

Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

SUSANNE GOTTWALD

Die Wissensberechnungsmaschine Wolfram|Alpha

Die Wissensberechnungsmaschine Wolfram|Alpha Report

Susanne Gottwald (gottwald@zib.de)
Konrad-Zuse-Zentrum für
Informationstechnik Berlin
Unit Scientific Computing
Scientific Information Systems
Takustr. 7, 14195 Berlin
www.zib.de

Abstract: Im Mai 2009 wurde Wolfram|Alpha gestartet, ein Service, der seinen Namen von seinem Entwickler, dem britischen Mathematiker Stephen Wolfram, ableitet. Dem Benutzer soll nicht nur eine Liste von Webseiten als Ergebnis auf Anfragen geliefert werden, sondern Antworten auf konkrete Fragen geben. In diesem Report soll gezeigt werden, warum sich Wolfram|Alpha von Suchmaschinen abgrenzt und was die Berechnung von Antworten auf natürlichsprachliche Fragen möglich machen kann.

1 Einführung

Wolfram|Alpha wurde bereits vor dem Start eine enorme mediale Aufmerksamkeit gewidmet. Stephen Wolfram kündigte seinen Dienst in einem Blogeintrag schließlich auch mit den Worten an, dass Wolfram|Alpha das “neue Paradigma zur Benutzung von Computern und des Webs“ sei und dass die Vision des Computers von vor 50 Jahren wahr wird: Computer als Wissensberechnungsmaschinen[Wol09b]. In Wolframs Unternehmen Wolfram Research wird der neue Dienst am Ende der geschichtlichen Entwicklung berechenbaren Wissens eingeordnet. Am Ende einer Timeline, die 20.000 Jahre v.Chr. mit der frühen Arithmetik beginnt und Höhepunkte auflistet wie die Turingmaschine, Künstliche Intelligenz, relationale Datenbanken, das Web und darüber hinaus andere Wolfram Research Produkte wie Mathematica[Res09b].

Wolfram|Alpha als Wissensberechnungsmaschine erhält als Eingabe eine natürlichsprachliche Frage. Wolfram ist in [Wol09b] der Meinung, da Menschen miteinander in natürlicher Sprache kommunizieren, sei dies bei dem Umgang mit dem ganzen Spektrum an Wissen der einzig realistische Weg, auch mit Computern zu kommunizieren. Er schreibt weiter „[...] nimm Fragen, die Menschen in natürlicher Sprache stellen und repräsentiere sie in einer präzisen Form, die in die Berechnungen passt.“. Die Grundlagen dafür nennt er einen „Mix aus vielen cleveren Algorithmen und Heuristiken, sehr vielen linguistischen Entdeckungen und Kurationen und einigen ernststen theoretischen Durchbrüchen“. Sogenannte NLQs (natural language queries) wurden bereits vor Wolfram|Alpha untersucht und zur Nutzung bei Suchmaschinen in Betracht gezogen (siehe [Kor97]). Sie werden dort

definiert als „Anfragen, in denen der Benutzer eine Frage oder eine Beschreibung formuliert.“ Die Suchmaschine muss dann alle indexierten Terme für eine Suche extrahieren. Manche Wörter werden eliminiert, da sie keine Bedeutung tragen. Das Problem bei diesem ersten Ansatz war allerdings, dass „Computer es schwer haben, einen Term mit seinem syntaktischen, semantischen und pragmatischen Wissen korrekt zu extrahieren.“ [Kor97] Das bedeutet, dass der Term seinen Kontext nach der Extraktion verliert. Wolfram|Alpha zeigt dem Benutzer immer an, wie bzw. mit welchem Kontext die Eingabe verstanden wurde und gibt in manchen Fällen Hinweise auf den alternativen Gebrauch der erkannten Wörter. In den FAQs [Res09c] steht, dass man nicht notwendigerweise eine Frage formulieren muss, sondern am besten so wenig Stichworte wie notwendig verwendet.

Aus der Eingabe wird dann eine einzige Antwort aus der vorhandenen Datenbasis berechnet. Wolfram|Alpha beantwortet Fragen über „alle Arten von methodischem objektivem Wissen“ [Res09c]. Hauptschwerpunkt bildet aber nach eigenen Angaben die Mathematik. Darüberhinaus soll Wolfram|Alpha alles kennen, das bekannt und veröffentlicht ist [Res09c]. Grenzen lägen einzig bei Meinungen: „It only deals with facts, not opinions.“ Die Liste der abfragbaren Themen ist dem entsprechend vielfältig und wächst seit dem Start stetig, wie in Tabelle 1 zu sehen ist.

Mathematik	Statistik und Datenanalyse	Physik
Chemie	Materialien, Farben	Entwicklung
Astronomie	Geowissenschaften	Biowissenschaften
Technik	Transport	Informatik
Web und Rechnersysteme	Einheiten und Maße	Geld und Finanzen
Datum und Zeit	Orte und Geographie	Sozioökonomie
Wetter	Gesundheit und Medizin	Nahrung und Ernährung
Kultur und Medien	Menschen und Geschichte	Bildung
Organisationen	Sport und Spiele	Musik

Tabelle 1: Bereiche, aus denen Wolfram|Alpha Antworten berechnen kann

1.1 Suchmaschinen

In den FAQs bei Wolfram|Alpha [Res09c] beantwortet Stephen Wolfram die Frage „Ist Wolfram|Alpha eine Suchmaschine?“ mit „No. It’s a computational knowledge engine: it generates output by doing computations from its own internal knowledge base, instead of searching the web and returning links.“ Wolfram|Alpha ist deshalb keine Suchmaschine, weil die Ergebnisse auf ganz andere Weise erzeugt werden als bei herkömmlichen Suchmaschinen.

In [Bri09] wird definiert, was eine Suchmaschine ist: „ein Computerprogramm, das Ant-

worten auf Anfragen in einer Sammlung von Informationen finden soll. Die Sammlung kann ein Bibliothekskatalog, eine Datenbank oder das World Wide Web sein. Eine Suchmaschine für das Web generiert eine Liste von Seiten, also Dokumenten im Web, die die Terme der Anfrage enthalten.“ Abbildung 1 zeigt, was Wolfram|Alpha über eine Suchmaschine weiß. Bei der Anfrage konnte beobachtet werden, dass die komplette Frage als

The image shows the Wolfram|Alpha interface for the query "what is a search engine?". At the top is the Wolfram|Alpha logo with the tagline "computational knowledge engine". Below the logo is a search bar containing the text "what is a search engine?". Underneath the search bar, the input is interpreted as "search engine (English phrase)". A definition is provided: "search engine" is defined as "a computer program that retrieves documents or files or data from a database or from a computer network (especially from the internet)". Below the definition, narrower terms are listed: "Ask Jeeves | Google | Yahoo". Broader terms are listed as: "computer program | computer programme | program | programme".

