

Stefan Waidele
Ensisheimer Straße 2
79395 Neuenburg am Rhein
Stefan@Waidele.info

AKAD University
Immatrikulationsnummer: 102 81 71

Modul WIN03 — Innovative Themen der Wirtschaftsinformatik
Assignment

WEB 3.0

DATEN SIND DAS ÖL DES 21. JAHRHUNDERTS

Betreuer: Prof. J. Anton Illig

24. November 2014



AKAD University

Inhaltsverzeichnis

Abbildungsverzeichnis	ii
1 Einleitung	1
1.1 Begründung der Problemstellung	1
1.2 Ziele dieser Arbeit	1
1.3 Abgrenzung	2
2 Grundlagen	3
2.1 Vergleichskriterien	3
2.2 Betrachtung von Erdöl	3
2.3 Definition: Daten, Information, Wissen	4
2.4 Definition: Web 1.0, Web 2.0	4
2.5 Web 3.0 = Web 2.0 + Semantik = Semantisches Web	5
2.6 Die Bedeutung der Daten	6
3 Web 3.0 — Das Semantische Web	8
3.1 Daten als Handelsware	8
3.2 Einschätzungen von Analysten und der Politik	8
3.3 Auswirkungen auf die Gesellschaft	9
3.4 Auswirkungen auf die Wirtschaft	11
3.5 Vergleich der Auswirkungen mit denen des Öls	12
4 Fazit & Ausblick	14
4.1 Fazit	14
4.2 Ausblick	15
Literatur– und Quellenverzeichnis	iii
Index	vi

Abbildungsverzeichnis

1	Der Stack des semantischen Web	7
2	Internetnutzung in Deutschland 2014	10

1 Einleitung

1.1 Begründung der Problemstellung

Technologien rund um das Schlagwort *Big Data* sind laut Gartner die großen Triebfedern in der Informationstechnologie.¹ Durch Entwicklungen wie *Internet of Things*, *Ubiquitous Computing* und *Life Tracking* werden die in naher Zukunft die generierte Datenmenge als auch die Anzahl der verarbeitenden Instanzen in den nächsten Jahren deutlich zunehmen.

Diese Entwicklung macht eine systematische Aufbereitung der anfallenden Daten hin zu repräsentiertem Wissen möglich und notwendig. Das semantische Web, oder auch *Web 3.0*, verspricht Struktur in die Datenmenge zu bringen. Die Erwartungen, aber auch die Befürchtungen gegenüber den entsprechenden technischen Fortschritten sind immens.

1.2 Ziele dieser Arbeit

Ziel dieser Arbeit ist es, die momentanen Entwicklungen der Datenbeständen hin zum semantischen Web und deren Auswirkungen auf Wirtschaft und Gesellschaft mit denen durch die Industrielle Nutzung des Öls im 20. Jahrhundert zu vergleichen.

Hierzu werden zunächst im Kapitel 2 *Grundlagen* die für diese Arbeit relevanten Begriffe, Konzepte und Vergleichskriterien definiert, bevor im Kapitel 3 *Web 3.0 — Das Semantische Web* ebendiese Kriterien untersucht und betrachtet werden.

¹vgl. [Bet14]

1.3 Abgrenzung

Der Augenmerk dieser Arbeit liegt auf den behandelten Konzepten und Technologien der Informationstechnologie. Die Entsprechungen in der Ölindustrie werden nicht in der gleichen Tiefe erörtert und belegt, sondern beschränken sich auf generelle Aussagen und Ansichten.

Die Themenbereiche *Internet of Things*, *Ubiquitous Computing* und *Life Tracking* werden in dieser Arbeit unter dem Begriff *Ubiquitous Computing*² zusammengefasst, auch wenn damit unterschiedliche Techniken, Geräte und Anwendungsfälle beschrieben werden können. Ebenfalls werden diese Gebiete als stetig wachsende Datenquellen angesehen, ohne dass sie genauer untersucht werden.³

²ubiquitous: Englisch für „allgegenwärtig“

³Siehe hierzu auch Abschnitt 4.2 *Ausblick*

2 Grundlagen

2.1 Vergleichskriterien

Da es sich bei Öl und Daten um gänzlich verschiedene Dinge — bzw. eben „Nicht-Dinge“ — handelt, ist ein direkter Vergleich weder möglich noch zielführend. Für einen Vergleich bieten sich die folgenden Kriterien an: Die Eignung als Handelsware, die Einschätzung von Politik und Analysten, sowie die Auswirkungen auf Wirtschaft und Gesellschaft. Außerdem kann der jeweilige Herstellungsprozess zur Betrachtung herangezogen werden.

