

## Diskussion

Der Zusammenhang zwischen der nachfolgenden Diskussion und dem Informationsagenten ist der, daß aus dieser Diskussion letztendlich die Idee zu dem Informationsagenten geboren wurde. Die Diskussion war quasi die Ursache für die Entstehung des Konzepts für den Informationsagenten.

Falls Sie an den im Vorfeld gemachten Gedanken nicht interessiert sind, so können Sie auch direkt zur Beschreibung des Informationsagenten [springen](#).





In dieser Diskussion soll eine Antwort auf die Frage "**Gibt es alternative Lösungen, die möglicherweise besser sind als die von mir entwickelte Lösung ?**" gefunden werden ?. Die Diskussion dient der Suche nach alternativen Lösungen und der kritischen Auseinandersetzung mit deren tatsächlichem Nutzen.

### Hinweis


Für die abschließende Diskussion einer beliebigen Lösung gibt es eine [Mustervorlage](#). Eine gute Übung für eine solche Diskussion ist, wenn Sie die nachfolgende Diskussion erst einmal genau durchlesen und dann mit Hilfe der Mustervorlage einmal für sich selbst die Diskussion durchführen.

Nachfolgend sollen nun die **Vor- und Nachteile meiner Lösung** diskutiert werden. Die **Vor- und Nachteile sind durch ihre Reihenfolge gewichtet**, das heißt, die Punkte mit hohem Gewicht werden zuerst aufgeführt, die Punkte mit geringem Gewicht folgen weiter unten. Danach werden die **Vor- und Nachteile gegeneinander abgewogen** und es wird nach Möglichkeiten zur Verbesserung meiner Lösung gesucht.

### Diskussion

-  Vorteile meiner Lösung ↓
-  Nachteile meiner Lösung ↓
-  Suche nach Möglichkeiten zur Verbesserung/Erweiterung meiner Lösung ↓
-  Kritische Betrachtung der gefundenen Möglichkeiten zur Verbesserung/Erweiterung meiner Lösung ↓

#### Vorteile meiner Lösung: ↑

-  Es werden [Lehr-/Lernmethoden](#) angewandt, die der Art und Weise, wie Menschen eigentlich von Natur aus lernen, bestmöglich entspricht - damit wird das Lernen maximal vereinfacht.

**Konkret bedeutet das, lernen durch:**

-  **kritisches und möglichst vorurteilsfreies Fragen**  
(Ich kann nur wissen wonach ich fragen kann !)

praktische Anwendung der Methode in meinem Lehrgang:

Meine unzähligen Frage-Antwortpaare und die vielfältigen, von mir im Rahmen des Lehrgangs beschriebenen Möglichkeiten, wie Sie mit Hilfe des Internets Antworten auf die vielfältigsten Fragen finden können.

Dieses Fragen ist am Anfang, wo Sie zum Lernen motiviert werden und sich ein ganzheitliches Bild von den Lerninhalten, den Lernzielen und dem Lernweg machen genauso wichtig, wie bei der anschließenden Beobachtung und Nachahmung meiner Lerninhalte, im Rahmen derer Sie sich das Grundwissen aneignen, sowie bei der praktischen Anwendung dieses Grundwissens, indem Sie sich durch Versuch und Irrtum aktiv im Rahmen meiner Methoden und Übungsaufgaben mit den Lerninhalten auseinandersetzen und dabei Ihre natürliche Neugierde (Ihre vielen Fragen) so weit als möglich ausleben. Das Fragen geht weiter indem Sie am Ende die gemachten Erfahrungen, das von Ihnen konstruierte Wissen

und letztendlich Ihren gesamten Lernerfolg hinterfragen.

### ● **Beobachtung und Nachahmung**

praktische Anwendung der Methode in meinem Lehrgang:

Meine Lerninhalte, insbesondere aber die vielen praktischen Beispiele und die praxisrelevanten Inhalte, im Rahmen derer Sie sich das Grundwissen aneignen.

### ● **Versuch und Irrtum**

praktische Anwendung der Methode in meinem Lehrgang:

Meiner Methoden und Übungsaufgaben im Rahmen derer Sie das erlernte Grundwissen praktisch anwenden und dabei eigene Erfahrungen und Erkenntnisse sammeln.

Wenn ich die eben genannten Lernmethoden mit der Art und Weise vergleiche, wie ich im Rahmen des traditionellen Bildungssystems (Frontalunterricht) lernen musste, so kann ich nur sagen, daß keine der Methoden auch nur annähernd zufriedenstellend angewendet wurde. Fragen nur wenn es unbedingt sein muß und auf keinen Fall dumme Fragen bzw. Fragen die andere dumm finden könnten. Beobachten ja, nachahmen (mit sofortiger Rückfragemöglichkeit) nein. Das Lernen durch Versuch und Irrtum war meist auf das theoretische (und damit vorhersehbare) beschränkt, gar nicht auf das einzigartige Wesen des einzelnen Menschen ausgerichtet und praktisch, wenn überhaupt nur sehr eingeschränkt, beispielsweise im Rahmen von genau vorgegebenen Übungszeiten, Übungsplänen und Lerngruppen möglich.

### ● **Es geht darum sich Kompetenzen anzueignen, anstatt nur Wissen ohne eigene Erfahrungen und Erkenntnisse.**

Während Sie mit Hilfe Ihres Wissens nur entsprechende Fragen beantworten können, **können Sie mit Hilfe Ihrer Kompetenzen** entsprechende Fragen beantworten, **Probleme und Aufgaben lösen, richtige Entscheidungen treffen, Sie können also mit Hilfe Ihres Wissens richtig handeln.**

Dabei helfen neue Lehrmethoden, bei denen die Aktivität des Lernenden im Vordergrund steht (viele Praxisbeispiele die als Mustervorlage für die praxisnahe und alltagsrelevante Selbsterkundung des Internets dienen, Methoden, die in verschiedenen Situationen angewendet werden können, diskutieren, Fragen stellen und beantworten, das eigene Wissen zur Diskussion stellen, um es beispielsweise auf seine Richtigkeit zu überprüfen,...). Diese Lernmethoden, **bei denen die Möglichkeiten, die der Computer mit Internetzugang bietet, sinnvoll genutzt werden**, helfen dem Lernenden dabei, daß er sich mit den gefundenen und ausgewählten Informationen aktiv auseinandersetzen und so sein gerade benötigtes Wissen selbst konstruieren bzw. anstehende Probleme lösen kann.

### **! Hinweis**

Die neuen Medien erfüllen übrigens die zwei wichtigsten Forderungen zur praktischen Umsetzung des von mir entwickelten Lehrkonzepts, das auf der konstruktivistischen Lehre beruht: Zum einen eröffnen sie dem Lernenden den Zugang zu beliebigen Informationen, die letztendlich die Bausteine sind, mit denen er **sein Wissen selbst konstruieren** kann, und zum anderen ermöglichen sie die zeit- und ortsunabhängige **Zusammenarbeit (Fragen stellen und beantworten, diskutieren,...) mit anderen (andere Lernende, Lehrer, Experten, Fachinteressierte,...)**, die vor allem deswegen von so hoher Bedeutung ist, weil der Lernende nur dadurch, daß er sein selbst konstruiertes Wissen anderen zur Diskussion stellt, eine Rückmeldung auf sein Wissen erhält und so mögliche Fehler erkennen und korrigieren kann.

### ● **Die Lerninhalte sind gut strukturiert, praxisorientiert und in einer gut verständlichen Sprache formuliert und können somit schnell erlernt werden.**

### ● **Meine Lösung kann völlig flexibel und unkompliziert an die individuellen Bedürfnisse und Anforderungen des Benutzers angepaßt werden.**

Dies gilt sowohl für die auf die persönlichen Informationsbedürfnisse abgestimmte Internetquellensammlung als auch für den Zugriff und vor allem die Erweiterung des in HTML dokumentierten Wissens.

### ● **Die Kompetenzen, das sich die Lehrgangsteilnehmer im Rahmen des Lehrgangs**

aneignen, sind universell einsetzbar, das bedeutet, sie können fach- und themenunabhängig dazu genutzt werden, um Aufgaben schneller zu erfüllen, Probleme leichter zu lösen, Ideen schneller in die Tat umzusetzen, Marktlücken zu erkennen und effektiv zu nutzen und gute Entscheidungen zu treffen.

- Die im Lehrgang gezeigten Methoden (Suchmethoden, Bewertungsmethoden,...) sind flexibel nutzbar und können weitestgehend auf die persönlichen Vorlieben angepaßt werden; die starren, für Maschinen geschaffene Universalmethoden dagegen lassen nur einen geringen Handlungsspielraum zu.

Die individuelle Internetquellensammlung ermöglicht dem Benutzer beispielsweise, das Internet seinen individuellen Bedürfnissen entsprechend zu nutzen und schützt ihn gleichzeitig vor der Informationsflut.

Die individuelle Internetquellensammlung hat als Zielgruppe genau eine Person und ist damit an Zielgenauigkeit nicht zu übertreffen.

- Durch den hohen manuellen Anteil, vor allem bei der Auswahl der Internetquellen und der Bewertung der Relevanz und Qualität der gefundenen Informationsquellen/Informationen, ist die Informationsbeschaffung weitestgehend transparent für den Benutzer.

Der hohe manuelle Anteil der Informationsbeschaffung hat noch weitere Vorteile: Der Benutzer kann seine Aktivitäten weitestgehend selber bestimmen und verbessert mit jeder Recherche seine Netzwerkkompetenz.

- Der Lehrgang zeigt, wie man das Internet effektiv und sicher dazu nutzen kann, um qualitativ hochwertige und günstige Produkte und Dienstleistungen zu erwerben. Dieses Wissen vermeidet nicht nur ganz allgemein eine Menge Ärger, sondern fördert auch, daß alle Anbieter, auch kleine, die qualitativ hochwertige Produkte und Dienste anbieten, die gleiche Chance haben, gefunden zu werden wie große mit einem entsprechenden Marketing-Budget.

Nur ein fairer und freier Wettbewerb ermöglicht qualitativ hochwertige und preisgünstige Produkte und Dienstleistungen für den Kunden. Darüber hinaus gibt es (zum Glück) keinen Internetdienst, der als entsprechendes Universalwerkzeug eingesetzt werden könnte. Aus dem gesamten Internetangebot müssen deshalb die für den persönlichen Bedarf am besten geeigneten Diensteanbieter ausgewählt und in der jeweils passenden Kombination zur Erfüllung der anstehenden Aufgabe genutzt werden.

- Meine Lösung ermöglicht es dem Benutzer, das von mir konstruierte Grundwissen um eigene Informationen (nützliche Methoden und Werkzeuge, Tipps und Tricks,...) zu erweitern bzw. zu ergänzen und es so abzuspeichern, daß er es später schnell und gezielt wieder finden kann.

- Meine Lösung:

- ist fach-/themenunabhängig
- kann sowohl auf Einzelplatzrechnern als auch bei Netzwerk-Computern eingesetzt werden
- ist unabhängig vom verwendeten Betriebssystem und auch unabhängig von den verwendeten Anwendungsprogrammen
- ist mobil, kann also auch unterwegs genutzt werden

---

Durch die geschickte Kombination aus menschlicher Intelligenz und der richtigen Anwendung der für die jeweilige Aufgabenstellung geeigneten Methoden und Werkzeuge ermöglicht meine Lösung dem Anwender **eine äußerst flexible, weitestgehend Anbieter-unabhängige Nutzung der Internetdienste und Werkzeuge und bietet zudem größtmöglichen Schutz für die vertraulichen Daten des Benutzers und vor Manipulationen, etwa daß bestimmte Informationen gefiltert oder zahlungskräftige und mächtige Diensteanbieter bevorzugt werden** (Stichwort: [globales und konkurrenzloses Suchsystem](#)).

Die Lösung ermöglicht es dem Anwender, **selbst zu entscheiden, welche Internetquellen er für seinen jeweiligen Bedarf nutzen möchte und welche Quellen und Informationen er als tatsächlich relevant und qualitativ hochwertig beurteilt**. Darüber hinaus stellt das von mir entwickelte [Lehr-/Lernkonzept](#) dem Anwender eine Anleitung bereit, wie er sich aktiv mit den jeweils ausgewählten Informationen auseinandersetzen und so das gerade benötigte Wissen selbst konstruieren kann. Dieses Wissen kann er dann produktiv nutzen, beispielsweise um [Probleme zu](#)

[lösen](#) oder [gute Entscheidungen zu treffen](#). Die von mir erdachte Lösung baut wie das Internet selbst auf die konstruktive Zusammenarbeit zwischen Menschen, also auf deren Einfallsreichtum und deren Bereitschaft zur offenen Wissensteilung, und nutzt Maschinen, in deren Abläufe die Anwender keinen oder nur einen sehr begrenzten Einblick haben, nur dort, wo es sinnvoll ist.

Durch die Nutzung vieler unterschiedlicher Internetdienste (Stichwort: [Spezielle Such- und Bewertungsdienste](#)) die von Menschen bereitgestellt werden, die voneinander unabhängig sind (denken Sie hierbei an die [Mängel der aktuellen maschinellen Bewertungssysteme](#)), kann **ein Mensch mit grundlegender Netzwerkkompetenz wirksam dazu beitragen, das die Anbieter die sich ernsthaft anstrengen, um innovative und qualitativ hochwertige Produkte und Dienste auf dem Markt anzubieten, auch angemessen dafür belohnt werden** - und zwar weitestgehend unabhängig von ihrer Größe, ihrer finanziellen Ausstattung und ihren "einflussreichen" Beziehungen (Stichwort: [multinationale Konzerne](#)). Denken Sie hierbei an die **hervorragenden, technischen Möglichkeiten, die das Internet bietet, um die unterschiedlichsten Produkte und Dienste zu finden, und sich darüber zu informieren, von wem das Produkt/der Dienst stammt** (siehe Abschnitt "[Suche nach Informationen von und über Unternehmen](#)"). **Nicht zuletzt ist ein fairer und freier Wettbewerb auch die sicherste Methode um dauerhaft ein vielfältiges Angebot an innovativen und qualitativ hochwertigen Produkten und Diensten auf dem Markt sicherzustellen.**

---

### **Weitere Vorteile bietet das webbasierte Lernen, auf dem der Lehrgang letztendlich basiert:**

Das webbasierte Lernen bietet dem Lernenden die Möglichkeit, **an beliebigen Orten** (Arbeitsplatz, zuhause,...) und **bei Bedarf** (zeitunabhängig) sich **das gerade benötigte Wissen** anzueignen. Die zeitliche und räumliche Unabhängigkeit ermöglicht es, daß man **weitestgehend selbst bestimmen kann, wann** (morgens, abends,...), **wo und mit welchem individuellen Lerntempo** (sehr lange am Stück, nur kleine Häppchen) **man lernt**. Um an dem Lehrgang teilzunehmen, müssen die Lernenden keinen Urlaub nehmen und können sich die Zeit und die Kosten für die Anreise und die Übernachtung sparen.

Um komplexe Sachverhalte zu verstehen bzw. generell um Wissen aufzunehmen, ist eine gewisse Zeit notwendig. Ein oder mehrere Tage, in denen umfangreiche und komplexe Lerninhalte über Workshops vermittelt werden, sind meist viel zu wenig für eine **nachhaltige Wissensvermittlung**.

Da **die Teilnahme am Lehrgang zeitlich nicht begrenzt ist** und die Möglichkeiten der modernen Informations- und Kommunikationstechnologien optimal genutzt werden, haben die Teilnehmer über einen Zeitraum, den sie selbst bestimmen, die Möglichkeit, jederzeit Fragen zu stellen und Probleme mit den anderen Lehrgangsteilnehmern oder externen Experten auszudiskutieren.

Eine solch intensive und über einen längeren Zeitraum andauernde Betreuung wäre bei traditionellen Lernangeboten (direkte Betreuung durch einen Lehrer) schon aus Kostengründen undenkbar. Dieses Problem wird beispielsweise bei den häufig zur Weiterbildung von Mitarbeitern veranstalteten Workshops deutlich, da hier eine große Menge an Wissen in sehr kurzer Zeit vermittelt wird; die Wissensaufnahme ist für den Lernenden in der Regel jedoch nicht sehr nachhaltig. Häufig ist es auch so, daß viele wichtige Fragen erst nach dem Workshop oder Seminar auftreten, wo man dann den betreffenden Workshop- bzw. Seminarleiter nicht mehr fragen kann.

Selbst eine einfach gestaltete Plattform zum webbasierten Lernen ermöglicht es, die Lernenden **bei geringerem Kostenaufwand weitaus effizienter (weil über einen längeren Zeitraum und mit individueller Zuwendung) zu betreuen als bei den Präsenz-Workshops**.

### **Nachteile meiner Lösung: ↑**

Das [Lehr-/Lernkonzept](#) fordert vom Lernenden, daß er sich **aktiv mit Informationen auseinandersetzt** und in **oft mühevoller Arbeit seine Netzwerkkompetenz Schritt für Schritt selbst und in Zusammenarbeit mit anderen konstruiert** - vom Lernenden wird also **ein hohes Maß an selbstgesteuertem und selbstorganisiertem Denken und Handeln gefordert**. Vor allem beim Selbstlernkurs ist der Lernende selbst für seine Motivation und die Beurteilung seines Lernerfolgs verantwortlich.

### **Konkret bedeutet das, der Lernende muß :**

- sich selbst motivieren können, was am besten funktioniert, wenn man einen echten Sinn, hinter dem was man lernt, erkennen kann
- Wissenslücken selbständig erkennen und sich das jeweils fehlende Wissen aus eigenem Antrieb

heraus aneignen

- Probleme selbständig erkennen, die Probleme formulieren und nach geeigneten Lösungen suchen
- die Lerninhalte weitestgehend selbständig durcharbeiten, das bedeutet, er muß in der Lage sein, sich wichtige Fragen zum Erarbeiten der Lernziele selbst zu stellen, um so die entsprechenden Antworten zu finden
- selbst entscheiden, in welchem Umfang er selbst bzw. mit anderen zusammenlernen möchte
- die Antworten auf Fragen selbst in relevanten Dokumenten finden oder über entsprechende Personen erfragen

Das bedeutet, daß er die für seine Fragestellung relevanten Personen oder Dokumente (Fachartikel, Fachbücher, Diplomarbeiten,...) selbst im Internet suchen, finden und deren grundsätzliche Relevanz und Qualität bewerten muß. Weiterhin muß er eine entsprechende Lese- bzw. Schreibfertigkeit besitzen, um mit Hilfe der gefundenen Dokumente bzw. Personen tatsächlich eine Antwort auf seine Frage zu finden.

- die Lernzeiten selbst einteilen und vor allem einhalten
- festlegen, welche Lerninhalte er in welcher Reihenfolge lernen möchte
- den vielfältigen Versuchungen und Ablenkungen des Informationsdschungels "Internet" widerstehen und sich auf die tatsächlich notwendigen Inhalte konzentrieren können
- ...

**Es ist ganz deutlich zu erkennen, daß vieles was der Lernende eigentlich erst lernen bzw. trainieren soll, bereits zum Lernen notwendig ist.**

Dieser Widerspruch an sich ist jedoch bei näherer Betrachtung keiner mehr. Im Grunde trainiert man mit jeder praktischen Anwendung der im Lehrgang beschriebenen Methoden, Werkzeuge und Internetdienste allgemein seine Netzwerkkompetenz - Übung macht den Meister (Learn by doing). Mit jeder praktischen Erfahrung baut der Lernende seine Fähigkeiten als Problemlöser und Autodidakt weiter aus.

---

**Die individuelle Internetquellensammlung muß vom Benutzer selbst und in Zusammenarbeit mit anderen Anwendern gleicher Aufgaben- und Interessengebiete erstellt, optimiert und aktualisiert werden.** Das bedeutet, vom Anwender wird auch hier ein hohes Maß an selbstgesteuertem und selbstorganisiertem Denken und Handeln gefordert, das erforderlich ist, um seine Internetquellensammlung kontinuierlich und bestmöglich an die aktuellen Anforderungen anzupassen.

**Konkret bedeutet das, der Nutzer muß selbst entscheiden:**

- welche Informationen und Dienste er für seinen persönlichen Bedarf benötigt und dann selbständig die geeigneten Anbieter in dem umfassenden Informations- und Dienstangebot des Internets finden und die jeweils passenden auswählen
- in welchem Umfang er alleine lernen oder dazu mit anderen zusammenarbeiten möchte
- welche Informationen und Dienste er mit welcher Intensität fortlaufend beobachten möchte, um seine IQS auf dem neuesten Stand zu halten

### **! Hinweis**

Wird die individuelle Internetquellensammlung beruflich eingesetzt, so reicht die aufmerksame Beobachtung der jeweils relevanten Informationen und Dienste alleine nicht aus.

Hier geht es vor allem darum, daß die Informationen und Dienste produktiv genutzt werden, beispielsweise um eine aktuelle Beschaffungsmethode zu verbessern oder die aktuelle Neuentwicklung eines Produkts voranzutreiben und, falls die Information/der Dienst (auch) einem anderen Kollegen nützen könnte, sie/ihn an diesen weiterzuleiten.

● ...

**Der Anwender bestimmt also mit seiner Netzwerkkompetenz, seiner inneren Bereitschaft zur konstruktiven Zusammenarbeit mit anderen und seinem Engagement die Qualität seiner individuellen Internetquellensammlung und damit, falls sie beruflich eingesetzt wird, auch, wie produktiv er sie nutzen kann.**

Dies ist mit Blick auf die Zukunft sogar ein Vorteil, da in der [Antwort](#) zur einführenden Frage "Warum lohnt es sich für jeden Einzelnen, den Zeit- und Arbeitsaufwand zu investieren, um die kompetente



Nutzung der neuen Medien zu erlernen und aktiv am weltweiten Netzwerk "Internet" teilzunehmen?" bereits ausführlich beschrieben wurde, daß die innere Bereitschaft jedes Einzelnen, **seine Lebensaufgabe zu finden und auszuführen und mit anderen konstruktiv zusammenzuarbeiten**, eine wichtige Grundvoraussetzung ist, um zum einen die Grundlagen für den nächsten wirtschaftlichen Aufschwung zu schaffen und zum anderen den Zeitraum des jetzigen Abschwungs möglichst kurz zu halten.

Außerdem wurde im Fallbeispiel zum Abschnitt "[Ministerien, Behörden und Hochschulen](#)" festgestellt, daß **selbstgesteuertes und selbstorganisiertes Denken und Handeln** Fähigkeiten sind, denen in naher Zukunft vor allem im Beruf, aber auch im gesamten gesellschaftlichen Leben eine sehr hohe Bedeutung zukommen wird.

**Meine Lösung hilft Ihnen nicht nur, die grundlegenden Kompetenzen hierfür zu lernen, sondern liefert gleichzeitig das Übungsmaterial, um sie zu trainieren.**

## Fazit:

Mir fallen auch nach längerem Nachdenken keine weiteren wesentlichen Nachteile meiner Lösung mehr ein. Es gilt nun im nächsten Teilabschnitt darum, darüber nachdenken, ob es Möglichkeiten gibt, den Zeit- und Arbeitsaufwand zu reduzieren, der notwendig ist, um kompetent an Netzwerken teilzunehmen bzw. das Universalwerkzeug "Computer mit Internetzugang" effektiv und sicher als Wissens- und Problemlösungsquelle einzusetzen. Möglicherweise lassen sich auf diesem Wege nutzbringende Verbesserungen/Erweiterungen finden.

**Suche nach Möglichkeiten zur Verbesserung/Erweiterung meiner Lösung: ↑**

Dazu wird die Lösung als Aufgabe betrachtet, in ihre Teilaufgaben zerlegt, jede Teilaufgabe beschrieben und dann nachgedacht, wie man die jeweilige verbessern könnte. Grundsätzlich können diese Verbesserungen entweder darin bestehen, die Bearbeitung der jeweiligen Teilaufgabe durch maschinellen Einsatz stärker zu automatisieren oder aber die konstruktive Zusammenarbeit mit andern Anwendern gleicher Aufgaben-/Interessengebiete zu verstärken.

**Die wesentlichen Teilaufgaben sind:**

**1.**

**Mit Hilfe des Lehrgangs und der Lernanleitung (das Lehr-/Lernkonzept) die kompetente Teilnahme an Netzwerken trainieren bzw. die Kompetenzen aufbauen, mit denen das Universalwerkzeug "Computer mit Internetzugang" effektiv und sicher als Wissens- und Problemlösungsquelle genutzt werden kann.**

Das hierfür notwendige Grundwissen eignet sich der Lernende vor allem dadurch an, daß er die im Lehrgang beschriebenen Methoden (Suchmethoden, Bewertungsmethoden,...), Werkzeuge (Wissensdatenbank, Newsreader,...), Internetdienste und Informationen selbst praktisch anwendet, um alltagsrelevante Fragen zu beantworten (Wissen konstruieren) bzw. praxisrelevante Probleme zu lösen. **Das bedeutet, daß der Lernende seine Netzwerkkompetenz bzw. das dazu notwendige Wissen durch selbstbestimmtes Handeln und Entdecken, durch die Zusammenarbeit mit anderen** (etwa, um offene Fragen zu klären, die Fragen anderer zu beantworten, bestimmte Sachverhalte zu diskutieren oder das eigene Wissen, beispielsweise eine bestimmte Erkenntnis, zur Diskussion zu stellen) **und die dabei gemachten Erfahrungen und gewonnenen Erkenntnisse weitestgehend selbst konstruiert**. Mit diesem Wissen ist er dann in der Lage, die Internetdienste effektiv und sicher zur Informationsbeschaffung, Fernkommunikation, webbasierten Aus-/Weiterbildung und zum virtuellen Ein-/Verkauf zu nutzen. Den täglichen Umgang mit dem gelernten Wissen (die Anwendung der beschriebenen Methoden, Werkzeuge und Internetdienste, um Wissen zu konstruieren und Probleme zu lösen) nutzt der Lernende schließlich, **um seine grundlegende Netzwerkkompetenz auf seine berufliche Aufgabenstellung bzw. sein Interessengebiet hin zu optimieren** (individuelle Nutzung). Das bedeutet, daß er **die aktuell angewandten Methoden, Werkzeuge und Internetdienste fortlaufend in Frage stellt und gegebenenfalls verbessert (ergänzt, erweitert,...) oder gar durch neue, besser geeignete ersetzt**. Diese Verbesserungsmöglichkeiten können mit der entsprechenden Netzwerkkompetenz gezielt gesucht, genauso aber zufällig während einer Recherchetour gefunden werden. Auf jeden Fall müssen sie **leicht wiederauffindbar, also im richtigen**

**Zusammenhang in der individuellen bzw. allgemeinen Internetquellensammlung abgespeichert werden.**

**2.**

**Die Anforderungen an die auf den persönlichen Bedarf des Anwenders ausgerichtete individuelle Internetquellensammlung festlegen.**

Kurze und prägnante Beschreibung der Aufgabe der individuellen IQS und Formulierung der allgemeinen Anforderungen, die an sie zur optimalen Erfüllung der betreffenden Aufgabe/ganzheitlichen Abdeckung des gewählten Themas gestellt werden.

Danach wird der zur optimalen Erfüllung der Aufgabe/ganzheitlichen Abdeckung des Themas notwendige grobe Informations- und Dienstbedarf ermittelt.

**3.**

**Die grobe Zusammenstellung der individuellen Internetquellensammlung.**

Die für den im vorhergehenden Schritt ermittelten Bedarf in Frage kommenden Internetquellen suchen, finden und bezüglich ihrer Qualität bewerten und die ausgewählten Quellen nach Herkunft in die entsprechenden Kategorien einordnen. Die einzelnen Quellen kurz beschreiben und wichtige Einzelinformationen lokal (in der Wissensdatenbank) abspeichern.

**4.**

**Die Optimierung der zuvor zusammengestellten Internetquellensammlung.**

Die einzelnen Informationen und Dienste der zuvor nach Herkunft sortierten Internetquellen den jeweils relevanten Teilaufgaben/Einzelthemen zuordnen. Darüber hinaus werden im Rahmen dieser Teilaufgabe auch Überlegungen angestellt, beispielsweise, ob es sinnvoll ist, mit der vorhandenen Rechercheerfahrung die Informationen selbst zu suchen und auszuwerten oder ob es wirtschaftlicher ist, diese Aufgabe an einen externen Dienstleister (z. B. Information Broker) zu übertragen oder ob es Bereiche gibt, in denen die Informationsversorgung automatisiert werden kann.

**5.**

**Die kompetente und produktive Nutzung der Internetquellensammlung.**

Das bedeutet, die zuvor für den individuellen Bedarf zusammengestellten Internetquellen sollen effektiv und sicher für:

- die Informationsbeschaffung (suchen und informieren)
- die Fernkommunikation (fragen und diskutieren)
- die webbasierte Aus- und Weiterbildung (aus- und weiterbilden)
- den virtuellen Ein- und Verkauf (ein- und verkaufen)

genutzt werden.

Die individuelle Internetquellensammlung soll dabei vor allem zwei Kernfunktionen für den Benutzer erfüllen:

- **schnell und gezielt relevante Antworten auf konkrete Fragestellungen** hinsichtlich einer bestimmten Aufgabe/eines bestimmten Themas **finden**
- eine **umfassende und fortlaufende Informationsbeschaffung** bezüglich einer bestimmten Aufgabe/eines bestimmten Themas ermöglichen

Darüber hinaus muß die individuelle IQS **fortlaufend auf den neuesten Stand gebracht und an die aktuellen Anforderungen bezüglich des beruflichen Aufgabengebiets/des persönlichen Interessengebiets** angepaßt werden. Die Anpassung und Aktualisierung der individuellen IQS dient gleichzeitig der fortlaufenden Erweiterung bzw. Ergänzung des von mir konstruierten Grundwerks.

Da bei der Bearbeitung der einzelnen Teilaufgaben die konstruktive Zusammenarbeit mit anderen Anwendern gleicher Aufgaben-/Interessengebiete bereits eine wichtige Rolle spielt, soll nun im nachfolgenden überlegt werden, wie die Bearbeitung der Teilaufgaben durch einen höheren

maschinellen Einsatz optimiert werden könnte.

**Die Teilaufgaben könnten wie folgt verbessert werden:**

**1.**

**Mit Hilfe des Lehrgangs und der Lernanleitung (das Lehr-/Lernkonzept) die kompetente Teilnahme an Netzwerken trainieren bzw. die Kompetenzen aufbauen, mit denen das Universalwerkzeug "Computer mit Internetzugang" effektiv und sicher als Wissens- und Problemlösungsquelle genutzt werden kann.**

#### **Problem**

Der Aufbau der genannten Kompetenzen erfordert einen nicht unerheblichen Zeit- und Arbeitsaufwand.

#### **Lösung**

Das hierfür notwendige Wissen kann nur durch Lesen, Denken, praktische Übung, die konstruktive Zusammenarbeit mit anderen Anwendern gleicher Aufgaben- und Interessengebiete, das Ausprobieren eigener Ideen und vor allem durch die dabei gemachten praktischen Erfahrungen und die daraus gewonnenen Erkenntnisse gelernt werden. Diese zeit- und arbeitsaufwendige Teilaufgabe läßt sich nicht automatisieren, es können höchstens Lernhilfen eingesetzt werden, die ein effektiveres Lernen ermöglichen. Ein Beispiel hierfür wäre [Supermemo](#). Im wesentlichen hängt der Zeit- und Arbeitsaufwand, sich die Netzwerkkompetenz anzueignen, davon ab, wieviel Vorwissen und vor allem welche Einstellung man zum Lernen hat - wer gerne lernt, lernt schneller (siehe [Motivation](#)).

**2.**

**Die Anforderungen an die auf den persönlichen Bedarf des Anwenders ausgerichtete individuelle IQS festlegen.**

Die Anforderungen können nur durch den Benutzer selbst erkannt und formuliert werden.

**3.**

**Die grobe Zusammenstellung der individuellen IQS.**

#### **Problem**

Die Zusammenstellung der individuellen IQS erfordert eine grundlegende Netzwerkkompetenz und ist mit einem entsprechenden Zeit- und Arbeitsaufwand verbunden.

#### **Lösung**

Die relevanten Internetquellen könnte man grundsätzlich auch maschinell zusammenstellen.

**4.**

**Die Optimierung der individuellen IQS.**

#### **Problem**

Die Optimierung der individuellen IQS erfordert eine grundlegende Netzwerkkompetenz und ist mit einem nicht unerheblichen Zeit- und Arbeitsaufwand verbunden, der notwendig ist, um die Informationen und Dienste der einzelnen Internetquellen den jeweiligen Teilaufgaben/Einzelthemen zuzuordnen, für die sie benötigt werden.

#### **Lösung**

Da nur der Benutzer selbst genau wissen kann, welche Information er für welche Teilaufgabe/welches Einzelthema sinnvoll verwenden kann, sollte diese Aufgabe nicht automatisiert werden.

#### **Hinweis**

Hochwertig wird die individuelle IQS erst dann, wenn der Benutzer sie häufig verwendet und



die Informationen und Dienste der einzelnen Internetquellen penibel den jeweils relevanten Teilaufgaben/Einzelthemen zuordnet.

## 5.

### Die kompetente und produktive Nutzung der Internetquellensammlung.

#### Problem

Immer wieder gleiche Suchanfragen, die an immer wieder dieselben Internetquellen gerichtet werden, kosten unnötig Zeit und Arbeit. Auf die häufig benötigten Methoden, Werkzeuge und Dateien/Dokumente muss von unterschiedlichen Stellen des Computerarbeitsplatzes zugegriffen werden.

Die Quellen müssen zum großen Teil manuell aktualisiert werden, das bedeutet für den Benutzer einen gewissen Zeit- und Arbeitsaufwand, den er investieren muß, um sich bezüglich des betreffenden Themas/der Aufgabe auf dem Laufenden zu halten.

#### Lösung








Integration einer Suchfunktion, mit deren Hilfe man bestimmte, zuvor aus der persönlichen Sammlung ausgewählte, Informationsquellen mit der gleichen Suchanfrage in einer vorgegebenen Reihenfolge immer wieder durchsuchen kann. Darüber hinaus könnten die Suchmuster, die man häufiger benötigt, abgespeichert und so jederzeit wiederverwendet werden.