Abbildung 1: Frage an Wolfram|Alpha "What is a search engine?" vom 29.10.2009

allgemeines Thema verstanden wurde und man auf die Seite der Beispiele weitergeleitet wurde. Erst bei der Wahl der Alternative „Als Wort benutzen“ wurde das gezeigte Ergebnis produziert. Als „nähere Bezeichnung“ für Suchmaschine gibt Wolfram|Alpha Ask Jeeves, Google und Yahoo an. Dieser Teil der Antwort ist unvollständig, denn es werden lediglich diese drei individuelle Suchmaschinen aufgezählt, wobei sehr viel mehr existieren. Auch die Reihenfolge gibt keinen Hinweis auf die allgemeine Wichtigkeit der genannten Suchmaschinen. Laut [Gmb99] ist Google im Juni 2009 mit 81,22% die weltweit meistgenutzte Suchmaschine. Dies deckt sich auch mit den Ergebnissen anderer Statistiken.

Der Start von Wolfram|Alpha hat auch deshalb so viel Aufsehen erregt, weil in den Medien der kommende „Google-Killer“ erwartet wurde[Stö09]. Die letzten populären Versuche von Mitbewerbern, Googles Marktanteil zu senken, liegen ebenfalls nicht lange zurück. Im Juni 2009 startete Microsofts LiveSearch Nachfolger BING¹, im Juli 2008 versuchten es ehemalige Google-Mitarbeiter mit CUIL². Beide betonen, anders zu sein: CUIL habe den größeren Index und speichere keinerlei Nutzerdaten und BING sei keine Suchmaschine sondern eine Entscheidungsmaschine. Sie liefere gezieltere Ergebnisse beispielsweise für

¹<http://www.bing.de> gesehen: 20.10.2009

²<http://www.cuil.com> gesehen: 20.10.2009



Abbildung 2: Startseiten von CUIL und BING, die trotz Hintergrundfarben an Google erinnern

die Suche nach einem Flug an einem bestimmten Tag. Der Benutzer soll direkt auf Seiten weitergeleitet werden, auf denen er entsprechende Angebote findet. Allerdings wird dem Benutzer auch die Entscheidung abgenommen, bestimmte Anfragen zu stellen, da die Suche zum Beispiel nach „Beate Uhse“ lediglich mit folgendem Satz beantwortet wird „Der Suchbegriff führt möglicherweise zu sexuell eindeutigen Inhalten.“. Informationen über die Person an sich werden bei BING nicht zugänglich gemacht.

Bei beiden Konkurrenten erscheinen dennoch Benutzung, Layout sowie Präsentation der Ergebnisse wie bei jeder herkömmlichen Suchmaschine, darunter besonders Google, bekannt (siehe Abbildung 2).

Beide Services produzieren nach der bekannten Verfahrensweise einer Suchmaschine Ergebnisse. In [Kor97] wird aufgezeigt, wie diese aussieht. Zuerst wird ein Volltextindex der gesammelten Seiten erstellt. Dabei werden alle Wörter gefundener Dokumente in einem Index aufgenommen, bis auf Stoppwörter, also Wörter ohne Bedeutung, wie z.B. und, oder, so. Anfragen an die Suchmaschine werden von dieser an den Volltextindex weitergeleitet, der auf das Vorhandensein der gesuchten Wörter überprüft wird. Die gefundenen Seiten, die die Suchworte enthalten, werden als Ergebnisse nach einem bestimmten Verfahren geordnet („Ranking“). Die Verweise auf die Ursprungsdokumente werden schließlich in einer Liste an den Benutzer ausgeliefert. Google arbeitet ebenso nach diesem Prinzip und verfügt über einen Index, der im Juli 2008 die Marke von einer Trillion Webseiten überschritten haben

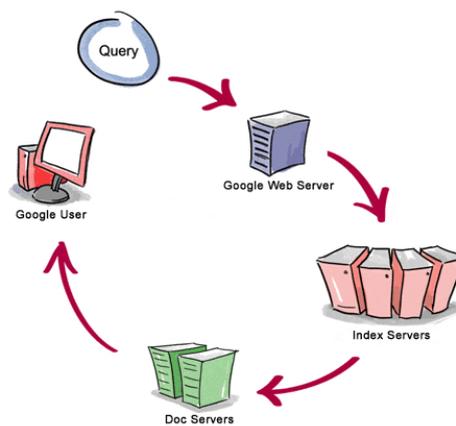


Abbildung 3: Ablauf einer Google-Suche

soll³. Abbildung 3 zeigt den schematischen Ablauf bei der Suche mit Google⁴. Eine Abfrage wird vom Google Web Server an die diversen Index Server weitergeleitet, bei denen überprüft wird, welche Seiten relevant für die gesuchten Begriffe sind. Diese werden schliesslich aus den Dokumentenservern herausgesucht, nach dem PageRank-Verfahren sortiert und dem Benutzer als Ergebnis angezeigt. Wenn man bei Google eine Frage eingibt, dann wird dem Benutzer also der Ort gezeigt, an denen er wahrscheinlich eine Antwort finden wird.

Such-string	Ergebnis Google.com	Ergebnis Wolfram Alpha
A) Sydney New York	1 - Local business results for Sydney near New York, NY (diverse Ärzte mit Namen Sydney) 2 - Cheap Flights to New York, USA from Sydney (airfaresflights.com) 3 - Flights New+york Sydney (edreams.com)	Input interpretation: Sydney, New South Wales, Australia New York, United States Populations, Path, Distance Current local time, Approximate elevations
B) php	1 - PHP: Hypertext Preprocessor php.net 2 - PHP: Downloads (php.net/downloads.php) 3 - PHP - Wikipedia, the free encyclopedia	Input interpretation: Php (Philippine peso) Local currency conversion for Php 1: 1.43c (eurocents) Exchange history Currency conversions
C) light bulb	1 - Image results 2 - Incandescent light bulb - Wikipedia, the free encyclopedia 3 - Light Bulb History - Invention of the Light Bulb (ideafinder.com)	Input interpretation: light bulb (English phrase) Definiton, Noun: electric lamp consisting of [..] Synonyms: bulb, electric light[.. Broader terms, Synonym network
D) Aspirin Tylenol	1 - WikiAnswers - Can you take Aspirin and Tylenol together 2 - WikiAnswers - Does Tylenol contain Aspirin 3 - Differences Between Aspirin , Tylenol and Advil (hubpages.com)	Input interpretation: aspirin acetaminophen Structure Diagrams, 3D structure Basic, Thermodynamic, Solid, Safety, Toxicity properties Chemical identifiers, NFPA label

Tabelle 2: Vergleich der Ergebnisse bei Google und Wolfram|Alpha

Vergleicht man diese Funktionsweise und die Ergebnisse in Tabelle 2 mit der von

³<http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html> gesehen: 21.10.2009

⁴Quelle: <http://www.google.de/intl/de/corporate/tech.html>

Wolfram|Alpha, kann man erkennen, dass Wolfram|Alpha keine Suchmaschine ist. Stephen Wolfram selber schreibt dazu in [Wol09b], dass "Suchmaschinen sehr effizient nach spezifischen Termen und Phrasen in in Billionen von Textseiten suchen können" aber "damit nicht berechnet werden kann." Wolfram|Alpha zeigt nicht, wo man die Antwort finden kann, sondern gibt im Ergebnis die Antwort auf die Frage. Das sind zwei gänzlich verschiedene Ansätze bezüglich ihrer Technologie und Vergleiche zwischen beiden dienen lediglich der Bewertung der Funktionalität von Wolfram|Alpha. Angelehnt an [Tal09] erfolgt in Tabelle 2 ein Gegenüberstellen der Ergebnisse beider Services auf dieselben Anfragen (durchgeführt am 21.10.2009). Da Wolfram|Alpha hauptsächlich englisch sprachliche Inhalte besitzt, wurde zur Suche auch Google.com verwendet. Es werden jeweils die ersten drei Ergebnisse von Google sowie eine Zusammenfassung der einzelnen "Pods"(einzelne Teile, Fenster einer Antwort) bei Wolfram|Alpha dargestellt.