2.2 Betrachtung von Erdöl

Bei Erdöl handelt es sich um den wichtigsten Energieträger des 20. Jahrhunderts bis hin zur Gegenwart. Durch Erdöl wurden viele technologische Entwicklungen begünstigt oder gar erst möglich gemacht. Durch systematische Aufbereitung in Raffinerien sind neben den offensichtlichen Anwendungen wie der Nutzung als Schmier- und Kraftstoff für Mobilität und Maschinen oder zur Wärmegewinnung sind viele weitere Anwendungen möglich, ohne die die moderne Gesellschaft nur schwer vorstellbar wäre.⁴

Die Wertschöpfungskette rund um Erdöl gliedert sich in die Phasen des Finden bzw. Förderns, des Sammelns bzw. Aufbereitens und in die anschließende Nutzung in diversen Endprodukten, vom Treibstoff bis hin zu diversen Kunststoffen. Dabei gilt, dass Erdöl zwar in großen Mengen⁵, aber nicht unbegrenzt zur Verfügung steht. Aufgrund der großen Bedeutung ist Erdöl ein strategisch wichtiges Handelsgut, welches genutzt werden kann, um wirtschaftlichen und politischen Druck auszuüben.

⁴z.B. Kunststoffe und Lacke, welche zum großen Teil auf Erdöl basieren, oder auch die Reichweite, die Fahrzeuge wegen der hohen Energiedichte von Benzin bzw. Diesel erzielen.

⁵Ob diese Mengen, speziell im Vergleich zum weltweiten Verbrauch, tatsächlich immer noch groß sind, wird z.T. bezweifelt

Die Auswirkungen des Erdöls auf Technologie, Wirtschaft, Gesellschaft und Politik des 20. Jahrhunderts sind enorm. Erdöl ermöglichte großen Reichtum von Unternehmen und Staaten, aber verursachte auch Krisen und Kriege.

2.3 Definition: Daten, Information, Wissen

In dieser Arbeit sollen die folgenden Definitionen gelten: Ein *Datum* ist eine formalisierte Sachverhaltsaussage, ohne inhärente Bedeutung (z.B. „23°C“). Durch Interpretation im Kontext kann daraus eine *Information* werden (z.B. „Die Außentemperatur beträgt 23°C“).⁶ Durch Vernetzung mehrerer Informationen miteinander, aber auch durch Erfahrung kann *informatives Wissen* entstehen (z.B. „Das Wetter ist schön“).⁷

In weiteren Verfeinerungsschritten entsteht dann *handlungsorientiertes Wissen*, (z.B. „Ich benötige beim Nachmittagsspaziergang keinen Pullover“) das dann zu einer konkreten Entscheidung führen kann (z.B. „Ich lasse den Pullover zu Hause.“).⁸

2.4 Definition: Web 1.0, Web 2.0

Unter *Web 1.0* versteht man das *World Wide Web* wie es ursprünglich entwickelt wurde: Eine Menge von statischen Daten, die miteinander auf willkürliche Weise verknüpft werden konnten. Die Auszeichnungssprache *HTML* ermöglicht es Autoren, bestimmte Abschnitte zu kennzeichnen. Schon hier gibt es unterschiedliche Informationsstufen der Auszeichnungen: Während `...` lediglich aussagt, dass der ausgezeichnete Abschnitt in fetter Schriftart angezeigt werden soll, ist eine mit `<h1>...</h1>` ausgezeichnete Überschrift tatsächlich als solche zu erkennen. Auch wenn die dadurch gewonnene Information für ein automatisch erstelltes Inhaltsverzeichnis schon nützlich sein kann, wird hier keine Aussage bzgl. des

⁶vgl. [FK00], Seite 40

⁷vgl. [Nik05], Seite 106

⁸vgl. [Tay82], Seite 342

eigentlichen Inhalts getroffen. Somit sind die Dokumente des Web 1.0 dem Bereich der *Daten* zuzuordnen. Darin enthalten Information bzw. darin enthaltenes Wissen ist erst zugänglich, wenn diese von Menschen gelesen und ausgewertet werden.⁹

Die Daten des *Web 2.0* werden i.d.R. in Datenbanken vorgehalten und die Webseiten erst bei Abruf generiert. Durch die Popularität von Werkzeugen wie Blogs und Wikis sind deutlich mehr Menschen an der Erstellung der Inhalte beteiligt. Weitere Daten werden durch Techniken rund um *Ubiquitous Computing* automatisch erfasst. Diese werden auch mit sogenannten *Meta-Daten* angereichert. Hierdurch wird in maschinenlesbarer Form angegeben, welche Informationen die Dokumente enthalten. Neben vom Autor selbst zugeordneten *Taxonomien* kommen auch automatisch generierte Meta-Daten hinzu. Beispiele hierfür sind etwa das Veröffentlichungsdatum, Beziehungen zu anderen Dokumenten¹⁰ oder Geoinformationen („Wo wurde das Dokument erstellt?“). Inzwischen werden auch die Stimmung des Autors erfragt (z.B. bei Runtastic-Aktivitäten bzw. Facebook-Einträgen) oder durch Textanalyse ermittelt (z.B. bei der Auswertung von Produktrezessionen¹¹). Die durch Daten und Meta-Daten erzielte Informationsstufe ist deutlich über der von *Web 1.0* einzuordnen, unterliegt aber je nach Dienst bzw. Nutzereingaben deutlichen Schwankungen.