Diese Suchfunktion sollte über eine zentrale Stelle, ein persönliches Portal, zugänglich sein; dies gilt natürlich auch für die anderen häufig benötigten Werkzeuge (die Internetquellensammlung, wichtige Anwendungsprogramme, nützliche Tools,...) und Methoden (universelle Suchstrategie, Bewertungsmethoden,...).

Grundsätzlich könnte man die Informationsbeschaffung auch vollständig automatisieren und zwar über einen sogenannten Informationsagenten, der an Stelle des Benutzers weitestgehend selbständig nach zuvor festgelegten Informationen recherchiert. Ein solcher Informationsagent wäre in der Lage, auch dem weniger versierten Benutzer die Informationen aus dem Internet zu beschaffen, die er benötigt.

#### Kritische Betrachtung der gefundenen Verbesserungsmöglichkeiten: ↑

Welche Erweiterungen meiner Lösung wären denkbar und wie könnten Sie technisch umgesetzt werden ?

-  Die Zusammenstellung und Aktualisierung der individuellen Internetquellensammlung könnte maschinell durchgeführt werden. ↓
-  Ein Werkzeug, im übertragenen Sinne eine individuelle Meta-Suchmaschine, mit der der Benutzer bestimmte, zuvor aus seiner individuellen Internetquellensammlung ausgewählte, Informationsquellen nach bestimmten Informationen maschinell durchsuchen kann. ↓
-  Ein persönliches Portal, das es dem Benutzer ermöglicht, über eine zentrale Stelle auf häufig benötigte Werkzeuge (die individuelle Internetquellensammlung, wichtige Anwendungsprogramme, nützliche Tools,...) und Methoden (universelle Suchstrategie, Bewertungsmethoden, Suchmuster,...) zuzugreifen. ↓
-  Ein Programm zur vollautomatischen Informationsbeschaffung, das anstelle des Benutzers bestimmte Informationen auffindet. ↓
  -  das "Semantic Web" als technische Grundlage für das Programm ↓
  -  kritische Betrachtung einer Maschine zur vollautomatischen Informationsbeschaffung ↓
  -  Fazit:  
Mit welchen grundlegenden (mehr oder weniger lösbaren) Problemen müssen sich die Entwickler einer Maschine zur vollautomatischen Informationsbeschaffung auseinandersetzen

 **Lösung**

**Die Zusammenstellung und Aktualisierung der individuellen Internetquellensammlung könnte maschinell durchgeführt werden. ↑**

Die von mir entwickelte manuelle Methode zur Erstellung einer individuellen Internetquellensammlung (IIQS) könnte teilweise automatisiert werden. Die technische Basislösung hierfür könnte ein (globales) Internetquellensuchsystem sein, das die jeweils relevanten Internetdienste für das entsprechende Aufgabengebiet aus dem vorhandenen Angebot auswählt. Die Aktualisierung des Internetquellensuchsystem, also das Einfügen neuer Quellen und das Löschen nicht mehr vorhandener Quellen, wird durch eine Experten-Redaktion und durch die Benutzer sichergestellt. Darüber hinaus kann sich natürlich jeder Anbieter eines relevanten Internetdienstes in das System kostenlos eintragen.

Die Aktualisierung der individuellen IQS könnte durch den regelmäßigen und automatischen Online-Abgleich mit dem Internetquellensuchsystem erfolgen.

Dem Benutzer steht damit ein Werkzeug bereit, mit dem er sich einfach und schnell maschinell eine individuelle Internetquellensammlung erstellen lassen und diese ohne eigenen Aufwand automatisch aktualisieren lassen kann. Mit dieser Quellensammlung kann er dann die Suche manuell mit den auf sein Themen-/Fachgebiet begrenzten Informationsquellen durchführen.

### Kritische Betrachtung der Lösung:

Die kompetente und produktive Nutzung der IQS setzt eine ausreichende Netzwerkkompetenz beim Benutzer voraus. Für jemanden, der über diese Kompetenz verfügt und den Lehrgang als Nachschlagewerk nutzen kann, ist es aber auch kein allzu großes Problem, seine IQS manuell zu erstellen und zu aktualisieren, so daß sich der Aufwand, der für die praktische Umsetzung einer solchen Maschine notwendig wäre, kaum lohnen wird.

Hier sollte man wissen, welche enorme Schwierigkeit Maschinen, die keine Zusammenhänge verstehen, haben, wenn es darum geht, die Relevanz und Qualität der Quellen eindeutig zu bewerten. Dies ist aber zwingend notwendig, wenn man eine qualitativ hochwertige IQS haben möchte. Eine Maschine kann natürlich auch keine individuellen Kommentare und Tipps zur jeweils bewerteten Informationsquelle abgeben, die für den Benutzer in der Regel sehr wertvoll sind.

Die beschriebenen Probleme gelten in noch viel stärkerem Maße für die Aktualisierung der IQS.

Ein großer Nachteil ist außerdem, daß man dem Anbieter sein Aufgaben-/Interessengebiet zur maschinellen Erstellung der persönlichen Internetquellensammlung preisgeben muß. Wie viele aktuell angebotene Dienste (MyLycos, My Yahoo, Personalize Search,...), für die maschinelle Zusammenstellung einer persönlichen Suchmaschine beweisen, bestimmt der Anbieter den Rahmen der Personalisierungsmöglichkeiten und zwar nicht nur das Layout der persönlichen Suchseite, sondern auch das Informations- und Dienstangebot, das mit der persönlichen Suchmaschine durchsucht werden kann.

 **Lösung**

**Ein Werkzeug, im übertragenen Sinne eine individuelle Meta-Suchmaschine, mit der der Benutzer bestimmte zuvor aus seiner individuellen Internetquellensammlung ausgewählte Informationsquellen nach bestimmten Informationen maschinell durchsuchen kann. ↑**

Viele Fragestellungen beziehen sich auf Themen, in denen die Erkenntnisse rasch zunehmen. Das bedeutet, daß zu diesen Themen mit den selben Suchanfragen in regelmäßigen Zeitabständen nach neuen Veröffentlichungen gesucht werden muß. Diese Internetseiten mit den Suchergebnissen kann man beispielsweise als Lesezeichen abspeichern und bei Bedarf wieder verwenden und sich so den

Aufwand sparen, gleiche Suchanfragen immer wieder einzugeben. Die andere Möglichkeit ist, die Suchanfrage aus der Zwischenablage in die entsprechenden Eingabefelder der Suchdienste zu kopieren.

Noch komfortabler wäre ein Werkzeug, mit dem man diese Recherchen maschinell durchführen kann.

Nachdem die Suche gestartet wurde, richtet das Programm die betreffende Suchanfrage gleichzeitig an die zuvor ausgewählten Quellen. Dazu müssen natürlich entsprechende Schnittstellen zu den betreffenden Quellen (Fachdatenbanken, lokale Suchmaschinen von Websites, allgemeine Suchmaschinen,...) vorhanden sein. Diese Schnittstellen übersetzen die Suchanfrage des Benutzers in die spezifische Abfragesprache des betreffenden Internetdienstes und legen fest, wie die Suchergebnisse der abgefragten Dienste dargestellt werden sollen. Die Suchergebnisse bzw. Treffer müssen vom Benutzer selbst ausgewertet werden.

Am besten wäre, wenn der Anwender die Schnittstellen, beispielsweise mit Hilfe einer einfachen Agentensprache, selbst programmieren könnte, die Programmiersprache sollte von einem Fachkundigen in kurzer Zeit erlernt werden können.

Je nachdem, welche Kombination von Suchdiensten abgefragt werden soll, kann auch eine der zahlreichen allgemeinen und speziellen Meta-Suchmaschinen, die im World Wide Web angeboten werden, für die Suche eingesetzt werden. Bei Patenten und Ausschreibungen beispielsweise werden von entsprechenden Datenbank Anbietern bereits automatisierte Überwachungsrecherchen angeboten. Um jedoch die Recherche-Werkzeuge flexibler auswählen zu können, wird man um die lokale Installation eines entsprechenden Tools, beispielsweise ein entsprechender **Suchagent**, nicht herumkommen.

### **Kritische Betrachtung der Lösung:**

Ein solches Werkzeug wäre eine feine Sache, da man mit dessen Hilfe die Suche nicht mehr selbst und vor allem konsequent zum erforderlichen Zeitpunkt durchführen muß, sondern daß dies automatisch von der Maschine erledigt wird. Darüber hinaus spart man sich möglicherweise auch den Aufwand, die neuen Treffer immer wieder mühsam von den bereits bekannten zu trennen, und muß sich nur noch mit den tatsächlich neuen Informationen (in zusätzlichen oder aktualisierten Internetseiten) auseinandersetzen, wobei natürlich die Hauptarbeit, nämlich das Be- und Auswerten der neuen Informationen, nicht durch die Maschine, sondern (zum Glück) eigenhändig durchgeführt werden muß.

Allerdings wäre der technische Aufwand, um die dazu erforderlichen Schnittstellen zu den jeweiligen Internetdiensten zu schaffen, enorm, wenn nicht gar unmöglich. Ohne diese Schnittstellen aber hätte man nur einen sehr geringen Vorteil, da erst sie die komfortable Auswertung der Suchergebnisse unterschiedlicher Internetdienste ermöglichen würden.

Da grundsätzlich die meisten Tätigkeiten, die bei der Informationsbeschaffung anfallen, von einem Menschen mit einer entsprechenden Netzwerkkompetenz schneller und wesentlich flexibler ausgeführt werden können wie von einem Tool, stellt sich die Frage, ob sich der Kostenaufwand, ein solches Tool zu programmieren, lohnen würde.



**Interessante Suchwerkzeuge kann man beispielsweise mit Hilfe des speziellen Software-Suchdienstes "Winload.de" unter der Rubrik "Internet" bzw. "Bookmark/Suche" finden.**

- Auktionstools
- [Bookmark/Suche](#)
- Browser
- Downloadtools
- eMail
- File Sharing
- FTP
- HTML-Editoren
- Internettools
- IRC/Chat
- News
- PlugIns
- Sonstiges
- Webmaster

01 02 03 04 05 06 [nächste Seite](#) ▶

Sortieren nach: [Datum](#) | [Alphabet](#) | [Userranking](#) | [Dateigrösse](#)  
[Downloads](#) | [Nur Freeware](#) | [Nur Shareware](#)

Es wurden **68 Einträge** unter der gewählten Kategorie gefunden:

### CUB meta 0.95

Oberkategorie: [Internet](#) Kategorie: [Bookmark/Suche](#)

Kurzbeschreibung				
Explorer Erweiterung zur Metasuche über mehrere hundert Suchmaschinen und Themen				
Sprache	Downloads	Userranking	Betriebssystem	
	 592	 1,3		
Datum	Dateigrösse	Lizenz	Kosten	Screenshots
25.11.2003	1,31 MByte	Adware	--	ja
<a href="#">details</a> ▶ <a href="#">download</a> ▶				

- .
- .
- .

### equero Personal Edition 1.12 1.12

Oberkategorie: [Internet](#) Kategorie: [Bookmark/Suche](#)

Kurzbeschreibung				
Meta-Suchmaschine zum Selberbauen				
Sprache	Downloads	Userranking	Betriebssystem	
	 362	 3,0		
Datum	Dateigrösse	Lizenz	Kosten	Screenshots
28.03.2002	6,30 MByte	Freeware	0,00	nein
<a href="#">details</a> ▶ <a href="#">download</a> ▶				

# equero Personal Edition 1.12 1.12

## Programmbeschreibung

Das im nicht-kommerziellen Einsatz kostenlose Programm basiert auf der equero Technologie zur Kommunikation mit beliebigen Webdiensten. Mit Hilfe der equero-Workbench können nach kurzer Einarbeitungszeit beliebige Websites angeschlossen und abgefragt werden. Im Gegensatz zu herkömmlichen Suchmaschinen ist die Suche auch in dynamischen, auf Datenbanken basierenden Websites möglich. Damit können z.B. Preisvergleiche durch Abfragen mehrerer Online-Shops oder Nachrichten-Übersichten realisiert werden. Im Lieferumfang ist eine einfache, vorkonfigurierte Meta-Suchmaschine enthalten.






Sprache	Userranking	Betriebssystem		
	 3,0	    		
Datum	Dateigröße	Lizenz	Kosten	Downloads
				
28.03.2002	6,30 MByte	Freeware	0,00	362
 <a href="#">Programm-Homepage</a>			<a href="#">download</a> 	

Abb. Winload.de - Bookmark/Suche - Equero Personal Edition 1.12

Es empfiehlt sich, die nachfolgenden, vom Hersteller des vorher genannten Suchwerkzeugs "Equero Personal Edition 1.12" bereitgestellten Hintergrundinformationen zu lesen, da die darin enthaltenen Problembeschreibungen für das tiefere Verständnis der Probleme, die mit der Realisierung dieses Verbesserungsvorschlags und des folgenden verbunden sind, von hoher Bedeutung sind.

Home

Hintergrund

Vision

Mission

Deep Web

Technologie

Einsatz

Produkte

Download

Jobs

Unternehmen

Presse

Kontakt



# Hintergrund

## Das Internet enthält alle Informationen der Welt

Die im weltweiten Datennetz verfügbare Informationsmenge wächst unaufhörlich. Herkömmliche Suchmaschinen können nur einen kleinen Teil dieser Informationen finden: Die meisten Informationen - und fast immer auch die entscheidenden - sind in Datenbanken gespeichert, auf die man nur über dynamisch generierte Webseiten zugreifen kann. Sie verschließen sich damit dem Zugriff der gängigen Suchmaschinen.

## Es existieren keine offenen Schnittstellen

Verschiedene Standards und offene Schnittstellen wurden bereits beschlossen. Leider ist der größte Teil der im Internet vorhandenen Informationen nicht über diese Schnittstellen zugänglich. Mit equero-Technologie lassen schnell und einfach Schnittstellen zu allen Webdiensten einrichten - für komfortablen Datenaustausch mit den gewünschten Webdiensten!

## Datenextraktion und Datendistribution

Neben der Datenextraktion aus Webdiensten zur Auswertung oder weiteren Verarbeitung der Daten entsteht im Internet ein zunehmender Bedarf an Verteilung von Daten auf verschiedene Webdienste. Für die meisten Anwendungsfälle werden für den Datenaustausch spezielle Schnittstellen zwischen den unterschiedlichsten Systemen programmiert, obwohl die für den Datenaustausch nötige Funktionalität ohnehin über die Website zur Verfügung gestellt wird. equero nutzt die vorhandene Funktionalität zur Erstellung von Schnittstellen zu einem Bruchteil der Kosten herkömmlicher Schnittstellenprogrammierung.



Home

Hintergrund

Vision

Mission

Deep Web

Technologie

Einsatz

Produkte

Download

Jobs

Unternehmen

Presse

Kontakt



# Vision

## Vision - Accessing the Deep Web

Das Internet ist bereits jetzt die zentrale Informationsplattform der Menschheit. Dieser Trend wird sich in den nächsten Jahren noch weiter verstärken. Entscheidend ist deshalb der gezielte und automatisierte Zugriff auf die gewaltigen Informationsmengen, die über das Internet verfügbar sind.

Vereinfacht gesehen besteht das Internet aus einfachen, statischen Seiten ("Surface Web") und aus dynamisch generierten Seiten ("Deep Web"), hinter denen sich oft hochkomplexe Systeme mit integrierten Datenbanken verbergen. Die Zukunft des Internet ist dabei mehr und mehr durch das "Deep Web" charakterisiert.

Während statische Seiten durch herkömmliche Technologien wie Crawler, Spider etc. automatisch katalogisiert und in Suchmaschinen angeboten werden können, scheitert dieser Ansatz bei dynamisch generierten Seiten.

Die Vision von equero ist es, die weltweit führende Technologie zum Zugriff auf Informationen des Deep Web weiterzuentwickeln und in möglichst vielen Anwendungen einzusetzen.

So sind wir in der Lage, unseren Kunden einen entscheidenden Wettbewerbsvorteil zu bieten, indem wir ihnen den Zugriff auf beliebige Informationen, die bisher nur auf dynamischen Websites verfügbar sind, ermöglichen.

<b>Home</b>	
<b>Hintergrund</b>	<h1>Deep Web</h1>
Vision	
Mission	
Deep Web	<b>Das Deep Web - Unsere Arbeit</b>
<b>Technologie</b>	equero hat das Ziel, die Informationen in den Datenbanken, auf denen dynamische Websites basieren, zur maschinellen Weiterverarbeitung nutzbar zu machen. Da hierbei im Allgemeinen viele Quell-Websites einzubeziehen sind, ist der einzige praktikable Weg, die Websites genau so zu nehmen, wie sie sind. Deswegen mussten und müssen wir uns mit einigen unangenehmen Realitäten über das heutige Web abfinden, die hier kurz beschrieben werden.
<b>Einsatz</b>	
<b>Produkte</b>	
<b>Download</b>	
<b>Jobs</b>	
<b>Unternehmen</b>	<b>Web-Standards? Werden nicht eingehalten.</b>
<b>Presse</b>	In der Theorie hört es sich schön an: Das W <sup>3</sup> C definiert für jede mögliche Technologie und Sprache einen Standard, der von den Betreibern von Websites dann eingehalten wird. Praktisch sieht es jedoch leider ganz anders aus: Die großen Browser-Hersteller sind offenbar der Ansicht, dass ihr Browser dann besser ist, wenn er noch "schlechtere" (technisch) Websites irgendwie halbwegs anzeigen kann. Die Webmaster dagegen entwickeln üblicherweise unter Zeitdruck eine WebSite. Sobald die von den aktuellen Browser-Versionen dargestellt werden kann, ist für sie die Arbeit getan. Diese Praxis führt zu immer schlechterem Code, der von Web-Servern generiert wird. Zum Teil ist es wirklich haarsträubend, wie stark der HTML Code auch großer Websites den Standard verletzt.
<b>Kontakt</b>	
   	

### **Offene Daten-Schnittstellen? Nicht verfügbar.**

Obwohl es Datenformate gibt, die das Erfragen von Informationen wesentlich erleichtern würden (wie z.B. XML), gibt es leider in der Realität so gut wie keine offenen Schnittstellen, um Daten in einem solchen Format zu beziehen. Wenn überhaupt, werden diese Standards für eine Kommunikation zwischen zwei Systemen benutzt, bei der beide Partner mitarbeiten. Einerseits scheuen die Betreiber von WebSites den Aufwand, solche modernen Datenschnittstellen in ihre Systeme einzubinden, andererseits fehlt ihnen oft auch jegliche Motivation dafür.

### **Usertracking? Eine beliebte Unsitte.**

Viele wesentliche WebSites haben offenbar ein großes Interesse daran, das Benutzerverhalten möglichst genau zu protokollieren. Die zweifelhafte Motivation ist oft eine anschließende Vermarktung dieser Daten zu Werbezwecken. Dazu setzen sie Usertrackingmethoden wie Cookies oder Session-ID's ein., Um Daten von einer solchen WebSite beziehen zu können, müssen diese natürlich simuliert werden. Bei vielen Zugriffen sind auch Optimierungen wie Cookiepools etc. erforderlich.

### **Gleiches System = gleiche Webausgabe? Mitnichten.**

Schon durch geringe Änderungen an der Konfiguration oder durch kleinere Anpassungen, die in der Praxis fast immer vorgenommen werden, wird die Ausgabe von WebSites durch ein Standardsystem sich grundlegend ändern. Wir mussten feststellen, dass der Ansatz, spezielle Webadapter für die wichtigsten großen Systeme zu entwickeln, schon allein an diesem Punkt scheitert. Aber schon auf Grund der Tatsache, dass neue Content-Management-Systeme und ähnliche Werkzeuge in einem enormen Tempo auf dem Markt erscheinen, ist es vollkommen illusorisch, die großen Systeme generisch anzuschließen, um damit jede wesentliche Website abfragen zu können.

### **Teilweise kaum nachvollziehbares Verhalten von Webservern? Leider üblich.**

Während unserer Arbeit haben wir die merkwürdigsten Verhaltensweisen von Webservern kennen gelernt. Es kommt sehr häufig vor, dass extrem unsinnig erscheinende http-Kommunikation zwischen Browser und Webserver abläuft, ohne dass ein Benutzer etwas davon merkt. Für ein System wie equero ist es jedoch zwingend erforderlich, diese Kommunikation mit einem Webserver nach einem identischen Muster generisch für verschiedene Parameter durchführen zu können.

### **Wir haben diese Probleme gelöst!**





Gerade wegen dieser unangenehmen Eigenschaften vor allem des Deep Web sind wir stolz, dass wir es trotzdem geschafft haben, Webserver in wenigen Minuten an equero anschliessen zu können. Dafür mussten wir sehr viel Know How aufbauen und etliche Mannjahre Arbeit in die Entwicklung des Systems, vor allem aber auch der Tools, die uns beim Anschluss von WebSites unterstützen, investieren.





### **Abb. Equero.de - Hintergrund**


---

Die nachfolgenden Inhalte beschreiben näher, was die equero-Technologie kann. Internetdienste, die die Equero-Technologie in der Praxis einsetzen, sind der Live-Preisvergleich bei ShoppingScout24 "[www.shoppingscout24.de](http://www.shoppingscout24.de)", die aktuellen Finanznachrichten bei FinAgent "[www.finagent.de](http://www.finagent.de)" oder die aktuellen Computer-Nachrichten bei CompAgent.de "[www.compagent.de](http://www.compagent.de)".

---

<b>Home</b>	
<b>Hintergrund</b>	
<b>Technologie</b>	<h1>equero - Technologie</h1>
Funktionen	<b>Weltweit werden immer komplexere Dienste über das Internet angeboten. Die Angebote reichen von einfachen Informationsseiten über Suchmaschinen und Kataloge bis hin zu komplizierten Shop-Systemen.</b>
Flexibilität	
System	
<b>Einsatz</b>	
<b>Produkte</b>	Dabei ist allen Systemen gemeinsam, dass keine offenen Schnittstellen zum Datenaustausch existieren, obwohl entsprechende Standards wie XML schon vor längerer Zeit verabschiedet wurden.
<b>Download</b>	
<b>Jobs</b>	<b>Dieses Problem löst die equero-Technologie.</b>
<b>Unternehmen</b>	Die equero Technologie wurde entwickelt, um alle existierenden Webdienste automatisch steuern und auswerten zu können. Sie gestattet es, jede Information, die über Webdienste verfügbar ist, in eigenen Software-Systemen weiter zu verwenden, sowie Informationen aus den eigenen Systemen in beliebige existierende Webdienste automatisch zu übernehmen.
<b>Presse</b>	
<b>Kontakt</b>	
   	

<b>Home</b>	
<b>Hintergrund</b>	
<b>Technologie</b>	<h1>Funktionen</h1>
Funktionen	<b>Die equero-Technologie bietet im Wesentlichen folgende Funktionen:</b>
Flexibilität	
System	
<b>Einsatz</b>	- Steuerung von Webdiensten unabhängig von deren Komplexität
<b>Produkte</b>	- Überführung von Daten in die speziellen, von den jeweiligen Webdiensten benötigten Formate und umgekehrt
<b>Download</b>	
<b>Jobs</b>	- Auswertung und Interpretation empfangener Daten mit anschließender fallweiser Ausführung bestimmter Aktionen
<b>Unternehmen</b>	- Einfacher, schneller und damit kostengünstiger Anschluss neuer Webdienste
<b>Presse</b>	
<b>Kontakt</b>	- Verwaltung einer großen Anzahl unterschiedlicher Webdienste
   	

Home	
Hintergrund	
Technologie	<b>Flexibilität</b>
Funktionen	
Flexibilität	<b>Die equero-Technologie wurde entwickelt, um flexibel mit den vielseitigen Schwierigkeiten im Internet umgehen zu können. Dabei sind folgende Probleme gelöst worden:</b>
System	
Einsatz	<b>Abfrage von Informationen aus dynamischen, zur Laufzeit generierten Webseiten</b>
Produkte	
Download	Damit ist der Zugriff auch auf die oftmals zu Grunde liegenden Datenbanken möglich. Herkömmliche Suchmaschinen können Informationen lediglich aus statischen Webseiten beziehen.
Jobs	
Unternehmen	<b>Nachbildung von User-Tracking-Methoden</b>
Presse	Viele Websites personalisieren die dargestellten Informationen für die einzelnen Benutzer mittels Cookies, Session-IDs oder Login-Vorgängen (z.B. beim Warenkorb in Online-Shops). Damit ist ein einfacher Zugriff auf diese Informationen nicht mehr möglich. equero bildet die Session-Verwaltung nach und gewinnt dadurch Zugriff auf die relevanten Informationen.
Kontakt	
	
	<b>SSL-Verschlüsselung</b>
	Der Zugriff auf sensible Daten ist oftmals nur über eine verschlüsselte Verbindung möglich. equero beherrscht auch die Kommunikation mit Websites über verschlüsselte Internet-Verbindungen.
	<b>Auswertung der Informationen</b>
	equero führt im Unterschied zu herkömmlichen Suchmaschinen keine einfache Volltextsuche in den einzelnen Internet-Seiten durch, sondern interpretiert an Hand der HTML-Struktur die gewonnenen Daten auf vordefinierte Weise. Damit ist equero in der Lage, die gewonnenen Daten geordnet und strukturiert in einer Datenbank ab zu legen.

**Abb. Equero.de - Technologie**

Ebenfalls sehr interessant waren die Ideen, wie man eine individuelle Meta-Suchmaschine sinnvoll einsetzen kann - darunter war auch die für diesen Verbesserungsvorschlag relevante Idee der "themenspezifischen Recherchertools".

Home

Hintergrund

Technologie

Einsatz

Marktplätze

Informationen

Transaktionen

Referenzen

Produkte

Download

Jobs

Unternehmen

Presse

Kontakt



# elektronische Marktplätze

**Die wesentliche Eigenschaft eines elektronischen Marktplatzes für einen Käufer ist die Tatsache, dass er sich hier ein besonders guten Überblick über Produkte und Preise verschaffen kann.**

Da die Informationen über Produkte und Preise üblicherweise über dynamische Websites der Anbieter oder Vermittler verfügbar sind, kann ein elektronischer Marktplatz nur dann sinnvoll funktionieren, wenn er eine Technologie wie equero zum Einbeziehen dieser Informationen einsetzt. Hier einige Beispiele:

## Shopping-Portale

Viele Benutzer kaufen einige Produkte bereits heute häufig im Web, dieser Trend wird sich noch verstärken. Eine Übersicht und ein Preisvergleich kann ein equero-basiertes System ideal bieten.


## Immobilien-Portale

Auch für Wohnungssuchende ist es von enormer Wichtigkeit, alle Angebote von Maklern aus ihrer Region sehen zu können. Da die Makler diese Angebote über ihre eigenen dynamischen Websites publizieren, kann ein equero - System hier optimal eine Übersicht generieren.

## B2B - Marktplätze

Gerade im Business - to - Business - Bereich entstehen immer wichtigere Marktplätze. Auch diese können nur funktionieren, wenn alle relevanten Informationen für das entsprechende Gebiet zugreifbar sind. Dies lässt sich am einfachsten durch ein equero - System realisieren, das einfach die existierenden Systeme der einzelnen Marktteilnehmer einbezieht.



<b>Home</b>	
<b>Hintergrund</b>	<h1>fachliche Informationen</h1>
<b>Technologie</b>	
<b>Einsatz</b>	
Marktplätze	
Informationen	
Transaktionen	
Referenzen	
<b>Produkte</b>	
<b>Download</b>	
<b>Jobs</b>	
<b>Forum</b>	
<b>Presse</b>	
<b>Kontakt</b>	
	

**In etlichen Bereichen der Wirtschaft, aber auch des Privatlebens besteht ein hoher Bedarf an fachspezifischen Informationen.**

Da diese jedoch vermehrt über dynamische Websites verfügbar sind, scheitert die Benutzung herkömmlicher Suchmaschinen häufig. Hier kann equero ideal eingesetzt werden, um die relevanten Informationen aus spezifischen Websites zu beziehen und weiterzuverarbeiten.

### **Newsticker**

Die Entwicklung einiger Gebiete auf der Welt vollzieht sich so schnell, dass stündlich neue Informationen über bestimmte Bereiche veröffentlicht werden. Da hier dynamische Websites als Publikationsmittel ideal sind, entstehen immer mehr Newsticker in den verschiedensten Gebieten. Hier werden kurze aktuelle Nachrichten veröffentlicht.

Da dies auf vielen verschiedenen Newsportalen geschieht, wird dringend eine Technologie wie equero benötigt, die es ermöglicht, sehr mächtige Newsticker zu realisieren, die dann alle wesentlichen aktuellen Nachrichten anbieten können.

### **Themenspezifische Recherchertools**

Immer umfangreicheres Wissen aus immer mehr Bereichen wird in großen Datenbanksystemen abgelegt, die häufig eine Webschnittstelle anbieten. Eine sinnvolle Recherche sollte in den wichtigsten dieser Systeme durchgeführt werden. Dies ist nur dann mit vertretbarem Aufwand möglich, wenn eine Technologie wie equero genutzt wird, um diese Systeme zu verbinden, und damit alle wesentlichen Informationen in die Recherche einbeziehen zu können.









Home	
Hintergrund	<h1>Transaktions-Systeme</h1>
Technologie	
Einsatz	
Marktplätze Informationen Transaktionen Referenzen	<p><b>Viele Systeme ermöglichen es, über eine Webschnittstelle Transaktionen aus verschiedensten Bereichen durchzuführen.</b></p> <p>Wenn solche Systeme in eigene Anwendungen integriert werden sollen, ist eine Technologie zum Anschluß dynamischer WebSites unumgänglich.</p>
Produkte	<b>Börsen-Tradingsysteme</b>
Download	Denkbar wäre zum Beispiel ein System, das über Online-Broker Aktienkäufe und -verkäufe realisiert. Hier könnte equero optimal eingesetzt werden.
Jobs	
Unternehmen	
Presse	<b>Fahrkarten- und Ticketbuchungssystem</b>
Kontakt	Ein System zur Buchung einer Reise ist darauf angewiesen, bei verschiedenen Anbietern (Bahn, Mietauto, Flugzeug, Hotel, ÖPNV etc.) automatisiert eine Ticketbestellung vornehmen zu können. Auch dies ist nur realisierbar, wenn eine Technologie wie equero zum Anschluss der benutzen dynamischen Websites verwendet wird.
   	

Abb. Equero.de - Einsatz

## Lösung

**Ein persönliches Portal, das es dem Benutzer ermöglicht, über eine zentrale Stelle auf häufig benötigte Werkzeuge (die individuelle Internetquellensammlung, wichtige Anwendungsprogramme, nützliche Tools,...) und Methoden (universelle Suchstrategie, Bewertungsmethoden, Suchmuster,...) zuzugreifen. ↑**

Integration der individuellen Internetquellensammlung in ein persönliches Zugangsportal, über das man von zentraler Stelle aus auf die häufig benötigten Arbeitsmittel schnell und gezielt zugreifen kann, beispielsweise auf:

-  die Anleitungen zu wichtigen Methoden (universelle Suchstrategie, universelle Methode zur Bewertung von Informationsquellen,...)
-  häufig benötigte webbasierte Werkzeuge (Such- und Informationsdienste, Kommunikationsdienste,...)
-  häufig benötigte lokale Werkzeuge (Wissensdatenbank, Textverarbeitungsprogramm, Verschlüsselungs- und Signierprogramm,...)
-  häufig benötigte lokal gespeicherte Dateien und Dokumente (aktuelle Projektdateien,...)

Wie ein solches [persönliches Zugangsportal mit dem Internet Explorer](#) erstellt werden kann, wurde in der Lektion "Such- und Informationsdienste" im Abschnitt "Web-Portale" demonstriert.

## Hinweis

Da die meisten der Lesezeichensysteme der gängigen Web-Browser (Internet Explorer, Netscape Navigator, Opera, Mozilla,...) Verweise auf beliebige Internetseiten, lokal gespeicherte Dokumente und Programmdateien ermöglichen, kann ein solches Portal natürlich auch mit Hilfe des Lesezeichensystems realisiert werden. Allerdings muß man dann über hierarchisch organisierte Kategorien auf die Inhalte zugreifen und bekommt diese nicht auf einer Bildschirmseite übersichtlich

präsentiert.

### Kritische Betrachtung der Lösung:

Da gerade der Internet Explorer mit seiner ActiveX-Technologie immer wieder durch neu gefundene Sicherheitslücken auffällt, könnten die vertraulichen Portal-Inhalte, die sehr viel über das Benutzerverhalten verraten, durch entsprechende Spionagefunktionen ausgelesen werden.

Andererseits bietet ein persönliches Zugangportal seinem Benutzer aber sehr große Mehrwerte, deshalb wäre es sinnvoll, eine professionelle Lösung zu suchen, die zum einen wesentlich einfacher und flexibler zu bedienen ist als die mit dem Internet Explorer realisierte, und zum anderen einen ausreichenden Schutz für die Portal-Inhalte bietet. Zum Aufbau von Unternehmensportalen gibt es zahlreiche professionelle Lösungen.

Wie der Test der Wissensdatenbank "askSam" gezeigt hat, könnte man auch mit den Hyperlink-Funktionen dieses Informationssystems ein [persönliches Zugangportal](#) erstellen. Die [Hypertext-Funktionen von askSam](#) ermöglichen wie der Internet Explorer, **aber mit wesentlich besseren Schutzmöglichkeiten**, vielfältige Verknüpfungen zwischen einem Text und beliebigen lokalen Dokumenten, Dateien und Textstellen sowie externen Webinhalten.

Weitere, für viele möglicherweise ausreichende Möglichkeiten zum Aufbau eines [persönlichen Zugangsportals](#) bietet der [MindManager](#) oder das nachfolgend kurz beschriebene Werkzeug "URLsave" ([www.urlsave.de](http://www.urlsave.de)). Mit Hilfe von URLsave könnte die Internetquellensammlung komfortabler und sicherer genutzt werden als mit einem bestimmten Web-Browser.

