2006): Luis von Ahn. In: <http://www.cs.cmu.edu/~biglou/>, aufgerufen am 06.08.2012.

WikiMedia Foundation (2012): About Wikipedia. In: <http://en.wikipedia.org/wiki/Wikipedia:About>, aufgerufen am 06.08.2012.

Zandy, Ashley (2012): Facebook Reports Second Quarter 2012 Results. In: <http://investor.fb.com/releasedetail.cfm?ReleaseID=695976>, aufgerufen am 06.08.2012.

Audiovisuelle Medien

Eric Schmidt at Techonomy (2010): Hopkins, Debby / Kelly, Kevin / Kirkpatrick, David / Randall, Lisa / Schmidt, Eric. YouTube Video, 38:20 Min. In: <http://www.youtube.com/watch?v=UAcClSrAq70>, aufgerufen am 06.08.2012.

Human Computation (2006): Von Ahn, Luis. Google TechTalks Video, 51:31 Min. In: <http://video.google.com/videoplay?docid=-8246463980976635143>. Aufgerufen am 06.08.2012.