2.5 Web 3.0 = Web 2.0 + Semantik = Semantisches Web

„*Semantik*, auch *Bedeutungslehre*, nennt man die Theorie oder Wissenschaft von der Bedeutung der Zeichen. Zeichen können in diesem Fall Wörter, Phrasen oder Symbole sein. Die Semantik beschäftigt sich typischerweise mit den Beziehungen zwischen den Zeichen und den Bedeutungen dieser Zeichen.“¹²

⁹vgl. [Alk12], Seite xvi

¹⁰Realisiert durch sog. Backtracks – „Wer verlinkt auf dieses Dokument?“

¹¹vgl. [Spr11], Seite 11ff

¹²[Wik08]

Im *Web 3.0* werden die Daten bzw. Informationen des Web 2.0 durch Beifügung von Bedeutung zu Information bzw. informativem Wissen veredelt¹³. Hierdurch soll es möglich werden, die schnell steigenden Datenmengen sinnvoll zu nutzen¹⁴. Die Bezeichnung *semantisches Web* ermöglicht eine Abgrenzung gegenüber anderen Interpretationen des Begriffs *Web 3.0*, wie sie z.T. im Marketing¹⁵ oder in der Politikwissenschaft¹⁶ zu finden sind.

2.6 Die Bedeutung der Daten

Den vorhandenen oder neu gesammelten Daten kann auf verschiedene Arten Bedeutung beigelegt werden. Sogenannte *Meta-Daten* beschreiben Dateien durch Schlagworte oder Zuordnung von Werten zu definierten Schlüsselworten. Die Meta-Daten können in separaten Dateien gespeichert werden.¹⁷ Ebenfalls möglich ist die Speicherung direkt in der entsprechenden Datei.¹⁸ Diese Daten werden i.d.R. automatisch¹⁹ oder halbautomatisch²⁰ den Daten zugeordnet. Manche Metadaten können auch bei Abruf direkt ermittelt werden.²¹

In textbasierten Dateiformaten wie XML werden hingegen einzelne Elemente durch Markup-Tags maschinenlesbar mit Bedeutung versehen. So sind Zeichenketten als Namen, Adressen, Telefonnummer, etc. identifizier- und auswertbar. Die Meta-Daten sind hier Teil des Datenstroms.

Durch Techniken wie Gesichtserkennung oder Textanalyse können Meta-Daten auch automatisch aus den Daten generiert werden. Hierbei bestimmt die Menge

¹³vgl. [Mar06]

¹⁴vgl. [Tol07]

¹⁵z.B. „Web 3.0 marketing is the convergence of new technologies and rapidly changing consumer buying trends.“ in [Tas10], Abschnitt „What is Web 3.0 Marketing?“

¹⁶z.B. „Is this Web 3.0? Not a tech-upgrade, a smarter algorithm, slicker fibre optic or better Bluetooth beam. Instead, Web 3.0 as in an outcome, the demonstrated consequences of being able to access information?“ in [Pon12], Abschnitt „Web 3.0: Regime Change“

¹⁷z.B. in den .INFO-Dateien in der Benutzeroberfläche des Commodore Amiga „Workbench“, oder in der Bilderdatenbank der Applikation „Adobe Lightroom“

¹⁸z.B. EXIF-Daten in Bilddateien oder Informationen zum Werk in MP3-Dateien

¹⁹z.B. durch die Kamera beim fotografieren.

²⁰Einmalig manuell, dann Bereitstellung für andere Nutzer per CDDb oder andere Dienste.

²¹z.B. Dateigröße

der auszuwertenden Daten und die Qualität der bereits vorhandenen Metadaten die Qualität der Ergebnisse. Diese werden attraktiver, je mehr die Entwicklung von *Big Data* zu *Big Information* und gar *Big Knowledge*²² geht.²³

Durch Regeln, Logik und Verknüpfungen kann aus den Informationen auch Wissen erzeugt werden: Person A ist auf einem Bild zusammen mit Person B zu sehen (Gesichtserkennung). Die Geoinformation und Uhrzeit (EXIF-Tags) zeigen, dass das Foto auf einer Veranstaltung aufgenommen wurde, das zu diesem Zeitpunkt an diesem Ort stattfand (Semantisches Markup der Veranstaltung). Daraus kann auf politische Gesinnung, Musikgeschmack und/oder Trinkfreudigkeit von Person A geschlossen werden.