### Die Verwendung von URLsave bietet folgende wesentliche Vorteile gegenüber der Nutzung der Lesezeichensysteme der gängigen Web-Browser, insbesondere des Internet Explorer's:

- es bietet einen gewissen Schutz davor, daß die Lesezeichen durch einen anderen PC-Benutzer oder Schad- und Spionageprogramme ausgelesen oder gar verändert werden können, da alle URL's passwortgeschützt in einer Access- Datenbank gespeichert sind
- Lesezeichen können von den unterschiedlichsten Web-Browsern (Internet Explorer, Opera, Netscape Navigator, Mozilla, Mozilla FireFox) in die URL-Save-Datenbank importiert bzw. aus dieser in die Lesezeichensysteme der genannten Web-Browser exportiert werden; URLsave ermöglicht so auch den problemlosen Wechsel von einem zum anderen oder die Benutzung mehrerer Web-Browser
- zu jedem Lesezeichen können beliebige Kommentare und ein Bild als Zusatzinformation gespeichert werden, die gespeicherten URL's bzw. die zuvor eingegebenen Kommentare können dann nach bestimmten Stichwörtern durchsucht werden, um einen bestimmten Link aus der Linksammlung schnell und gezielt wieder zu finden
- innerhalb eines Netzwerks können die jeweils vorgesehenen Anwender auf eine gemeinsame Lesezeichensammlung zugreifen
- das im MS-Access-Datenbankformat vorliegende Lesezeichensystem kann um individuelle Funktionen erweitert werden.
- ....



## URLsave 2.53

Oberkategorie: **Internet** Kategorie: **Bookmark/Suche**

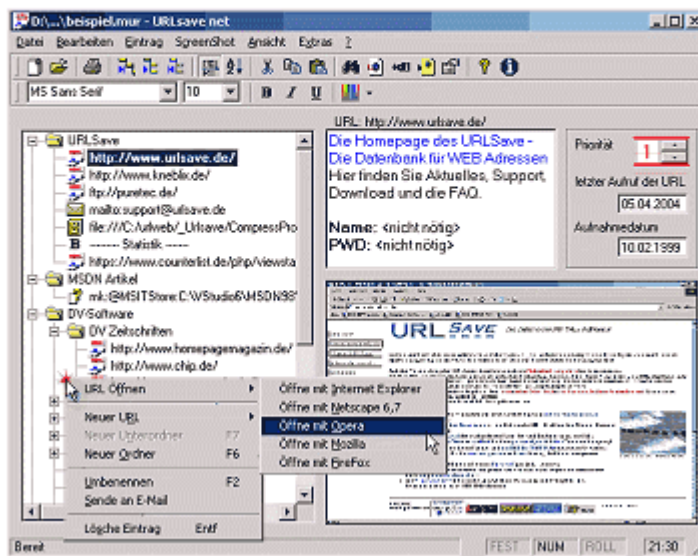
### Kurzbeschreibung

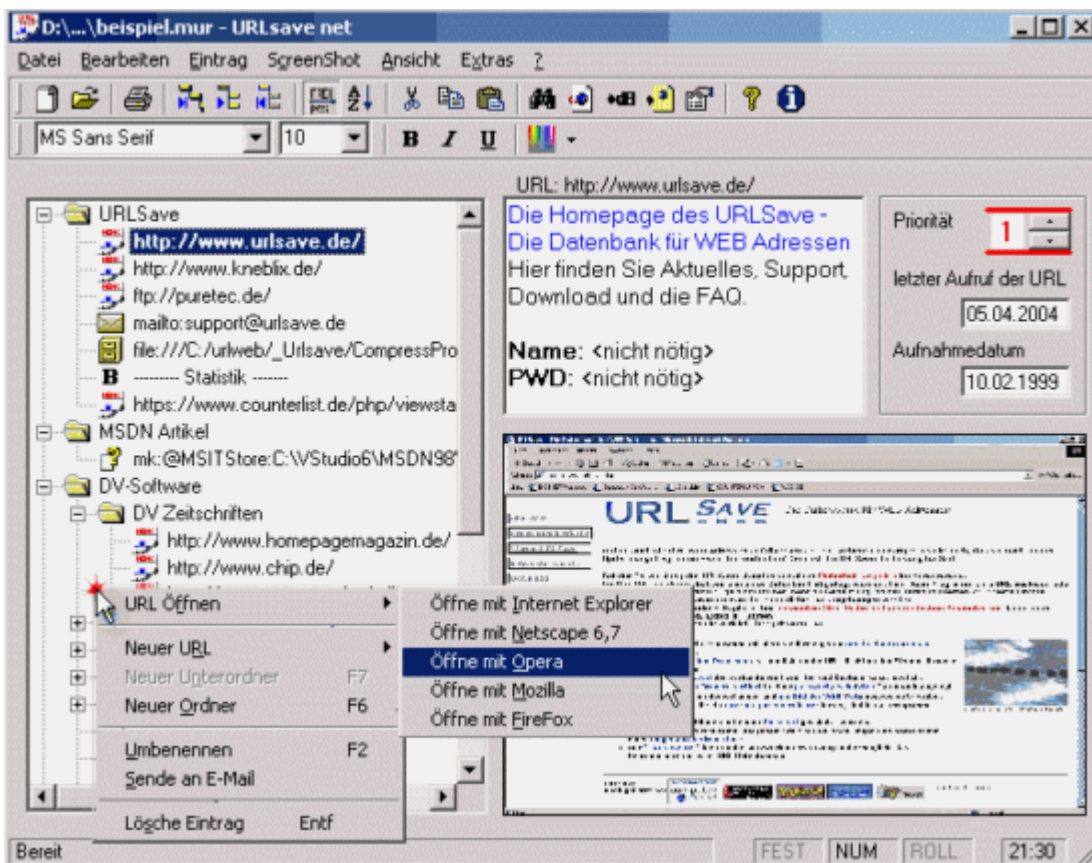
Die Datenbank für WEB Adressen

Sprache	Downloads	Userranking	Betriebssystem	
	 <b>2080</b>	 <b>1,5</b>	 <b>9x</b>	 <b>ME</b>  <b>2000</b>
Datum	Dateigrösse	Lizenz	Kosten	Screenshots
<b>01.10.2004</b>	<b>11,30 MByte</b>	<b>Shareware</b>	<b>15.00 EUR</b>	<b>ja</b>
<a href="#">details</a> ▶ <a href="#">download</a> ▶				

Rubrik: **Internet - Bookmark/Suche**

## URLsave 2.53





## Programmbeschreibung

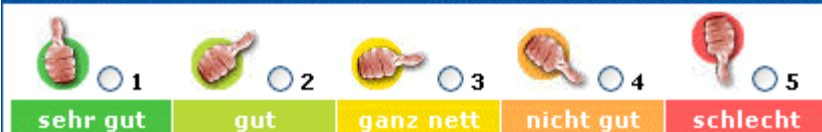
Der URLsave ist eine Datenbank, in der Sie Ihre URL sicher und komfortabel verwalten können. Zu jeder URL können eigene Kommentare und ein Bild der Web-Seite gespeichert werden. Die URL werden in Ordnern abgelegt und durch einen Doppelklick in Ihrem Wunschbrowser angezeigt. Der URLsave enthält einen Passwortschutz, Import- und Exportfunktionen, einen Internetspurenvernichter für den IE. Alle Browser können parallel benutzt werden.

Sprache	Userranking	Betriebssystem		
	1,5			
Datum	Dateigrösse	Lizenz	Kosten	Downloads
01.10.2004	11,30 MByte	Shareware	15.00 EUR*	2080

[Programm-Homepage](#) [download](#)

**\*Der Download dieser Software ist kostenlos. Der angegebene Preis gilt für die registrierte Vollversion.**

## Ihre Wertung für das Programm



Das Programm wurde von **34 Usern** im Durchschnitt mit der **Note 1,5** bewertet.

[vote](#)

Abb. Winload.de - Bookmark/Suche - URLsave 2.53



Ein Programm zur vollautomatischen Informationsbeschaffung, das anstelle des Benutzers bestimmte Informationen auffindet. ↑

### Was muß ein Programm für die vollautomatische Informationsbeschaffung grundsätzlich können ?

Nachdem der Benutzer der Maschine mitgeteilt hat, welche Informationen er benötigt, führt diese die Informationsbeschaffung weitestgehend selbständig an seiner Stelle durch. Im Unterschied zur bereits beschriebenen Lösung, mit der man die Informationsbeschaffung teilweise und unter bestimmten Voraussetzungen automatisieren könnte, würde dieses Programm nicht nur die Quellen aus der individuellen Internetquellensammlung nutzen, sondern zusätzlich sowohl die relevanten lokalen als auch die in Frage kommenden externen Internetquellen. Außerdem muß dieses Programm auch in der Lage sein, selbständig zu entscheiden, welche Informationsquellen (Datenbanken, Kataloge, Websites, Datenarchive von Newslettern oder Diskussionsforen,...) die jeweils gewünschte Information am ehesten liefern können und die betreffende Quelle dann auch dazu verwenden, um Antworten auf bestimmte Fragestellungen zu finden. Vor allem aber muß die Maschine in der Lage sein, die Relevanz und Qualität der gefundenen Information/Informationsquellen im Sinne des Benutzers zu bewerten. Am Ende der Recherche listet die Maschine die gefundenen Treffer nach Relevanz und Herkunft geordnet im Suchergebnis auf und ist darüber hinaus in der Lage, grob zu beschreiben, warum sie den Treffer für relevant bewertet hat.

Im Grunde sollte die Maschine die Internetdienste genauso effektiv und sicher nutzen können wie ein medienkompetenter Anwender mit Hilfe seiner allgemeinen und individuellen Internetquellensammlung. Das bedeutet aber, daß sie über die Informationsbeschaffung hinaus auch in der Lage sein müßte, mit anderen Anwendern zu kommunizieren, ihre Wissenslücken selbst zu erkennen und mit maschinengerechten Weiterbildungsangeboten zu schließen und im Dienste des Benutzers qualitativ hochwertige und preisgünstige Produkte/Dienstleistungen sicher einzukaufen. Im Rahmen der Diskussion sollen aber nur die Möglichkeiten eines Programms zur vollautomatischen Informationsbeschaffung betrachtet werden. Da dieses Programm, wie ein Agent, eine Aufgabe im Auftrag eines anderen erledigt, wird es auch als Informationsagent bezeichnet. Dieses Programm arbeitet eng mit seinem Benutzer zusammen, das bedeutet, es lernt von ihm, stellt Rückfragen, macht ihm Vorschläge und versucht, möglichst alles so zu machen, wie er es machen würde.

### Basiseigenschaften von Agenten

Das Konzept eines Software Agenten läßt sich vor allem durch die Eigenschaften Anpassungsfähigkeit, Kooperation und Autonomie charakterisieren.

- **Anpassungsfähigkeit**  
Agenten passen sich an ihre Umgebung und ihre Benutzer an. Sie reagieren auf Veränderungen und berücksichtigen die Wünsche und Interessen ihrer Benutzer. Der Agent lernt aus Erfahrung.
- **Kooperation**  
Agenten benutzen Standard-Sprachen und -Protokolle, damit die Kommunikation mit anderen Agenten und auch Benutzern gewährleistet ist. Das Ziel ist Teamfähigkeit.
- **Autonomie**  
Agenten verfolgen selbständig ihren Auftrag. Sie sind proaktiv, können also von sich aus aktiv werden, und natürlich reaktiv sowie zielgerichtet. Dies schließt nicht aus, daß sie gegebenenfalls die Hilfe des Benutzer bzw. anderer Agenten einholen können.

### Kritische Betrachtung der Lösung:

#### Welche grundsätzlichen Probleme stellen sich bei der vollautomatischen Informationsbeschaffung für die Maschine ?

Die Probleme sollen nun beschrieben werden anhand eines praktischen Beispiels, in dem typische Situationen genannt werden, mit denen der Mensch und damit letztendlich auch der Agent bei der Nutzung des Internets konfrontiert werden. Das Beispiel soll verdeutlichen, wie einfach viele der Situationen für einen Menschen mit Netzwerkkompetenz zu bewältigen sind und welche enormen Schwierigkeiten sich für eine Maschine auftun, die das gleiche leisten soll.

Dazu werden die einzelnen Aktionen und Ergebnisse in der entsprechenden Reihenfolge kurz



beschrieben und jeweils untersucht, welche Probleme sich für eine Maschine stellen würden, wenn sie diese Aufgabe ausführen müßte, außerdem wird jeweils die Notwendigkeit des Einsatzes einer Maschine in Frage gestellt.

**Suche nach einem sehenswerten Reiseziel für einen kurzen Wanderurlaub in Deutschland und, falls ein solches ausfindig gemacht wird, soll der dazu geeignete Reiseplan erstellt werden.**

### **Hinweis**

Die nachfolgend genannten Probleme sind bei weitem nicht vollständig beschrieben.

### **Aktion**

Eingabe von sinnvollen Internetadressen direkt in Adresseingabezeile des Browsers ([www.freizeitipps.de](http://www.freizeitipps.de), [www.reisetipps.de](http://www.reisetipps.de), [www.deutschlandreise.de](http://www.deutschlandreise.de), ...)

### **Resultat**

Kein vernünftiges Internetangebot gefunden.

#### **Welche Probleme stellen sich der Maschine ?**

Zur Eingabe der sinnvollen Internetadressen muß die Maschine kreativ sein; da sie das nicht ist, kann sie diese Teilaufgabe auch nicht vernünftig erfüllen. Es wird eifrig an Programmen gearbeitet, die mit Hilfe aufwendiger Algorithmen die Sprache in ihrer Bedeutung aufschlüsseln; diese können jedoch mit einem Menschen, der das gut beherrscht, keinesfalls mithalten, schon gar nicht, wenn es um die unterschiedlichsten sinnverwandten Wörter geht.

#### **Wäre hierfür tatsächlich eine Maschine notwendig ?**

Geht manuell viel einfacher und besser. Außerdem macht das Finden von sinnverwandten Wörtern nicht nur Spaß, sondern schult auch die sprachlichen Fertigkeiten. Im Internet gibt es zudem zahlreiche Werkzeuge, die helfen können, sinnverwandte Wörter zu finden.

### **Aktion**

Suche mit einem Suchdienst für spezielle Suchdienste und zwar nach einem speziellen Suchdienst für Deutschlandreisen

### **Resultat**

Keinen vernünftigen Suchdienst gefunden.

#### **Welche Probleme stellen sich der Maschine ?**

Hierfür müßte eine extra Schnittstelle eingerichtet werden, mit der zum einen die Verbindung zum Suchdienst hergestellt und zum anderen die Suchanfrage in die jeweilige Abfragesprache des Suchdienstes übersetzt wird. Darüber hinaus muß die Maschine die Relevanz der Treffer, falls solche vorliegen, bewerten, die relevanten Treffer entsprechend formatieren und zur weiteren Recherche heranziehen; das bedeutet, mit Hilfe des gefundenen Suchdienstes eine Antwort auf die eigentliche Fragestellung zu finden, wobei eine solche ja noch gar nicht existiert, weil der Anwender sein Reiseziel ja erst finden möchte. Wie soll die Aufgabe, die am meisten Zeit in Anspruch nimmt, nämlich die Suchdienste zu bewerten und auszuwählen, automatisiert werden: mit einer Maschine, die nicht selbst denken kann ?

Darüber hinaus muß der Agent die Erlaubnis vom Betreiber des Suchdienstes bekommen, daß er die Seite überhaupt durchsuchen darf. Dies ist keine Selbstverständlichkeit, weil die automatischen Abfragen verhindern, daß ein Anwender persönlich die Website besucht. Viele Anbieter stellen dort allerlei nützliche Informationen und Nachrichten zur Verfügung, die natürlich auch vom Besucher gelesen werden sollen. Manche finanzieren ihre Website zumindest zum Teil durch die Veröffentlichung von Werbebannern; wenn die Seite nicht direkt angeklickt wird, gibt es auch kein Geld vom Werbenden.

#### **Wäre hierfür tatsächlich eine Maschine notwendig ?**

Geht manuell viel einfacher und besser. Warum sollte der spezielle Suchdienst automatisch durchsucht werden - das manuelle Suchen dauert wenige Minuten und das Auswerten des Suchergebnisses muß der Anwender sowieso selbst durchführen.

### **Aktion**

Suche mit einer allgemeinen Suchmaschine (Suchanfrage: **reise deutschland tipp**)



Nach längerem Durchforsten des Suchergebnisses wurde eine interessante Website gefunden, auf der unter anderem mehr als 60 von den Community-Mitgliedern geschriebene Reisetipps für Deutschland zu finden waren.

#### **Welche Probleme stellen sich der Maschine ?**

Hier stellen sich grundsätzlich die gleichen Probleme wie beim vorhergehenden Suchdienst.

#### **Wäre hierfür tatsächlich eine Maschine notwendig ?**

Geht manuell viel einfacher und besser. Warum sollte das Community-Archiv automatisch durchsucht werden, das manuelle Suchen dauert wenige Minuten und das Auswerten des Suchergebnisses muß der Anwender sowieso selbst durchführen ?



Durchsicht der Community-Beiträge



Nach der Durchsicht einiger Tipps war ein interessanter Reisevorschlag dabei, der auf eine andere Website führte, auf der die Reise näher beschrieben wurde - leider war das Reiseziel zu weit entfernt. Die ebenfalls dort genannten Reiselinks führten schließlich zu einem interessanten Reiseportal, auf dem deutsche Nationalparks und Biosphären genannt wurden - die betreffenden Seiten wurden auf den lokalen Rechner kopiert. Danach ging es zurück zur Community, auf der die Tipps genannt wurden; ein in einem weiteren Tipp genannter Link führte schließlich zu einer Website, auf der letztendlich das Reiseziel beschrieben wurde, auf das die Entscheidung fiel - die sächsische Schweiz (die Stadt Wehlen).

#### **Welche Probleme stellen sich der Maschine ?**

Die Aufgabe für den Informationsagenten - finde ein noch unbekanntes Reiseziel - ist unmöglich in ein Programm umsetzbar. Der Informationsagent braucht genaue Vorgaben, damit er handeln kann. Darüber hinaus müßte eine spezielle Schnittstelle zur Community programmiert werden. Da es hier um Beiträge geht, die von den unterschiedlichsten Menschen geschrieben wurden, steht die Maschine, wenn es um die Auswertung der Inhalte geht, vor einem unlösbaren Problem, denn sie ist nicht in der Lage, Sprache wirklich kompetent zu verarbeiten.



Durchsicht der Website, auf der eine Frau ihre Reise in die sächsische Schweiz (die Stadt Wehlen) beschreibt.



Auf der Website wurden einige sehr schöne Wandertouren (Wanderung von Wehlen über Amselsee zur Bastei,...) und eine Stadtbesichtigung von Dresden (Zwinger, Semper Oper,...) mit Bildern ausführlich und interessant beschrieben, so daß die betreffenden Seiten als Wanderführer verwendet, also lokal abgespeichert und ausgedruckt wurden.

Nachdem nun das Reiseziel und im Prinzip ein hervorragender Wanderführer gefunden waren, galt es, die entsprechenden Reiseinformationen zu suchen:

- einen **Anfahrtsplan**,
- einen **Umgebungsplan** der Stadt Wehlen,
- einen **Stadtplan von Dresden**,
- ein **Verzeichnis über Gästehäuser und Hotels der Stadt Wehlen**

Die relevanten Reiseinformationen wurden, nachdem sie gefunden wurden, von den betreffenden Websites in ein HTML-Dokument kopiert, das am Ende dann ausgedruckt wurde und als Reiseführer diente.



Erstellen eines **Anfahrtsplans** mit einem geeigneten Routenplaner, Eingabe der entsprechenden

Daten (Start, Ziel,...) und Starten des Routenplaners



Ein Anfahrtsplan, der anschließend mit Strg + C bzw. Strg + V direkt in das HTML-Dokument kopiert wurde.

### Welche Probleme stellen sich der Maschine ?

Was hier vom Menschen in Sekundenschnelle umgesetzt werden kann, wäre für einen Informationsagenten eine Mammutaufgabe. Die Maschine könnte nur die Routenplaner auswählen, die in einem ihr bekannten Format vorliegen. Alles, was die Maschine irgendwo eingeben soll, muß vom Benutzer vorher festgelegt werden (ohne daß er weiß, welches genaue Routenplaner-Formular ihn erwartet).

Was passiert, wenn sich der Anwender kurzfristig dazu entschließt, anstatt mit dem Auto mit der Bahn zu reisen ? Die Website der Bahn hat ein anderes Formular als der Routenplaner, damit müßten auch die Eingabedaten neu festgelegt werden.

Das sekundenschnelle Kopieren von Bildern aus dem Web in das geöffnete HTML-Dokument eines Homepage-Editors ist für einen Agenten allein vom Aufwand, der dazu notwendig wäre (Schnittstellen zu den unterschiedlichsten Homepage-Editoren,...), unvertretbar.

### Wäre hierfür tatsächlich eine Maschine notwendig ?

Geht manuell viel einfacher und besser.



Suche nach einem **Umgebungsplan für die Stadt Wehlen**. Dieser kann beispielsweise über den YellowMap-Cityguide und der Eingabe des Ortes, in diesem Fall "Wehlen", gefunden werden. Da es mehrere Orte mit diesem Namen in Deutschland gibt, mußte die Suche noch verfeinert werden (Region: sächsische Schweiz).



Umgebungsplan der Stadt Wehlen, der anschließend mit Strg + C bzw. Strg + V direkt in das HTML-Dokument kopiert wurde.

### Welche Probleme stellen sich der Maschine ?

Hierfür müßte eine extra Schnittstelle programmiert werden, damit der Informationsagent den Suchdienst maschinell abfragen, die relevanten Treffer speichern und zur Beantwortung der eigentlichen Fragestellung verwenden kann. Eine Mammutaufgabe wäre es, den Informationsagenten so zu programmieren, daß er die Suchanfrage für beliebige Anwendungsfälle selbständig und richtig verfeinern kann.

### Wäre hierfür tatsächlich eine Maschine notwendig ?

Geht manuell viel einfacher und besser.



Suche nach einem **Stadtplan von Dresden**, der über Landkartenindex.de und der Auswahl "deutsche Städte" gefunden wurde. Der Stadtplan konnte leider nicht nach Sehenswürdigkeiten (Stichwort "Semper Oper") durchsucht werden, sondern nur nach Straßennamen, die leider in der zuvor gefundenen Website mit den Ausflugstipps nicht angegeben waren.

Das Problem konnte mit einem kurzen Wechsel zu einer allgemeinen Suchmaschine und der Suchanfrage "Semper Oper Adresse" gelöst werden. Ein Treffer des Suchergebnisses lieferte die betreffende Adresse - die vorangegangenen Suchanfragen, wie etwa "Semper Oper Strasse" führten zu keinem brauchbaren Treffer, da die Oper am Theaterplatz steht.

Zurück bei Landkartenindex.de wird schließlich nach der Eingabe der Adresse der betreffende Stadtteil von Dresden angezeigt, gleich daneben befindet sich auch der Zwinger, so daß sich eine weitere Suche erübrigt.



Der relevante Teil des Stadtplans von Dresden, der anschließend mit ALT Gr + Druck in die Zwischenablage kopiert und von dort mit Strg + V in eine mit einem Bildbearbeitungsprogramm geöffnete Datei kopiert wird. Danach wurde er etwas nachbearbeitet (Größe auf 550 Pixel reduziert, so daß es nachher keine Druckprobleme gibt, weil die Bilddatei aufgrund ihrer Größe

nur teilweise auf dem Papier abgebildet wird) und schließlich im Gif-Format abgespeichert und in das HTML-Dokument eingefügt.

### Welche Probleme stellen sich der Maschine ?

Ein Formular, das Eingaben verlangt, die man mit den verfügbaren Informationen nicht machen kann, ist ein unvorhersehbares Ereignis; auf dieses kann ein Programm, wenn überhaupt, nur sehr schwer richtig reagieren. Das bedeutet, es muß im Programm genau beschrieben werden, was für welches Ereignis zu tun ist. Wie soll man aber Reaktionen auf ein Ereignis beschreiben, das man erst kennt, nachdem es aufgetreten ist ?

Nicht weniger problematisch ist, den Informationsagenten so zu programmieren, daß er eine Suchanfrage durch eigenes Nachdenken so ändert, daß er richtige Lösungen im Suchergebnis erhält (hier im Beispiel die Änderung der Suchanfrage von "Semper Oper Strasse" auf "Semper Oper Adresse"). Das Programm müßte in der Lage sein, das Suchergebnis zu analysieren und den logischen Schluss zu ziehen, daß die Semper Oper möglicherweise nicht an einer Strasse steht, sondern beispielsweise an einem Platz, einem Park oder in einer Allee und hierfür einen Überbegriff, nämlich Adresse, finden. Diese Situation müßte sie natürlich für die unterschiedlichsten Anwendungsfälle meistern. Hier hat die Maschine mit einem grundsätzlichen Problem zu kämpfen, Sprache ist programmtechnisch nur sehr eingeschränkt erfaßbar und eine universelle Sprache wie die IPC, die Patente in einzelne technische Bereiche unterteilt, gibt es hier nicht.

Ein Mensch kann für beliebige Probleme dieser Art Lösungen finden, einem Programm, das nicht selbst denken und deshalb auch keine eigenen Lösungen finden kann, muß für jeden Fall genau beschrieben werden, was es tun soll. Die möglichen Kombinationen können aufgrund ihres Umfangs unmöglich vollständig erfaßt und im Programm beschrieben werden.

### Wäre hierfür tatsächlich eine Maschine notwendig ?

Geht manuell viel einfacher und besser.



Suche nach einem **Verzeichnis über Gästehäuser und Hotels der Stadt Wehlen**, das wohl am ehesten über die Website der Stadt zu finden sein wird. Die Eingabe [www.wehlen.de](http://www.wehlen.de) lieferte zwar die Homepage der Stadt Wehlen, aber leider nicht die der richtigen, diese wurde dann mit Hilfe einer allgemeinen Suchmaschine (Suchanfrage "Wehlen sächsisch schweiz") unter [www.wehlen-online.de](http://www.wehlen-online.de) gefunden. Dort führte dann der entsprechende Link im Hauptmenü zum gesuchten Verzeichnis.



Ein Verzeichnis mit Gästehäusern und Hotels der Stadt Wehlen, das markiert, mit Strg +C kopiert und in das HTML-Dokument eingefügt wurde; das gleiche wurde dann mit dem Stadtplan der Stadt Wehlen und einem interessanten Reistipp gemacht.



### Auswahl eines geeigneten Gästehauses und die Reservierung des Zimmers

Über den Stadtplan wurde dann ein Gästehaus in guter Lage (nahe der Elbe) ausgewählt und im elektronischen Terminkalender überprüft, ob eventuell wichtige Termine im geplanten Reisezeitraum vorliegen. Leider war auf dem Gästehausverzeichnis keine Vorwahl angegeben, so daß diese erst ermittelt werden mußte. Hier half die Telefonnummern-Datenbank "Teleinfo.de" (Eingabe des Namens, der Stadt und der Region) und lieferte schließlich die benötigte Vorwahlnummer. Danach wurde über einen Unified Messaging Service bei der Familie des ausgewählten Gästehauses angerufen und nach einem freien Doppelzimmer gefragt.



Ein geeignetes Gästehaus gefunden und ein Doppelzimmer reserviert.

### Welche Probleme stellen sich der Maschine ?

Die Maschine müßte in der Lage sein, den HTML-Text gezielt zu durchsuchen, die Position zu finden, an der die Gästehäuser angegeben sind, deren Preise zu vergleichen und ihre Lage zu erkunden. Es müssen also Internetseiten mit sehr unterschiedlichem Aufbau korrekt maschinell analysiert werden, was programmtechnisch nur schwer, wenn überhaupt, umsetzbar wäre. Es ist zwar möglich, für bestimmte Websites Übersetzungsprogramme, sogenannte Wrapper, zu programmieren; da diese aber nur die Inhalte der betreffenden Websites bzw. einzelne Internetseiten korrekt extrahieren können, kann die wertvolle (manchmal auch erdrückende)

Vielfalt des Internets nicht genutzt werden, zumal man die Seiten für die man Wrapper programmieren müßte, vorher noch gar nicht kennt.

Ein weiteres Manko ist, daß sich die Websites vom Layout her nicht ändern dürfen, da sonst der Wrapper die darin enthaltenen Daten nicht mehr korrekt interpretieren kann. Mit Hilfe eines Wrappers können also nur die Inhalte von Internetseiten mit einem bestimmten Layout in lokale Anwendungen eingebunden werden.

Für die Überprüfung des elektronischen Terminkalenders, ob eventuell wichtige Termine im geplanten Reisezeitraum vorliegen, müßte eine Schnittstelle zwischen dem betreffenden Anwendungsprogramm und dem Informationsagenten vorhanden sein oder, wenn es überhaupt möglich ist, eine eigene programmiert werden. All diese Angaben (die man im Vorfeld nicht weiß, sondern erst im Rahmen der Recherche klar definieren kann) müßten natürlich dem Informationsagenten vorher vom Benutzer mitgeteilt werden, ansonsten würde der Agent ständig Rückfragen stellen und man könnte letztendlich gleich selber suchen.

Angenommen, der Agent soll ein preisgünstiges Hotel mit mindestens 3 Sternen ausfindig machen, so wäre dies im aktuellen World Wide Web eine gewaltige Aufgabe. Die Sterne können beispielsweise so aussehen "\*\*\*" oder so "3 Sterne", genauso können sie aber auch als einzelnes Bild mit der entsprechenden Anzahl von Sternen oder mehrmals dasselbe Bild von einem Stern nebeneinander dargestellt werden. Mit viel Glück könnte man die Bilddateien über ihren Namen erkennen, etwa stern.gif oder stern1.gif, eine höchst komplizierte Aufgabe wäre die Erkennung des Bildinhalts. Die Entwicklung diesbezüglicher Software ist zwar, alleine wegen der dringenden Nachfrage nach der Bilderkennung in Bezug auf jugendgefährdende Bilder und solche, die unerlaubt kopiert und veröffentlicht werden (Urheber- und Markenrecht), in vollem Gange, aber wäre mit enorm hohen Lizenzierungskosten verbunden.

Für den paßwortgeschützten Zugang zum Unified Messaging Service müßte mit hohem Aufwand eine Schnittstelle eingerichtet werden. Hier ist erstens fraglich, ob der Anbieter dies erlauben würde und zweitens, wie eine Maschine via IP-Telefon mit dem Gästehausanbieter kommunizieren soll.

#### **Wäre hierfür tatsächlich eine Maschine notwendig ?**

Geht manuell viel einfacher und besser.

#### **Das genannte Suchbeispiel zeigt einige von vielen Situationen, die sich bei der Informationsbeschaffung im Internet stellen können, wie beispielsweise, daß:**

- die gesuchte Information häufig am Anfang der Suche noch nicht genau definiert werden kann, sondern erst im Rahmen der Recherche
- häufig nicht die Suche zeit- und arbeitsaufwendig ist, sondern die Bewertung der Relevanz und Qualität der gefundenen Informationen, vor allem aber müssen die tatsächlich brauchbaren Informationen erst mühsam zu Wissen verarbeitet werden
- Suchverfahren, die auf der Volltextsuche beruhen, häufig zu viele oder zu wenig Treffer auf eine Suchanfrage liefern und deshalb in geeigneter Weise verfeinert werden müssen, dazu gehört ein gewisses Sprachverständnis
- die Verfeinerung einer Suchanfrage zusätzlich erschwert wird, indem die Maschine schon bei kleinsten Veränderungen der Suchanfrage häufig völlig andere Suchergebnisse liefert, die nur mit einem entsprechenden Aufwand richtig analysiert werden können
- Suchergebnisse, die keine relevanten Treffer liefern, richtig analysiert werden müssen, um so entsprechende Ideen für Suchanfragen zu entwickeln, mit denen man die gewünschte Lösung finden kann
- sich die folgenden Suchschritte meistens aus dem vorhergehenden Suchergebnis ergeben, das aber nicht vorhersehbar ist
- gleiche Informationen auf unterschiedlichste Art und Weise beschrieben werden, sie liegen also in den verschiedensten Formulierungen, Formatierungen und Layouts vor
- die Informationsbeschaffung häufig aus dem Zugriff auf die unterschiedlichsten Internetdienste (Datenbanken, Websites, Communities, Unified Messaging Service,...) besteht
- die Informationen von den unterschiedlichsten Personen angeboten werden, die unterschiedliche Motive dazu bewegt haben, aus fremden Kulturen stammen und verschiedene Sprachen (hier sind

auch nationale Modewörter und weitere sprachliche Unterscheidungen einer Nation gemeint) verwenden

...

All die genannten Situationen können von Maschinen bisher entweder nicht zufriedenstellend oder überhaupt nicht gemeistert werden.

Das Beispiel, das auch auf viele andere Fälle übertragen werden kann, demonstriert weiterhin, wie man mit einer entsprechenden Netzwerkkompetenz das Internet ohne großartige maschinelle Unterstützung effektiv zur Informationsbeschaffung nutzen kann. Die gesamte Recherche mit all ihren unterschiedlichen Aktionen dauerte einschließlich der Reservierung des Zimmers und dem Ausdrucken der im HTML-Dokument gespeicherten Reiseinformationen etwa 1 Stunde.

**Das wichtigste aber ist, daß dadurch die flexible Nutzung der unterschiedlichsten Internetangebote möglich ist und damit jeder Anbieter, der gute Dienste leistet, zum Zuge kommt.** Wie hätte beispielsweise mit einem Informationsagenten die Website gefunden werden sollen, die zum einen die Entscheidungsgrundlage für die Reise in die sächsische Schweiz war und zum anderen als guter Reise- und Wanderführer diente ?