Tauscht man bei Suche A) Sydney gegen Berlin aus, nimmt Wolfram|Alpha einen Ort in der Nähe von New York an und andere Orte werden nicht vorgeschlagen. Sucht man nur nach Berlin, erhält man Daten der deutschen Hauptstadt wie "Location, Populations, Time", aber der Vergleich mit New York kommt erst dann zustande, wenn man zwischen den beiden Städten noch ein "vs." einfügt, was bei dem Vergleich "Sydney New York" nicht notwendig war. Es erscheint, dass Wolfram|Alpha keinen einheitlichen Umgang mit Eingaben betreibt. Vergleich B) zeigt, dass bei der semantischen Analyse der Eingabe noch gearbeitet werden muss bzw. die Datenbasis erweitert werden sollte. PHP als Skriptsprache kennt Wolfram|Alpha auch nach mehreren umformulierten Suchanfragen nicht. Dass seit Mai 2009 aber an der Erweiterung der Datenbasis gearbeitet wurde, kann man in Beispiel C sehen. Im ersten Test kurz nach dem Start des Services lautete das Ergebnis dieser Anfrage *Wolfram|Alpha isn't sure what to do with your input.*⁵ In aktuellen Test kann Wolfram|Alpha den Begriff grob erläutern. Stephen Wolfram schreibt in [Wol09a], dass der Anteil der nicht verstandenen Abfragen im August 2009 bei 10% lag. Die Ursache dafür läge aber nicht bei fehlenden Daten oder fehlenden Berechnungsmethoden, sondern daran, dass die Maschine nicht versteht, was gefragt wurde.

Aus dem Vergleich wird ebenfalls die unterschiedliche Funktionalität zwischen Google und Wolfram|Alpha sichtbar. Während Google dem Benutzer als Ergebnis die Seiten aus seinem Index auflistet, die die Keywords der Suchanfrage und damit eventuell die Antwort auf die Frage enthalten, reagiert Wolfram|Alpha mit der zu den interpretierten Keywords passenden Anzahl von vorhandenen Fakten, was in der Gesamtheit als eine einzige Antwort verstanden werden muss.

Im weiteren Verlauf des Kapitels sollen Ziele und der Nutzen von Wolfram|Alpha aufgezeigt werden. In Kapitel 2 werden der mögliche Aufbau und die Funktion von Wolfram|Alpha unter Beleuchtung seiner Komponenten und Arbeitsweise skizziert. Außerdem wird in diesem Teil die nicht außer Acht zu lassende Problematik der Datengrundlage dargelegt. Diesem Kapitel schließt sich ein Ausblick an, bei dem der aktuelle Stand bewertet und eine zukünftige Entwicklung abgesehen werden sollen.

⁵Mittlerweile lautet die Meldung: "Wolfram|Alpha isn't sure how to compute an answer from your input."

1.2 Ziele

Das Unternehmen Wolfram Research benennt seine Ziele auf seiner Webseite [Res09a] ganz direkt. Das langfristige Ziel sei es, das ganze methodische Wissen berechenbar und für jeden zugänglich zu machen. Wolfram|Alpha ziele darauf ab, alle objektiven Daten zu sammeln und zu kuratieren (lt. Duden: betreuen) und jedes bekannte Modell, jede Methode und jeden Algorithmus zu implementieren. Was auch immer berechnet werden kann, soll berechnet werden können. Basierend auf wissenschaftlichen Ergebnissen und methodischem Wissen soll Wolfram|Alpha als Datenquelle entwickelt werden, der von jedem vertraut werden kann und die definitive Antworten auf sachbezogene Anfragen gibt. In den FAQs [Res09c] wird das Ziel mit „Expertenwissen für jedermann“ benannt. Peet Morris schreibt in [Mor09], das „ultimative Ziel von Wolfram|Alpha ist es, eine natürlichsprachliche Wissensberechnungsmaschine zu werden“.

Stephan Wolfram erwähnt in seiner Wolfram|Alpha Ankündigung [Wol09b] mit Mathematica und NKS (A New Kind of Science) zwei Wolfram Research eigene Produkte derart häufig als Grundlage von Wolfram|Alpha, dass durchaus geschlussfolgert werden kann, gezielte Werbung und Erhöhung des Bekanntheitsgrades für diese beiden Projekte sei ebenfalls eines der verfolgten Ziele. Alleine in diesem Artikel gibt es 13 Nennungen seiner „anspruchsvollen Projekte“ [Wol09b], Wolfram|Alpha wird lediglich 3 mal erwähnt, davon einmal in der Überschrift und einmal als Link auf die Startseite. In [Wol09b] erläutert Wolfram, dass er mit Mathematica eine einerseits symbolische Sprache besäße, um alles Mögliche zu repräsentieren und andererseits das algorithmische Vermögen, jede denkbare Art von Berechnung durchzuführen. NKS [Wol02] würde für Wolfram|Alpha das Paradigma darstellen, dass jegliche Komplexität aus einfachen Regeln entstehen könnte. In den FAQs [Res09c] schließlich steht, „Wolfram|Alpha is the first „killer app“ for NKS.“

Wolfram Research verfolgt mit Wolfram|Alpha auch finanzielle Ziele. Seit Anfang Oktober 2009 kann Wolfram|Alpha auch per Webservice genutzt werden, wobei die Nutzung das Beantragen einer „App-ID“ voraussetzt und bestimmte Nutzungskontingente kostenpflichtig sind. Die Einnahmen aus der Webservice-API sind eine von derzeit drei vorhandenen oder angekündigten Möglichkeiten zur Finanzierung. „Enterprise & Custom Solutions“ sowie die angekündigten „Corporate & Professional Versions“ richten sich an Unternehmen, die die Technik von Wolfram|Alpha kostenpflichtig auf ihre internen Daten anwenden können.

1.3 Nutzen

Jedermann kann und soll nach eigenen Angaben Wolfram|Alpha anwenden. Darin liegt ein großer Nutzen: eine weitgehend kostenlose Frage-Antwort-Maschine für viele Themenbereiche. Der Benutzer erhält seine Antworten direkt, ohne wie bei herkömmlichen Suchmaschinen auf den angezeigten Seiten danach suchen zu müssen. Die Natürlichsprachlichkeit sorgt dafür, dass die Benutzer es möglichst einfach haben, mit Wolfram|Alpha zu kommunizieren. Wolfram schreibt in [Wol09a], dass es interessant sei, nach der ersten Laufzeit die Anfragen zu untersuchen, die gestellt werden und vor allem wie sie formuliert wer-

den. Er stellte fest, dass sich eine neue menschliche Sprache entwickelt hat, die auf einer einfachen Sprache basiert und Feinheiten außen vor lässt. Die gestellten Anfragen würden wahrscheinlich mehr die Art widerspiegeln, wie Menschen denken. Sicherlich hat sich über die Jahre der Web- und Suchmaschinennutzung ein selbstverständlicher Umgang mit solchen Dienste entwickelt. In dem zitierten Beispiel von Abbildung 1 wurde allerdings sichtbar, dass es für Wolfram|Alpha irrelevant war, ob eine Frage oder nur die Keywords „search engine“ eingegeben wurden, die Antwort war in beiden Fällen dieselbe. An diesem Punkt geht ein gewisser Teil der Natürlichsprachlichkeit verloren, was sicherlich auch durch das Verhalten der Benutzer maßgeblich beeinflusst wird.