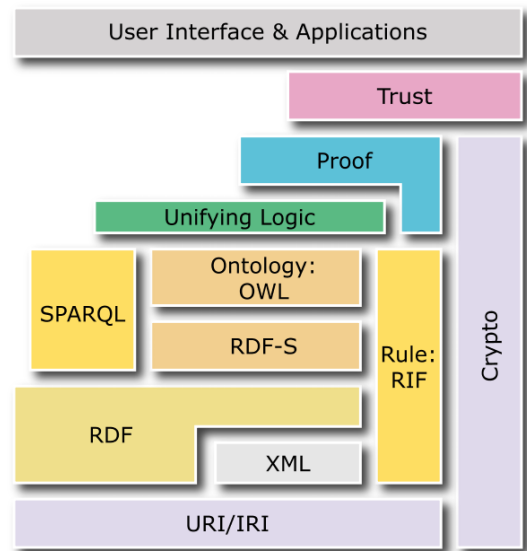


Abbildung 1: Der Stack des semantischen Web²⁵

Dieser hier nur grob skizzierte, auf den Spezialfall zugeschnittene Pfad von den Daten zum Wissen kann durch verschiedene Elemente des Web 3.0 Werkzeugkastens²⁶ allgemein formuliert und abgebildet werden. Hier sind beginnend von den datennahen Schichten zur Identifikation und Beschreibung von Objekten über den Datenaustausch, Abfragesprachen und Regelwerke bis hin zu den obersten Schichten der vereinheitlichten Logik, automatischer Beweisführung und dem darauf aufbauenden Vertrauen in das generierte Wissen alle relevanten Bereiche abgedeckt.

²²vgl. [Gra14]

²³vgl. [Sch09]

²⁵[Bra07], Seite 24

²⁶Web 3.0 Stack, siehe Abbildung 1

3 Web 3.0 — Das Semantische Web

3.1 Daten als Handelsware

Sowohl Daten als auch Information und Wissen werden ähnlich wie physische Güter gehandelt. Hierbei bestehen aufgrund der Intangibilität und der Irreversibilität von Transaktionen deutliche Unterschiede zum Handel mit Erdöl.²⁷ Da jedoch ein realer Markt für virtuelle Güter besteht und diese dort einen Preis erzielen, stellen Daten und die daraus gewonnenen Produkte auch Handelswaren dar.²⁸

Der Wert dieser Ware lässt sich jedoch nicht beliebig erhalten, da dieser durch Weiterverbreitung bzw. Veränderungen in durch die Daten beschriebenen Realität abnehmen kann.²⁹

3.2 Einschätzungen von Analysten und der Politik

Gartner rechnet mit einem Anstieg der mit dem Internet verbundenen Geräte um das Fünffache von fast fünf auf über 25 Milliarden Geräten im Jahr 2020.³⁰ Zusätzlich zu den durch Herstellung und Vertrieb dieser Geräte erzielten Wertschöpfung wird in 2020 die Inanspruchnahme von Dienstleistungen im Wert von 263 Milliarden Dollar erwartet.³¹

Von Politikern auf verschiedenen Ebenen wird der Themenbereich *Big Data* als sehr wichtige Wachstumsfeld eingeschätzt. So erklärt z.B. die EU-Kommissarin für die Digitale Agenda Needie Kroes: „Daten sind Antrieb und Grundlage für die Wirtschaft der Zukunft. Organisationen jeder Art [...] benötigen Daten als Bausteine, um leistungsfähiger zu werden.“³² und rechnet mit 100.000 neuen Arbeitsplätzen in der Datenverarbeitungsbranche. Durch ein Förderprogramm in

²⁷vgl. [Wie52], Seite 117

²⁸vgl. [Wie52], S115

²⁹vgl. [Wie52], Seite 113 bzw. 122

³⁰vgl. [Gar14], Tabelle 1

³¹vgl. [Gar14], Absatz 3

³²Zitiert nach [Zie14]

Höhe von 2,5 Milliarden Euro soll die Entwicklung von Techniken und Diensten zur Verarbeitung großer Datenmengen durch Einsatz von künstlicher Intelligenz gefördert werden.³³

Das Umwelt-Bundesamt kümmert sich um Nachhaltigkeit nicht nur bei der Hardware, sondern auch bei Software und den erhobenen Daten. „Wachsende Datenmengen und überdimensionierte Programme erfordern den beständigen Ausbau von IT-Netzen, Datenspeichern und Rechenkapazitäten. Höchste Zeit also, die Diskussion über Nachhaltige Software anzustoßen.“³⁴

Somit werden in der Politik die gleichen Ziele verfolgt wie im Bezug auf die Ölbranche: Sicherung bzw. Steigerung des Marktanteils, Erzielung von möglichst positiven Effekten auf den Arbeitsmarkt sowie die bestmögliche Vermeidung der negativen Auswirkungen auf Umwelt und Gesellschaft.