Ein Informationsagent könnte mit großem Aufwand so programmiert werden, daß er einige Teilaufgaben der genannten Aufgabe meistern könnte, dann aber nur mit Hilfe von ganz bestimmten Internetanbietern, der Rest wird wegen fehlender Schnittstellen ausgeschlossen. **Gerade aber die freie Auswahl aus vielen Angeboten und die einfachen Vergleichsmöglichkeiten sind große Mehrwerte, die das Internet gegenüber dem traditionellen Angeboten bietet und die auch genutzt werden sollten.**

**Wir Menschen können, im Gegensatz zu Maschinen, beispielsweise:**

- unsere Werte kritisch hinterfragen und bei Bedarf ändern (Maschinen haben keine Gefühle und deshalb auch keine Werte)
- auf unvorhersehbare Ereignisse reagieren
- angemessen mit Unsicherheit umgehen, das bedeutet, wir können fehlende oder unzureichende Informationen und Anweisungen sinnvoll ergänzen und Widersprüche tolerieren (dies ist auch mit ausgeklügelter Fuzzy-Logik von Maschinen nicht zu bewerkstelligen, hierzu bräuchte sie menschliche Erfahrungen aus einer Welt, in der sie nie gelebt hat)
- nachdenken, also Sachverhalte kritisch bewerten, Zusammenhänge erkennen, Informationen in persönliche Kategorien einordnen, Schlußfolgerungen ziehen, unklare Sachverhalte mit anderen diskutieren,...
- lernen, also die Ergebnisse des Nachdenkens im Gedächtnis speichern
- eigene Erfahrungen sammeln und eigene Erkenntnisse gewinnen und daraus eine erfolgversprechende Einstellung oder Verhaltensweise ableiten
- ein neues Problem formulieren, ein Problem analysieren und Lösungsstrategien entwerfen
- mit Sprachen kompetent umgehen
- .....

**Bei welche Aufgaben ist die Maschine dem Menschen tatsächlich überlegen ?**

- umfangreiche Berechnungen mit Daten durchführen
- große Datenmengen speichern und verwalten und damit dem Menschen als Gedächtnishilfe sehr wertvolle Dienste leisten (der biologische Speicher des Menschen speichert nur die häufig benötigten Informationen)
- immer wiederkehrende Aufgaben mit exakt gleichbleibendem Handlungsablauf, also Routinearbeiten wie beispielsweise Überwachungsrecherchen zuverlässig, präzise und kostengünstig durchführen
- auf vorhersehbare Ereignisse schnell und zuverlässig reagieren
- .....

Im Grunde fehlen einer Maschine also all die Fähigkeiten, die notwendig sind, um das Internet selbständig, effektiv und sicher zur individuellen Informationsbeschaffung zu nutzen. Die Maschine kann nämlich weder denken und lernen, noch kreative Lösungen für Probleme finden.

---

**Andererseits sind** angesichts der Unmengen an Informationen, die von Menschen tagtäglich erzeugt



werden, **Maschinen unentbehrliche Helfer, um aus der wachsenden Informationsflut die für den Einzelnen tatsächlich relevanten Informationen zu beliebigen Themen und Fragestellungen herauszufischen.**

Allerdings müssen die Maschinen dazu die Inhalte der im Web verteilten Dokumente verstehen. Nur so ist es möglich, daß Suchmaschinen einen kontextbezogenen Index aufbauen, die Beziehungen zwischen den Inhalten von Katalogen maschinell verständlich formuliert werden können (um so das Problem mit den nicht eindeutigen Kategorien zu entschärfen) und letztendlich ein Informationsagent die Bedeutung von Informationen/Informationsquellen in Bezug auf die Suchanfrage verstehen und aufgrund dessen zur vollautomatischen Informationsbeschaffung eingesetzt werden kann.

Es sollen nun nachfolgend die Gründe für die aktuellen Probleme der maschinellen Informationsbeschaffung näher und das "Semantic Web" als mögliche Lösung der Probleme beschrieben werden.


## **Problem**

**Maschinen sind bisher nicht in der Lage, den Inhalt von Texten zu analysieren und inhaltsbezogen einem bestimmten Themengebiet zuzuordnen** - der Begriff "Hochspannung" bedeutet beispielsweise in der Elektrotechnik das Vorhandensein einer hohen Spannung und kann beim Menschen auch im übertragenen Sinne eine höchst angespannte mentale Verfassung sein. **Für Maschinen dagegen ist der Begriff "Hochspannung" nur ein aus bestimmten Buchstaben zusammengesetztes Wort ohne Bedeutung. Das hat zur Folge, daß Maschinen keinen Zusammenhang herstellen können zwischen der Information, die der Benutzer mit Hilfe seiner Suchanfrage beschreibt, und dem tatsächlichen Informationsgehalt der Internetdokumente, die sie gespeichert und indiziert haben. Sie können die Worte innerhalb von Informationen nicht in die Beziehungen setzen, die von Menschen automatisch erkannt werden.**

Für den Anwender äußert sich das Problem beispielsweise, indem er von (Volltext)-Suchmaschinen auch auf genaue Suchanfragen viele Treffer erhält, die in keinem Zusammenhang mit der eigentlich gesuchten Information stehen. Ein anderes Problem, das vom Anwender möglicherweise gar nicht erkannt wird, entsteht durch die Verwendung des NOT-Operators. Diese Suchfunktion dient normalerweise dazu, die Internetseiten aus dem Suchergebnis auszuschließen, in denen der nach dem NOT-Operator angegebene Suchbegriff vorkommt. Tatsächlich ist es aber so, daß die Maschine, weil sie den wahren Zusammenhang, in dem ein Wort innerhalb einer Information verwendet wurde, nicht kennt, auch relevante Internetseiten (möglicherweise sogar mit dem gesuchten Inhalt) aus dem Suchergebnis ausschließt. Hier nützt dann auch ein komplizierter Ranking-Algorithmus, mit dem die einzelnen Internetseiten bezüglich ihrer Relevanz zur Suchanfrage bewertet werden, nichts mehr, da die eigentlich relevante Internetseite beim Ranking nicht mehr dabei ist. Wie genau die Relevanz und vor allem die Qualität von Informationen/Informationsquellen mit Hilfe maschineller Bewertungssysteme (Ranking-Algorithmen), die auf bloßen Zähl- und Berechnungsverfahren von Informationen ohne Kontext beruhen, tatsächlich bewertet werden kann, wurde bereits [erläutert](#).

Viele Probleme der aktuellen Informationsbeschaffung mit Hilfe von Suchmaschinen und mögliche Lösungen wurden bereits im Fazit der [Diskussion](#) zur Lektion "Such- und Informationsdienste" beschrieben. Die wichtigsten dort genannten Probleme und weitere relevante Einzelheiten, die im Laufe der Entwicklung des Lehrgangs erkannt wurden, sollen nun nachfolgend zusammengefaßt werden.

**Die aktuellen Probleme der maschinellen Informationsbeschaffung haben folgende wesentlichen Gründe:**

-  **Es fehlt an einem gemeinsamen Wortschatz, der sowohl von Autoren und Anwendern als auch von Maschinen gemeinsam verstanden und verwendet werden kann.**

Der Suchende beschreibt über seine Suchanfrage, welche Information er sucht und der Autor beschreibt mit seinem Wissen, das sich in jedem Fall von dem des Suchenden unterscheidet, in welchem Zusammenhang er seine Information veröffentlicht. Die Maschine hat die Aufgabe, mit Hilfe der genannten, in der Regel sehr unterschiedlichen und damit schwer vergleichbaren Informationen den Suchenden mit dem richtigen Autor bzw. dessen Dokument zusammenzubringen. Dies kann der Maschine, die Sprache nur sehr eingeschränkt verarbeiten kann, nicht gelingen, da sie nur die in der Suchanfrage enthaltenen Wörter, ohne deren wahre Bedeutung zu kennen, mit den Wörtern vergleicht, die sie in ihrem Index gespeichert hat. Ein wahllos zusammengestellter Index bietet der Maschine keine Möglichkeit, den Zusammenhang zwischen einer Suchanfrage und den von ihr gespeicherten Informationen herzustellen. Es fehlt also an einem gemeinsamen Wortschatz, den sowohl die Autoren der Internetdokumente, die

Suchenden und die Maschinen verstehen und verwenden.

**Es fehlt an Möglichkeiten, mit denen der Autor die Struktur und einzelne Inhalte seines Dokuments näher beschreiben kann.**

Mit der Formatsprache HTML bzw. mit den darin festgelegten Tags wird vorwiegend die Darstellung der Inhalte (Texte, Bilder, Tabellen,...) einer Internetseite beschrieben. Ein Beispiel hierfür ist, an welcher Stelle auf der Internetseite eine bestimmte Tabelle positioniert werden soll, wie viele Zeilen und Spalten diese enthalten soll sowie deren Länge, Breite und Inhalte.

Mit Hilfe von Metatags können Informationen über das Dokument, sogenannte Metadaten, wie etwa der Name des Autors oder eine Kurzbeschreibung des Inhalts, in die HTML-Seite eingefügt werden.

Der wohl wichtigste Bestandteil von HTML sind die Hyperlinks, mit denen Verweise zu anderen Internetseiten des eigenen Internetangebots, aber auch zu anderen Websites erstellt werden können.

Was bei HTML durch die fest vorgegebenen Tags fehlt, sind Möglichkeiten, mit denen der Autor die Struktur und einzelne Inhalte seines Dokuments näher beschreiben kann, beispielsweise in welchem Zusammenhang ein mehrdeutiges Wort (z. B. Hochspannung) innerhalb einer bestimmten Information verwendet wird. Die so gemachten Angaben (Metadaten) müssen natürlich von Maschinen verstanden und verwendet werden können, da sie sonst nicht in der Lage sind, die Metadaten und damit den Inhalt des Dokuments mit der jeweiligen Suchanfrage in Verbindung zu bringen.

Da bei HTML-Dokumenten die Beschreibung des Inhalts und des Layouts nicht getrennt voneinander stattfindet, muß sich die Maschine durch den gesamten HTML-Code wühlen, um bestimmte Daten und Informationen zu erfassen und auszuwerten. Bei HTML können zwar Formatierungen (Cascading Style Sheets oder kurz CSS) und Scripts (z. B. Javascript), die für die ganze Website gelten sollen, ausgelagert werden, nicht aber die Metainformationen, mit denen die Inhalte der Seite beschrieben werden können.

**Die Erkenntnisse von Anwendern (Leser eines Dokuments, Benutzer einer Informationsquelle, Betrachter eines Videos,...) können nur ungenügend in die eigene Bewertung von Informationen/Informationsquellen miteinbezogen werden.**

Wenn jemand eine Datenbank gefunden hat, in der bestimmte, normalerweise sehr schwer auffindbare Information mit Hilfe bestimmter Suchanfragen aufgefunden werden können, so wäre es vorteilhaft, wenn er die diesbezügliche Erkenntnis anderen Anwendern zugänglich machen könnte. Das gleiche gilt natürlich für alle Objekte (Dokumente, Textausschnitte, Bilder,...). Hierfür bietet HTML selbst keine Möglichkeiten.

**Die im Internet verteilten Informationen über ein bestimmtes Objekt werden nicht zur Bewertung herangezogen.**

Das bedeutet, Hyperlinks können nicht nur dazu verwendet werden, um relevante Inhalte (die von beliebigen, öffentlich zugänglichen Websites bzw. Web-Servern stammen können) miteinander zu verbinden, sondern man könnte durch eine entsprechende Kombination der verlinkten Informationen auch Schlußfolgerungen ziehen, um vorhandenes Wissen über bestimmte Objekte abzuleiten.

Die vorher genannten Probleme treten vor allem in unstrukturierten Informationsbeständen auf, das bedeutet in **Dokumenten (Internetseiten, PDF-Dokumente,...) ohne Metadaten**, die im Volltext erfaßt werden.

**Hinweis**

Neben unstrukturierten Informationsbeständen kann noch zwischen zwei weiteren Formen unterschieden werden:

strukturierte Informationsbestände  
Es handelt sich dabei um Datenbanken, bei denen die Inhalte der Datensätze mehr oder weniger genau mit entsprechenden Datenfeldern (Metadaten) beschrieben werden. Die maschinelle Suche in strukturierten Informationsbeständen liefert bei einem kompetenten Rechercheur hochwertige Suchergebnisse.

semistrukturierte Informationsbestände  
Es handelt sich dabei um Internetseiten, bei denen sowohl Volltext als auch Metadaten erfaßt wurden. Da die Metadaten von den Autoren freiwillig und mit unterschiedlicher Qualität erstellt werden, kann die Suche keine vergleichbar hochwertigen Ergebnisse liefern wie in den strukturierten Informationsbeständen.

## Lösung

Grundsätzlich gibt es zwei Möglichkeiten, um die vorher beschriebenen Probleme zu lösen: **Entweder wird eine Technologie geschaffen, die es ermöglicht, daß die natürliche Sprache von Maschinen verstanden und verarbeitet werden kann, oder die Menschen beschreiben die Inhalte ihrer natürlichen Sprache in einer Form, die auch Maschinen verstehen können.** Da Maschinen bekannterweise die natürliche Sprache nicht verstehen und es auch keine Technologie gibt, die das ermöglicht, bleibt nur die zweite Lösung, die allerdings mit einem Mehraufwand für die Informationsaufbereitung und -bereitstellung verbunden ist.

Die dazu notwendige technische Weiterentwicklung des World Wide Web wird als "**Semantic Web**" bezeichnet, mit dem die bisherigen Probleme bei der Informationsbeschaffung und -bereitstellung im WWW gelöst werden sollen. Das Konzept und die technischen Standards für das "**Semantic Web**" wurden von [Tim Berners-Lee](#), dem Erfinder des World Wide Web und dem World Wide Web Consortium (W3C) in Zusammenarbeit mit IT-Unternehmen entwickelt.

Tim Berners-Lee stellt übrigens auch eine kurze Slide-Show mit Informationen über das Semantic Web unter "[www.w3.org/...](#)" bereit. Besonders interessant war dabei die nachfolgend abgebildete Grafik über die grundlegende Architektur des Semantic Web ([www.w3.org/...](#)).



# Semantic Web - XML2000 by Tim Berners-Lee

## Table of contents

- . *Semantic Web on XML*
- . *Semantic Web*
- . *Philosophy - "Web"*
- . *Philosophy - "Semantic"*
- . *Semantic links*
- . *Graphs...*

- . Architecture
- . RDF+Schema layer
- . Ontology layer
- . Logic layer
- . Semantic Web bus
- . Practical deployment
- . Killer App for the SWeb
- . Thank you - Questions

## Architecture

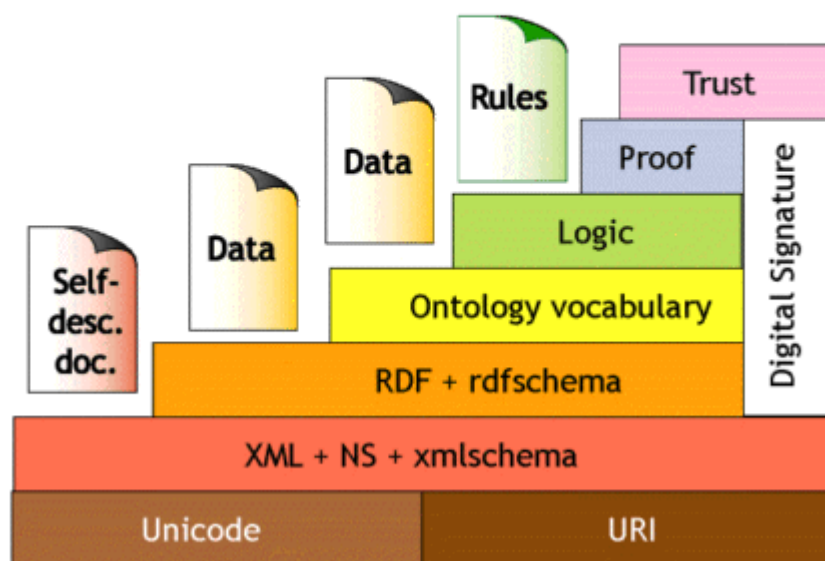


Abb. W3C.org - Semantic Web - XML2000 von Tim Berners-Lee

Die Maschine kann den Zusammenhang, in dem die Information vom Autor veröffentlicht wurde, nur dann verstehen, wenn die betreffende Information vom Autor mit Hilfe der für das Semantic Web entwickelten Technologien maschinenlesbar beschrieben wird. Als weitere Bewertungshilfen können von Anwendern an die Information annotierte Daten und Informationen und weitere im Internet verteilte Ressourcen (Websites, Webseiten, Texte, Bilder,...) dienen.

### Hinweis

Grundsätzlich kann man mit Metadaten bestimmte Daten und Informationen oder deren Struktur maschinenlesbar beschreiben. Metadaten können unterschieden werden nach ihrer Formalisierung (Stichwörter, Schlagwörter, Textbeschreibungen) und ihrer Anknüpfung an die zu beschreibenden Daten (unmittelbar in das Dokument integriert, separat in einer eigenen Datei gespeichert).

Aus den vorhandenen Metadaten kann die Maschine Informationen erzeugen, die wiederum mit Hilfe formalisierter Wissensstrukturen weiterverarbeitet werden können. Dadurch sollen Maschinen weniger und vor allem wirklich relevante Treffer im Suchergebnis liefern. Darüber hinaus sollen sie nicht nur

Verweise auf Dokumente beschaffen, sondern auch Antworten auf konkrete Fragestellungen.

## **Hinweis**

Die Semantik beschäftigt sich übrigens mit der Bedeutung von Wörtern und Sätzen (die Information in einen Zusammenhang bringen).

Ein weiterer wichtiger Begriff im Zusammenhang mit der Semantik ist die Syntaktik - das ist die Ebene, auf der Maschinen arbeiten - die Syntaktik beschäftigt sich mit der Verbindung von Wörtern zu einem Satz (Daten in Informationen umsetzen).

Über der Semantik steht die Pragmatik, bei der es darum geht, die Informationen dazu zu verwenden, richtig handeln zu können. Aus der Pragmatik leitet sich eine philosophische Lehre ab, der sogenannte Pragmatismus, **die das Handeln über die Vernunft stellt und die Wahrheit und Gültigkeit von Ideen und Theorien allein nach ihrem Erfolg bemisst.**

das "Semantic Web" als technische Grundlage für das Programm ↑

Das "Semantic Web" soll die Inhalte der im Web verteilten Dokumente mit folgenden Technologien besser zugänglich machen, durch:

- Eine Formatsprache (XML = Extensible Markup Language), die es dem Autor ermöglicht, die Struktur seiner Inhalte (Aufbau des Dokuments) maschinenlesbar zu beschreiben und bei der die Inhalte und das Layout streng voneinander getrennt sind. ↓
- Eine standardisierte Metadatensprache (RDF = Resource Description Framework), die es dem Autor ermöglicht, die einzelnen Inhalte seines Dokuments maschinenverständlich zu beschreiben und die von den unterschiedlichsten Personen und Gruppen anwendungs- und themenübergreifend ausgetauscht und genutzt werden kann. ↓
- Eine Technologie (Ontologie), die es ermöglicht, komplexe Wissensstrukturen der unterschiedlichsten Themen-/Fachgebiete maschinenverständlich zu formulieren. ↓

## Zusammenfassung ↓

Eine Formatsprache (XML = Extensible Markup Language), die es dem Autor ermöglicht, die Struktur seiner Inhalte (den Aufbau des Dokuments) maschinenlesbar zu beschreiben und bei der die Inhalte und das Layout streng voneinander getrennt sind. ↑

Der Aufbau von XML ist dem von HTML sehr ähnlich, was nicht verwundert wenn man weiß, das beide Formatsprachen aus der gleichen textbasierten Meta-Auszeichnungssprache SGML (Standard Generalized Markup Language), aber mit unterschiedlichen Zielen, zusammengestellt wurden. Wie bereits erwähnt ging es bei der Zusammenstellung von HTML vorwiegend darum, mit festgelegten Tags die Darstellung der Inhalte (Texte, Bilder, Tabellen,...) einer Internetseite zu beschreiben. XML dagegen soll die beschränkten Möglichkeiten von HTML so gut wie möglich erweitern ohne jedoch zu kompliziert zu werden.

Das Aussehen der Daten kann der Anwender bei XML mit Hilfe der sogenannten XSL Style Sheets selbst festlegen, ohne dabei den Inhalt des Dokuments zu verändern. **Damit ist es möglich, den gleichen Inhalt auf verschiedene Art und Weise darzustellen.** Bei XML ist die Darstellung der Daten strikt von deren Inhalt getrennt.

Diese Trennung des Inhalts der Daten von deren Darstellung ermöglicht neben dem effektiven Datenaustausch, bei dem **die verschiedensten (in einer Datenbank gespeicherten und strukturierten ) Inhalte (Testberichte, Erfahrungsberichte, AGB's, Rechnungen,...) aus unterschiedlichen Quellen (Textverarbeitungsprogramm, Grafikbearbeitungsprogramm, Tabellenkalkulation,...) auf einfache Weise in einem (neuen) XML-Dokument zusammengeführt werden können,** auch die Möglichkeit, selbst zu entscheiden, wie die Daten am besten präsentiert werden sollen. Die Trennung von Layout und Inhalt hat gerade für die maschinelle Suche einen großen Vorteil: Die Maschinen müssen sich nicht, wie bei HTML, durch den gesamten Code wühlen, sondern können die Suche auf den relevanten Teil beschränken.

Der wohl bedeutendste Unterschied zwischen XML und HTML ist der das mit XML eigene Tags



**definiert werden können. Dies ermöglicht es eine eigene (maschinenlesbare) Sprache für die jeweils anfallenden Daten und Informationen zu entwickeln.**

So können beispielsweise Universitäten eine DTD für die schnelle und gezielte Suche in Doktorarbeiten entwickeln oder Unternehmen eine DTD erstellen, mit der die Inhalte der speziell in ihrem Aufgabenbereich anfallenden Daten und Informationen maschinenverständlich beschrieben werden können, das die Mitarbeiter und Partner des Unternehmens (Kunden, Lieferanten,...), mit maschineller Hilfe, schnell zu den jeweils benötigten Daten und/oder Informationen gelangen.

Mit der Formatsprache XML kann der Autor zwar die Struktur seiner Inhalte maschinenlesbar beschreiben, die tatsächliche Bedeutung der Inhalte kann aber von Maschinen noch nicht ausgewertet werden. Hierfür wurden weitere Technologien geschaffen, nämlich das Resource Description Framework (RDF) und Ontologien.

**Eine standardisierte Metadatensprache (RDF = Resource Description Framework), die es dem Autor ermöglicht, die einzelnen Inhalte seines Dokuments maschinenverständlich zu beschreiben und die von den unterschiedlichsten Personen und Gruppen anwendungs- und themenübergreifend ausgetauscht und genutzt werden kann. ↑**

Damit relevante Inhalte von Maschinen gezielt gefunden, miteinander verknüpft und ausgewertet werden können, müssen **die Inhalte (eine Person, eine Institution, eine Organisation, eine Website, eine Internetseite, ein Textabschnitt, die Bedeutung eines Wortes innerhalb eines Satzes, Bilder, Videos,...) mit Hilfe von Metadaten vom Autor so beschrieben werden**, daß die Maschine deren Bedeutung verstehen kann. Nur so kann sie die Relevanz und Qualität eines Inhalts (vom Autor) in Bezug auf eine konkrete Fragestellung (vom Suchenden) besser bewerten.

Mit RDF können ganze Websites, einzelne Webseiten oder aber auch einzelne Textabschnitte beschrieben werden. Die Metadaten können dabei entweder in den Kopfteil einer HTML-Seite eingefügt oder auch in einer eigenen Datei abgespeichert werden. Dazu muß dann aber eine Verbindung zwischen der RDF-Datei und dem betreffenden XML- oder HTML-Dokument hergestellt werden.

**Die mit RDF erstellten Metadaten können anwendungs- und themenübergreifend ausgetauscht werden.** Deshalb können in den RDF-Metadaten die unterschiedlichsten im Internet verteilten und relevanten Inhalte angegeben werden, etwa eine interessante Internetseite, eine Internetadresse oder ein relevanter Textabschnitt, die der Maschine helfen, die Bedeutung eines bestimmten Dokumenteninhalts besser zu verstehen und möglicherweise zur weiteren maschinellen Auswertung herangezogen werden können.

Dazu müssen die Metadaten und die damit beschriebenen Inhalte nicht nur für eine bestimmte Fragestellung gesammelt und zusammengeführt werden, sondern **es muß sich dabei um Inhalte handeln, deren Herkunft eindeutig bestimmbar ist, also einem konkreten Urheber zugeordnet werden können.** Die eindeutige Identifizierung von im Netz verteilten Inhalten ist von sehr hoher Bedeutung, da nur die Inhalte für Entscheidungen herangezogen werden können, denen man vertrauen kann.

Für die Identifizierung (es handelt sich tatsächlich um das Original des gewünschten Inhalts) und Adressierung (Orte an denen der Inhalt im Netz gespeichert ist) eines Inhalts dient als einheitliche Angabe der "Uniform Resource Identifier" (URI), der als Lösung für das nachfolgend kurz beschriebene Problem entwickelt wurde.

### **Problem**

Mit Hilfe von Hyperlinks kann man auf beliebige elektronische Inhalte bzw. Objekte (ein Text, ein Bild, eine Video-Datei, die Startseite einer Website, ein Dokument,...) verweisen und diese bei Bedarf abrufen. Die Inhalte sind aber nur dann verfügbar wenn deren Internetadresse (die URL) zum Zeitpunkt des Aufrufs gültig ist, ansonsten erscheint die Fehlermeldung "No such file or directory" (HTTP 404 - Datei nicht gefunden).

Gründe hierfür sind beispielsweise, das der Inhalt in ein anderes Verzeichnis verschoben oder auf einen anderen Web-Server verlegt wurde (Änderung der URL) oder der Inhalt ganz gelöscht wurde.

Als kurzfristige Maßnahme können Web-Server so konfiguriert werden, daß sie den Anwender automatisch beim Aufruf des Inhalts von der alten (nicht mehr gültigen) Adresse auf die neue umleiten. Eine andere Möglichkeit wäre es die jeweils zum aktuellen Zeitpunkt gültige Internetadresse (den aktuellen Speicherort des Inhalts) mit dem Web-Server und einem entsprechenden Script aus einer Datenbank auszulesen und den Anwender so dynamisch auf die jeweils gültige Adresse umzuleiten.



Leider sind das noch keine optimalen Lösungen. Was nützt die Weiterleitung zu einem Inhalt der auf einem Web-Server gespeichert ist, der nicht oder nur schlecht erreichbar ist ?. Was passiert wenn Änderungen auf dem Server vorgenommen müssen, beispielsweise der notwendige Austausch des Datenbanksystems ?

Aber selbst wenn der Inhalt abrufbar ist, so ist noch lange nicht sichergestellt, daß es sich um den Original-Inhalt oder unverfälschte im Web verteilte Kopien des Originals handelt. URL's alleine sind also keine langfristige Lösung für den sicheren Zugriff auf die im Netz zugänglichen Inhalte (Objekte).

**Es bedarf eines Adressierungs- und Identifizierungssystems mit dem der Zugriff auf beliebige im Netz verfügbare (zugängliche) Texte, Bilder, Video-Dateien, Internetseiten, Dokumente, usw. eindeutig, dauerhaft und mit hoher Glaubwürdigkeit (es handelt sich um den Original-Inhalt oder unverfälschten Kopien des Originals) gewährleistet werden kann.**



**Eine für das beschriebene Problem vom W3C entwickelte Lösung ist die URI.**

Wird der Inhalt im Internet veröffentlicht so besteht die vollständig URI aus der URL, der optional ein "#" und die ID des Objekts folgen kann. Während die Adressierung (URL) notwendig ist, ist die Identifizierung optional, das bedeutet der Inhalt kann von seinem Anbieter auch nur durch eine URL beschrieben werden (der häufigste Spezialfall eines URI ist übrigens ein URL). Die genaue Syntax der URI wird übrigens in der "RFC 2396" beschrieben.

**Möglichkeiten für die Beschreibung von Metadaten gab es übrigens auch schon, bevor das Resource Description Framework entwickelt wurde.**

Eine bereits erwähnte Möglichkeit sind die Metatags in HTML, mit denen beispielsweise Metadaten wie der Name des Autors, wichtige Stich- und Schlagwörter und eine kurze Inhaltsbeschreibung maschinell lesbar (aber nicht sehr gut auswertbar) in das betreffende HTML-Dokument eingefügt werden können.

Eine weitere Möglichkeit bietet der von Bibliothekaren, Informationswissenschaftlern und Informatikern entwickelte DUBLIN CORE ELEMENT SET, der es auch Laien ermöglicht, ihre Webinhalte mit Metadaten (Thema, Anbieter, Kurzbeschreibung des Inhalts, Dokumentenformat, Sprache,...) genau zu beschreiben und damit für Menschen und Maschinen besser zu erschließen. Die Metadaten werden dabei in dem betreffenden HTML-Dokument als XML-Code oder mittels RDF abgespeichert. Falls der DUBLIN CORE ELEMENT SET für PDF- oder Postscript-Dateien verwendet werden soll, müssen diese auf einem gesonderten Internetdokument gespeichert werden, das die Metadaten enthält und eindeutig mit dem betreffenden Dokument in Verbindung gebracht werden kann.

Die Möglichkeiten des Resource Description Framework gehen allerdings weit über die bloße Beschreibung von Metadaten und deren Austauschbarkeit über Anwendungs- und Themengrenzen hinweg hinaus. Mit RDF ist es möglich, Daten, die eine URI besitzen, im World Wide Web zu referenzieren und aus diesen Daten durch ein entsprechende Verknüpfung Aussagen (Informationen) zu bilden. Eine Aussage besteht in der Grammatik grundsätzlich aus Subjekt (Satzgegenstand), Prädikat (Satzaussage) und Objekt (Satzergänzung). Ähnlich sind auch die Aussagen in RDF aufgebaut. Das Subjekt (eine Person, eine Gruppe, eine Website, eine Internetseite,...) steht bei einer solchen Aussage in einem bestimmten Zusammenhang zu einem Objekt (einer anderen Person, Gruppe, Website, Internetseite,...), der durch das Prädikat (ist Autor von, ist Mitarbeiter bei,...) näher beschrieben wird. Alle Elemente der Aussage sind dabei durch einen Uniform Resource Identifier (URI) bezeichnet. Durch unterschiedliche Formulierungen des Prädikats können so die unterschiedlichsten Zusammenhänge zwischen Subjekt und Objekt hergestellt werden.

Praktische Beispiele zu RDF-Aussagen findet man mit Hilfe entsprechender Suchanfragen bei allgemeinen Suchmaschinen. Ein interessanter Artikel über das Semantic Web mit dem Titel "Mein Computer versteht mich" wurde von den W3C-Mitgliedern Tim Berners-Lee, James Hendler und Ora Lassila verfaßt und unter anderem in der August-Ausgabe 2001 der wissenschaftlichen Fachzeitschrift "Spektrum der Wissenschaft" ab Seite 42 veröffentlicht.

Das bloße Vorhandensein von maschinenverständlichen Metadaten und Aussagen befähigt die Maschine aber noch nicht, diese Daten und Informationen auch weiter zu verarbeiten, beispielsweise Beziehungen zu anderen Begriffen bzw. damit in Verbindung stehenden Informationen herzustellen, diese richtig zu kombinieren und einfache logische Schlußfolgerungen zu ziehen. Erst durch diese Fähigkeit kann die Maschine wichtige Aufgaben der Informationsbeschaffung (eingeschränkt) erfüllen, etwa vorhandenes und für die Relevanz- und Qualitätsbewertung einer Information/Informationsquelle nutzbares Wissen für den Suchenden ableiten oder eine konkrete Antwort auf eine Suchanfrage geben.

Um diese Aufgabe ansatzweise zu erfüllen, wurde RDF um die Technologie RDF Schema erweitert. Damit ist es möglich, einfache Zusammenhänge zwischen Daten (Name einer Person oder Gruppe, die Adresse einer Website oder Internetseite,...) bzw. (Fach-)Begriffen herzustellen, die dann eine sehr eingeschränkte Weiterverarbeitung von Informationen (Fakten) ermöglichen.

Diese in starren Schemata vereinbarten Relationen können nicht einfach an die sich fortlaufend ändernden Bedingungen angepaßt werden und sind damit nicht flexibel innerhalb von Dokumenten und Anwendungsprogrammen einsetzbar.

Für komplexe Aufgabenstellungen bezüglich der Weiterverarbeitung von Informationen wurden sogenannte Ontologien geschaffen, mit denen nicht nur Zusammenhänge zwischen Daten/Begriffen hergestellt, sondern auch Regeln erstellt werden können, die wiederum die Basis für einfache logische Schlußfolgerungen sind. Darüber hinaus bieten Ontologien die Möglichkeit, den Anwender aktiv in die Erweiterung und Anpassung der Wissensbasis mit einzubeziehen.

**Eine Technologie (Ontologie), die es ermöglicht, komplexe Wissensstrukturen der unterschiedlichsten Themen-/Fachgebiete maschinenverständlich zu formulieren. ↑**

Eine Ontologie (ein von Menschen und Maschinen verarbeitbares objektorientiertes Wissensmodell) dient zur formalen Beschreibung (Begriffshierarchien, Relationen, Regeln) eines Wissensgebiets (Themen-/Fachgebiet, berufliches Aufgabengebiet,...). Das Ziel einer Ontologie ist es, eine für Anbieter, Anwender und Maschinen gleichermaßen verständliche und gemeinsam verwendbare Wissensbasis bereitzustellen. Das Basiswissen der Ontologie wird von Experten modelliert, die mit den Aufgaben und den wichtigen Fragen, die sich im jeweiligen Themen-/Fachgebiet stellen, vertraut sind. Die gemeinsame Wissensbasis kann dann von Ontologie-Anbietern (Autoren) und Anwendern maschinenverständlich erweitert bzw. an die aktuellen Erfordernisse angepaßt werden.

#### **Eine Ontologie:**

- **stellt die Beziehungen zu anderen Begriffen eines bestimmten Wissensgebiets in einer sogenannten Begriffshierarchie (Taxonomie) her**

Personen können beispielsweise nach Rollen und Funktionen oder entsprechend ihrer Gruppenzugehörigkeit eingeordnet werden. Dokumente können beispielsweise nach Autoren, Thema, Zielgruppe und Erstellungsdatum erschlossen werden.