 **WolframAlpha**[™] computational knowledge engine

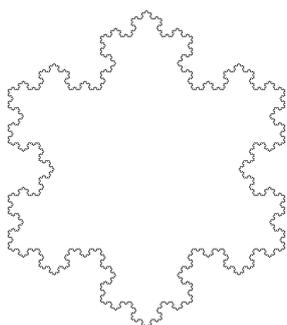
Input interpretation: <input type="button" value="Koch snowflake"/>	Iteration rule: 								
Input value: iterations <input style="width: 40px;" type="text" value="5"/>	L system grammar representation: <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">axiom</td> <td style="padding: 2px;">$F \rightarrow F--F$</td> </tr> <tr> <td style="padding: 2px;">rules</td> <td style="padding: 2px;">$F \rightarrow F+F--F+F$</td> </tr> <tr> <td style="padding: 2px;">rotation angle</td> <td style="padding: 2px;">60° (degrees)</td> </tr> </table>	axiom	$F \rightarrow F--F$	rules	$F \rightarrow F+F--F+F$	rotation angle	60° (degrees)		
axiom	$F \rightarrow F--F$								
rules	$F \rightarrow F+F--F+F$								
rotation angle	60° (degrees)								
Result: 	Dimensions after 5 iterations: <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">length</td> <td style="padding: 2px;">$\frac{1024}{81} \approx 12.642$ (times initial side length)</td> </tr> <tr> <td style="padding: 2px;">area</td> <td style="padding: 2px;">$\frac{31288}{19683} \approx 1.5896$ (times initial area)</td> </tr> </table>	length	$\frac{1024}{81} \approx 12.642$ (times initial side length)	area	$\frac{31288}{19683} \approx 1.5896$ (times initial area)				
length	$\frac{1024}{81} \approx 12.642$ (times initial side length)								
area	$\frac{31288}{19683} \approx 1.5896$ (times initial area)								
Limiting properties: <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">length</td> <td style="padding: 2px;">∞</td> </tr> <tr> <td style="padding: 2px;">area</td> <td style="padding: 2px;">$\frac{8}{5} \approx 1.6$ (times initial area)</td> </tr> <tr> <td style="padding: 2px;">capacity dimension of curve</td> <td style="padding: 2px;">$\frac{2 \log(2)}{\log(3)} \approx 1.26186$</td> </tr> <tr> <td style="padding: 2px;">capacity dimension of enclosed area</td> <td style="padding: 2px;">2</td> </tr> </table>		length	∞	area	$\frac{8}{5} \approx 1.6$ (times initial area)	capacity dimension of curve	$\frac{2 \log(2)}{\log(3)} \approx 1.26186$	capacity dimension of enclosed area	2
length	∞								
area	$\frac{8}{5} \approx 1.6$ (times initial area)								
capacity dimension of curve	$\frac{2 \log(2)}{\log(3)} \approx 1.26186$								
capacity dimension of enclosed area	2								

Abbildung 4: Schneeflocken-Fraktal der Stufe 5

Im mathematisch-naturwissenschaftlichen Bereich ist Wolfram|Alpha durch Verwendung ihres seit den 80er Jahren stetig weiterentwickelten Programms Mathematica besonders ausgeprägt. Mathematica ist ein Softwarepaket, das unter anderem die die symbolischen Verarbeitung von Gleichungen, das numerische Lösen und Auswerten von Gleichungen und die Darstellung von Graphen ermöglicht. In [Res09c] wird die Frage „Wieviel Mathematik kennt Wolfram|Alpha?“ mit „So ziemlich alles mögliche und die derzeit besten Algorithmen.“ beantwortet. Der Benutzer kann sich schnell und einfach Lösungen selbst für komplizierte Gleichungen und Methoden ausrechnen lassen, so z.B. wie in Abbildung 4 gezeigt. Die Frage bleibt, wer und wie häufig solch geartete Informationen abrufen möchte. [Fel09] schreibt zur Einschränkung von Wolfram|Alpha, das es darauf ankommt, wie man es dem Durchschnittsnutzer ermöglicht, von dieser Technologie zu profitieren. [Spi09] bemerkt, dass Wolfram|Alpha für und von Menschen mit einem IQ im Bereich von Stephen Wolfram entwickelt wurde und dass es fraglich sei, ob der durch-

schnittliche Benutzer auf die Möglichkeit einer komplexen Fragemaschine gewartet hat und Wolfram|Alpha deshalb so selbstverständlich wie eine Suchmaschine benutzt werden wird. Man kann ergänzen, dass Dienste dann dauerhaft genutzt werden, wenn sie möglichst alltagstauglich sind bzw. den Alltag vereinfachen.

Der wohl größte Nutzen von Wolfram|Alpha und Vorteil gegenüber anderen Suchmaschinen ist die Darstellung als zentraler Zugangspunkt zu einer großen Menge an Wissen aus diversen Bereichen. Suchmaschinen bauen ihren Index mit einem Crawler beginnend bei einer Initialmenge von Webseiten über die Verlinkung zu anderen Webseiten auf. Damit werden aber lediglich Informationen des so genannten „Surface Web“ erfasst. Nach Ergebnissen in [Ber01] ist dieses ca. 400 bis 550 mal kleiner als das „Deep Web“, dessen Informationen zu 95% in durchsuchbaren und zugreifbaren Datenbanken liegen, die nur auf direkte Anfragen hin dynamische Inhalte erzeugen. Aktuelle Suchmaschinen erreichen diese Inhalte einerseits nicht über Links und andererseits kennen sie nicht die Bezeichner und Schnittstellen der Datenbanken. Ein automatisiertes Erschließen dieser riesigen Datenmenge ist deshalb immer noch problematisch. [Ber01] gibt an, dass 85% der Benutzer Suchmaschinen zur Informationssuche bedienen und ein genauso hoher Teil sei frustriert, weil selten die gewünschten Informationen gefunden würden. Wolfram|Alpha ermöglicht dies und schließt die Lücke zwischen „Deep Web“ und Benutzer, was den eigentlichen Mehrwert ausmacht. Es spielt dabei gar keine so große Rolle, dass Ergebnisse berechnet werden sollen. In [Res09c] gibt Wolfram Research an, mehrere Trillionen automatisch erfasste und kuratierte Daten durchsuchen zu können. Man kann davon ausgehen, dass die automatisch erfassten Daten sich auf Informationen des „Surface Web“ beziehen und kuratierte Daten manuell aus dem „Deep Web“ gesammelt wurden. [TODO: WARUM?]