3.3 Auswirkungen auf die Gesellschaft

Schon heute drehen sich die meisten der regelmäßig ausgeführten Aktivitäten im Internet um *Informationen* oder gar *Wissen*, nicht um *Daten*. So sind 56% der im Digitalindex 2014 befragten Deutschen überzeugt, im Internet die aktuellsten Informationen zu finden, 60% sucht benötigte Informationen zuerst im Netz³⁵. Die Erwartungen bzgl. Aktualität und Informationsgehalt an das *World Wide Web* sind also sehr hoch.

Unter den Top 10 der regelmäßig durchgeführten Tätigkeiten der Befragten im Web finden sich die informationsorientierten Tätigkeiten „nach Inhalten/Informationen suchen“ auf Platz eins, „über aktuelle Ereignisse des Wohnorts informieren“ auf Platz 6 und die Erfassung eigener Daten auf dem 10. Platz.³⁶ Weitere Tätigkeiten wie „Soziale Netzwerke nutzen“ (Platz 7) bzw. „Online-Videos ansehen“ (Platz 3)

³³vgl. [Zie14]

³⁴[Köh14]

³⁵vgl. [Ini14a], Seite 6

³⁶vgl. [Ini14a], Seite 37

sowie „Online-Shopping“ (Platz 2) nutzen Internetangebote, die per Design sehr gut mit Taxonomien, Meta-Daten und Verknüpfungen ausgestattet sind.



Abbildung 2: Internetnutzung in Deutschland 2014³⁷

In den TV-Werbespots von Google und Apple werden Verbraucher gezielt auf entsprechend formulierte Anfragemöglichkeiten hingewiesen.³⁸ Somit werden die Möglichkeiten wie Routenplanung, Musiktiterkennung, Wettervorhersage und andere, die schon heute durch einzelne Anwendungen zur Verfügung gestellt werden in einer Oberfläche Gebündelt und in das Bewusstsein der Verbraucher gebracht.

Neben dem offensichtlichen Nutzen des *World Wide Web* bringt die Entwicklung des Webs auch negative Auswirkungen auf die Gesellschaft. Dies äußert sich in der Besorgnis von 60% der Nutzer über die im Internet möglicherweise verfügbaren persönlichen Daten.³⁹ Bei der Nutzung von Webdiensten der öffentlichen Verwaltung haben 65% der Befragten Angst vor Datendiebstahl, und 73% der Deutschen haben ein starkes Interesse daran, wie Behörden mit den Daten der Bürger umgeht.⁴⁰ Sowohl fiktive⁴¹ also auch reale⁴² Unrechtsstaaten basieren

³⁷[Ini14a], Seite 37

³⁸z.B. „Ok Google, zeig mir mal den schnellsten Weg in den Münchner Tierpark?“, [Goo14]

³⁹vgl. [Ini14a], Seite 6

⁴⁰vgl. [Ini14b], Seite 9 bzw. Seite 34

⁴¹z.B. Ozeanien in [Orw49]

⁴²z.B. die DDR, vgl. [Wik14]

auf der intensiven Erfassung und Auswertung von personenbezogener Daten der Bürger, was entsprechende Befürchtungen nährt und erklärt.

Man kann davon ausgehen, dass mit weiter wachsenden Datenmengen, aber auch durch entsprechenden Wachstum an generierten und erfassten Informationen und Wissen sowohl die positiven Erwartungen als auch die Befürchtungen und Ängste in der Bevölkerung zunehmen werden.

3.4 Auswirkungen auf die Wirtschaft

Schon heute haben Unternehmen erkannt, dass es nicht ausreicht, immer mehr Daten anzuhäufen. Der Schritt von *Big Data* zu *Big Information* bringt Benutzerfreundlichkeit und echten Mehrwert für Unternehmen.⁴³ In großen Unternehmen wird unter den Schlagwort *Business Intelligence* bzw. *Online Analytical Processing* mit unterschiedlichen Technologien aus den im *Data Warehouse* gespeicherten Daten Informationen und informatives Wissen zu extrahieren, auf dem dann Entscheidungen und Handlungen basieren können.

Schon in der Vergangenheit haben Fortschritte in der Daten- und Informationsbearbeitung erhebliche Auswirkungen auf die Wirtschaft gehabt. So war die Entwicklung der Telegraphie ein wichtiger Schritt für die Meteorologie, da hiermit zeitnahe Datensammlung und Auswertung möglich wurde. Als Resultat waren Wettervorhersagen für die Schifffahrt möglich. Mit steigender Datenmenge und immer noch wachsender Verarbeitungsgeschwindigkeit werden im Zusammenspiel mit dem Fortschritten des semantischen Webs auch in anderen Bereichen immer treffendere Analysen und Vorhersagen möglich werden.