Die Begriffshierarchie eines Themen-/Fachgebiets kann ähnlich erstellt werden wie bei einem Thesaurus (Synonym-Wörterbuch oder ein thematisch begrenztes Wörterverzeichnis mit zahlreichen Querverweisen).

Ein Thesaurus kann in elektronischer Form, beispielsweise bei Suchsystemen, dazu verwendet werden, um die Suche nach Informationen zu einem bestimmten Themen-/Fachgebiet effektiver zu gestalten. Anders als bei einem (Volltext)-Index, in dem bis auf wenige Ausnahmen (der, die, wo,...) alle Wörter der indexierten Dokumente wahllos gespeichert werden, gibt der Thesaurus die Begriffe (Wörter/Wortkombinationen), die zum Indexieren von Dokumenten benutzt werden dürfen, vor. Für einen Suchbegriff können so über- oder untergeordnete Begriffe gefunden werden. Der Suchbegriff kann auch durch verwandte Begriffe erweitert oder durch Synonyme ersetzt werden. Einige wichtige, im Thesaurus definierte Beziehungen zwischen Begriffen können sein: Begriff A ist Oberbegriff für B, Begriff C ist Synonym für B, Begriff B ist Unterbegriff für A, Begriff A ist verwandter Begriff von B, für Begriff A benutze Synonym B, Begriff B wird benutzt für Synonym A oder für Begriff A benutze Kombination B und C.

Der Zusammenhang zwischen der vom Anwender mit der Suchanfrage beschriebenen Information und der jeweils relevanten Information in den in der Datenbank gespeicherten Dokumenten wird über diese im Thesaurus festgelegten Begriffsbeziehungen hergestellt. Das bedeutet, die Maschine vergleicht die in der Suchanfrage vorkommenden Begriffe mit den von den Experten im Thesaurus vorgegebenen Begriffsbeziehungen und überträgt bei einer Übereinstimmung die jeweils zugeordneten Inhalte (Fachartikel, Pressemitteilungen, Verweise auf Internetseiten,...) in das Suchergebnis.

Die Verwendung eines Thesaurus erfordert zwar einen gewissen Einarbeitungsaufwand für die korrekte Verwendung der Syntax und den Umgang mit dem jeweiligen Funktions- und Leistungsumfang, erleichtert aber die Auswahl der Suchbegriffe aus der sprachlichen Vielfalt, mit der Informationen eines Themen-/Fachgebiets beschrieben sein können, entscheidend.

Die formale Beschreibung eines Wissensgebiets mit Hilfe einer Ontologie geht aber weit über die Funktionen eines Thesaurus und die Erstellung von Begriffshierarchien hinaus.

## **erlaubt es, Regeln und Relationen festzulegen, die es einer Maschine ermöglichen, einfache logische Schlußfolgerungen aus bekannten Fakten zu ziehen**

Dazu werden zuerst Relationen (Beziehungen) zwischen Sachverhalten hergestellt, also Fakten in das System eingegeben. Danach werden Regeln beschrieben, die sich aus den Beziehungen zwischen den Sachverhalten ergeben. Mit Hilfe dieser Regeln und Relationen kann das System dann einfache logische Schlußfolgerungen ableiten und so auch Antworten auf konkrete Fragen geben. Eine Suchmaschine, die mit Ontologien arbeitet, kann somit, wie bereits erwähnt, im Gegensatz zur ausschließlichen Verwendung der Volltext- oder Feldsuche nicht nur Verweise auf Dokumente und Personen liefern, in denen der Anwender nach einer möglichen Antwort suchen muß, sondern, indem sie einzelne Daten und Informationen von verschiedenen Quellen verknüpft und in einen bestimmten Zusammenhang bringt, auch konkrete Antworten auf Fragen geben.

Solche Relationen können folgendermaßen aussehen:

- `pfl egt_ die_ Informationsbestände_ der` (Universitätsbibliothek, Fachdatenbank): Eine Universitätsbibliothek pflegt die Informationsbestände von verschiedenen Fachdatenbanken
- `liefert_ Informationen_ zu` (Fachdatenbank, Fachgebiet): Eine Fachdatenbank liefert Informationen zu einem bestimmten Fachgebiet.

In diesem Fall lautet die über die Relationen definierte Regel: Pflegt eine Universitätsbibliothek X eine Fachdatenbank Y und behandelt die Fachdatenbank Y das Fachgebiet Z, dann hat die Universitätsbibliothek X Informationsbestände zum Fachgebiet Z.

Auf die Frage »Wer hat Informationsbestände zum Fachgebiet Z ?« würde die Antwort lauten »Universitätsbibliothek X«, obwohl diese Aussage in keiner Quelle explizit formuliert ist.

Diese einfachen Relationen und Regeln befähigen eine Maschine natürlich nicht dazu, wie Menschen Sprache zu verarbeiten, zu denken, zu lernen und vor allem neues Wissen zu schaffen, aber sie ermöglichen es der Maschine, einfache logische Schlüsse zu ziehen - beispielsweise eine bestimmte, für die jeweilige Fragestellung geeignete Informationsquelle in eine Recherche zu integrieren oder eine unbekannte Informationsquelle mit einer bekannten (formal) zu vergleichen und so Rückschlüsse auf deren Relevanz und Qualität zu ziehen. Von Leistungen, die das menschliche Gehirn vollbringen kann, sind die einer Maschine (glücklicherweise) weit entfernt.

## **ermöglicht, daß sich Anwender aktiv an der fortlaufenden Erweiterung der Wissensbasis beteiligen können**

Eine Möglichkeit hierzu können formularbasierte Schnittstellen bieten, über die aktuelle Informationen (Erkenntnisse, Meinungen, Erfahrungen, Tipps, relevante Internetseiten, Forenbeiträge,...) maschinenverständlich in die Wissensbasis eingebunden werden können.

Eine weitere Möglichkeit bieten ontologiebasierte Annotationssprachen in Verbindung mit einem entsprechenden Annotierungswerkzeug. Damit können wichtige Daten und Informationen (Erkenntnisse, Kommentare, Links,...) von Anwendern an das betreffende Objekt (eine Internetseite, ein Bild, eine Datenbank, ein Textabschnitt,...) angeheftet (annotiert) werden. Das Objekt, die annotierten Informationen sowie Metadaten über den Annotierer, die Zeit der Annotierung und möglicherweise weitere Daten werden dann in der zugrundeliegenden Wissensbasis gespeichert. Durch die Annotierung werden Ontologie-konforme Metadaten erzeugt, die von der Maschine mit weiteren relevanten Metadaten in Verbindung gebracht werden können, um so mit Hilfe entsprechender Regeln wichtige Schlußfolgerungen zu ziehen. Die Maschine kann dem Anwender so helfen, vorhandenes Wissen transparent zu machen, beispielsweise, um ihm die Entscheidung bezüglich der Relevanz und Qualität einer Information/Informationsquelle zu erleichtern.

## **ermöglicht die Kommunikation zwischen Mensch und Maschine und zwischen Maschinen und zwar unabhängig von der jeweils eingesetzten Technologie**

Damit ist es möglich, relevante (ontologiekompatible), im WWW verteilte und gespeicherte Informationsquellen, beispielsweise eine lokale Datenbank, ein XML Schemata, Fachdatenbanken, Suchmaschinen oder themenorientierte Websites mit der Ontologie zu verknüpfen und damit für die effektive Recherche nutzbar zu machen.

Als gemeinsame Wissensbasis ermöglichen Ontologien die einheitliche Verwaltung von unternehmensweiten Daten- und Informationsquellen. Theoretisch ist dies auch bei weltweit im Internet verteilten Informationsquellen denkbar. Ein solches Projekt, das versucht, einen großen Teil des menschlichen Weltverständnisses in Ontologien zu fassen, ist "OpenCyc", die OpenSource-Version von Cyc.

**Mit Ontologien wird unter anderem folgendes möglich:**

- Aus verschiedenen Dokumenten können die jeweils relevanten Suchergebnisse herausgefiltert und in einem neuen Dokument zusammengestellt werden - damit wird ein kontextabhängiger Zugriff auf die Wissensbasis möglich.
- Die zielgruppenorientierte Präsentation von Informations- und Wissensinhalten, also unterschiedliche Sichten auf eine bestimmte Wissensbasis wird ermöglicht, damit kann der Einstieg in ein Wissensgebiet für Laien vereinfacht werden.
- Im Gegensatz zu starren Schemata kann eine flexible Wissensbasis aufgebaut werden.
- Der Anwender kann durch flexible Navigationsmöglichkeiten beim Stellen der Suchanfrage unterstützt werden.
- Die Qualität des Suchergebnisses kann durch die Strukturen der Ontologie (Unterbegriffe, Synonyme,...) und die Nutzung von Beziehungen und Regeln verbessert werden.
- Es können Antworten auf konkrete Fragestellungen gegeben werden, etwa eine Liste mit den Vor- und Nachnamen und den E-Mail-Adressen von bestimmten Personen auf die Suchanfrage "Gib mir zu allen Ingenieuren der Entwicklungsabteilung den Vor- und den Nachnamen und die E-Mail-Adresse"
- Mit Hilfe von Computern und globalen Netzwerken wie dem Internet wird der weltweite Wissensaustausch von Gruppen mit ähnlichen Interessen und Vorlieben ermöglicht.
- ...

Mit Hilfe von Ontologien kann also das Basiswissen zu bestimmten Themen-/Fachgebieten von Experten so beschrieben werden, daß es von Anbietern (Autoren), Anwendern und Maschinen gleichermaßen verstanden und verwendet werden kann. Das Basiswissen kann dann von Autoren, die ihr Dokument der Ontologie konform mit Metadaten beschreiben, dazu genutzt werden, um das in ihrem Dokument enthaltene Wissen für Anwender besser zugänglich zu machen und gleichzeitig das Basiswissen der Ontologie zu erweitern. An der Erweiterung des Wissens können aber nicht nur Autoren mitwirken, sondern auch aktive Anwender mit Hilfe von ontologiekonformen Annotierungen, wie etwa eigene Kommentare, Verweise auf Internetseiten, Beiträge aus Webforen oder Bilddateien.

**Das Wissen zu einem Themen-/Fachgebiet kann also von Autoren und Anwendern gemeinsam erweitert werden und schafft so das Potential, mit maschineller Hilfe den Zugriff auf das betreffende Wissen für jeden Einzelnen zu vereinfachen. Dadurch, daß die Ontologie von vielen, mit jeweils unterschiedlicher Sichtweise auf das Wissen, erweitert wird, kann die maschinelle Verarbeitung des Wissens bisher noch ungeahnte Möglichkeiten für den persönlichen Zugang des Einzelnen zu diesem Wissen schaffen. So könnte beispielsweise nur das Wissen eines bestimmten Zeitraums oder nur aus einer bestimmten Region präsentiert werden.**

Im Semantic Web wird es wohl nicht eine Ontologie geben, sondern sehr viele, die jeweils einen kleinen Anwendungsbereich abdecken. Möglicherweise werden auch mehrere Ontologien für den gleichen Anwendungsbereich von unterschiedlichen Personen erstellt. Damit gibt es mehrere Wissensstrukturen für ein und dasselbe Themen-/Fachgebiet; das wiederum bedeutet, man muß Möglichkeiten bereitstellen, mit denen die Relevanz und Qualität der Ontologien für den eigenen Bedarf bewertet werden kann.

Die Ontologie wird dabei in einer eigenen Datei gespeichert und im Web veröffentlicht. Der Autor einer Website kann dann von seinen Internetseiten aus eine Verbindung zu der Ontologie-Datei herstellen.

Eine Sprache, mit der Ontologien erstellt werden können und die auf RDF Schema (RDFS) aufbaut, ist beispielsweise DAML+OIL (DARPA Agent Markup Language + Ontology Inference Layer, Referenz siehe "[www.w3.org/TR/2001/NOTE-daml+oil-reference-20011218](http://www.w3.org/TR/2001/NOTE-daml+oil-reference-20011218)"). Auf DAML+OIL basiert auch die vom "World Wide Web Consortium" entwickelte Ontologiesprache OWL (Web Ontology Language, Überblick siehe "[www.w3.org/TR/2004/REC-owl-features-20040210/](http://www.w3.org/TR/2004/REC-owl-features-20040210/)").

### Zusammenfassung: ↑

XML ermöglicht es, die Struktur der Inhalte maschinenlesbar zu beschreiben und die Angaben zu den Inhalten (was sagen die Daten aus) von den Angaben zum Layout (wie sollen die Daten dargestellt werden) zu trennen. Damit wird es möglich dem Leser ein und denselben Inhalt in unterschiedlichen Zusammenhängen darzustellen, dem Leser also einen Sachverhalt aus verschiedenen Blickwinkeln zu präsentieren. Die Technologie hilft dem Leser so entscheidend dabei Dinge kritisch zu hinterfragen und richtige Entscheidungen zu treffen. Mit RDF können beliebige Daten aus einer bestimmten Website oder aber auch vom gesamten Web in einen ganz bestimmten Zusammenhang gebracht und als Aussage (Subjekt, Prädikat, Objekt) formuliert werden. Mit RDF können also Informationen

konstruiert werden. Ontologien dienen dazu, beliebige, mit RDF formulierte Aussagen (Informationen) in einen ganz bestimmten Zusammenhang zu bringen. Mit Hilfe von Ontologien kann also Wissen konstruiert werden.

Das Semantic Web erweitert das bisherige World Wide Web im Wesentlichen um eine von Autoren, Anwendern und Maschinen gemeinsam verstandene und verwendete Sprache, die ermöglicht, daß Informationsanbieter ihre Dokumente und Informationssuchende ihre Suchanfragen maschinenverständlich beschreiben können. Die von den Informationsanbietern erstellten Metadaten werden an die eigentlichen Inhalte (Websites, Webseiten,...) angehängt. So entstehen zwei Ausgaben eines Dokuments: eine, die sich an Menschen richtet und eine, die für Maschinen bestimmt ist.

**Das für eine Maschine verständlich abgebildete menschliche Expertenwissen, die gemeinsame Weiterentwicklung dieses Basiswissens durch aktive Autoren und Anwender ist die Grundlage dafür, daß Maschinen qualitativ hochwertige Suchergebnisse liefern können und der Zugriff auf das jeweilige Wissensgebiet für jeden einfacher wird.**

Grundsätzlich fungiert die Maschine zwischen dem Suchenden, der eine bestimmte Information sucht, und dem Autor, der eine bestimmte Information anbietet - die Maschine soll letztendlich den Suchenden mit dem richtigen Autor bzw. dessen Information zusammenbringen. Dazu verwendet sie die Ontologie, die praktisch das durch Experten, Autoren und Anwender geschaffene Gehirn der Maschine repräsentiert. Die Aufgabe (den Suchenden mit dem richtigen Autor zusammenzubringen) kann der Maschine umso besser gelingen, je eindeutiger der Autor die von ihm angebotene Information beschreibt (Metainformation) und je klarer der Suchende beschreiben kann, welche Information er sucht.

Dadurch, daß die Ontologie von vielen mit jeweils unterschiedlicher Sichtweise auf das Wissen erweitert wird, kann die maschinelle Verarbeitung des Wissens bisher noch ungeahnte Möglichkeiten für den persönlichen Zugang des Einzelnen zu diesem Wissen schaffen.

Auf die gemeinsame Wissensbasis können die Anwender dann mit Hilfe von webbasierten oder lokal installierten Werkzeugen (Suchmaschinen, Informationsagenten) zugreifen. Suchmaschinen verwenden dabei sowohl herkömmliche Methoden, wie die Volltext- und Feldsuche, als auch ontologiebasierte Methoden. Für erstere gelten natürlich dann auch die eingangs genannten Probleme.

Wie ein Informationsagent auf der Basis der "Semantic Web"-Technologien funktionieren könnte, soll nun im nachfolgenden Abschnitt beschrieben werden.

### **kritische Betrachtung einer Maschine zur vollautomatischen Informationsbeschaffung ↑**

**Damit eine Maschine zur vollautomatischen Informationsbeschaffung eingesetzt werden kann, muß sie, genauso wie der Mensch auch, folgende Teilaufgaben zufriedenstellend erfüllen:**

**Für die erfolgreiche Suche muß die Maschine:**

**a.) vor der Suche wissen:**

- was sie genau sucht ↓
- wo sie sucht ↓
- wie sie mit der gewählten Informationsquelle sucht ↓

**b.) während und nach der Suche wissen:**

- wie sie auf Fehler reagieren kann, die während der Suche auftreten können ↓
- wie sie die Relevanz und Qualität der gefundenen Informationen/Informationsquellen richtig bewertet ↓

Dem Benutzer eine einfache Möglichkeit bieten, die relevanten Informationen/Informationsquellen so abzuspeichern, daß er sie später leicht wieder auffinden kann. ↓

Für eine sichere Nutzung der Internetdienste muß sie weiterhin geeignete Schutzmaßnahmen treffen,



um den Benutzer vor den Sicherheitsrisiken zu schützen, die mit dem Anschluß an das weltweite Datennetz "Internet" verbunden sind. ↓

Die vorher erwähnten Fähigkeiten gehören zum Grundlagenwissen; reicht dieses aber nicht mehr aus, so muß die Maschine bzw. der Agent in der Lage sein, sich selbständig weiteres Wissen anzueignen, das notwendig ist, um die gewünschte Information zu beschaffen. ↓

### Hinweis

Welche **Probleme** sich bei den jeweiligen Teilaufgaben stellen und wie Menschen mit Netzwerkkompetenz und ihrer persönlichen Internetquellensammlung bzw. die allgemeine Suchmaschine Google die dabei auftretenden Probleme lösen, wurde bereits in der "**Einführung zum Lehrgang**" in der Antwort zur Frage "Welche großen Probleme stellen sich für Internetnutzer, die nicht über eine ausreichende Netzwerkkompetenz verfügen, und wie können sie gelöst werden ? ↑" beschrieben.

**Nachfolgend soll - nun aufbauend auf die dabei gewonnenen Erkenntnisse - eine mögliche Maschine zur vollautomatischen Informationsbeschaffung (ein Informationsagent) auf der technischen Basis des "Semantic Web" beschrieben werden.**

---

Für die erfolgreiche Suche muß die Maschine, wie der Internetanwender auch:

### Hinweis

Die nachfolgenden Probleme sind allgemein, aus der Sicht des Internetanwenders beschrieben und stellen sich so auch für die Maschine.

a.) vor der Suche wissen:

was sie genau sucht ↑

### Problem

Problembeschreibung

### Lösung

Wie kann ein Informationsagent das Problem lösen ?

Der Benutzer kann dem Agenten seinen Informationsbedarf mitteilen, indem er:

- eine Menge von relevanten Suchbegriffen angibt, die eine bestimmte Information oder ein langfristig interessierendes Themengebiet aussagekräftig beschreiben

### Hinweis

Vorteilhaft wäre hier, wenn der Agent dem Benutzer entsprechende Suchassistenten bereitstellt, beispielsweise für die Suche nach Informationen

- zu einem bestimmten Themen-/Fachgebiet
- zu einer bestimmten Institution/Organisation
- zu einem bestimmten Produkt/einer bestimmten Dienstleistung
- zu einem bestimmten Aus-/Weiterbildungsangebot
- ...

Falls der Suchende über wenig oder gar kein Vorwissen zu dem, was er sucht, besitzt, sollte der Suchassistent hierfür geeignete Suchfunktionen oder zumindest Verweise auf hilfreiche Quellen bereitstellen.

Eine Suchanfrage, mit der nach bestimmten Produktinformationen gesucht werden soll, könnte beispielsweise den Produktnamen, Stich- und Schlagwörter zur Beschreibung der Qualität des Produkts (technische Angaben, Materialangaben,...), den Preis, die gewünschten Service-Leistungen (Garantie, Lieferzeit,...), den Zeitrahmen, innerhalb dem der Agent das gewünschte Produkt suchen soll, und eventuell benötigte Zusatzprodukte enthalten. Natürlich können die Angaben auch ungenauer sein, wobei dann die Gefahr besteht, daß man auch viele irrelevante



Suchergebnisse erhält.

Darüber hinaus sollte es auch möglich sein, einen eigenen Suchassistenten für die Suchanfrage (Suchfelder, Suchoptionen, Auswahl bestimmter Datenquellen,...) zu erstellen.

-----

- zusätzlich oder alternativ zu einer bestimmte Information oder einem langfristig interessierenden Themengebiet **relevante Dokumente in das Agentensystem importiert und dem Agenten durch die Auswahl der jeweiligen Wörter und Textpassagen mitteilt, welche Suchbegriffe relevant sind**

### **! Hinweis**

Um die Möglichkeiten des "Semantic Web" zu nutzen, müßten die in der Suchanfrage verwendeten Begriffe auch in der Ontologie definiert sein, sonst kann die Maschine die Möglichkeiten, die eine Ontologie zur verbesserten Informationsbeschaffung bereitstellt, nicht nutzen.

Der Informationsagent bietet dem Benutzer die Möglichkeit, genau vorzugeben, bei welchen Quellen (bestimmte Ministerien, Hochschulen, Fernsehsender,...) er nach den benötigten Informationen suchen und nach welchen grundsätzlichen Bewertungskriterien er die gefundenen beurteilen soll. Zusätzlich kann auch der Datenbestand auf dem eigenen PC oder Netzwerk durchsucht werden, etwa dort vorhandene Wissensdatenbanken. Der Benutzer hat die Möglichkeit, für das Suchergebnis ein bestimmtes Layout festzulegen. Falls der Benutzer es wünscht, macht der Agent Vorschläge, welche Internetquellen für die Suchanfrage relevant sein könnten und nennt Beschreibungen und Praxisbeispiele für die genannten Quellen.

Doppelt gefundene Suchergebnisse werden automatisch entfernt. Die genaue Auswahl der Treffer erfolgt durch den Benutzer. Die Suchanfrage kann vom Benutzer durch die Auswertung der gefundenen Dokumente weiter verfeinert werden oder aber durch den Agenten selbst, indem er das Nutzerprofil des Anwenders auswertet (siehe ... **muß der Agent in der Lage sein, sich selbständig weiteres Wissen anzueignen, das notwendig ist, um die gewünschte Information zu beschaffen ... ↓** ). Falls neue Informationsquellen gefunden werden, können diese einfach in das Agentensystem eingefügt werden.

Ist der Benutzer mangels Hintergrundwissen nicht in der Lage, dem Agenten seinen Informationsbedarf genügend präzise mitzuteilen oder die Suchanfrage weiter zu verfeinern, so muß der Agent ihn diesbezüglich mit entsprechenden Hilfsfunktionen unterstützen.

Grundsätzlich könnte der Agent für das Thema der Suche beispielsweise nützliche Katalogkategorien, spezielle Suchkataloge oder Thesauri nennen, über die der Benutzer sich die für die Suchanfrage relevanten Begriffe erschließen kann.

Besser wäre, die in der betreffenden Ontologie (zum Themen-/Fachgebiet, in das die gewünschte Information eingeordnet werden kann) erstellten Begriffshierarchien grafisch darzustellen. Der Benutzer bekommt so einen einfachen und geordneten Überblick über die relevanten Begriffe, erfährt, wie diese zusammenhängen, kann sich den Sinn von Begriffsketten erschließen, damit die relevanten Begriffe auswählen und sich so eine geeignete Suchanfrage zusammenbasteln oder eine bestehende Suchanfrage verfeinern. Zusätzlich können den einzelnen Begriffshierarchien entsprechende Sichtweisen auf den gesamten Wissensbestand erzeugt werden. Mit Hilfe der Ontologie wäre aber auch denkbar, ein bestimmtes Wissensgebiet so darzustellen, daß es für den Laien besser erschlossen werden kann. Damit werden dem Benutzer, vor allem dem mit wenig Hintergrundwissen, hilfreiche Möglichkeiten geboten, genauer zu definieren, welche Informationen die Maschine ihm beschaffen soll.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

### **Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:**

Die relevanten Inhalte des [universellen Suchsystems](#), das neben den speziellen Suchdiensten und Suchmethoden, auch die anderen im Rahmen des Lehrgangs beschriebenen Suchmöglichkeiten mit einbezieht.

- die entsprechenden Lerninhalte der "[universellen Suchstrategie](#)"
- die "[Methoden zum Einstieg in ein Thema und zur Suche nach relevanten Stich- und](#)

Schlagwörtern" mit der geeignete **Stich- und Schlagwörter** gefunden werden können



...

### Kritische Betrachtung:

Die Maschine benötigt genaue Anweisungen, um "richtig" handeln zu können; da sie nicht wirklich denken kann, ist sie auch nicht in der Lage, fehlende oder unzureichende Informationen und Anweisungen sinnvoll zu ergänzen und Widersprüche zu tolerieren. Das bedeutet, der Benutzer des Agenten muß die Information, die er sucht, möglichst genau beschreiben. Daß dies offensichtlich vielen Menschen schwer fällt, beweisen die zahlreichen Einwort-Suchanfragen bei großen Suchmaschinen.

Vor allem dann, wenn der Agent die Suchanfrage selbständig aufgrund des Nutzerprofils verfeinern soll, wird es Schwierigkeiten geben. In einem noch sehr neuen Nutzerprofil wird der Agent möglicherweise keine Informationen für die weitere Vorgehensweise finden, in einem alten dagegen viel zu viele. Es wird so nicht selten der Fall sein, daß der Agent das eigentliche Handeln des Benutzers, würde er die Suche selbst ausführen, falsch interpretiert.

wo sie sucht ↑



### Problembeschreibung



### Wie kann ein Informationsagent das Problem lösen ?

Mit Hilfe eines maschinellen, vom Benutzer einstellbaren Suchsystems hat der Agent ein grobes Konzept, wo er eine bestimmte Information, am besten finden kann. Über ein entsprechendes Dialogfenster kann der Anwender, dem Agenten die von ihm bevorzugten, in einer Datenbank hinterlegten Internetquellen bzw. speziellen Such- und Informationsdienste mitteilen.

Die Anbieter der speziellen Such- und Bewertungsdienste müssen dazu alle Internetquellen, beispielsweise wie im Lehrgang gezeigt, beschreiben und die Ergebnisse als entsprechende (vom Agenten) auswertbare Metadaten bzw. Such- Ontologien in einer Datenbank abspeichern. Der Agent kann dann auf die von den Anbietern, der speziellen Such- und Bewertungsdienste zu den betreffenden Informationen/Informationsquellen erstellten Metadaten zurückgreifen und verwendet für die Suche, die vom Anwender genannten Internetquellen bzw. speziellen Such- und Informationsdienste. Falls der Anwender keine bestimmte Quellen vorgibt, dann nutzt der Agent einfach die Standard-Quellen, die automatisch in möglichst kurzen Zeitabständen aktualisiert werden.

Damit der Agent also erkennen kann, ob eine Quelle für die Erfüllung seiner Aufgabe nützlich ist, müssen die Quellen mit Schemata oder, noch besser, mit Ontologien beschrieben werden. Durch die formale Beschreibung der Quellen wird es für den Agenten möglich, mit den in der Ontologie eingebundenen Internetquellen zu arbeiten. Das bedeutet, mit den dort vereinbarten Relationen und Regeln können dann bestimmte, für die jeweilige Fragestellung geeignete Informationsquellen in die Recherche miteinbezogen werden.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

### Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:

Das universelle Suchsystem, das neben den speziellen Suchdiensten und Suchmethoden, auch die anderen im Rahmen des Lehrgangs beschriebenen Suchmöglichkeiten mit einbezieht.

Relevante Methoden sind:



die Methode "[Eine individuelle Internetquellensammlung erstellen und nutzen](#)"

- die flexibel einsetzbare **universelle Suchstrategie**, die als Wegweiser durch den Informationsdschungel dient und unabhängig von der gesuchten Information mit hoher Wahrscheinlichkeit zu relevanten Informationen führt
- die **universelle Einkaufsstrategie**, die als wertvolle Anleitung dient, um qualitativ hochwertige und günstige Produkte/Dienstleistungen zu finden und sicher online zu erwerben
- ...

### Kritische Betrachtung:

**Dem Informationsagenten geht es wie den allgemeinen Suchmaschinen auch, ihm bleibt der Zugang zu vielen Quellen verwehrt.** Es gibt jede Menge hochwertiger Fachdatenbanken und Websites, die weder für Robots von allgemeinen Suchmaschinen noch für den Informationsagenten zugänglich sind. Ein anderes Problem sind die vielen unterschiedlichen Formate, in denen die Informationen vorliegen; der Informationsagent kann nur die ihm bekannten Datenformate lesen. Die Erfassung all dieser im sogenannten "Deep Web" (Invisible Web) befindlichen Quellen und deren Einbindung in das Semantic Web wäre, wenn überhaupt, nur mit erheblichem Aufwand zu meistern.

Bestimmte Internetanbieter haben möglicherweise auch gar kein Interesse daran, daß der Agent Zugang zu ihren Websites hat. Soll der Agent beispielsweise für ein bestimmtes Produkt verschiedene Angebote (Preis, Serviceleistungen,...) von Online-Shops einsammeln, so werden ihm hier viele Shop-Anbieter den Zugang verwehren. Während die meisten Käufer wohl begeistert wären, mit Hilfe des Agenten den billigsten (und besten) Händler für das betreffende Produkt zu finden, wäre der Agent für viele Verkäufer ein zu mächtiges Instrument in den Händen des Kunden. Das kommt daher, daß der Käufer und der Verkäufer teilweise gegenteilige Geschäftserwartungen haben: Der Käufer möchte beispielsweise alle Angebote für ein Produkt vergleichen, der Verkäufer hingegen möchte die Kunden an sich binden und ihnen lediglich eine Auswahl der eigenen Produkte anbieten. Die oben beschriebene Reaktion der Online-Shop-Anbieter ist also verständlich, da die Käufer, die jetzt bequem die billigsten (und besten) Produkte über einen Agenten ermitteln könnten, nur noch dann zum Händler kommen, wenn er entsprechende Schnäppchen anbietet.

Interessant für beide Seiten dagegen wäre ein Agent für den Handel mit Sonder- oder Restposten, da der Verkäufer hier froh ist, die Ware überhaupt an den Mann oder die Frau zu bringen und der Käufer so eine bessere Möglichkeit bekommt, überhaupt erst von diesem Schnäppchen zu erfahren.

Ein weiteres Problem ist, daß es bereits viele Millionen Internetseiten gibt, die nicht den technischen Standards des "Semantic Web" entsprechen und bei denen es sich einfach vom Zeit- und Arbeitsaufwand her nicht lohnen würde, ihn nachträglich einzuführen.

Es ist sogar insgesamt fraglich, ob es sich lohnen würde, den Informationsagenten durch die riesigen Informationshalden im Internet zu schicken, obwohl zahlreiche kostenpflichtige Fachdatenbanken existieren, in denen der Benutzer, der hochwertige Informationen benötigt, bequem und einfach auf die von ihm gewünschte Information zugreifen kann.

Viele Fachdatenbanken findet man unter dem Dach von Fachinformationsdiensten, die häufig auch eine direkte Schnittstelle zum lokalen Intranet anbieten.

Es sind umfangreiche Kenntnisse und Erfahrungen notwendig, um entscheiden zu können, mit welcher Quelle welche Information am ehesten gefunden werden kann; dies kann man nicht exakt in Relationen und Regeln fassen.

Eine vollständige Suche über alle relevanten Informationsquellen ist mit der Maschine nicht denkbar. Dies ist zwar auch beim Menschen ausgeschlossen, aber er kann auf manuellem Wege wesentlich flexibler mit den Informations- und Kommunikationssystemen umgehen als eine Maschine. Die große Gefahr bei der vollautomatischen Informationsbeschaffung besteht darin, daß viele an sich sehr nützliche Quellen ausgeschlossen werden.

**wie sie mit der gewählten Informationsquelle sucht ↑**



### Problembeschreibung

## Lösung

### Wie kann ein Informationsagent das Problem lösen ?

Damit der Agent die lokalen und im Internet verteilten Informationsquellen (Datenbanken, Kataloge,...) abfragen kann, müssen **standardisierte Schnittstellen** zwischen dem Agenten und den betreffenden Internetangeboten vorhanden sein. Das bedeutet, die Quellen müssen für die effektive Abfrage dem Semantic Web gemäß mit Metadaten beschrieben sein und die Suchanfrage des Benutzers muß in die Abfragesprache der jeweiligen Quelle übersetzt werden.

Das Suchergebnis muß dann, dem Themen-/Fachgebiet der gefundenen Informationen entsprechend, für den Benutzer einfach hinsichtlich seiner Relevanz bewertbar präsentiert werden. Falls die Information für den Benutzer relevant ist, sollte sie in einem Format vorliegen (XML), das gut für die elektronische Weiterverarbeitung und zum generellen Datenaustausch geeignet ist.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

### Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:

keine vorhanden

### Kritische Betrachtung:

Da es bei Datenbanken nicht nur Volltext zu durchsuchen gilt, sondern auch strukturierte Inhalte, gibt es, abhängig vom Inhalt der Datenbank, die unterschiedlichsten Suchfelder. Auch die Abfragesprachen sind sehr verschiedenartig. Es wird mit enormem Aufwand verbunden sein, die unterschiedlichen Suchfelder und Abfragesprachen im Agentensystem zu berücksichtigen. Um die Richtigkeit dieser Aussage zu erkennen, braucht man sich nur die Hilfeseiten und die darin mehr oder weniger verständlichen Informationen über die Bedeutung der Suchfelder und die meist zahlreichen Abfragemöglichkeiten einer Reihe von Fachdatenbanken anzusehen. Gerade hinsichtlich der häufig sehr speziellen Suchfelder wird es notwendig sein, daß der Agent Rückfragen an den Benutzer stellt. Eine andere Möglichkeit wäre, daß der Agent anhand der Suchbegriffe automatisch erkennt, welche ihm bekannten Quellen für die Suche in Frage kommen und den Benutzer nur über die Besonderheiten (spezielle Suchfelder und Abfragetechniken) informiert.