2 Aufbau und Funktion

In diesem Teil sollen der Aufbau und die Funktionsweise von Wolfram|Alpha näher beleuchtet werden. Dabei geben Hinweise aus [Res09c] und Aussagen Stephen Wolframs in seinem Blog Aufschluss, denn explizite Informationen werden von Wolfram Research nicht bekannt gegeben. Deshalb beruhen viele Annahmen im kommenden Teil auf Recherche und dem Wissen der Autorin über Eigenschaften und dem Funktionieren von modernen Informationssystemen.

In [Res09c] wird die Kerntechnologie von Wolfram|Alpha mit vier Schlüsselbereichen beschrieben: „data curation pipeline, the algorithmic computation system, the linguistic processing system and the automated presentation system“. Bei der Datenkuration sind nach eigenen Angaben viele Mitarbeiter damit beschäftigt, vertrauliche Datenquellen zu finden und deren Informationen bei Wolfram|Alpha zu integrieren. Das algorithmische Berechnungssystem besteht hauptsächlich aus der Software Mathematica. Das linguistische Verarbeitungssystem ist verantwortlich für die Interpretation der natürlichsprachlichen Benutzereingaben. Das automatisierte Präsentationssystem übernimmt die Darstellung der Ergebnisse und die Generierung downloadbarer Informationen.

2.1 Komponenten

Abbildung 5 zeigt, welche Komponenten in Wolfram|Alpha hauptsächlich zum Einsatz kommen könnten. Im Folgenden werden sie näher beschrieben.

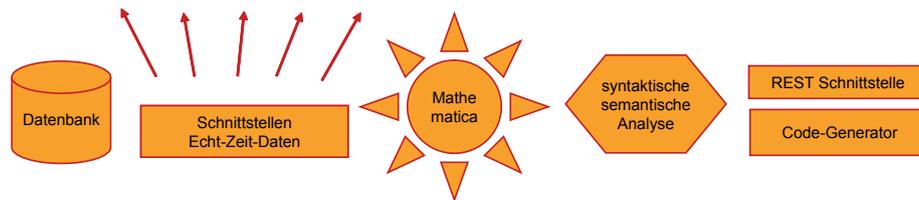


Abbildung 5: mögliche Komponenten

Datenbank Wolfram gibt an, neben kuratierten Daten auch Echtzeit-Daten zur Beantwortung der Anfragen zu verwenden. Auch wenn es nicht explizit so genannt wird, werden die kuratierten Daten als statische da historische Daten in einer Datenbank gespeichert. Denkbar wäre ein Data Warehouse für jeden Themenbereich, der direkt die Verknüpfung zwischen Datensätzen und eine Analyse gespeicherter Informationen zulässt. Über Schnittstellen zu „Deep Web“ Quellen extrahieren Mitarbeiter manuell Informationen und fügen sie zu den kuratierten Daten hinzu. Dadurch entsteht eine Datenbank, die mehr Wissen enthält, als über eine herkömmliche Suchmaschine, die mittels der Linkstruktur im Web agiert, abgefragt werden kann.

Echtzeit-Schnittstellen Zum Abfragen von Echtzeit-Daten erfüllt Wolfram|Alpha die Schnittstellen von anderen Services und leitet während der Laufzeit die angenommenen Anfragen entsprechend weiter. Die Antworten der Services werden dann an den Benutzer ausgeliefert. Auf diesem Weg erhält dieser immer die aktuellsten Informationen auf Gebieten wie Finanzen oder Wetter durch den Besuch einer einzigen Quellseite.

Mathematica Im Kern und zur Berechnung von mathematisch-naturwissenschaftlichen Anfragen wird Mathematica als Werkzeug benutzt. Das Softwarepaket Mathematica enthält unter anderem ein Computeralgebrasystem zur symbolischen Verarbeitung von Gleichungen, eine Numerik-Software zum numerischen Lösen oder Auswerten von Gleichungen, ein Visualisierungs-Tool zum Darstellen von Graphen und eine Programmiersprache, die Elemente des prozeduralen, objektorientierten, funktionalen und regelbasierten Programmierens in sich vereint. Wolfram|Alpha stellt die die Web-Benutzerschnittstelle für Mathematica dar. Anfragen, die nicht mit Mathematica beantwortet können, werden aus den anderen Komponenten abgeleitet und infolgedessen nicht einer tatsächlichen Berechnung unterzogen.

Eingabe-Analyse Wie in [Mor09] erwähnt wird, muss die komplette Eingabe des Benutzers einer syntaktischen und semantischen Analyse unterzogen werden. So kann Wolfram|Alpha verstehen, welche Zeichen eingegeben werden und was diese bedeuten sollen. Es ist vorstellbar, dass an diesem Punkt Semantic-Web-Technik zum Erkennen der Semantik von Keywords zum Einsatz kommt. Wolfram schreibt in [Res09c], zwar nicht an welcher Stelle aber dass neben einer „dynamischen algorithmischen Version einer Ontologie auch viele statische Ontologien bestimmter Domänen“ verwendet werden. Denkbar wäre die Nutzung z.B. zur Modellierung von Maßeinheiten, länderspezifischen Bezeichnungen oder Wörterbüchern, da für diese andauernd ontologische Konzepte entwickelt und verbessert werden.

REST-API Seit Oktober 2009 bietet Wolfram|Alpha Schnittstelle für ihre Webservice-API an. Diese Schnittstelle beruht auf dem Softwarearchitekturstil REST (Representational State Transfer) für das Web, bei dem jede Ressource über einen eindeutigen Bezeichner verfügt. Dabei werden auch dynamische Inhalte erfasst, die eine eigene URL besitzen sollen.

Mit einer vorhandenen App-ID kann dann über eine URL eine Anfrage kodiert werden. Das Ergebnis wird per XML zurück an den Benutzer geschickt. Darüber können beispielsweise automatisierte Abfragen realisiert werden, die in bestimmten Abständen selbstständig aktualisiert werden.

Code-Generator Ein Code-Generator wird benötigt, um die Anfragen einerseits für Mathematica in die eigene Sprache zu übersetzen und andererseits um Anfragen für die Datenbank zu formulieren. Außerdem sorgt der Code-Generator auch für die Kommunikation mit den Schnittstellen zu Echt-Zeit-Daten.

2.2 Ablauf und Arbeitsweise

[Mor09] geht davon aus, dass die Arbeitsweise von Wolfram|Alpha folgendermaßen aussieht:

- (a) Anfrage an einen linguistischen Parser geben
- (b) auf diese Ausgabe weitere Regeln und Manipulationen anwenden
- (c) Zugriff auf Deep-Web Daten und Antwort generieren

Dies wird dem generellen Ablauf entsprechen, auch wenn es für ein besseres Verständnis noch weiterer Details bedarf. Abbildung 6 zeigt den schematischen Ablauf bei einer Anfrage an Wolfram|Alpha.

Ein Benutzer kann derzeit über zwei verschiedene Wege Anfragen an Wolfram|Alpha stellen: einerseits für die Mensch-Maschine-Kommunikation über die Webseite und andererseits für die Maschine-Maschine-Kommunikation über den Zugriff per REST Webservice.