Auch jenseits von Vorhersagen sind durch die Datenverarbeitung im großen Stil neue bzw. verbesserte Produkte möglich. Als Beispiele ist hier die Jacht „BMW Oracle“ zu nennen, die durch Datenverarbeitung in Echtzeit zu einem hocheffi-

⁴³vgl. [Hub12]

zienten Segelschiff wurde.⁴⁴ Herzschrittmacher, die Patientendaten beobachten um nur bei Bedarf als Taktgeber oder Defibrillator zu handeln,⁴⁵ sowie der ganze Bereich der Ferndiagnostik und –wartung bei Maschinen und auch Menschen im Umfeld von *Ubiquitous Computing* sind weitere Produkte, die nur durch intensive Daten– und Informationsverarbeitung möglich bzw. sinnvoll werden.

Die wie auf Seite 10 beschriebenen, weit verbreiteten und zum Teil großen Befürchtungen stellen Herausforderungen an die Unternehmen dar. Neue Technologien und Anwendungen ebendieser sollten so entworfen und kommuniziert werden, dass Verbraucher ihnen vertrauen können.

Es sind also auch erhebliche Auswirkungen auf die Wirtschaft zu erwarten, sowohl durch Optimierung vorhandener Produkte und Dienstleistungen aber auch durch Neuentwicklungen.

3.5 Vergleich der Auswirkungen mit denen des Öls

Daten und Öl bilden die Grundlage für Produkte bzw. Dienstleistungen. Große Teile der Wirtschaft sind direkt von ihnen abhängig. Noch größer dürfte die indirekte Abhängigkeit vom Öl sein, die sich durch sämtliche Branchen und auch auf private Haushalte erstreckt. Eine solche Abhängigkeit ist für Daten bereits zu erahnen.

Rohöl kann als Roh–, Hilfs– und Betriebsstoff bei der Erzeugung unterschiedlichster Produktarten wie z.B. Kunststoff, Schmierstoffe, Kosmetika oder Treibstoffe. Gleiches gilt für Daten: Oft sind die mit Hilfe künstlicher Intelligenz⁴⁶ erzielten Informationen direkt das gewünschte Produkt. Diese Informationen können aber auch nur wie ein Betriebsstoff dafür sorgen, dass andere Prozesse besser ablaufen.

⁴⁴vgl. [Ken13], Seite 59f

⁴⁵vgl. [Ger06], Seite 146

⁴⁶im Englischen Wortsinn von Informationsbeschaffung

Auch die Entwicklung der zur Verfügung stehenden Menge stellt sowohl beim Öl als auch bei Daten ein Problem dar: Während Ölvorkommen endlich sind, und somit nach Alternativen geforscht werden muss, werden Daten in ihrer Unendlichkeit⁴⁷ mit steigender Menge schwerer zu verarbeiten.⁴⁸ Somit muss die Wirtschaft lernen mit weniger Öl, aber mit immer mehr Daten zurechtzukommen.

Eine weitere Gemeinsamkeit zeigt sich bei historischer Betrachtung: Sowohl Öl als auch Daten und ihre Verarbeitung sind schon lange vor ihrer industriellen Nutzung genutzt worden. Dies gilt für Rohölprodukte (z.B. als Bau und Brennstoff⁴⁹) als auch für Daten (z.B. zur Wettervorhersage durch den hundertjährigen Kalender⁵⁰ oder die frühen Verschlüsselungstechniken⁵¹). Jedoch bedurfte es einem bestimmten Stand der Technik, ab dem die Einsatzzwecke dann sprunghaft anstiegen.

⁴⁷Mit jedem erfasstem Datum entstehen weitere erfassbare Daten, wie z.B. Speicherort und Erhebungszeitpunkt

⁴⁸Sowohl aufgrund der Datenmengen und den daraus steigenden Anforderungen an Hard- und Software, sowie auch aufgrund der wachsenden Entropie bzw. sinkenden relativen Informatinosgehalt (vgl. [Wie52], Seite 118)

⁴⁹vgl. [Pre00], Abschnitt 2.1

⁵⁰vgl. [Deu14] bzw. [Kna], Seite 1: „Erklärungen über die Beschaffenheit, Gestalt und Bewegung unserer Erde und die anderen Weltkörper, über die besonderen Naturerscheinungen, und über das ganze Weltwissen überhaupt; dann über die Witterungsvorhersagung nach den besten Bauernregeln und sichersten Wetteranzeigen,...“

⁵¹[Sue08], Abschnitt 56.6: „... wenn etwas Geheimes zu überbringen war, schrieb er in Zeichen, das heißt, er ordnete die Buchstaben so, dass kein Wort gelesen werden konnte: Um diese zu lesen, tauscht man den vierten Buchstaben, also D für A aus und ebenso mit den restlichen.“

4 Fazit & Ausblick

4.1 Fazit

In der Analogie zu Erdöl lassen sich die *Daten* wie in der These beschrieben mit Rohöl vergleichen. Die Techniken rund um *Big Data* gemeinsam mit den datenbankbezogenen Schichten des *semantischen Webs* entsprechen somit den Ölspeichern, welche den Rohstoff bereit halten. Die Weiterverarbeitung erfolgt dann in den logischen Schichten des *Web 3.0*, welche somit am ehesten mit den Raffinerien vergleichbar sind, die aus den klebrigen schwarzen Rohöl der Daten die verschiedenen Informations- und Wissensprodukte erzeugt. Ob diese dann mit Schmierstoffen, schwerem Schiffsdiesel, hocheffizientem Kerosin oder Kunststoffen vergleichbar sind hängt sowohl von den Ausgangsdaten, aber auch wie beim Öl von den Zielsetzungen und Anforderungen ab.