Noch problematischer wird es dann, wenn der Agent die für die Suchanfrage relevante Quelle erst ausfindig machen muß und bei der Abfrage der Quelle dann feststellt, daß ihm der Zusammenhang zwischen den Suchbegriffen und den vorhandenen Suchfeldern nicht klar ist.

### b.) während und nach der Suche wissen:

wie sie auf Fehler reagieren kann, die während der Suche auftreten können ↑

## Problem

### Problembeschreibung

## Lösung

### Wie kann ein Informationsagent das Problem lösen ?

Damit der Agent auf unvorhergesehene, aber bekannte Ereignisse richtig reagieren kann, muß die für das Ereignis notwendige Reaktion im Programm, beispielsweise innerhalb einer entsprechenden Funktion, programmiert werden.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

## Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:

- das [Handlungswissen](#) um Probleme, die während der Suche auftreten können, kompetent zu lösen
- ...

## Kritische Betrachtung:

Damit der Agent richtig handeln kann, müssen alle möglichen Ereignisse und die notwendigen Reaktionen vom Programmierer mit entsprechenden Anweisungen im Rahmen des Programms beschrieben werden. Tritt ein im Programm nicht definiertes Ereignis ein, so bleibt nur die Möglichkeit, daß der Agent seinen Benutzer um Rat fragt.

Aber auch die bekannten Ereignisse sind zahlreich und vor allem die darauf zu erfolgenden Reaktionen. Wie soll der Agent beispielsweise auf ein Suchergebnis reagieren, das sehr wenig oder gar keine Treffer erhält ? Das kann die unterschiedlichsten Gründe haben, beispielsweise ein falsch geschriebenes Suchwort, die Verwendung einer falschen oder gar nicht vorhandenen Suchfunktion oder ein zu spezielles Suchwort. Bei den genannten Fehlern müßte die Suchmaschine den Fehler erst erkennen und dann richtig reagieren. Die Fehlermöglichkeiten sind hier mannigfaltig; alleine für die Rechtschreibung müßte die Maschine über ein umfassendes mehrsprachiges Wörterbuch verfügen, das falsch geschriebene Wort ausfindig machen und durch das richtig geschriebene ersetzen.

Ganz nebenbei stellt sich auch die Frage, wo ist die Grenze zu setzen, bei der zu wenig Treffer gefunden wurden.

**wie sie die Relevanz und Qualität der gefundenen Informationen/Informationsquellen richtig bewertet ↑**



**Problem**

## Problembeschreibung



**Lösung**

## Wie kann ein Informationsagent das Problem lösen ?

Mit Hilfe eines maschinellen, vom Benutzer einstellbaren [Bewertungssystems](#) kann der Agent die Relevanz und Qualität der gefundenen Informationen/Informationsquellen bewerten. Dazu gibt es ein entsprechendes Dialogfenster über das der Anwender dem Agenten die von ihm bevorzugten, in einer Datenbank hinterlegten, [speziellen Bewertungskriterien und -methoden](#) mitteilen kann.

Die Anbieter der [speziellen Such- und Bewertungsdienste](#) müssen dazu alle Informationen/Informationsquellen mit Hilfe des im Lehrgang beschriebenen [Bewertungssystems](#) bewerten und die Ergebnisse als entsprechende (vom Agenten) auswertbare Metadaten in einer Datenbank abspeichern. Der Agent kann dann die von den Anbietern, der [speziellen Such- und Bewertungsdienste](#) zu den betreffenden Informationen/Informationsquellen erstellten Metadaten, mit den Vorgaben des Anwenders vergleichen und mit den technischen Möglichkeiten des "[Semantic-Web](#)" darstellen. Falls der Anwender keine bestimmten Bewertungskriterien und -methoden vorgibt, dann nutzt der Agent einfach die Standard-Bewertungsvorgaben, die automatisch in möglichst kurzen Zeitabständen aktualisiert werden.

Beispielsweise könnten die Ergebnisse dieses Vergleichs in eine entsprechende Bewertungs- [Ontologie](#) eingeordnet werden, die dann vom Benutzer grafisch ausgewählt werden kann. Dies hat gegenüber den herkömmlichen Such- und Informationssystemen den Vorteil, **daß der Benutzer auf einen Blick erkennt, in welchem konkreten Zusammenhang die betreffende Information/Informationsquelle tatsächlich veröffentlicht wurde und wie gut sie seinen individuellen Bewertungskriterien entspricht** und erlaubt so eine schnelle und gezielte und vor allem qualitativ hochwertige Auswahl.

Bei der Standard-Trefferausgabe (ohne Benutzervorgaben) werden die gefundenen Internetdokumente (Internetseiten, PDF-Dokumente, Musikdateien,...) mit ihrem Titel, einer kurzen Beschreibung, der

Quellangabe und gegebenenfalls weiteren wichtigen Daten im Suchergebnis nach Relevanz und Herkunft (Ministerien, Bibliotheken, Suchkataloge, Newslettern, Webforen,...) sortiert aufgelistet. Weiterhin lassen sich die einzelnen Treffer vergleichen und nach den individuellen Bewertungskriterien sortieren. Die besonders relevant erscheinenden Treffer können markiert und durch einen Befehl von der entsprechenden Quelle auf die lokale Festplatte heruntergeladen werden, wo sie offline betrachtet werden können. Die komplette Trefferliste kann dauerhaft auf dem lokalen Rechner abgespeichert und jederzeit ausgedruckt oder per E-Mail versendet werden.

In jedem Treffer werden neben wichtigen Basisangaben auch zusätzliche Informationen wie etwa themenrelevante Kommunikationsdienste (Webforen, Newsgroups,...) genannt, um so nicht nur nutzbringende Dokumente, sondern auch andere [Menschen](#) zu finden, die ebenfalls an der betreffenden Information bzw. dem jeweiligen Thema interessiert sind. Durch die Kontaktaufnahme mit diesen Menschen wird es möglich, gerade komplexere Informationen gemeinsam zu verarbeiten, beispielsweise im Rahmen einer Diskussion. Genauso kann auf diese Weise natürlich bereits erarbeitetes Wissen zur Diskussion gestellt werden, um so mögliche Fehler zu finden und entsprechend zu korrigieren.

Falls der Agent die gewünschte Information nicht finden kann, könnte er stattdessen in Frage kommende Menschen nennen und so die gemeinsame Suche ermöglichen. Diese Menschen könnte der Agent über die Nutzerprofile anderer Agenten ausfindig machen.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

**Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:**

Das [universelle Bewertungssystem](#), das neben den verschiedenen Bewertungskriterien und -methoden auch die anderen im Rahmen des Lehrgangs beschriebenen Bewertungsmöglichkeiten mit einbezieht.

Relevante Methoden sind:

- die universelle [Methode zur Relevanz- und Qualitätsbewertung von Websites](#), mit der man **schnell und gezielt die Relevanz und Qualität einer Website oder Internetseite in Bezug auf die gesuchte Information** bewerten kann
- die universelle [Methode zur Qualitätsbewertung von Informationen und beliebigen Internetquellen](#), mit der man systematisch herausfinden kann, **wer (Einzelpersonen, Institutionen, Organisationen,...) und wie viele eine Informationsquelle oder ein Internetdokument bzw. die darin enthaltenen Informationen immerhin so gut finden, daß sie mit einem Hyperlink auf einer ihrer Internetseiten oder im Rahmen einer elektronischen Nachricht (Webforum, Newsgroup,...) darauf verweisen**
- die [Bewertungskriterien für Websites](#) bzw. die [Bewertungskriterien](#), mit denen die **allgemeine Qualität der gefundenen Information bzw. der betreffenden Informationsquelle** (Website oder Teilangebote einer Website wie etwa Suchmaschinen bzw. Datenbanken, Newslettern, Webforen, Webchats, Online-Shops,...) **bewertet werden kann**
- [Methode zur gezielten Auswahl eines geeigneten Weiterbildungsangebots](#), die eine wertvolle Hilfestellung gibt, **um das für das persönliche Lernziel am besten geeignete** aus der Vielzahl der möglichen **Weiterbildungsangebote auszuwählen**
- die universelle [Methode zur Relevanz- und Qualitätsbewertung von Online-Handelsplattformen](#), mit der man **schnell und gezielt die Relevanz und Qualität einer Online-Handelsplattform in Bezug auf das gesuchte Produkt/die Dienstleistung** bewerten kann
- ...

**Kritische Betrachtung:**



Relevantere Suchergebnisse kann der Agent nur dann liefern, wenn ihm maschinell einwandfrei auswertbare Metadaten zu den jeweiligen Informationen/Informationsquellen auf der technischen Grundlage des Semantic Web zur Verfügung stehen.

Das bedeutet, alle Internetanbieter müssen ihre Inhalte gewissenhaft mit Metadaten beschreiben. Welche Schwierigkeiten hier zu erwarten sind, erfahren Sie im [Fazit](#), also am Ende der theoretischen Beschreibung des Informationsagenten.

Darüber hinaus müssen die Quellen in einem zur Ontologie kompatiblen Format erstellt sein. Viele Dokumente liegen in anderen Formaten als XML vor, etwa Word-Dokumente. Sie können zwar mit entsprechenden Programmen in das XML-Format umgewandelt werden, ihnen fehlen dann aber trotzdem die Metainformationen.

Der Agent bewertet die Informationen rein objektiv, der Mensch dagegen auch subjektiv. Diese subjektive Bewertung ist für den Einzelnen sehr wichtig und ist auch von zentraler Bedeutung, denn es kennzeichnet ihn als Individuum und ermöglicht erst die Vielfalt an Informationsangeboten (wo keine Nachfrage, da kein Angebot). Durch eine rein objektive Bewertung kann es beispielsweise passieren, daß wichtige Informationen, die nicht im Hauptinteresse des Benutzers liegen, ausgeblendet werden. Da der Agent das Suchverhalten des Benutzer kopiert, könnte er beispielsweise aus den bisherigen Erfahrungen schließen, daß für den Benutzer die ein oder andere Information/Informationsquelle nicht relevant ist und ihm so möglicherweise wichtige Informationen/Informationsquellen vorenthalten.

Noch bedenklicher wird es, wenn der Agent dem Benutzer gezielt wichtige Informationen vorenthält, also bestimmte Themen ausschließt oder Dokumente ausfiltert. Ein entsprechend weit verbreiteter Informationsagent könnte somit sehr gut als (globale) Zensurmaschine mißbraucht werden. Der Agent müsste also, um einem Mißbrauch vorzubeugen, als Open-Source-Projekt entwickelt werden.

Letztendlich wird ein Agent die Aufgabe, Suchergebnisse aus verteilten Quellen zu bewerten und zusammenzufassen, nur sehr grob erfüllen können. Die sehr zeit- und arbeitsaufwendige Feinauswertung muß der Benutzer weiterhin selbst durchführen.

**Dem Benutzer eine einfache Möglichkeit bieten, die relevanten Informationen/Informationsquellen so abzuspeichern, daß er sie später leicht wieder auffinden kann. ↑**

## **Problem**

### [Problembeschreibung](#)

## **Lösung**




### **Wie kann ein Informationsagent das Problem lösen ?**

Diese Teilaufgabe muß nicht vom Informationsagenten gelöst werden; es wäre aber sinnvoll, wenn Schnittstellen zu wichtigen lokalen Informationssystemen entweder bereits vorhanden sind und nur aktiviert werden müssen oder zumindest nachträglich eingerichtet werden können. Damit kann die Wissensbasis effektiv erweitert werden.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

### **Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:**

-  das hierarchisch organisierte [Ordnungssystem](#) (die [gemeinsamen Internetquellensammlung](#))
-  die Methode "[Eine individuelle Internetquellensammlung erstellen und nutzen](#)" (die [individuelle Internetquellensammlung](#))
-  das [Handlungswissen zum schnellen Wiederauffinden von Informationen](#), das zeigt, wie man mit Hilfe geeigneter Werkzeuge (Web-Client, News-Client, Bookmark-Manager, Windows-Explorer,...) die im Internet gefundenen Informationsquellen und Informationen leicht wiederauffindbar auf dem lokalen Rechner abspeichern kann

**Für eine sichere Nutzung der Internetdienste muß sie weiterhin geeignete Schutzmaßnahmen treffen, um den Benutzer vor den Sicherheitsrisiken zu schützen, die mit dem Anschluß an das weltweite Datennetz "Internet" verbunden sind. ↑**

## **Problem**

### Problembeschreibung

## **Lösung**

### **Wie kann ein Informationsagent das Problem lösen ?**

Ein Informationsagent verhindert keine Gefahren, die mit der Nutzung der Internetdienste verbunden sind, sondern er schafft vielmehr neue Risiken.

Da der Agent anstelle des Benutzers im Internet nach den gewünschten Informationen sucht, benötigt er nicht nur personenbezogene Daten seines Benutzers (die Identität des Benutzers und wichtiger Kommunikationspartner, Geburtsdatum, Benutzerkenn- und Paßwörter für die unterschiedlichsten Internetdienste,...), sondern speichert in dem sogenannten Nutzerprofil auch eine Vielzahl an Informationen (Vorlieben, Absichten, aufgetretene Fehler und relevante Reaktionen, einzelne Suchanfragen und entsprechende Suchergebnisse,...), die er für Entscheidungen während der Recherche benötigt. Da der Agent vom Suchverhalten seines Benutzers lernt, protokolliert er all seine Aktivitäten mit, also auch die Recherchen, die er nicht selber durchführt.

Darüber hinaus kommuniziert der Agent mit anderen Agenten, beispielsweise um zu erfahren, mit welcher Informationsquelle er eine bestimmte Information am besten finden kann.

Da Agenten, die mit entsprechenden Schad- und Spionagefunktionen ausgestattet sind, auch in der Lage wären, die personenbezogenen Daten und die Interessenprofile des Benutzers eines anderen Agenten zu aususpionieren und/oder manipulieren, tun sich Gefahren auf, die zwar den Benutzer direkt betreffen, zu denen er aber möglicherweise selbst keine Schutzmaßnahmen treffen kann. Was nützt in diesem Fall ein gefahrenbewußter Benutzer, wenn dessen Agent alle Regeln zum Schutz und zur Sicherheit seiner Daten mißachtet.

Man denke hier beispielsweise an den Entwickler in einem Unternehmen, dessen Informationsbedarf dem Konkurrenzunternehmen offengelegt wird oder dessen Informationen, die in der Regel zu wichtigen Entscheidungen herangezogen werden, manipuliert werden. Auch mit den personenbezogenen Daten in Verbindung mit den Interessenprofilen privater Anwender können Datenhändler viel Geld verdienen.

Die Vertraulichkeit und Integrität der vom Agenten im Nutzerprofil gespeicherten oder ausgetauschten Daten, insbesondere der personenbezogenen Daten, müssen durch entsprechende Schutzmaßnahmen sichergestellt werden. Darüber hinaus muß der Agent die Identität der Server und Agenten, mit denen er kommuniziert, eindeutig nachweisen können.

Die vorher genannten Sicherheitsanforderungen könnten am besten durch die Integration eines hochwertigen Verschlüsselungs- und Signiersystems in den Agenten erfüllt werden. Die elektronische Verschlüsselung soll sicherstellen, daß die (per Internet) übertragenen Daten nur vom rechtmäßigen Empfänger gelesen werden können (Vertraulichkeit der Information). Die elektronische Signatur dient dem Absender/Empfänger dazu, zum einen die eindeutige Identität des Kommunikationspartners (Benutzer, Server, andere Agenten) nachzuweisen und ferner ist sie erforderlich, um die Unversehrtheit der übertragenen Information beim Überprüfen der Signatur zu erkennen (Authentizität und Integrität der Information).

Eine weitere für den Agenten interessante, vom WWW-Consortium entwickelte Technologie, mit der einerseits die Privatsphäre des Anwenders geschützt und andererseits die Interessen der Online-Wirtschaft gewahrt werden soll, ist das Protokoll P3P.




Damit der Benutzer dem Agenten grundsätzlich vertrauen kann, muß dessen Funktionsweise so transparent wie möglich sein. Das bedeutet, der Benutzer kann so weit wie möglich auf die Aktivitäten des Agenten Einfluß nehmen, insbesondere in Bezug auf die Nutzung der personenbezogenen Daten, der Agent stellt Rückfragen oder macht Vorschläge, wo diese notwendig sind und legt seine Aktivitäten so weit wie möglich offen. Dies gilt übrigens nicht nur für sicherheitsrelevante Aktionen

des Agenten, sondern ganz allgemein für alle Aufgaben der Informationsbeschaffung.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe der Informationsbeschaffung löst und wie das Ganze eine Maschine (ein Informationsagent, Google) macht.**

[Mensch](#) | [Informationsagent](#) | [Google](#)

**Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:**

-  das [Handlungswissen zur sicheren Nutzung der Internetdienste](#)
-  der **Lehrgang "Sichere Nutzung von IT-Systemen mit Internetzugang"**
-  ...

### **Kritische Betrachtung:**

Grundsätzlich wird eine wirklich sichere Nutzung von Agenten erst durch die elektronische Signatur sichergestellt werden können. Leider ist es bisher nicht gelungen, die elektronische Signatur flächendeckend einzuführen, da bis jetzt keine Lösung gefunden wurde, die Ausgabe und Verwaltung der privaten und öffentlichen Schlüssel, die für die elektronische Signatur notwendig sind, für eine große Masse von Anwendern einfach, sicher und kostengünstig zu realisieren. Wie lange es noch dauern wird, kann man nicht genau voraussagen.

Aber selbst wenn die elektronische Signatur flächendeckend verfügbar ist, bleiben Sicherheitslücken bestehen. So ist es beispielsweise möglich, mit der elektronischen Signatur die Herkunft und die Authentizität eines Programmcodes, beispielsweise eines anderen Agenten, zu überprüfen, es kann aber dadurch nicht sichergestellt werden, daß dieser tatsächlich keine Schad- und/oder Spionagefunktionen enthält.

Ein weiteres Problem ist, daß zahlreiche im Internet nach Informationen stöbernde und miteinander kommunizierende Agenten das sowieso schon sehr belastete Netz zusätzlich stark beanspruchen würden.

### **Problem**

Die vorher erwähnten Fähigkeiten gehören zum Grundlagewissen, reicht dieses aber nicht mehr aus, so **muß der Agent in der Lage sein sich selbständig weiteres Wissen, das notwendig ist um die gewünschte Information zu beschaffen, aneignen.** ↑

### **Lösung**

**Wie kann ein Informationsagent das Problem lösen ?**

Damit ein Agent lernen kann, muß er aufgrund fehlender Eigenintelligenz das Rechercheverhalten seines Benutzers protokollieren und auswerten und verwendet schließlich die Ergebnisse als Anleitung für die eigenen Suchaktivitäten. Der Agent paßt also seine Wirkungs- und Funktionsweise selbständig an die individuellen Notwendigkeiten und Ziele des Benutzers an. Damit diese Anpassung wirklich im Sinne des Benutzers geschieht, hat dieser die Möglichkeit, kontrollierend und, falls notwendig, korrigierend in den Anpassungsprozeß einzugreifen.

So kann er beispielsweise, um die aktuelle Suchanfrage zu verfeinern, die bisher gespeicherten Suchanfragen (einschließlich der dazugehörenden Verfeinerungen) nach ähnlichen Anfragen durchsuchen und, falls er fündig wird, diese als Vorlage benutzen. Im Zweifelsfall kann er den Benutzer fragen, der dann an seiner Stelle die Suchanfrage verfeinert. Eine weitere denkbare Aktion wäre, daß er zur Relevanz- und Qualitätsentscheidung bezüglich der gefundenen Informationen ältere ähnliche Anfragen und deren Ergebnisse heranzieht oder die Reaktionen des Benutzers auf die von ihm empfohlenen Treffer beobachtet.

Dabei kann der Agent beispielsweise auf folgende Merkmale achten und im Nutzerprofil speichern:



- Welche Treffer des Suchergebnisses wurden lokal gespeichert
- Wie häufig wurde auf die jeweiligen lokal abgespeicherten Treffer zugegriffen
- Wie lange wurden die Dokumente betrachtet
- Wurden im Dokument angegebene Links verfolgt
- ...

Falls ähnliche Suchanfragen noch nicht vorliegen, können vom Benutzer besonders wichtige Inhalte von relevanten Dokumenten dem Agenten verständlich gemacht werden, beispielsweise durch das Markieren wichtiger Suchbegriffe. Weiterhin lernt der Agent durch die Beobachtung und Speicherung der Reaktionen. Darüber hinaus kann er durch die Zusammenarbeit mit anderen Agenten lernen. Das bedeutet, er kann die relevanten Inhalte der Nutzerprofile der Benutzer dieser Agenten für seine Recherche nutzen. Dazu muß er die verschiedenen Nutzerprofile, beispielsweise in Bezug auf den Inhalt der Suchanfrage und das Vorwissen der jeweiligen Benutzer, vergleichen und bei einer Übereinstimmung kann er die von den anderen Benutzern verwendeten Informationsquellen bzw. die ausgewählten Informationen seinem Benutzer vorschlagen. Für den dazu notwendigen Informationsaustausch müssen alltagstaugliche (auf den Technologien des Semantic Web basierende) und vor allem sichere Kommunikationsprotokolle zum Einsatz kommen. Die jeweiligen Informationen müssen in maschinenlesbarer Form gespeichert werden und so abrufbar zur Verfügung stehen. Die wichtigste Aufgabe des Agenten ist es aber, mit seinem Benutzer zu kommunizieren, sein Suchverhalten zu beobachten und ihm sinnvolle Vorschläge zu machen.

Eine einfachere Möglichkeit wäre, dem Agenten den Zugriff auf eine Datenbank zu ermöglichen, in der sich Benutzer oder deren Agenten mit Angabe der benötigten Information eintragen können und die dann im Bedarfsfall vom Agenten abgefragt werden kann. Der Agent nennt dann, falls er fündig wird, den entsprechenden Anwender und der Benutzer kann mit ihm dann beispielsweise per E-Mail Kontakt aufnehmen. Weitere Möglichkeiten bietet die Suche in archivierten Newsgroup- und Webforenbeiträgen.

**Direkter Vergleich zwischen dem wie ein Mensch mit grundlegender Netzwerkkompetenz die genannte Teilaufgabe löst und wie das Ganze eine Maschine (ein Informationsagent) macht.**

[Mensch](#) | [Informationsagent](#)

**Relevante (automatisierbare bzw. in Software umsetzbare) Methoden und Lerninhalte:**

keine vorhanden

**Kritische Betrachtung:**

Um den Agenten zu trainieren, muss er aufgrund fehlender Eigenintelligenz das Suchverhalten des Benutzers kopieren. Das bedeutet, daß die selbständige Suche des Agenten nur dann von wirklichem Erfolg gekrönt sein kann, wenn der Benutzer die Informationsbeschaffung im Internet beherrscht. Welchen Nutzen kann aber ein Meister von seinem Lehrling ziehen ? Letztlich kann der Lehrling nur ein guter Gehilfe sein, der Standardaufgaben erledigt - dies kann zwar auch nützlich sein, viel wichtiger aber wäre, wenn der Informationsagent wirklich effektive, dem Benutzer noch unbekanntes Suchstrategien für die unterschiedlichsten Informationen vorschlagen und zumindest eine grobe Vorbewertung des gefundenen Informationsmaterials vornehmen könnte. Vorschläge für relevante Suchstrategien könnte der Agent möglicherweise auch von anderen Agenten erfragen.

Problematisch in diesem Zusammenhang wäre es für die Maschine langfristig sicher auch, aus der Vielzahl der protokollierten Suchaktivitäten die tatsächlich richtigen für die aktuelle Suchsituation auszuwählen. Diese Entscheidung kann der Benutzer sicher schneller und besser treffen als sein Agent, wobei dies möglicherweise im Einzelfall anders sein könnte - aber nur um diesen Einzelfall zu nutzen, müßte sich der Benutzer mit allen vom Agenten vorgeschlagenen Suchstrategien auseinandersetzen - die ja letztendlich auch viele bereits gemachte und korrigierte, aber vom nichtsahnenden Agenten kopierte Fehler beinhalten. Es wäre sehr zeitaufwendig für den Benutzer, sich mit bereits erkannten und eigentlich aus dem Gedächtnis gestrichenen Fehlern immer wieder auseinander setzen zu müssen.

Darüber hinaus könnte die Maschine ihren Benutzer höchstens als Objekt erfassen, nicht als Subjekt. Das bedeutet, in die Suche können keine Begabungen, Leidenschaften und Werte des Benutzers mit einspielen, die aber für die meisten eine wichtige Rolle spielen dürften. Die objektive Protokollierung und Speicherung des Suchverhaltens des Benutzers und die maschinelle Auswertung dieser Daten wird zwangsläufig dazu führen, daß der Agent seinen Benutzer häufig fehlinterpretiert. Die einfachen

logischen Schlußfolgerungen des Agenten werden daher eher den Charakter eines Vorschlages haben als den einer konkreten Entscheidung im Sinne des Benutzers. Das vorher genannte Problem dürfte sich noch verschärfen, wenn der Benutzer nicht ausschließlich Informationen zu einem bestimmten Themengebiet benötigt, sondern zu häufig wechselnden Fachbereichen und Fragestellungen.

Die Zusammenarbeit mit anderen Agenten bzw. der Vergleich mit deren Nutzerprofilen wird allerdings nur dann nützliche Vorschläge für die weitere Suche liefern, wenn die Suchanfrage ein beliebtes Thema betrifft, da er bei sehr speziellen Themen kaum ein geeignetes Nutzerprofil finden wird.

**Fazit: ↑**

**Mit welchen grundlegenden (mehr oder weniger lösbaren) Problemen müssen sich die Entwickler einer Maschine zur vollautomatischen Informationsbeschaffung auseinandersetzen ?**

**Da das Internet bzw. die Menge der dort veröffentlichten Informationen mit hoher Geschwindigkeit weiterwächst, wird die Bedeutung von Maschinen, die den Menschen beim Auffinden von bestimmten Informationen unterstützen, ebenfalls zunehmen.** Dazu müssen die Maschinen allerdings die Bedeutung der veröffentlichten Informationen verstehen. Dies wiederum setzt voraus, daß die Autoren ihre Dokumente gewissenhaft mit Hilfe der Technologien des "Semantic Web" beschreiben.

Die Theorie von durch das Netz streifenden Agenten, die im Dienste ihres Auftraggebers wie ein Rechercheprofi nach relevanten Informationen suchen, klingt verlockend. Solche Agenten (mit einer entsprechenden Wissensbasis) könnten es auch weniger versierten Anwendern ermöglichen, die Informationen im Internet zu finden, die sie benötigen.

Damit der Agent aber nicht nur selbständig durch das Internet reist, sondern auch fündig wird, muß er allerdings:

- Suchanfragen selbständig verfeinern können
- mit seinem Benutzer oder anderen Agenten kommunizieren und vor allem von ihnen lernen können
- selbständig die für die Suchanfrage relevanten Internetdienste auswählen können
- die Suchanfrage mit der jeweils spezifischen Anfragesprache an die ausgewählten Internetdienste stellen können
- die Treffer der Suchergebnisse zumindest grob bewerten können (Informationsfilter)
- in der Lage sein, einfache Schlußfolgerungen aus Informationen zu ziehen
- auf unvorhergesehene Ereignisse (Website nicht erreichbar,...) oder plötzlich veränderte Bedingungen (Aktualisierung und/oder Veränderung des Internetangebots,...) in seiner Umgebung reagieren und entscheiden können, was dann zu tun ist
- die persönlichen Daten und Nutzerprofile des Anwenders wirklich schützen
- über ein bestimmtes Grundwissen verfügen, es effektiv nutzen können und in der Lage sein, sich weiteres Wissen anzueignen, lernfähig sein bzw. das Suchverhalten seines Benutzers analysieren und kopieren können
- ....

Das sind eine Menge an Anforderungen, die mit Sicherheit bisher von keinem Informationsagenten auch nur ansatzweise alle zufriedenstellend erfüllt werden. An den Ideen zur Entwicklung effektiver Informationsagenten wird zwar gearbeitet, wie eine ["Stand-der-Technik-Recherche" bei Depatist.net](http://www.depatist.net) ergab, aber von den wenigen Informationsagenten, die bisher auch realisiert wurden, konnte sich offensichtlich noch keiner wirklich am Markt durchsetzen. Dies ist aufgrund der vielfältigen Schwierigkeiten, mit denen ein Programm zur vollautomatischen Informationsbeschaffung konfrontiert ist, auch wenig verwunderlich.

Dies bestätigen auch die Projekterfahrungen einiger Agentenentwickler, wie beispielsweise Henry Liebermann, der Entwickler von Letizia oder die Entwickler von WebWatcher ([www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-6/web-agent/www/ijcai97.ps](http://www.cs.cmu.edu/afs/cs.cmu.edu/project/theo-6/web-agent/www/ijcai97.ps)).

**Mit welchen grundlegenden (mehr oder weniger lösbaren) Problemen müssen sich die Entwickler einer Maschine zur vollautomatischen Informationsbeschaffung auseinandersetzen ?**

- Problematik, daß Computer nicht in der Lage sind, Sprache zu verarbeiten ↓
- Das Grundproblem (oder vielmehr Glück), daß Computer nicht selbständig denken, lernen und kreativ sein können ↓
- Die Grundprobleme, mit denen das "Semantic Web" zu kämpfen hat:
  - die Problematik der Wissensteilung ↓
  - die Problematik der Wissensvermittlung ↓

Aus den genannten Gründen wird es noch viele Jahre dauern, bis das "Semantic Web" für die Menschheit im Sinne der Erfinder nutzbar wird. ↓

<b>Problematik, daß Computer nicht in der Lage sind, Sprache zu verarbeiten ↑</b>
---

Um die Teilaufgaben, die zur Informationsbeschaffung notwendig sind, zufriedenstellend zu erfüllen, **müßte die Maschine wie der Mensch die natürliche Sprache verstehen und verarbeiten können**; nur so könnte sie grundsätzlich auch [denken, lernen und kreativ sein](#).

Die Maschine benötigt zur Kommunikation mit anderen Maschinen neben Protokollen, die den Daten- bzw. Informationsaustausch regeln, vor allem eine gemeinsame Sprache, mit der sie tatsächlich in der Lage wäre, benötigte Informationen zu formulieren, die Antwort des anderen Agenten zu verstehen und für die weitere Recherche nutzbar zu machen. Die Maschine, die zwischen dem Informationssuchenden auf der einen Seite und dem Informationsanbieter auf der anderen Seite steht, müßte darüber hinaus mit dem Benutzer kommunizieren können, beispielsweise um Rückfragen zu stellen, und zwar in einer Sprache, die weit über unverständliche Code-Angaben hinausgeht.

Der Benutzer müßte der Maschine (die menschliche Sprache nicht versteht) eindeutig verständlich und weitestgehend unabhängig von der Art der Information mitteilen können, nach welchen Informationen er sucht. Der Agent müßte weiterhin die benötigten Informationen (die auch in verschiedenster Weise artikuliert werden, da es kein einheitliches Sprachsystem gibt) mit den jeweils relevanten Informationsquellen aus Millionen von möglichen (die nur teilweise und wenn, dann in häufig unterschiedlicher Weise mit unterschiedlicher Sprache beschrieben sind) in Zusammenhang bringen. Die universelle Ontologie, in der das menschliche Weltverständnis erfaßt ist, müßte alle vorhandenen Daten und Informationen, unabhängig von den verschiedenen Sprachen, Kulturen, Epochen und persönlichen Vorlieben präzise erfassen und eine semantische Verbindung zu den jeweils relevanten Quellen herstellen.

Daß dies weniger auf technischem Wege als vielmehr auf dem Wege der Bildung erreicht werden kann, ist einleuchtend - warum sollen von Menschen geschaffene Maschinen das erreichen, was ihre Schöpfer nicht schaffen ? **Das Problem kann im Grunde nur entschärft werden, wenn Menschen im Umgang mit ihrer Muttersprache und einer Weltsprache (Englisch) bestmöglichst geschult werden, also gut schreiben, lesen und sprechen können. Dies ist die Voraussetzung, um die Informationen von anderen zu verstehen und eigene Informationen klar zu formulieren.**

Grundsätzlich könnte aber der Zugriff auf Faktenwissen durch die Verwendung eines weltweit gültigen Systems, das die verschiedensten Themen-/Fachgebiete in einzelne Bereiche unterteilt, für jeden Einzelnen erleichtert werden.

Ein praktisches Beispiel, das zeigt, daß so etwas möglich ist, bietet die Internationale Patentklassifikation "IPC" (International Patent Classification), die Patentdokumente in über 60.000 technische Bereiche ordnet. Jedem Bereich wird dabei ein bestimmter, von der Maschine verständlicher Code, zugeordnet. Der Code "G06F 17/30" beispielsweise steht für den technischen Bereich "Digitale Rechen- oder Datenverarbeitungsanlagen oder -verfahren, besonders angepaßt an spezielle Funktionen - Wiederauffinden von Informationen; Struktur der Datenbasis dafür".

Damit ist es möglich, bestimmte technische Informationen durch die Eingabe des IPC-Codes sprachunabhängig zu finden. Aber auch bei der IPC gibt es das Problem, daß nicht jeder ein bestimmtes Patent in dem jeweiligen technischen Bereich vermuten würde, in den es vom zuständigen Patentanwalt eingeordnet wurde.



**Das Grundproblem (oder vielmehr Glück), daß Computer nicht selbständig denken, lernen und kreativ sein können ↑**

Wenn man sich die Ergebnisse der Automatisierungstechnik in der industriellen Produktion ansieht, könnte man meinen, es handle sich um hochintelligente Maschinen, die dem Menschen weit überlegen sind. In Wirklichkeit aber können diese Maschinen nur eine bestimmte Aufgabe ausführen und nur für exakt vorher bestimmbare Produktionsverfahren eingesetzt werden. Es muß sich also um stabile Abläufe handeln, denn die Maschine kann nur das tun, was ihr mit dem Programm bzw. den darin enthaltenen Anweisungen befohlen wird - sie ist nicht in der Lage, selbständig Probleme zu analysieren und mögliche Lösungen zu finden. Jede Reaktion auf ein Ereignis muß eigens peinlichst genau programmiert werden; die Maschine reagiert dann auf ein Ereignis genauso, wie es vorher vom Programmierer festgelegt wurde. Dem Programmierer müssen also sowohl die Ereignisse selbst als auch die bei deren Eintreten zu erfolgenden Reaktionen bekannt sein. Für Aufgaben mit nicht exakt vorhersehbaren Handlungsabläufen und vor allem für die selbständige Kommunikation in menschlicher Sprache sind Maschinen völlig unbrauchbar.