Die Parameter dafür werden in einer URL kodiert und der Webservice greift auf die notwendigen Komponenten zur Ergebnis-Generierung zurück. Der Ablauf des Webservice soll für diesen Report allerdings nicht weiter beleuchtet werden.

Die Eingabe eines Benutzers wird syntaktisch und semantisch dahingehend analysiert, was mit welcher Bedeutung gefragt wird und welche der drei Kernkomponenten Echtzeit-Schnittstellen, Mathematica oder Datenbank gefragt werden müssen. An dieser Stelle wird auch die Natürlichsprachlichkeit implementiert. Kann eine Anfrage nicht erkannt werden, wird dem Benutzer direkt die bereits genannte Fehlermeldung angezeigt und, falls vorhanden, über diverse Alternativen eine Lösung gesucht. „Input interpretation“ zeigt dem Benutzer bei Erkennen einer Anfrage direkt an, wie seine Eingabe verstanden wurde. Die syntaktisch und semantisch analysierte Eingabe wird dann an den Code-Generator weitergeleitet. Anfragen, die einer Berechnung bedürfen, also z.B. Anfragen aus einem mathematisch-naturwissenschaftlichen Umfeld, werden von dieser Komponente in Mathematica-Code übersetzt und zur Berechnung weitergeleitet. Anfragen, die Daten aus Datenbank benötigen, werden in datenbankspezifische Queries übersetzt. Anfragen, die sich auf Echt-Zeit-Daten beziehen, werden syntaktisch an die entsprechenden entfernten Services angepasst und umformuliert.

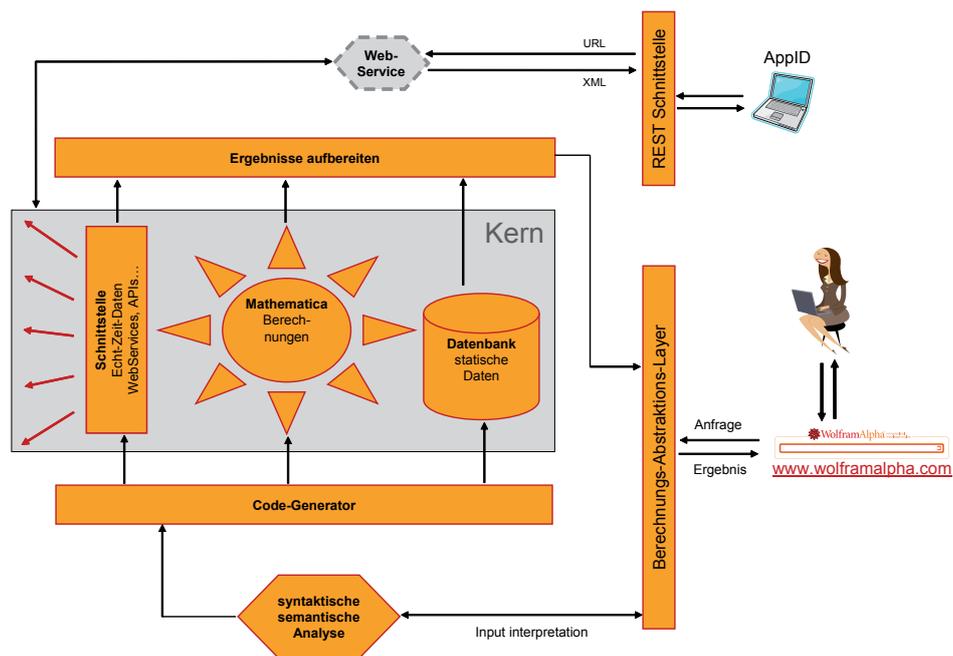


Abbildung 6: Ablauf und Interaktion der Komponenten

Beispiel 2.2 (05.11.09): Bei der Anfrage „50 digits of e “ erkennt Wolfram|Alpha, dass die Euler-Zahl e mit 50 Nachkommastellen angezeigt werden soll. Die Anfrage wird übersetzt in Mathematica-Code `N[E, 50]` und im Ergebnis `2.7182[.]70937000` ausge-

geben. Bei näherer Überprüfung fällt auf, dass es nur 49 Nachkommastellen sind, weil auf die 50. Stelle gerundet wurde. Die Zahl, die man als Text zum Kopieren bekommt, lautet 2.7182[...]709369996 und enthält zwar 50 Stellen, aber bei einer weiteren Überprüfung kam heraus, dass erneut gerundet wurde, was zu einer falschen 50. Stelle führt: e besitzt an der 50. Nachkommastelle eine 5⁶. Bei der veränderten Anfrage „50th digit of e “ erkennt Wolfram|Alpha zwar richtig, dass die 50. Nachkommastelle von e gesucht wird, aber das Ergebnis lautet 9 anstatt 5. Für die Antwort wurde offensichtlich nicht Mathematica benutzt, trotz des Hinweises „Computed by Wolfram Mathematica“, der auf jeder Seite in der Fußzeile steht. Ebenfalls fehlen sämtliche Quellenangaben.

Denkbar wäre es, Ergebnisse von komplizierten Berechnungen, also Berechnungen, die mindestens eine bestimmte Dauer benötigen, in der Datenbank abzuspeichern und beim nächsten Mal einfach auszuliefern. Das bringt aber nur dann einen Vorteil, wenn der Lookup in der Datenbank tatsächlich unaufwändiger ist als die Neuberechnung. Die 50. Stelle von e wurde eventuell auch aus der Datenbank ausgelesen, was einen Zeitvorteil gegenüber der Berechnung der 50 Stellen bedeutet, wenn doch nur eine Stelle benötigt wird. Die Ergebnisse von Echt-Zeit-Anfragen müssen allerdings jedes Mal frisch berechnet bzw. beim entsprechenden Service neu angefragt werden. Wolfram|Alpha könnte allerdings bestimmte Informationen zwischenspeichern und dem Benutzer dann ausliefern, wenn keine Kommunikation mit dem entfernten Service zustande kommt.

Das Ergebnis der Berechnung schließlich wird mit Grafiken, downloadbaren pdf-Dokumenten, Quellen-Angaben, Mathematica-Formeln aufbereitet und als Ergebnis an den Benutzer zurück gegeben. Teilweise befinden sich auch 3D-Modelle z.B. von der Struktur chemischer Elemente zur Betrachtung mit Live-Mathematica in der Antwort. Die pdf-Dokumente bestehen aus vektorisierten Grafiken und Texten.

2.3 Datenproblematik

Wie bereits angedeutet, ist es wohl der größte Nutzen von Wolfram|Alpha, seinen Benutzern einen zentralen Zugang zu Wissen aus vielen Bereichen zu bieten und vor allem Wissen aus dem „Deep Web“. Dieser Nutzen wird allerdings erst erreicht, wenn die Glaubwürdigkeit der Daten gewährleistet werden kann. Darüber hinaus hängt ebenfalls die Qualität der Antworten stark von den gesammelten Informationen ab. Der Benutzer erfährt aus den FAQs [Res09c] lediglich, dass das Wolfram Team die Daten kuratiert und die Richtigkeit garantiert, weil die Daten auf wissenschaftlichen Erkenntnissen beruhen. Es wird außerdem erwähnt, dass Experten verschiedenster Domänen zum Datenkurations-Team gehören.