Datenschutz ist in diesem Bild vergleichbar mit Umweltschutz, der auf allen Ebenen dafür kämpft, dass die negativen Auswirkungen der neuen Technologie gemindert bzw. eliminiert werden. Wo beim Öl Strände, Seevögel, die Atmosphäre und das Klima geschützt werden, sind es im Umfeld der des semantischen Webs die Bürger-, Grund- und Persönlichkeitsrechte, die vor übermäßigem und falschem Einsatz der Technologie geschützt werden sollen. Die Achtung und der Schutz dieser Rechte ist wie auch der Umweltschutz beim Öl wichtig, wenn die Verbreitung der neuen Techniken nicht durch gestörtes Vertrauen der Verbraucher gehemmt oder gar verhindert werden soll.

Auch die Einschätzungen von Politik und Analysten sowie der rege Handel sowohl mit Daten als auch mit Öl und den daraus gewonnenen Produkten stützen die hier untersuchte These.

4.2 Ausblick

Nach dieser Einordnung des *Web 3.0* bietet sich für weitere Arbeiten die genauere Untersuchung der einzelnen Aspekte an, die hier zum großen Teil nur genannt werden konnten. Speziell die automatische Informationserkennung durch Mustererkennung in Texten, Grafiken und Videos und die daraus entstehenden Möglichkeiten sind eine nähere Betrachtung wert.

Ebenso die höheren Ebenen des Web 3.0 Stacks, in denen es darum geht, Zusammenhänge rechnergestützt zu erkennen. Durch die Automatisierung der Informationsgewinnung wird diese skalierbar, so wie es die Datenspeicherung heute schon ist. Dies wird die in dieser Arbeit beschriebenen Entwicklungen nochmals deutlich beschleunigen und in ihrer Tragweite vergrößern.

Aufgrund der immensen Datenmenge, die von Technologien rund um das *Internet of Things*, *Ubiquitous Computing* und *Life Tracking* erfasst werden können bieten sich diese Themengebiete ebenfalls für genauere Studien mit dem Fokus auf der Datenqualität an.

Weiterhin empfiehlt sich auch eine gründliche Untersuchung der gesellschaftlichen und politischen Aspekte des Semantischen Webs.

Literatur– und Quellenverzeichnis

- [Alk12] Ghazi Alkhatib. *Models for Capitalizing on Web Engineering Advancements*. IGI Global, 2012.
- [Bet14] Betsy Burton et al. *Gartner’s Hype Cycle Special Report for 2014*. Gartner inc., Stamford, CT, USA, 2014.
- [Bra07] Steve Bratt. Semantic Web: Linked Data on the Web. [http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#\(24\)](http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#(24)), abgerufen am 07.11.2014, Januar 2007.
- [Deu14] Deutscher Wetterdienst (Hrsg.). Hundertjähriger Kalender. http://www.deutscher-wetterdienst.de/lexikon/index.htm?ID=H&DAT=Hundertjaehriger_Kalender, abgerufen am 11.11.2014, 2014.
- [FK00] Klaus Fuchs-Kittowski. *Wissens-Ko-Produktion — Organisationsinformatik*. Number 2. Auflage. Gesellschaft für Wissenschaftsforschung, 2000.
- [Gar14] Gartner (Hrsg.). Gartner Says 4.9 Billion Connected Things will be in Use in 2015. <http://www.gartner.com/newsroom/id/2905717>, abgerufen am 14.11.2014, November 2014.
- [Ger06] Gerd Fröhling et al. *Herzschrittmacher– und Defibrillator–Therapie*. Georg Thieme Verlag, Stuttgart, Stuttgart, 2006.
- [Goo14] Google. TV–Werbespot: Google Bonobo. http://www.youtube.com/watch?v=Hay3d5g4_fI, abgerufen am 11.11.2014, Juni 2014.
- [Gra14] Tom Graves. Big-data, Big-Information, Big-Knowledge. <http://weblog.tetradian.com/2014/03/04/big-data-big-information-big-knowledge/>, abgerufen am 08.11.2014, März 2014.
- [Hub12] Dan Hubbart. Enough about Big Data. Let’s talk Big Information. <http://blog.opendns.com/2012/05/15/enough-about-big-data-lets-talk-big-information/>, abgerufen am 08.11.2014, Mai 2012.