Daß die Maschine mit Hilfe der aus Fakten gebildeten Beziehungen und Regeln einfache logische Schlußfolgerungen ziehen kann (Stichwort: [Ontologien](#)), bedeutet noch lange nicht, daß die Maschine denken, lernen und kreative Lösungen finden kann.

Es gab in der Vergangenheit schon viele Versuche, den Maschinen menschliche Intelligenz einzuhauchen. Einer davon war, das Spezialwissen von Produktionsarbeitern in Expertensystemen abzubilden; da dies aber nicht zufriedenstellend gelang, ging die Softwareindustrie dazu über, Systeme zu entwickeln, die dem Menschen mögliche Lösungen für ein Problem präsentieren und deren Konsequenzen aufzeigen. Die Entscheidung darüber, welche Lösung letztendlich gewählt wird, muß dann der Mensch treffen. Diese analogie- und/oder fallbasierten Systeme beruhen auf Studien aus den Kognitionswissenschaften, die gezeigt haben, daß **Experten zum Finden von Lösungen oft auf frühere Fälle des betreffenden Themen-/Fachgebiets zurückgreifen (fallbasiertes Schließen). Wird darüber hinaus auch in anderen Wissensgebieten nach entsprechenden Fällen gesucht, so spricht man von analogiebasiertem Schließen.** Beide Möglichkeiten zum Finden von Lösungen für neue Probleme beruhen auf der gleichen Vorgehensweise. **Man sucht nach bereits erdachten Lösungen für ähnliche Probleme, die sich in der Vergangenheit gestellt haben. Je ähnlicher die gefundenen Probleme und damit auch die Lösungen sind, desto einfacher ist es dann im nächsten Schritt, die alte Lösung auf das neue Problem anzupassen.** Der Computer kann hierbei als wertvolles Hilfsmittel dienen, um die Fälle leicht wiederauffindbar und thematisch sortiert abzuspeichern. Die Maschine dient **also als maschinelle Gedächtnishilfe des Menschen und der Mensch selbst ist dazu da, kreative Lösungen für Probleme zu finden.** Dies ist auch der Basisgedanke meiner Lösung. Allerdings ist die Methode des analogie- bzw. fallbasierten Schließens auf vorhandene und brauchbare Fälle angewiesen.

**! Hinweis**

**Falls solche nicht verfügbar sind, kann übrigens das Internet sehr wertvolle Dienste leisten, um sie ausfindig zu machen,** beispielsweise durch die Suchanfrage "Fallbeispiel Problem" (das Problem muß durch geeignete Stichwörter näher beschrieben werden) an eine allgemeine Suchmaschine.

Die vorher beschriebene Methode des analogie- bzw. fallbasierten Schließens stützt sich auf konkret gemachte Erfahrungen, diese Lösungen haben also den Vorteil, daß sie, wenn sie entsprechend dokumentiert wurden, eindeutig nachvollziehbar sind und ihre Gültigkeit in einem bestimmten Rahmen beibehalten. **Bei Regeln werden oft nur angenommene, unrealistische Fälle abgedeckt, das bedeutet, sie können unterschiedlich interpretiert werden.** Die in den Ontologien zusammengestellten Regeln beruhen auf Fakten. Nur wenn alle Fakten vorhanden sind, kann die Maschine anhand der vorgegebenen Regeln logische Schlußfolgerungen richtig ziehen. Im Alltag ist es aber selten der Fall, daß alle für eine Entscheidung notwendigen Informationen verfügbar sind. Darüber hinaus wäre es in einer ganzen Reihe von Fällen aufgrund der Vielzahl an notwendigen Fakten vom Zeit- und Arbeitsaufwand her gar nicht möglich, alle diese Fakten zu verarbeiten und daraus dann die Entscheidung abzuleiten. Es darf zudem angezweifelt werden, daß es möglich ist, für alle Fälle entsprechende Regeln zu finden, da diese von einer Vielzahl von Fakten abhängen und deren Richtigkeit möglicherweise wiederum von vielen Umständen abhängt, die man auf die Schnelle gar nicht erfassen könnte.

**Das bedeutet, man muß zwangsläufig mit den unvollständigen und möglicherweise auch falschen Fakten versuchen, richtige Entscheidungen zu treffen. Dies gelingt umso besser, je mehr Erfahrung man auf dem jeweiligen Gebiet hat.** Der Mensch kann also vor allem aufgrund seiner Erfahrung mit Unsicherheiten umgehen. Ein guter Autofahrer reagiert in einer gefährlichen

Verkehrssituation auf Grund seiner Erfahrung so, daß ein Unfall vermieden wird, was einem Anfänger höchstwahrscheinlich nicht gelingen würde. Ein erfahrener Autofahrer handelt nicht mit seinem Bewußtsein, sondern unterbewußt. Das bedeutet, er muß für seine Handlung keine Fakten mehr in seinem Bewußtsein verarbeiten, denn das würde in einem Moment, in dem innerhalb von Bruchteilen von Sekunden entschieden werden muß, was zu tun ist, viel zu lange dauern, sondern zieht für das betreffende Problem bereits die fertige Lösung aus dem Unterbewußtsein. In der Regel kann die betreffende Handlung später nicht mehr genau rekonstruiert werden, das bedeutet, Erfahrungswissen kann nicht in formallogische Regeln gefaßt werden, **da Erfahrungen ganz eng an einen bestimmten Menschen gebunden sind - also subjektiv und nicht objektiv sind. Die Erfahrungen die ein Mensch deshalb macht, weil er von der Natur mit ganz einzigartigen Wesensmerkmalen ausgestattet wurde (die Maschine hat keine einzigartigen Wesensmerkmale und ist deshalb beliebig kopierbar) und in einem ganz bestimmten Umfeld lebt (die Maschine lebt nicht, sie macht nur genau das was der Mensch ihr beispielsweise durch ein Programm vorgegeben hat), sind es die den Menschen eindeutig als Subjekt und als Individuum kennzeichnen.**

**Am stärksten zeigen sich die Grenzen maschineller Intelligenz bei der Fähigkeit, neues Wissen zu schaffen. Grundsätzlich muß man dazu bekanntes Wissen in Frage stellen; dies kann man für sich alleine machen oder in dem man es anderen zur Diskussion stellt.** Da eine Maschine aber nicht selbständig denken kann, über keine eigenen Erfahrungen verfügt und sich damit auch keine eigene Meinung bilden kann, ist sie nicht fähig, Ideen zu entwickeln oder diese gar im Rahmen einer Diskussion mit anderen zu finden. Die Kreativität eines Menschen ist damit wohl der schwierigste Teil der Aufgabe, das menschliche Gehirn maschinell abzubilden. Das liegt daran, daß **die Kreativität unmittelbar an einen bestimmten Menschen gebunden ist**, ihn also eindeutig als Subjekt kennzeichnet - eine Maschine dagegen ist und bleibt ein Objekt.

**Nicht zuletzt hängt die Qualität der Ergebnisse des Denkens und Lernens und das Zustandekommen kreativer Gedanken wesentlich von der Einstellung des betreffenden Menschen ab.** Die innere Einstellung zu einer Aufgabe hat viel mit den Werten bzw. Gefühlen eines Menschen zu tun und die können grundsätzlich nicht programmiert werden. **Mit Herz und Verstand bei der Sache können eben nur Menschen mit ihren von der Natur ausgestatteten und einzigartigen Wesensmerkmalen sein und keine Maschinen. Nur Menschen die ihren Verstand mit dem Herzen einsetzen, können auf ihre [Intuition](#) zugreifen, die wie die Weltgeschichte zeigt, den Erfindern, Visionären und Künstlern als unverzichtbarer Wegweiser diente, um sie zu den für ihre Mitmenschen so wertvollen Denkergebnissen zu führen.**

Der Informationsagent, der im Dienste des Benutzers relevante Informationen herbeischafft, kann keinesfalls einen kompetenten menschlichen Rechercheur für die schnelle und unmittelbare Informationsbeschaffung ersetzen, ist aber ein sinnvoller Ansatz, um einen gewissen Zeit- und Arbeitsaufwand bei der Informationsbeschaffung einzusparen und es so auch weniger versierten Anwendern zu ermöglichen, das Internet als Informationsquelle zu verwenden. Vor allem in Unternehmen, bei denen es innerhalb der Organisation möglich ist, sich auf eine gemeinsame Kultur und Sprache zu einigen, könnten Agenten auch für die effektive Verteilung von Wissen an die Mitarbeiter genutzt werden.

Die Aufgaben, die mit dem höchsten Zeit- und Arbeitsaufwand verbunden sind:

- die Informationen auf ihre tatsächliche Relevanz und Qualität für die jeweilige Fragestellung zu bewerten,
- aus den Informationen das richtige Handeln ableiten, die Informationen also zu praktisch nutzbarem Wissen zu konstruieren
- das konstruierte Wissen in Frage zu stellen und so möglicherweise neues Wissen zu schaffen
- ...

müssen weiterhin vom Menschen selbst durchgeführt werden. Das wird hoffentlich auch so bleiben, denn Technologien sollen den Menschen helfen und nicht ihnen schaden.

Diesen Eindruck vermittelt der Anblick von hochtechnisierten Kriegen im Fernsehen zwar nicht, auf der andern Seite stehen aber auch Technologien wie der Helikopter beim Katastrophenschutz oder die Waschmaschine für den Haushalt, die dem Menschen wirklich wertvolle Dienste erweisen. Gerade am Beispiel des Helikopters kann man aber auch sehen, daß Technologien gleichzeitig zum Nutzen (Katastrophenschutz) und zum Schaden (Krieg) der Menschen eingesetzt werden können. Dies ist genau der Grund, warum die künstliche Intelligenz, mit der untersucht wird, wie menschlich intelligentes Verhalten von Computern erfaßt und nachvollzogen werden kann, ein sehr zweifelhaftes Wissenschaftsgebiet ist und bleiben wird, wobei das menschliche Denken wohl eher nicht auf

elektronischer Ebene (Transistoren, integrierte Schaltkreise,...) nachgebildet werden kann, sondern viel mehr direkt auf biologischer Ebene.

Was die Kombination an Erkenntnissen aus der Gehirnforschung, der Gentechnologie und der Informatik (Identifizierung von Schrift und Sprache, die Darstellung von Wissen und daraus abgeleiteten Schlußfolgerungen,...) bringen wird, kann man nur sehr schwer vorhersagen. Eines ist jedenfalls sicher: Die Menschheit benötigt bestimmt keinen beliebig programmierbaren und austauschbaren Menschen, um ihre Probleme zu lösen oder sich selbst gar überflüssig zu machen. Eine Maschine, die denken, lernen und kreativ sein kann, wäre nicht nur in der Lage, selbständig effektiv und sicher die Internetdienste im Sinne des Benutzers zu verwenden, sondern könnte ihn auch von beliebigen sonstigen Denkaufgaben befreien - und das 24 Stunden, 7 Tage die Woche und mit permanent hoher Konzentration. Dies wäre dann der Moment, wo Menschen ihre Kompetenz an Maschinen abgeben oder vielmehr verlieren würden.

Es ist zudem sehr zweifelhaft, warum die Kopie das schaffen soll, was das Original nicht schafft, oder anders formuliert: **Wie soll etwas Vollkommenes (die Schöpfung) von etwas Unvollkommenem (der einzelne Mensch) geschaffen werden ?**

Die wahre Vollkommenheit kann die Menschheit meiner Ansicht nach nur als Ganzes erreichen, wenn also jeder einzelne (oder zumindest ein sehr großer Teil) [seine Lebensaufgabe ausführt](#) und so konstruktiv mit seinen Mitmenschen zusammenarbeitet - letztendlich also jeder seine Stärken bestmöglich einbringt und dadurch die Schwächen des jeweils anderen kompensiert. Dieser Wohlfühl-Zustand ist denke ich das [Endziel](#) der Schöpfung.

**Der auf den Technologien des Semantic Web basierende Informationsagent besitzt keine Eigenintelligenz, sondern macht im Grunde die von den Autoren und Anwendern bereitgestellten Metadaten nutzbar, um seinem Benutzer die Informationsbeschaffung im unendlichen Web zu erleichtern. Dies ist auch der Grund, warum dieser Agent, wenn er richtig entwickelt wird, großen Nutzen stiften kann.**

Ein sehr wichtiger Aspekt, der ebenfalls gegen die vollautomatische Informationsbeschaffung spricht, ist die Tatsache, daß der Mensch durch die häufige oder gar ausschließliche Verwendung von Maschinen wichtige Kompetenzen abgibt. Nach dem Motto "Übung macht den Meister" wäre es besser, wenn jeder die Informationsbeschaffung, insbesondere die Teilaufgaben, wo man sucht und wie man die gefundenen Informationen/Informationsquellen hinsichtlich ihrer Qualität und Relevanz bewertet, weitestgehend selbst durchführt; so erweitert man nicht nur mit jeder Recherche die eigene Kompetenz, sondern verhindert zudem, daß man sich aufgrund fehlender eigener Kompetenz in die Abhängigkeit der Maschine stürzt. Man sollte sich dieses Argument vor allem im Hinblick auf die jetzige und vor allem künftige Bedeutung von Informationen überlegen. **Da Informationen die Grundbausteine sind, mit denen das Wissen im Kopf des Menschen konstruiert wird, kann durch eine rein maschinelle Informationsversorgung das Wissen und damit auch das Leben jedes Menschen (gezielt) beeinflußt werden.**

## Erkenntnis

Am besten kommen Menschen und Computer jeweils dort zum Einsatz, wo sie die besten Leistungen erbringen. **Der Mensch kann nachdenken, lernen und neues Wissen schaffen und der Computer dient dem Menschen als technische Grundlage für die zeit- und ortsunabhängige:**

-  Informationsbeschaffung,
-  Kommunikation,
-  Aus- und Weiterbildung

**und als Werkzeug dafür um zeit- und ortsunabhängig in den jeweiligen Online-Shops und virtuellen Marktplätzen Produkte und Dienste einzukaufen bzw. zu verkaufen.**

**Nicht zuletzt dient der Computer als maschinelle Gedächtnishilfe, in dem er es ermöglicht, Informationen leicht wiederauffindbar abzuspeichern und als Rechenmaschine, ohne die viele wertvolle wissenschaftliche Erkenntnisse, die auf umfangreichen Berechnungen beruhen, undenkbar wären.**

Außerdem kann der Computer immer wiederkehrende Aufgaben mit exakt gleichbleibendem Handlungsablauf, also Routinearbeiten wie beispielsweise Überwachungsrecherchen zuverlässig, präzise und kostengünstig anstelle des Menschen durchführen.

Gerade im Hinblick auf die Informationsbeschaffung können Maschinen dem Menschen weitere wertvolle Dienste leisten, in dem sie die formale Grundlage bereitstellen, um sehr komplexe Wissensstrukturen der unterschiedlichsten Themen-/Fachgebiete zu erstellen und zu speichern und damit das vorhandene Wissen der Menschheit für jeden Einzelnen besser als bisher nutzbar zu machen. Das so gespeicherte Wissen kann dann teilweise auch mit natürlichsprachlichen Fragen nach den betreffenden Antworten durchsucht werden. Die Bewertung der gefundenen Antworten kann durch Maschinen dadurch vereinfacht werden, daß sie die im Netz verteilten, von anderen bereitgestellten und für die Bewertung relevanten Fakten bei Bedarf zugänglich machen. **Den verantwortungsbewußten Umgang mit Technologien vorausgesetzt, wird die Zukunft eine geeignete Kombination aus menschlicher und maschineller Informationsbeschaffung sein.**

**Die Grundlage hierfür liefert das "Semantic Web", das sich aber weniger mit technischen, als vielmehr mit menschlichen Schwierigkeiten auseinandersetzen muß.**

Dies sind vor allem:

- die Problematik der Wissensteilung
- die Problematik der Wissensvermittlung

### Problematik der Wissensteilung ↑

Das Semantic Web baut auf verantwortungsvolle, technisch versierte Autoren, die ihre Dokumente gewissenhaft und technisch korrekt verfassen und mit Metadaten beschreiben. Nur bei den Dokumenten, bei denen die Maschine Metadaten vorfindet, kann sie diese auswerten. Weiterhin setzt das Semantic Web auf aktive Anwender, die durch das Annotieren von relevanten Informationen an die Ontologie anderen Anwendern wichtige Vorinformationen liefern.

**Wer aber das aktuelle Internet kennt, der weiß, daß die sinnvollen Grundgedanken, die hinter dem Semantic Web stehen, erst von vielen Internetnutzern erkannt und gelebt werden müssen.** Praktische Beispiele hierfür liefern die geringe Zahl an Internetdokumenten, die von Autoren mit Metadaten beschrieben werden und die große Zahl an Einwort-Suchanfragen. Daß grundsätzlich wenig Metadaten verfügbar sind und das Vertrauen in die korrekten Inhalte der verfügbaren Metadaten gering ist, beweist auch die Bedeutung, die den Metadaten von Suchmaschinen beigemessen wird. Während viele die Metatags völlig ignorieren, werden sie von einigen anderen zwar indiziert und als Bewertungskriterium beim Ranking herangezogen, aber nur wenige bieten auch die Möglichkeit, gezielt über Suchfelder nach bestimmten Angaben in den Metatags zu suchen. Ein praktisches Beispiel hierfür ist die allgemeine Suchmaschine Fireball, die darüber hinaus auch einen sogenannten Metatag-Generator bereitstellt, der es dem Autor erleichtert, seine Internetdokumente mit HTML-Metatags zu beschreiben.

Betrachtet man diese Tatsachen unter dem Aspekt, daß die Angabe von Metadaten mit Hilfe der seit langem bekannten HTML-Metatags wesentlich einfacher ist als die Beschreibung von Metadaten mit den (völlig neuen) Technologien des Semantic Web, so kann man folgende Erkenntnis ableiten: **Es müssen sich nicht nur die Technologien fortentwickeln, sondern auch die Menschen, die sie nutzen und das dauert in der Regel wesentlich länger.** Es werden noch viele Jahre vergehen, bis das "Semantic Web" von den Menschen tatsächlich verstanden und im Alltag gelebt wird.

**Die Angabe von Metadaten bedeutet vor allem für technisch weniger versierte Autoren einen enormen Mehraufwand für die Informationsaufbereitung und -bereitstellung.** Zum einen muß der Autor die nicht einfachen technischen Grundlagen lernen, zum anderen muß er sich genau überlegen, was er nun warum schreibt und dies dann auch noch in maschinenverständlicher Form angeben. Natürlich hängt es auch von der Art der Informationen ab, die mit Hilfe von Metadaten beschrieben werden sollen. Einfacher ist es beispielsweise, Bedienungsanleitungen (Kapitel mit Überschriften, Absätzen, Listen, Tabellen, technische Parameter, Anleitungen mit Arbeitsschritten, Formeln, Maßeinheiten,...) und Produktbeschreibungen (technische Daten, Eigenschaften, Produkthierarchie,...) maschinenlesbar zu beschreiben, als beispielsweise den Inhalt der hier vorliegenden Diskussion oder die Inhalte literarischer Werke, die häufig eine ganz subjektive Interpretation erlauben.

Weiterhin sollte man nicht vergessen, daß die Angabe von Metadaten auf freiwilliger Basis geschieht und daß viele wertvolle Inhalte von den jeweiligen Autoren sowieso schon mit viel Zeit- und

Arbeitsaufwand kostenlos im Internet veröffentlicht werden. Möglicherweise könnte dieses Problem durch entsprechende webbasierte Hilfen (Methoden, Werkzeuge) entschärft werden.

Gerade bei Metadaten, die Beschreibungen enthalten, kommt das bereits beschriebene Problem mit der Sprache ins Spiel - die Beschreibungen können unterschiedlich interpretiert werden. Hier könnte nur ein gemeinsamer Wortschatz Abhilfe schaffen oder entsprechende Werkzeuge. Für Dublin Core gibt es beispielsweise das Dublin-Core-Metadata-Template; das Formular definiert die Minimalanforderungen, aber auch gleichzeitig eine breite Palette von Standardeintragungen für alle Elemente und gegebenenfalls weitere Unterfelder.

Neben den grundsätzlichen Problemen, die sich für Autoren bei der Vergabe von Metadaten stellen, gibt es auch solche, die vermeidbar wären. Eines davon ist, daß Autoren **absichtlich falsche Angaben in den Metadaten machen, um so zu erreichen, daß ihre Seiten von der Suchmaschine als relevant erkannt werden und bei den Suchmaschinen ganz vorne auf der Rangliste stehen, obwohl deren Inhalt mit der Suchanfrage nichts oder nur sehr wenig zu tun hat.** Bei der Manipulation gehen manche Informationsanbieter so geschickt vor, daß sie weit vor den eigentlich relevanten Internetdokumenten im Suchergebnis der Suchmaschinen gelistet werden.

**Die Beschreibung der Dokumenteninhalte durch Metadaten bedeutet also auch einen Mehraufwand (zunächst durch Schulungen, später durch die Kontrollfunktion) für die Such- und Informationsdiensteanbieter, die dafür sorgen müssen, das die Dokumente bei Ablieferung alle Anforderungen erfüllen und sich einwandfrei in die Systemumgebung integrieren lassen.**

Da Metadaten heute und auch in weiterer Zukunft die einzige Chance sind, die Inhalte unstrukturierter Daten/Informationen wenigstens halbwegs zu bewerten, müssen die Bemühungen, diese Erkenntnis auch in das Bewußtsein der Web-Autoren zu bringen, verstärkt werden.

#### Problematik der Wissensvermittlung ↑

Der Suchende beschreibt mit seiner Suchanfrage, welche Information er sucht und der Autor kann mit Metadaten genau beschreiben, in welchem Zusammenhang er seine Information veröffentlicht. Die Maschine hat die Aufgabe, mit Hilfe der genannten Informationen den Suchenden mit dem richtigen Autor bzw. dessen Dokument zusammenzubringen. **Die Voraussetzung hierfür ist, daß der Suchende die von ihm gewünschte Information genau so beschreibt wie der Autor die von ihm angebotene, ansonsten wird es der Maschine genauso ergehen, wie es vielen Menschen, die miteinander kommunizieren, tagtäglich ergeht. Man versteht einen Sachverhalt, weiß aber nicht, wie man ihn so wiedergeben soll, damit ihn der andere in den richtigen Zusammenhang bringen und so verstehen kann.**

Damit aus einer neuen Information Wissen entsteht, muß diese vom jeweiligen Menschen mit seinen vorhandenen Informationen, Kenntnissen, Fähigkeiten und Erfahrungen verglichen, verknüpft und bewertet werden, so daß dieser aus der Information letztendlich sein Wissen konstruieren kann. Weil aber jeder über unterschiedliche Informationen, Kenntnisse, Fähigkeiten und Erfahrungen verfügt und damit anders denkt als der andere, **entsteht bei jedem Menschen aus ein und derselben Information unterschiedliches Wissen.** Wenn also ein Mensch sein Wissen an einen anderen weitergibt, wird daraus wieder eine Information, die vom jeweils anderen erst wieder mit seinen Informationen, Kenntnissen, Fähigkeiten und Erfahrungen zu seinem Wissen konstruiert werden muß. Noch wesentlich problematischer wird die Wissensvermittlung von einem zum anderen Menschen, wenn diese aus verschiedenen Kulturen (Religion, Feste, Traditionen,...) stammen.

Als Umkehrschluß könnte man nun sagen, daß die Wissensvermittlung dann am besten funktioniert, wenn die betreffenden Personen über möglichst gleiche Informationen, Kenntnisse, Fähigkeiten und Erfahrungen verfügen, aus der gleichen Kultur stammen und die gleiche Sprache sprechen. Das wiederum bedeutet, die Maschine müßte im Optimalfall zwei Menschen zusammenbringen, bei denen die vorher genannten Eigenschaften weitestgehend übereinstimmen und von denen der eine eine Information sucht und der andere über genau dieses Wissen verfügt.

**Eine Maschine, die das könnte, müßte über das gesamte Wissen (das auf Grund seiner Menge niemals maschinell gespeichert und vor allem niemals auf dem aktuellen Stand gehalten werden könnte) jedes einzelnen Menschen verfügen und in Windeseile das benötigte Teilwissen finden und so umsetzen, daß es der jeweils andere verstehen kann. Da dies den Menschen schon nicht gelingt und die Maschinen letztendlich eine mehr oder weniger schlechte Kopie menschlichen Wissens sind, wird es eine solche Maschine niemals geben.** Außerdem baut die Natur aus gutem Grunde (mit der Erfahrung aus Millionen von Jahren



menschheitlicher Entwicklungsgeschichte) auf eine schier unendliche Vielfalt - nicht einmal eineiige Zwillinge sind identisch, was ihren Körper und vor allem ihre Begabungen, Leidenschaften und Werte betrifft. Das bedeutet, auch wenn die Maschine das ganze Wissen erfassen könnte, wäre sie nicht in der Lage, zwei oder gar mehrere Personen zu finden, die über (möglichst) gleiche Informationen, Kenntnisse, Fähigkeiten und Erfahrungen verfügen, aus der gleichen Kultur stammen und die gleiche Sprache sprechen.

**Aus den genannten Gründen wird es noch viele Jahre dauern, bis das "Semantic Web" für die Menschheit im Sinne der Erfinder nutzbar wird. ↑**

Da aber das "Semantic Web" die Grundlage dafür ist, daß Maschinen relevantere Suchergebnisse liefern oder gar konkrete Antworten auf bestimmte Fragestellungen geben können, werden die bisher im World Wide Web eingesetzten Such- und Informationssysteme auch weiterhin mit den bisher verwendeten Technologien arbeiten und deshalb auch mit den damit verbundenen Schwierigkeiten konfrontiert sein. Auch ausgereifte Agentensysteme werden zumindest im World Wide Web erst dann effektiv zum Einsatz kommen können, wenn die Technologien des "Semantic Web" von den Menschen im Sinne der Erfinder verstanden und genutzt werden. Erst dann ist es denkbar, daß die Maschine selbständig die im Sinne des Benutzers relevanten und qualitativ hochwertigen Informationsquellen bzw. Informationen ausfindig machen kann, derzeit kann sie nur als Quellenvorschlagssystem genutzt werden - dies setzt aber eine gewisse Netzwerkkompetenz zur richtigen Bewertung der Quellen und vor allem der damit auffindbaren Informationen voraus.

Eine Möglichkeit wäre, mit geeigneten Lösungen die zu erwartende langsame Entwicklung des "Semantic Web" zu beschleunigen. So könnte beispielsweise eine globale Suchmaschine die von den Robots gefundenen Internetdokumente grob in das jeweilige Themengebiet nach ihrer Herkunft vorsortieren. Diese Dokumente könnten dann von entsprechenden Expertenredaktionen ausgewertet und auf Grundlage der Technologien des "Semantic Web" mit Metadaten beschrieben werden. Je nachdem kann auch der Autor des jeweiligen Dokuments miteinbezogen werden. Da dies aber mit einem gewaltigen Arbeits- und damit Kostenaufwand verbunden wäre, wird es wohl keine gangbare Lösung sein. Außerdem ist es fraglich, ob es sich finanziell lohnen würde, eine teure Expertenredaktion damit zu beschäftigen, eine riesige Müllhalde an Informationen zu durchpflügen, obwohl zahlreiche Informationsangebote (kostenpflichtige Fachdatenbanken, Online-Angebote von Fachzeitschriften, Regierungsdokumente,...) existieren, über die der Benutzer, der hochwertige Informationen benötigt, auf die von ihm gewünschte Information zugreifen kann. Sinnvoller wäre es deshalb, für alle (professionellen) Informationsanbieter (Ministerien, Behörden, Bibliotheken, Verlage, Tageszeitungen,...) entsprechende Anreize zu bieten, ihre Informationen vor der Veröffentlichung im Sinne des "Semantic Web" mit Metadaten zu beschreiben. Dies könnte allerdings dazu führen, daß die teilweise sehr interessanten Internetangebote von vielen Anbietern, beispielsweise Privatanwendern, und die bisher veröffentlichten Informationen (nicht mit Metadaten beschrieben) aus den vorderen Plätzen der Suchergebnisse ausgeschlossen werden. Auch muß man sich die Frage stellen, ob sich bei den Informationen, die nur eine sehr begrenzte Zeit gültig sind (dies dürfte wohl bei den meisten der Fall sein), der Zeit- und Kostenaufwand zum Erstellen der Metadaten überhaupt rechnen würde.

---

**Es ist also anzunehmen, daß uns die bisher eingesetzten Such- und Informationssysteme auch noch viele Jahre begleiten werden.** Die maschinelle Informationsbeschaffung mit Hilfe von Suchmaschinen und die damit verbundenen Probleme wurden bereits beschrieben; es gibt jedoch noch viele weitere Möglichkeiten, die das Internet zur Informationsbeschaffung bietet.

**Nachfolgend sollen nun einige dieser Möglichkeiten und die jeweils damit verbundenen wesentlichen Probleme beschrieben werden:**

 **von bezahlten Experten betreute Suchkataloge, wie beispielsweise Yahoo**

Nachdem die Website vom betreffenden Informationsanbieter beim Katalog angemeldet wurde, wird diese von einem bezahlten Experten bewertet und, falls dieser die Website als nützlich einstuft, in die seiner Meinung entsprechende(n) Kategorie(n) einsortiert. Die zum jeweiligen Thema in Frage kommenden Websites können dann über die entsprechenden Haupt- und Unterkategorien des Katalogs abgerufen werden. Durch die Suche über Kategorien ist es für den Anwender, der wenig Vorwissen zur gewünschten Information besitzt, leichter, diese zu finden als mit Suchmaschinen, bei denen relevante Suchbegriffe (Stich- und/oder Schlagwörter) bekannt sein müssen.

**Hier stellen sich folgende wesentliche Probleme:**



- Die relevanten Internetseiten werden nicht immer in den Kategorien vermutet, in die sie von den Experten eingeordnet wurden.
- Die Kategorien können durch viele Verweise auf entsprechende Internetangebote unübersichtlich werden.
- Die Anzahl der in einem Suchkatalog verzeichneten Webseiten ist beschränkt, da sie nicht wie bei Suchmaschinen automatisch, sondern von Redakteuren manuell bewertet und einsortiert werden müssen.
- Die Aktualität von hierarchischen Suchsystemen läßt häufig zu wünschen übrig. Da hierarchische Suchsysteme redaktionell betreut werden, kann es einige Zeit dauern, bis dort angemeldete Webseiten auch in das Verzeichnis eingeordnet sind.
- Häufig werden Webseiten bei hierarchischen Suchsystemen zwar an-, aber nur selten auch wieder abgemeldet. Die Zahl der toten Hyperlinks (Fehlermeldung: Seite nicht gefunden) ist leider recht hoch.
- ...

---

### ● **von Anwendern betreute Suchkataloge, wie beispielsweise das Open Directory Project (ODP)**

Das Ganze funktioniert ähnlich wie bei den von Experten betreuten Suchkatalogen, allerdings mit dem Unterschied, daß hier die Websites nicht von bezahlten Experten bewertet und in die jeweiligen Kategorien einsortiert werden, sondern von freiwilligen Redakteuren, die den Grundgedanken des Internets "Wissen weitergeben und vom Wissen des jeweils anderen profitieren" aktiv leben.

Das ODP sollte ursprünglich die Probleme redaktionell betreuter Verzeichnisse wie Yahoo (lange Wartezeiten bei der Aufnahme neuer Websites, hohe Anzahl toter Links,...) lösen, hat letztendlich aber mit den gleichen Problemen zu kämpfen. Durch die große Zahl an Redakteuren, die mit dem unterschiedlichsten Vorwissen (Experte, Fachinteressierter,...) an der Katalogisierung im jeweiligen Themen-/Fachgebiet mitwirken, ist die Qualität der in einer Kategorie einsortierten Internetangebote recht unterschiedlich. Ähnliches gilt für die jeweils vom Redakteur gewählte Kategorie. Im Gegensatz zu Suchkatalogen wie Yahoo kann beim Open Directory aber jeder Benutzer per E-Mail mit dem jeweils zuständigen Redakteur in Kontakt treten und über die betreffende Katalogisierungsentscheidung diskutieren. Außerdem gibt es keinen kostenpflichtigen Eintragungsservice, mit dem die Wartezeit zur Aufnahme der Website in den Katalog zwar wesentlich beschleunigt wird, aber hierfür nicht die inhaltliche Qualität der Website zählt, sondern die finanziellen Möglichkeiten des Internetanbieters.

---

### ● **spezielle Suchdienste für bestimmte Internetdienste (Webringe, Newslettern, Webforen, Diskussionslisten,...)**

Diese Suchdienste stellen in der Regel eine Datenbank und ein redaktionelles Verzeichnis für den jeweiligen Internetdienst zur Verfügung. Der Diensteanbieter zum gewünschten Thema kann dann entweder durch die Eingabe entsprechender Suchbegriffe über eine Suchmaschine von dieser Datenbank abgerufen werden oder man hangelt sich über das Verzeichnis (Suchkatalog) von einer allgemeinen Kategorie zu spezielleren Themenbereichen bis hin zum gewünschten Diensteanbieter vor. Für jedes indexierte/verzeichnete Diensteanbieter gibt es eine kurze Beschreibung (Titel, Inhalt, ...) und einen Link zum entsprechenden Anbieter.

Bei Verwendung eines qualitativ hochwertigen, speziellen Suchdienstes könnte man sich die Bewertung der dort gefundenen Diensteanbieter wesentlich vereinfachen, wenn dieser nur solche in sein Angebot aufnehmen würde, die gewissen Qualitätskriterien genügen.