Fragen, woher die Daten genau kommen und ob Ergebnisse generell allgemein anerkannt sind, bleiben offen. Letzteres ist z.B. im nicht-naturwissenschaftlichen Bereich der Fall, wenn unter den Wissenschaftlern keine Einigkeit über bestimmte Theorien oder Ergebnisse besteht. [Eva09] schreibt, dass Stephen Wolfram auf einer Wolfram|Alpha Demonstration am 28.04.2009 im Berkman Center der Harvard Law School von Anwesenden damit

⁶Quelle: University of Utah

konfrontiert wurde, dass Kulturanthropologen sich nicht einig sind über die Eigenschaften von Verwandtschaftsverhältnissen. Als Beispiel nannten sie die Interpretation von „uncle’s uncle’s brother’s son“, die nach Ward Goodenoughs Ideen eine andere als ist als nach Clifford Geertz’s Thesen. Abbildung 7 zeigt, wie Wolfram|Alpha mit dem Thema umgeht. Es fehlen, wie auch schon bei anderen Ergebnissen, sämtliche Quellenangaben, so dass der Benutzer nicht erfährt, welche These die Berechnung beeinflusst hat. Die Personen aus

The screenshot shows the Wolfram|Alpha interface with the query "uncle's uncle's brother's son". The input interpretation is "genealogic relation uncle's uncle's brother's son". The genealogic tree shows a path from "self" (red square) to "parents" (orange circles), "grandparents" (orange triangles), "uncle" (orange circle), and finally to the result (red square). The relationship properties table is as follows:

Relationship properties:	first cousin once removed	father
difference in generations	1	1
generations to common ancestor	3	1
blood relationship fraction	$\frac{1}{32}$ 3.125%	$\frac{1}{2}$ 50%

Abbildung 7: Ergebnis einer Anfrage zu Verwandtschaftsbeziehung

Wolframs Team sind genau diejenigen, die Quellen bewerten und entscheiden, was in die Datenbasis eingetragen wird und was nicht. Damit bestimmt eine Gruppe von Menschen, wie das System Wolfram|Alpha denkt und mit Fragen und Antworten umgeht, wodurch die Objektivität der Datengrundlage verloren geht.

Bei der Recherche für den vorliegenden Report wurde festgestellt, dass Quellenangaben häufig fehlten oder keinen Aufschluss über die Datenherkunft gaben. Wenn die Informationen angegeben wurden, findet der Benutzer meistens eine Primärquelle und mehrere Hintergrundinformationen und Referenzen. Standardmäßig wird bei der Primärquelle „Wolfram|Alpha curated data, 2009“ angegeben und andere eigene Quellen, wie z.B. „Wolfram Mathematica ChemicalData“. Die Liste weiterer Quellen ist hingegen recht lang, wobei man aber nicht erfährt, welchen Quellen was entnommen wurde. Am Ende der Liste wird dem Benutzer folgender Hinweis gegeben: „This list is intended as a guide to sources of further information. The inclusion of an item in this list does not necessarily mean that its content was used as the basis for any specific Wolfram|Alpha result.“ In einem Review über NKS[Ruc03] kritisierte der Autor Wolframs mangelhaften Umgang mit Quellenangaben: „For scientists, two big problems with *A New Kind of Science* are that Wolfram consistently adopts an immodest tone and that he is ungenerous in his acknowledgements of other’ work.“ Es macht den Eindruck, dass zumindest das zweitgenannte Problem auch auf Wolfram|Alpha zutrifft. Wenig transparent wird mit den Urhebern des kuratierten Wissens umgegangen.

Eine Quelle, die im Zusammenhang mit geographischen und länderspezifischen Fragen des Öfteren erscheint, ist „CIA - The World Factbook“⁷. Dahinter verbirgt sich eine riesige Datensammlung über 266 Länder der Welt, die sehr detailliert beschrieben werden und

⁷online: CIA World Factbook

zusätzlich direkt miteinander verglichen werden können. Es folgt ein Beispiel, bei dem die Daten bei Wolfram|Alpha mit denen vom Factbook verglichen werden sollen.

Beispiel 2.3 (06.11.09): Wolfram|Alpha beantwortet die Anfrage „germany population 2000“ unter anderem mit 82,3 Mio Menschen. Wiederholt man die Anfrage für das Jahr 2008 oder 2009, antwortet Wolfram|Alpha lediglich mit „no data available“, da Daten nur bis 2007 vorliegen. Das CIA World Factbook hätte die Frage mit 82,329 Mio Menschen beantworten können, da der aktuelle Stand vom Juli 2009 ist. Daran kann man sehen, dass das Factbook nicht als Echt-Zeit-Service genutzt wird, sondern in den kuratierten Daten einfließt. Diese wurden aber vom Wolfram Team noch nicht auf den neuesten Stand gebracht.

Das ruft ein weiteres Problem bezüglich der Datenbasis hervor. Wie in [Eva09] geschrieben wurde, „wird es keinem Informations-Team möglich sein, mit den wichtigsten Denkern der wichtigsten Themen der wichtigsten Bereiche menschlichen Wissens mit einem ausreichenden Maß an Erfahrungheit zu sprechen“. Für jeden der erfassten Themenbereiche benötigt man Jahre, um sich ernsthaft mit ihnen beschäftigen zu können. Darüber hinaus entwickeln sich neue Gebiete, neue Ergebnisse werden hervorgebracht und daher wird es immer unvermeidbare Lücken in einer derartigen Wissensmaschine geben.

3 Ausblick

Basierend auf den Ergebnissen der Recherche, soll im kommenden Abschnitt der aktuelle Stand von Wolfram|Alpha zusammengefasst und bewertet werden und ein Ausblick auf zukünftige Entwicklungen gegeben werden.

3.1 Derzeitiger Stand

Das Ziel von Stephen Wolfram ist die Entwicklung einer Wissensberechnungsmaschine, die mit ihren Benutzern natürlichsprachlich kommuniziert und Antworten berechnet anstatt Webseiten aufzulisten. Dieses Ziel ist insbesondere für mathematisch-naturwissenschaftliche Themen weitestgehend erreicht. Die Natürlichsprachlichkeit besteht allerdings noch nicht in zufriedenstellendem Umfang. Häufig antwortet Wolfram|Alpha damit, dass es nicht sicher ist, wie eine Antwort aus der Eingabe berechnet werden soll. Es scheint, als würden viele natürlichsprachliche Fragen das System noch überfordern. Mit einzelnen Keywords erzeugt der Benutzer meistens hilfreichere Reaktionen und Alternativvorschläge auf unbekannte Eingaben. Das ist gleichzeitig ein Aspekt, den Wolfram in [Wol09a] als eine der ersten Erkenntnisse gewonnen hat; die Benutzer verwenden eine abstraktere Sprache ohne Feinheiten zur Kommunikation mit Wolfram|Alpha. Es ist möglich, dass Benutzer aufgrund langjähriger Nutzung von Suchmaschinen den selbstverständlichen Umgang gewöhnt sind und der Bedarf nach Natürlichsprachlichkeit nicht mehr in dem Maße besteht wie in der Vision des Computers von vor 50 Jahren.