- [Ini14a] Initiative D21 et al. (Hrsg.). *Die Entwicklung der digitalen Gesellschaft in Deutschland*. Initiative D21 e.V., Berlin, 2014.
- [Ini14b] Initiative D21 et al. (Hrsg.). *eGovernment Monitor 2014*. Initiative D21 e.V., Berlin, 2014.
- [Ken13] Kenneth C. Laudon et al. *Essentials of Management Information Systems*. Pearson, London, tenth, international edition, 2013.
- [Kna] Moriz Knauer. Großer hundertjähriger Kalender für das neunzehnte Jahrhundert. http://media05.regionaut.meinbezirk.at/2013/04/20/4081937_web.jpg, abgerufen am 11.11.2014.
- [Köh14] Marina Köhn. Einladung zum Fachgespräch Nachhaltige Software. http://oekotop100.de/software/wp-content/uploads/2014/10/Einladung_Fachgespraech_Nachhaltige_Software_am_28-11-2014.pdf, abgerufen am 11.11.2014, November 2014.
- [Mar06] John Markoff. Entrepreneurs See a Web Guided by Common Sense. <http://www.nytimes.com/2006/11/12/business/12web.html?pagewanted=all>, abgerufen am 05.11.2014, November 2006.
- [Nik05] Paul Nikodemus. *Wissensmanagement und Innovation*. Cuvillier Verlag, Göttingen, 2005.
- [Orw49] George Orwell. *Nineteen Eighty-Four*. Secker & Warburg, London, 1949.
- [Pon12] Simon Pont. *The Better Mousetrap*. Kogan Page, London, web edition, November 2012.
- [Pre00] Carola Preßler. Erdöl und Erdölprodukte. Protokoll zum Experimentalvortrag, Philipps-Universität Marburg, Januar 2000.
- [Sch09] Holger Schmidt. Das Internet der nächsten Generation. <http://www.faz.net/aktuell/wirtschaft/netzwirtschaft/web-3-0-das-internet-der-naechsten-generation-1796817.html>, abgerufen am 08.11.2014, Mai 2009.
- [Spr11] Michèle Sprejz. Extraktion und Klassifikation von bewerteten Produktfeatures auf Webseiten. Diplomarbeit, TU Dresden, Dezember 2011.
- [Sue08] Suetonius. *Divus Julius*. in aedibvs B. G. Tevbneri, 1908.

- [Tas10] Michael Tasner. *The Limiting Factors of Web 2.0 and How Web 3.0 Is Different*. FT Press, Upper Saddle River, NJ, USA, web edition, Juli 2010.
- [Tay82] Robert S. Taylor. Value-Added Processes in the Information Life Cycle. *Journal of the American Society for Information Science*, pages 341–346, 1982.
- [Tol07] Robert Tolksdorf. Web 3.0 – die Dimension der Zukunft. <http://www.tagesspiegel.de/zeitung/web-3-0-die-dimension-der-zukunft/1028324.html>, abgerufen am 05.11.2014, August 2007.
- [Wie52] Norbert Wiener. *Mensch und Menschmaschine*. Metzner, 1952.
- [Wik08] Wikipedia. Semantik. <http://de.wikipedia.org/wiki/Semantik>, abgerufen am 05.11.2014, Juli 2008.
- [Wik14] Wikipedia. Ministerium für Staatssicherheit. http://de.wikipedia.org/wiki/Ministerium_f%C3%BCr_Staatssicherheit, abgerufen am 11.11.2014, 2014.
- [Zie14] Peter-Michael Ziegler. 2,5 Milliarden Euro für die Big-Data-Forschung. *c't Magazin, Heft 24 vom 01.11.*, page 52, November 2014.

Index

- Bedeutungslehre, *siehe* Semantik
- Big Data, 1, 7, 14
- Big Information, 7, **11**
- Business Intelligence, 11

- Data Warehouse, 11
- Daten, **4**, 5, 9, 14
- Datum, *siehe* Daten

- Information, **4**, 9
- Internet of Things, *siehe* Ubiquitous Computing

- Life Tracking, *siehe* Ubiquitous Computing

- Meta-Daten, 5, **6**

- Semantik, 5

- Taxonomien, 5

- Ubiquitous Computing, 1, **2**, 5, 12, 15

- Web
 - 1.0, 4, 5
 - 2.0, 5, 6
 - 3.0, 1, 6, 14, 15
 - semantisch, *siehe* Web 3.0
 - World Wide, 4, 9, 10
- Wissen, 9
 - handlungsorientiert, **4**
 - informativ, **4**

Eidesstattliche Erklärung

Ich versichere, dass ich das beiliegende Assignment selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie alle wörtlich oder sinngemäß übernommenen Stellen in der Arbeit gekennzeichnet habe.

(Datum, Ort)

(Unterschrift)