#### **Hier stellen sich folgende wesentliche Probleme:**

- Die Qualität der speziellen Suchdienste schwankt stark, je nach dem, ob der Anbieter die betreffenden Internetangebote gewissenhaft überprüft und bewertet oder nicht. Während die Qualität einiger spezieller Suchdienste, beispielsweise für Weiterbildungsangebote dank tatkräftiger Unterstützung des Staates vergleichsweise hoch ist, gibt es etwa für Communities oder Webforen keine oder nur qualitativ schlechte Spezialsuchdienste.
- Die Anbieter von speziellen Suchdiensten könnten mehr dafür tun, daß ihre Datenbanken und Kataloge genutzt werden. Sie sollten es dem Benutzer mit leicht auffindbaren

Vorinformationen möglichst einfach machen, zu erkennen, ob er die gewünschte Information in der jeweiligen Datenbank oder im betreffenden Katalog in der gewünschten Qualität finden kann oder nicht. Bisher muß das häufig über die mühsame Trial & Error-Methode (Versuch & Irrtum-Methode) herausgefunden werden.

- Es gibt keine allgemein anerkannten Prüfungsstandards.

- ...

### **Hinweis**

Eine weiterführende Beschreibung des Problems und wie es gelöst werden könnte finden Sie [hier](#).

#### **• von Experten betreute Themenportale, wie beispielsweise Meome.de oder das nicht mehr existierende Clickfish.com**

Der Benutzer kann auf der Hauptseite des Informationsangebots das für die von ihm benötigte Information relevante Themenportal auswählen. Jedes dieser Portale wird von einem Experten betreut und enthält von ihm erstellte Artikel und kommentierte Links. Weiterhin stehen die Experten dem Anwender für spezielle Fragen als persönlicher Ansprechpartner entweder per E-Mail oder in einem Webforum zur Verfügung.

Wie das relativ schnelle Verschwinden von Clickfish.com und die starke Reduzierung der Aktivitäten von Meome.de (gehört mittlerweile zu Freenet.de) gezeigt haben, scheinen diese Themenportale bei den Anwendern keinen allzu großen Anklang zu finden.

#### **Hier stellen sich folgende wesentliche Probleme:**

- Die Experten arbeiten meist unentgeltlich oder für einen geringen Lohn, so daß gerade für hochkarätige Experten der Anreiz fehlt, ein solches Portal langfristig mit hoher Qualität zu betreuen.

- Bei Themenportalen, bei denen der Anwender nicht in die Gestaltung miteinbezogen wird, fehlt für viele Nutzer ein wichtiger Anreiz, dieses Portal zu nutzen.

- ...

#### **• virtuelle Gemeinschaften, wie beispielsweise das Expertennetzwerk "Wer-weiss-was.de" oder die virtuelle Verbrauchergemeinschaft "Ciao.de"**

Grundsätzlich dienen virtuelle Gemeinschaften dazu, Menschen mit gemeinsamen Interessen eine Möglichkeit bereitzustellen, weitestgehend zeit- und ortsunabhängig in Kontakt zu treten, beispielsweise um Fragen zu einem bestimmten Thema zu stellen bzw. diese zu beantworten, um über unklare Sachverhalte zu diskutieren oder sich gegenseitig zu beraten. Das hat den Vorteil, daß Fragen in gewohnter Weise in Satzform gestellt werden können, nicht beschränkt auf wenige Stichworte wie in Suchmaschinen. Der antwortende Experte, bei dem es sich um einen staatlich anerkannten Fachmann genauso handeln kann wie um einen Fachinteressierten ohne Qualifikationsnachweis, steht damit dem Anwender als persönlicher Ansprechpartner zur Verfügung. In der Regel ist der Dienst für den Anwender nach der Registrierung beim Community-Anbieter kostenlos. Es gilt das Prinzip der gegenseitigen Hilfe, nämlich eigenes Wissen zur Verfügung zu stellen und gleichzeitig vom Know how des anderen zu profitieren. Je nach Anzahl und Aktivität der Mitglieder kann so eine Community eine umfassende Informationsquelle zu einem bestimmten Themengebiet darstellen. Vor allem für Produkte bieten große Communities nicht nur eine Möglichkeit zur gegenseitigen Kaufberatung, sondern auch zur kollektiven Bewertung.

Die Qualität der Antworten wird durch ein Bewertungssystem sichergestellt, die Ergebnisse der Bewertung sind für den Anwender jederzeit abrufbar. Außerdem kann der Anwender sich über die berufliche Laufbahn und sonstige Daten des Experten informieren, wodurch dessen Identität, bei einer entsprechenden Registrierung nachweisbar ist.

#### **Hier stellen sich folgende wesentliche Probleme:**

- Community-Anbieter können, schwer nachvollziehbar sowohl für das Mitglied selbst als auch für rechtliche Instanzen, gegen die Datenschutzgesetze verstoßen.

- Die kollektive Bewertung von Informationen, Produkten und Diensten von einer großen

Masse an Anwendern hilft dem Einzelnen bei der Bewertung von deren Qualität und deren Relevanz - **allerdings nur bei beliebten Themen und Fragestellungen.**

- Die Bewertungssysteme haben einige gravierende Mängel:
  - Die Bewertungsergebnisse können durch Mitglieder mit falschen Identitäten bewusst manipuliert werden. Das bedeutet, ein und dieselbe Person kann sich unter verschiedenen Namen bei der Community anmelden. Theoretisch könnte das Mitglied, das sich mehrere Identitäten zugelegt hat, Fragen an sich selbst stellen, diese mit vorgefertigten Antworten selbst beantworten und sich dann auch noch selbst bewerten. Weiterhin könnte es die Antwort eines Experten mehrfach gut oder schlecht bewerten.
  - Man kann sich nur schwer ein Bild von der Glaubwürdigkeit des Bewerbers machen. Experten können zusammenarbeiten und sich gegenseitig hochwerten oder, noch schlimmer, andere unbeliebte Nebenbuhler schlecht bewerten.
- Bei großen Gemeinschaften kann es zu Problemen mit der Informationsflut und dem Erhalt des Gemeinschaftsgefühls kommen.

● ...

● ...

**All die genannten und auch die übrigen im Lehrgang näher beschriebenen Möglichkeiten zur Informationsbeschaffung haben ihre Vor- und Nachteile. Die Lösung müßte demnach eine Mischung aus den vorhandenen Möglichkeiten sein, die alle Vorteile bietet und die Nachteile weitestgehend ausschließt.**

An diesem Kunststück versucht sich Google. Die allgemeine Suchmaschine arbeitet mit dem Suchkatalog Yahoo zusammen: Liefert eine Suchanfrage bei der Suchmaschine des Katalogs keine Einträge, so wird sie automatisch an Google weitergereicht. Im Grunde ist das Page-Rank-System ein kollektives Bewertungssystem, an dem die Informationsanbieter teilnehmen. Für das Ranking der Suchergebnisse bewertet Google auch, ob das jeweilige Internetangebot beim Open Directory Project eingetragen ist. Je nach Suchanfrage findet man auch relevante Beiträge aus dem Experten-Netzwerk "Wer-weiss-was.de" oder bei Produktanfragen werden relevante Erfahrungsberichte aus virtuellen Verbrauchergemeinschaften unter den ersten zwanzig Treffern angezeigt.

**Wie aber bereits in der "Einführung" beschrieben wurde, erfüllt Google keine der Aufgaben, die sich bei der Informationsbeschaffung stellen, befriedigend, geschweige denn die Aufgaben, die mit der Informationsbeschaffung zusätzlich verbunden sind, wie etwa die Kommunikation, beispielsweise mit Informationsanbietern oder anderen Anwendern, der Ein-/Verkauf von Produkten und Dienstleistungen und die webbasierte Aus- bzw. Weiterbildung.** Wie ebenfalls bereits beschrieben wurde, stehen aber die einzelnen Teilaufgaben der Informationsbeschaffung wie in einem System in Wechselwirkung zueinander. Wird also eine Teilaufgabe nicht zufriedenstellend gelöst, so hat dies auch eine negative Auswirkung auf andere Teilaufgaben. Eine ungenaue Suchanfrage wird mit hoher Wahrscheinlichkeit zu einem schlechten Suchergebnis führen. Da man nicht genau weiß, was man eigentlich sucht, kann man die Relevanz der einzelnen Treffer des Suchergebnisses nur mit großer Mühe und sehr ungenau bewerten.

**Ferner sollte man die ebenfalls bereits genannten Risiken, daß eine konkurrenzlose globale Suchmaschine als globale Zensurmaschine und als Manipulationsinstrument mißbraucht werden kann, nicht vergessen. Darüber hinaus fehlt es bei allen Internetdiensten an Möglichkeiten für eine wirklich individuelle Informationsbeschaffung.** Es gibt zwar Personalisierungsangebote, aber diese gelten in der Regel nur für die Nutzung eines bestimmten Internetdienstes oder es bestehen ernsthafte Zweifel darüber, daß der Anbieter die personenbezogenen Daten und das Nutzerprofil nicht an Dritte weiterverkauft.

Da also (zum Glück) kein einzelner Internetdienst für sich alleine die Möglichkeit bietet, alle der genannten Aufgaben der Informationsbeschaffung und die möglicherweise damit verbundenen Aufgaben zufriedenstellend zu lösen, müssen dem Benutzer die jeweils für seinen persönlichen Bedarf nützlichen Dienste der einzelnen Internetangebote aus dem gesamten Angebot einfach und gezielt abrufbar zusammengestellt werden, so daß er diese in einer geeigneter Kombination für den gerade anstehenden Bedarf nutzen kann. Dazu muß dem Anwender im Bedarfsfall eine Anleitung zur Verfügung stehen, die es ihm ermöglicht, die jeweils benötigten Internetquellen effektiv und sicher zu nutzen. Diese soll vor allem dazu dienen, sich eine grundlegende Netzwerkkompetenz anzueignen und ihm darüber hinaus eine maschinelle Gedächtnishilfe sein, die ihm die jeweils benötigte Information, beispielsweise eine bestimmte Methode, die Kurzanleitung zur Bedienung eines bestimmten Werkzeugs oder die Beschreibung eines bestimmten Internetdienstes, leicht auffindbar bereitstellt. Die Anleitung muß es dem Anwender ermöglichen, selbst zu entscheiden, welche Internetquellen er für

seinen Bedarf nutzen möchte und welche Quellen und Informationen er als tatsächlich relevant und qualitativ hochwertig beurteilt.

Die vorher genannten Anforderungen erfüllt **meine Lösung**, bei der das Denken, Lernen und das Finden von kreativen Lösungen für Probleme die Aufgaben des Menschen sind und die Maschine als technische Grundlage und Werkzeug dient, die ihn bei bestimmten Teilaufgaben der Informationsbeschaffung unterstützt.

## Erkenntnis

Im Grunde wird die von mir entwickelte Lösung nicht nur derzeit, sondern auch langfristig die beste Lösung sein. **Langfristig nicht nur, weil es elektronisch nahezu unmöglich scheint, eine selbstdenkende Maschine zu entwickeln (wenn, dann wohl eher auf biologischer Ebene), sondern vor allem, weil selbstdenkende Menschen eine solche Maschine gar nicht benötigen.** Ein Mensch mit einer grundlegenden Netzwerkkompetenz, seinem kontextbezogenen Nachschlagewerk für den Lehrgang und seiner individuellen Internetquellensammlung wird einer Maschine in seinem Wirkungsbereich immer überlegen sein. Wer kompetent an Netzwerken teilnehmen kann und eine grundlegende Ahnung von den Möglichkeiten der Programmierung hat, der weiß, daß es keine Maschine geben wird, die ihn ersetzen kann. Weiterhin weiß er, daß er mit jeder praktischen Anwendung besser wird und seine Kompetenz erweitert.

**Vielmehr wird es darauf ankommen, das wir die technischen Möglichkeiten des Internets für die konstruktive Zusammenarbeit mit anderen nutzen, um unser eigenes Wissen mit Hilfe der Erfahrungen und Erkenntnisse unserer Mitmenschen fortlaufend zu erweitern und gemeinsam wichtige Probleme zu lösen. Zum Lösen von Problemen bedarf es keiner Maschinen, sondern Menschen, die den Grundsatz "Wissen geben und nehmen" verstanden haben und aktiv leben, an gemeinsame Ziele und Visionen glauben und sich mit Herz und Verstand dafür einsetzen können.**

Durch die geschickte Kombination aus menschlicher Intelligenz und der richtigen Anwendung der für die jeweilige Aufgabenstellung geeigneten Methoden und Werkzeuge, die Grundlagen meiner Lösung sind, wird eine äußerst flexible, weitestgehend systemunabhängige Nutzung der Internetdienste möglich und bietet zudem größtmöglichen Schutz für die persönlichen Daten und vor Manipulationen, etwa daß bestimmte Informationen gefiltert werden. Das System baut wie das Internet selbst auf die konstruktive Zusammenarbeit zwischen Menschen und nutzt Technologien nur dort, wo sie sinnvoll sind.

**Die Realisierung einer Maschine zur vollautomatischen Informationsbeschaffung wäre trotz der angesprochenen Schwierigkeiten eine sinnvolle Erweiterung meiner Lösung.**

## Kurzbeschreibung

### Die Idee:

Im Rahmen des Open-Source-Projekts soll ein Programm zur vollautomatischen Informationsbeschaffung entstehen, das anstelle des Benutzers bestimmte Informationen auffindet.

### Praxisnahe Beschreibung der Idee:

**Was muß ein Programm für die vollautomatische Informationsbeschaffung grundsätzlich können ?**

Nachdem der Benutzer der Maschine mitgeteilt hat, welche Informationen er benötigt, führt diese die Informationsbeschaffung weitestgehend selbständig an seiner Stelle durch. Das Programm greift dabei auf einen umfassenden Bestand an Informationsquellen (individuelle Internetquellensammlung, lokale auf dem Rechner gespeicherte Informationen/Quellen, externe auf einem Internet-Server gespeicherte Informationen/Quellen) zu. Das Programm muß in der Lage sein, selbständig zu entscheiden, welche Informationsquellen (Datenbanken, Kataloge, Websites, Datenarchive von Newslettern oder Diskussionsforen,...) die jeweils gewünschte Information am ehesten liefern können und die betreffende Quelle dann auch dazu verwenden, um Antworten auf bestimmte Fragestellungen zu finden. Vor allem aber muß die Maschine in der Lage sein, die Relevanz und Qualität der gefundenen

Information/Informationsquellen im Sinne des Benutzers zu bewerten. Am Ende der Recherche listet die Maschine die gefundenen Treffer nach Relevanz und Herkunft geordnet im Suchergebnis auf und ist darüber hinaus in der Lage, grob zu beschreiben, warum sie den Treffer für relevant bewertet hat.

Im Grunde sollte die Maschine die Internetdienste genauso effektiv und sicher nutzen können wie ein kompetenter Anwender mit Hilfe seiner allgemeinen und individuellen Internetquellensammlung. Da dieses Programm, wie ein Agent, eine Aufgabe im Auftrag eines anderen erledigt, wird es auch als Informationsagent bezeichnet. Dieses Programm arbeitet eng mit seinem Benutzer zusammen, das bedeutet, es lernt von ihm, stellt Rückfragen, macht ihm Vorschläge und versucht, möglichst alles so zu machen, wie er es machen würde.

---

## Nutzen

Da das Internet bzw. die Menge der dort veröffentlichten Informationen mit hoher Geschwindigkeit weiterwächst, wird die Bedeutung von Maschinen, die den Menschen beim Auffinden von bestimmten Informationen unterstützen, ebenfalls zunehmen. Wie die kritische Betrachtung der aktuell erfolgreichsten Suchmaschine "Google" gezeigt hat, erfüllt diese keine der Teilaufgaben zur Informationsbeschaffung wirklich zufriedenstellend und einige gar nicht. Die mit der Volltextsuche und universellen Bewertungsmethoden arbeitenden Suchmaschinen können das Problem der Informationsflut offensichtlich noch nicht optimal lösen, sondern verstärken es in vielen Fällen.

Eine erfolgversprechende Alternative bieten Informationsagenten, die auf der Basis der "Semantic Web"-Technologien arbeiten. Besonders interessant wäre ein funktionierender Informationsagent für weniger versierte Anwender, die sich die benötigten Informationen dann vom Agenten beschaffen lassen können, und für die automatische Informationsbeschaffung innerhalb eines Unternehmens.

Der Informationsagent könnte beispielsweise die Mitarbeiter automatisch mit den jeweils für sie relevanten Informationen (ausgewählt aufgrund eines individuellen Informationsbedarfsprofils) versorgen. Damit der Agent allerdings effektiv eingesetzt werden kann, bedarf es einer gemeinsamen Unternehmenssprache und -kultur (Stichwort: Wissensmanagement).

Das eingesetzte System sollte auf den Technologien des "Semantic Web" basieren. Dies hätte den großen Vorteil, daß die maschinelle Informationsbeschaffung aus Quellen des lokalen Intranets, wenn das Semantic Web entsprechend weit entwickelt ist, einfach von diesen auf die relevanten Quellen des Internet ausgedehnt werden könnte. Gegebenenfalls könnte der Agent dann auch die individuelle Internetquellensammlung automatisch zusammenstellen und aktualisieren.

---

## Für die Realisierung der Maschine zur vollautomatischen Informationsbeschaffung müssen folgende Grundbedingungen erfüllt sein:

- Der Agent muß auf den Technologien des Semantic Web basieren. Die Intelligenz muß von den Menschen erbracht werden, die den Maschinen die Inhalte ihrer Informationen verständlich und auswertbar machen.
- Der Benutzer kann genau vorgeben, in welchen Quellen der Agent nach den benötigten Informationen suchen und nach welchen Bewertungskriterien er die gefundenen beurteilen soll.
- Die personenbezogenen Daten und das Nutzerprofil des Benutzers sind ausreichend geschützt (Verschlüsselung und Signierung,...).

### Hinweis

Da der Quellcode bei Open-Source-Projekten allen zugänglich ist, können alle interessierten und fachkundigen Anwender den Code im Detail nach möglichen Sicherheitslücken überprüfen. Außerdem kann so bestmöglichst verhindert werden, daß der Agent als (globale) Zensurmaschine mißbraucht werden kann.

- Die Funktions- und Arbeitsweise des Agenten ist für den Benutzer so transparent wie möglich.

**Meinem Konzeptvorschlag entgegen** steht ein hinter verschlossenen Türen programmierter Informationsagent, der ohne Mitbestimmung des Benutzers selbständig nach Informationen im "Semantic Web" sucht. Das bedeutet, der Agent entscheidet, wo gesucht wird und welche



Informationen relevant und qualitativ hochwertig sind. Das birgt nicht nur die Gefahr, daß der Agent als effektive Zensurmaschine mißbraucht werden kann, sondern wäre ein Schritt mehr in eine Welt, in der Menschen ihre eigentlich vorhandenen Kompetenzen an Maschinen abgeben. Die Welt braucht aber, um ihre Probleme zu lösen, keine Maschinen, sondern selbstbestimmte und selbstorganisierte Menschen.

---

## Konkurrenzprodukte und Basistechnologien für den Informationsagenten

Bevor man ein eigenes Produkt mühsam entwickelt, ist es natürlich sinnvoll zu überprüfen ob es nicht schon ein fertiges Produkt, das den grundsätzlichen Anforderungen genügt, zu mieten bzw. zu kaufen gibt. Möglicherweise kann dieses Produkt dann als Basis für den Informationsagenten dienen und durch entsprechende Verbesserungen und Ergänzungen an die eigenen Vorstellungen angepaßt werden.

Bei einer Recherche (Stand: Dezember 2004), nach aktuell auf dem Markt befindlichen Informationsagenten, konnte ich keinen Agenten finden, der sich am Markt wirklich durchsetzen konnte. Auch eine Recherche nach patentrechtlich geschützten Ideen zur Umsetzung eines Informationsagenten lieferte keine brauchbaren Ergebnisse. Bereits vorhandene Technologien zur Informationsbeschaffung (Natural Language Search, Phonetische Analyse, Visualisierung des Hypertextgeflechts,...) sollten soweit wie möglich in den Agenten integriert und genutzt werden. Einige interessante Technologien, die zur Realisierung des Informationsagenten verwendet werden könnten, finden Sie [hier](#) und bei der [Beschreibung des Suchagenten "Bingoo"](#).

Das Agentenkonzept soll so ausgelegt werden, daß neue Technologien ergänzend integriert und veraltete einfach entfernt werden können (Stichwort:Module). Außerdem sollte auf Schnittstellen- und Syntaxänderungen der Suchdienstanbieter lokal ohne großen Aufwand reagiert werden können.

Die Auswahl und Entscheidung für bestimmte Technologien wurde noch nicht getroffen. Das bedeutet, daß bisher auch keine entsprechenden Spezifikationen oder Dokumentationen von Software- und Hardware-Herstellern oder anderen Organisationen, die wichtige, für den Informationsagenten relevante technische Standards entwickelt haben, vorliegen.

Dies ist jedoch kein grundsätzliches Problem, da die Basistechnologie des Informationsagenten auf den technischen Standards des "Semantic Web" beruht und damit die grundsätzliche Entwicklung des Agenten nicht von noch unbekanntem Technologien abhängig ist.

---

## Wichtige Teilaufgaben (Funktionen) des Informationsagenten und wie mein Lehrgang bei der Erfüllung der Teilaufgaben helfen könnte

Für die erfolgreiche Suche muß die Maschine (der Informationsagent):

**vor der Suche wissen:**

- [was sie genau sucht](#)
- [wo sie sucht](#)
- [wie sie mit der gewählten Informationsquelle sucht](#)

**während und nach der Suche wissen:**

- [wie sie auf Fehler reagieren kann](#) , die während der Suche auftreten können
- [wie sie die Relevanz und Qualität der gefundenen Informationen/Informationsquellen richtig bewertet](#)

...dem Benutzer eine einfache Möglichkeit bieten, die relevanten Informationen/Informationsquellen so abzuspeichern, daß er sie [später leicht wieder auffinden](#) kann

...für eine sichere Nutzung der Internetdienste geeignete Schutzmaßnahmen treffen, um [den Benutzer vor den Sicherheitsrisiken zu schützen](#), die mit dem Anschluß an das weltweite Netzwerk "Internet" verbunden sind

Die vorher erwähnten Fähigkeiten gehören zum Grundlagenwissen; reicht dieses aber nicht mehr aus,



so muß die Maschine bzw. der Agent in der Lage sein, [sich selbständig weiteres Wissen anzueignen](#), das notwendig ist, um die gewünschte Information zu beschaffen.

### Wie können meine Produkte und Dienste bei der Lösung der Teilaufgaben helfen ?

Der Lehrgang bzw. die darin formulierte Wissensstruktur, insbesondere aber die dort beschriebenen Suchmethoden können als Basis für eine themen-/fachübergreifende Ontologie dienen, mit deren Hilfe ein Informationsagent, der auf den Technologien des Semantic Web basiert, vorhandene Internetquellen selbständig oder wie vom Benutzer vorgegeben auswählen und für die Recherche nutzen kann.

Die im Lehrgang beschriebenen Bewertungsmethoden und -kriterien können ebenfalls als Basis für eine themen-/fachübergreifende Ontologie dienen, mit deren Hilfe der Agent neue Internetquellen suchen und selbständig oder wie vom Benutzer vorgegeben hinsichtlich ihrer Relevanz und Qualität bewerten kann. Für den Benutzer, der die Inhalte des Lehrgangs kennt, hat dies den Vorteil, daß er grundsätzlich weiß, was die Maschine tut.

Als optimale Startpunkte für tiefergehende Recherchen, die der Agent anstelle seines Benutzer durchführt, dienen vielfältige und qualitativ hochwertige spezielle Such- und Bewertungsdienste, die von vielen unterschiedlichen, voneinander unabhängigen Anbietern (und letztendlich auch Anwendern) betrieben werden (siehe meine [Existenzgründungsidee](#)). Mit Hilfe [gemeinsamer Web- Standards](#) und den für den effektiven Datenaustausch geeigneten, offenen Schnittstellen (Stichwort "XML"), könnte der Informationsagent auf eine qualitativ hochwertige und mit Metadaten (Stichwort: [Semantic Web](#)) beschriebene Auswahl von Internetquellen zugreifen.

Darüber hinaus dient der in den Agenten integrierte Lehrgang dem Benutzer als Berater, beispielsweise in dem der Agent nützliche Quellen für einen bestimmten Bedarf vorschlägt und als maschinelle Gedächtnishilfe. Die Vorschläge bezüglich relevanter Internetquellen könnten dabei direkt aus meinem Lehrgang entnommen werden.

Meine Produkte und Dienste liefern darüber hinaus weitere wichtige Erfahrungen und Erkenntnisse, die in die Entwicklung des Informationsagenten mit einfließen können.

---

### Layout der Programmoberfläche

Es steht noch nicht fest wie die Programmoberfläche zur Bedienung des Programms aussehen soll. Diese soll erst nach und nach während der Programmierung des Informationsagenten entwickelt werden.

---

**Bisher ist das grob beschriebene Konzept eines Informationsagenten nur eine Idee, die aber wenn die entsprechenden Voraussetzungen (freiwillige und kompetente Helfer) erfüllt sind, jederzeit in die Praxis umgesetzt werden kann.**

## Erkenntnis

**Viele Grenzen, die auch in Zukunft weder durch wissenschaftliche Erkenntnisse noch durch die damit entwickelten Technologien überschritten werden können, setzt der Mensch selbst.**

Nach der Entwicklung der beiden Lehrgängen "Effektive, sichere und individuelle Nutzung der Internetdienste" und "Sichere Nutzung von IT-Systemen mit Internetzugang" kam ich zu der wichtigen Erkenntnis, daß bestimmte Probleme eine fortlaufende Begleiterscheinung der modernen Informations- und Kommunikationstechnologien bleiben werden.

Die modernen Informations- und Kommunikationstechnologien können trotz rasanter technischer Weiterentwicklungen:

- **nicht garantieren, daß der Anwender immer die für ihn relevanten Informationen in gleicher Qualität und vollständig erhält**

- Nicht jeder Mensch erfüllt immer mit der gleichen Einstellung seine Aufgabe; während der eine mit Herz und Verstand bei der Sache ist, ist es für den anderen gerade eine

zwingende Notwendigkeit.

Ein möglicher Grund hierfür kann sein, daß Menschen im Alltag sehr erfreuliche Dinge (Kind gesund geboren, Lottogewinn,...) und weniger erfreuliche Dinge (Frau gestorben, Autounfall,...) erleben, die sie möglicherweise für eine gewisse Zeit sehr stark beschäftigen und sie deshalb nicht ihre volle Konzentration auf die gerade anstehenden Aufgaben lenken können.

- In vielen Fällen ist es aufgrund der Vielzahl an notwendigen Informationen vom Zeit- und Arbeitsaufwand gar nicht möglich, alle vorhandenen Fakten vollständig zu beschaffen und deren sachliche Richtigkeit zu überprüfen, geschweige denn, diese zu verarbeiten und aus dem Denkergebnis dann die Entscheidung abzuleiten.

Das bedeutet, man muß zwangsläufig mit den unvollständigen und möglicherweise auch falschen Fakten versuchen, richtige Entscheidungen zu treffen. Dies gelingt umso besser, je mehr Erfahrung man auf dem jeweiligen Gebiet hat.

- Jeder Mensch verfügt über andere Vorinformationen, Vorkenntnisse, Fähigkeiten und Erfahrungen und betrachtet deshalb ein und dieselbe Information aus einem anderen Blickwinkel; deshalb konstruiert jeder Mensch aus ein und derselben Information ein anderes Wissen als der jeweils andere.

● ...

- **nicht sicherstellen, daß die vom Anwender/Anbieter gespeicherten oder über das Internet übertragenen Daten/Informationen 100 %-ig sicher sind**

- Die **Programmierfehler der Hersteller** von sicherheitsrelevanter Hard- und Software und **Installations-, Konfigurations- und Bedienungsfehler der Anwender** können nicht ausgeschlossen werden.

- Der **Wirksamkeit von Sicherheitstechnologien** sind grundsätzliche **technische Grenzen** gesetzt.

- Der **Wirksamkeit von Gesetzen** sind ebenfalls Grenzen gesetzt, so kann beispielsweise die Einhaltung des **staatlichen Verbraucherschutz- und des Bundesdatenschutzgesetzes** nicht hundertprozentig sichergestellt werden und strenge Datenschutzgesetze im eigenen Land können die Daten nicht schützen, wenn sie in anderen Länder transferiert und auf den entsprechenden Servern gespeichert und weiterverarbeitet werden.

- Menschliche Schwächen (etwa Bestechlichkeit) können nicht gänzlich ausgeschlossen werden.

● ...

Ein analoges Beispiel hierfür sind die maschinellen Transportmittel (Auto, Motorrad,...); auch hier wurden zahlreiche Schutzmaßnahmen ergriffen:

- technische wie der Airbag oder das Antiblockiersystem
- gesetzliche wie das vorgeschriebene Mindestalter, die Straßenverkehrsordnung oder das Ablegen einer Führerscheinprüfung

Trotzdem gibt es jedes Jahr zahlreiche Tote und Schwerverletzte. Die Gründe hierfür sind fast immer menschliche Schwächen wie Unachtsamkeit, Rücksichtslosigkeit oder gar Unwissenheit oder auch schlicht Unfähigkeit.

Die genannten und viele weiteren Probleme können zwar durch geeignete Technologien und Gesetze entschärft, aber niemals vollständig aus der Welt geschafft werden. Die Probleme können allerdings auf ein Minimum reduziert werden, wenn die Entwickler und Anwender die innere Bereitschaft besitzen ihre Kompetenzen fortlaufend zu erweitern und wenn beide zusammen konstruktiv an der Entwicklung hochwertiger Technologien und Gesetze arbeiten.

Man sollte dabei bedenken, daß die vielfältigen und komplexen Technologien und teilweise abstrakten Gesetze mit denen wir es im Alltag zu tun haben, sehr konkrete Auswirkungen auf unser Leben

haben. So kann beispielsweise die übertriebene Überwachung und Kontrolle der Informations- und Kommunikationsaktivitäten der Bürger bzw. des Mitarbeiters durch entsprechende von der Regierung erlassenen Gesetze bzw. von Unternehmen eingesetzte Überwachungstechnologien die Bürger/den Mitarbeiter demotivieren sich aktiv an der Informationsgesellschaft zu beteiligen und damit dem Allgemeinwohl/der Wettbewerbsfähigkeit der Unternehmen schweren Schaden zufügen. Bürger/Mitarbeiter die sich hier im Vorfeld, also bevor die Gesetze erlassen/die Technologien eingesetzt werden, kritisch mit den Risiken befassen und entsprechend entgegensteuern, können hier aktiv zur Problemvermeidung beitragen.

**An der Verbesserung bestehender oder Entwicklung neuer Technologien und Gesetze kann jeder Einzelne aktiv mitwirken, indem er ihren tatsächlichen Nutzen für sich und andere hinterfragt** und die festgestellten Mängel beispielsweise anderen zur Diskussion stellt. Hierbei gefundene Lösungen oder berechtigte Kritiken können dann gemeinsam an die Entwickler der Technologien bzw. den Gesetzgeber weitergeleitet werden. Um die Motivation der Anwender, sich Gedanken über die angewendeten Technologien/Gesetze zu machen, aufrechtzuerhalten, müssen sich die Entwickler/Gesetzgeber auch mit einer von den Anwendern spürbaren Ernsthaftigkeit mit ihren Lösungsvorschlägen und Kritiken auseinandersetzen.

**Damit man jedoch ganz allgemein Probleme an aktuellen Technologien und Gesetzen erkennen, formulieren und lösen kann, muß man sich die Zeit nehmen, das dazu notwendige Wissen fortlaufend zu erweitern. Außerdem müssen Experten/Fachinteressierte unterschiedlicher Fachgebiete einen gemeinsamen Sprachstil (mit so wenig Fachwörtern wie möglich) finden, um die Probleme/Lösungen so zu formulieren, daß sie von allen Beteiligten verstanden werden können - das bedeutet, der Techniker formuliert seine technischen Probleme/Lösungen so, daß sie auch vom Juristen verstanden werden können und umgekehrt.**

In diesen Zusammenhang paßt auch die Erkenntnis von **Konrad Zuse** :

"Das Problem ist doch nicht, daß Maschinen eines Tages wie Menschen denken könnten, sondern daß Menschen wie Maschinen denken."

Menschen, die wie Maschinen denken:

- sind sehr vergangenheitsbezogen,
- nehmen ihre Umgebung gezielt so wahr, daß sie es möglichst bequem haben und wälzen dabei so viel Verantwortung wie möglich auf ihre Mitmenschen ab,
- möchten gerne Routinetätigkeiten ausführen, Tätigkeiten also, deren Handlungsablauf übersichtlich und voraussehbar ist,
- ...

Auch noch so hoch entwickelte Maschinen können nicht **ihre Werte kritisch hinterfragen und bei Bedarf ändern**, aus eigenem Antrieb heraus **Probleme erkennen**, selbständig denken und lernen oder **kreative Lösungen für Probleme finden** und möglicherweise im Rahmen dieser Lösungen ihre bisherige Einstellung zu wichtigen Fragen des Lebens ändern.

Genau das aber können Menschen, wenn sie sich von ihren Ängsten und Zweifeln befreien, an sich und ihre Fähigkeiten glauben und **innerlich bereit sind, sich fortlaufend weiterzuentwickeln**.

Menschen, die nicht wie Maschinen denken, sind also entwicklungsfähig, hören auf ihre innere Stimme, nutzen ihren Verstand mit ihrem Herzen, um alltagsrelevante Probleme und Aufgaben effektiv zu lösen und lassen Routinetätigkeiten, soweit es möglich ist, von Maschinen ausführen, es sei denn, es handelt sich um Routinetätigkeiten, die wirklich das Herz erfreuen.

© Martin Glogger

[www.onlinetechniker.de](http://www.onlinetechniker.de)

Der Artikel ist ein Teil aus meinem Lehrgang "Netzwerkcompetenz - wie ?" - Effektive, sichere und individuelle Nutzung der Internetdienste (mehr Infos unter "[www.onlinetechniker.de/?q=node/9](http://www.onlinetechniker.de/?q=node/9)").