In [Eva09] schreibt Evans „Wolfram Alpha is only as good as its data and the abili-

ty of its programmers to give that data context and meaning[...].“ Die Datenbasis von Wolfram|Alpha wird zu einem großen Teil gebildet von den kuratierten Daten, die durch das Wolfram|Alpha Team ausgewählt und eingetragen werden. Das Team garantiert zwar die Glaubwürdigkeit und Richtigkeit der gesammelten Daten, aber ohne konsequente Quellenangaben müssen diese Eigenschaften und insbesondere die Objektivität der Datenbasis in Frage gestellt werden. Eine begrenzte Menge an Mitarbeitern entscheidet darüber, welche Daten zur Beantwortung genutzt werden und welche nicht.

Die Antworten sind besonders im naturwissenschaftlichen Umfeld detailliert und komplex. Es bleibt offen, ob der durchschnittliche Benutzer ausreichend Verständnis für die Ergebnisse aufbringen kann und wie groß überhaupt der Bedarf an komplexen Berechnungen bei einem heterogenen Querschnitt der Bevölkerung ist. Zusätzlich gibt es und wird es immer Fragen aus ganz unterschiedlichen Domänen geben, die auch Wolfram|Alpha nicht beantworten kann, da jegliche Berechnungsgrundlage fehlt. Dazu gehören subjektive Fragen, wie „Welcher Job passt zu mir?“ und regelmäßig neu aufkommendes oder geändertes Wissen, die nicht nur kurzzeitig eine Lücke im System belassen.

Der Benutzer profitiert in erster Linie von der bereits teilweise erfolgten Erschließung des „Deep Web“. Mittels Wolfram|Alpha kann der Benutzer auf eine vielfach größere Menge an Wissen zugreifen als das mit herkömmlichen Suchmaschinen möglich wäre. Er muss weder die ganzen einzelnen Wissensdatenbanken kennen noch wie man darauf zugreift. Diesen Arbeitsschritt nimmt Wolfram|Alpha dem Benutzer ab.

3.2 Zukünftige Entwicklungen

Die Datenbasis wird auch zukünftig maßgeblich für die Qualität der Antworten verantwortlich sein. Es wird am Wolfram|Alpha Team liegen, diese Datenbasis so gut es geht zu pflegen, ständig zu erweitern und aktuell zu halten. Auf diesem Weg kann Sorge dafür getragen werden, dass Benutzer immer wieder Wolfram|Alpha zur Beantwortung von Fragen bedienen.

Die ersten fertiggestellten und angekündigten APIs werden die Basis bilden für die Bereitstellung von weiteren APIs. Sie werden in der Zukunft eine wichtige Rolle spielen, vor allem in Hinblick auf die Ausweitung des Services in Bereichen wie Unternehmen. Wolfram sollte über die kostenlose Nutzung dieser Schnittstellen nachdenken, um seinen Dienst weiter zu verbreiten.

Auch wenn es für den Benutzer nicht die entscheidende Rolle spielen mag, so wird die Natürlichsprachlichkeit von Wolfram|Alpha weiter entwickelt werden. Auch im Hinblick auf andere Forschungsrichtungen, wie z.B. der Computerlinguistik, kann eine gelungene Natürlichsprachlichkeit stimulierend für die Weiterentwicklung und den Erfolg von Wolfram|Alpha wirken.

Der durchschnittliche Benutzer wird sich einen Großteil des abfragbaren Wissen nicht zunutze machen, weil dessen Komplexität das durchschnittliche Verständnis übersteigt. Der Umfang und die Komplexität der Antworten kann allerdings besonders im wissenschaftlichen Bereich Anwendung finden, worin auch die Zielgruppe für Wolfram|Alpha gesehen

werden kann. Jedoch ist es gerade im wissenschaftlichen Kontext unabdingbar, dass der Umgang mit Datenquellen mit der ihnen zustehenden Integrität erfolgt. Ebenso wichtig ist die Exaktheit der Ergebnisse, was am genannten Beispiel 2.2 der Euler-Zahl sichtbar wird. Es kann auf die Frage der 50. Stelle der Euler-Zahl nur eine richtige Antwort geben. Jede andere Ziffer als Antwort ist eine falsche Antwort. Daher sollte Wolfram|Alpha besser zielgruppenorientierter arbeiten. Denn letztlich werden es wissenschaftlich wirkende Benutzer aus der Zielgruppe sein, die Wolfram|Alpha und seine komplexen Antworten dauerhaft nutzen.

Literatur

- [Ber01] Michael K. Bergman. The Deep Web: Surfacing Hidden Value. Bericht, BrightPlanet Corporation, 2001. online: brightplanet.com vom 02.11.2009.
- [Bri09] Encyclopædia Britannica. search engine, 2009. online: britannica.com vom 29.10.2009.
- [Eva09] Woody Evans. Wolfram Alpha - Semantic Search is born, 2009. online: infotoday.com vom 11.08.2009.
- [Fel09] Michael Feldman. Wolfram Alpha: A Web-Based Application That Embraced Supercomputers. *HPCwire*, 2009. online: hpcwire.com vom 24.07.2009.
- [Gmb99] Statista GmbH. Die meistgenutzten Suchmaschinen weltweit nach Anteil der Suchanfragen (Juni 2009), Juni 299. online: statista.com vom 29.10.2009.
- [Kor97] Robert R. Korfhage. *Information storage and retrieval*. Wiley Computer Pub., 1997.
- [Mor09] Peet Morris. Wolfram Alpha - how does it work (part I+II). *ComputerWeekly.com*, Mai, Juni 2009. online: computerweekly.com vom 03.08.2009.
- [Res09a] Wolfram Research. About Wolfram|Alpha, 2009. online: wolframalpha.com vom 16.07.2009.
- [Res09b] Wolfram Research. The Quest for Computable Knowledge - A Timeline, 2009. online: wolframalpha.com vom 03.09.2009.
- [Res09c] Wolfram Research. Wolram Alpha Frequently Asked Questions, 2009. online: wolframalpha.com vom 16.07.2009.
- [Ruc03] Rudy Rucker. Reviews: A New Kind of Science. *The American Mathematical Monthly*, November 2003. online: rudyrucker.com vom 19.10.2009.
- [Spi09] Nova Spivack. Wolfram Alpha is Coming – and It Could be as Important as Google (But It's Completely Different), 2009. online: twine.com vom 03.08.2009.
- [Stö09] Christian Stöcker. Software-Genie verspricht den „Google-Killer“. *SPIEGELonline*, März 2009.
- [Tal09] David Talbot. Wolfram Alpha and Google Face Off, Mai 2009. online: technologyreview.com vom 11.08.2009.
- [Wol02] Stephen Wolfram. *A New Kind of Science*. Wolfram Media, Inc., 2002.

[Wol09a] Stephen Wolfram. What We've Been Doing This Summer, August 2009. online: wolframalpha.com vom 27.10.2009.

[Wol09b] Stephen Wolfram. Wolfram|Alpha is coming - Wolfram Blog, März 2009. online: wolfram.com vom 03.08.2009.