

Der Weg zu einem effektiven Wissensmanagement

1 Zusammenfassung

Wissen können wir nur teilen, wenn wir eine gemeinsame Sprache sprechen. Diese Wahrheit gilt auch für das Wissensmanagement. Hier soll das gesamte Wissen einer Organisation inhaltlich richtig vernetzt von Menschen intuitiv erfragt und von Maschinen ausgewertet werden können. Davon sind wir nach dem Stand der Technik jedoch weit entfernt. Im Artikel werden drei wichtige Voraussetzungen diskutiert, um kollektives Wissen effektiv technisch verwalten zu können. Die Sprache logic lingua wird vorgestellt, die der natürlichen Sprache strukturell verwandt ist und die Wissen umfassend kommunizieren kann. Dabei ist sie gleichermaßen Übermittlungsprotokoll, Abfragesprache und Programmiersprache.

2 Was ist Wissen?

Wissen soll die Bedeutung oder die Semantik¹ allen relevanten Seins und Geschehens der realen Welt erfassen. Menschen und Maschinen müssen dieses Wissen verstehen, weiterverarbeiten und zur Planung neuer - die reale Welt ihrerseits verändernder - Maßnahmen nutzen können (Bild 1). Dazu brauchen wir eine Sprache zur Dateneingabe, Datenausgabe und Abfrage.

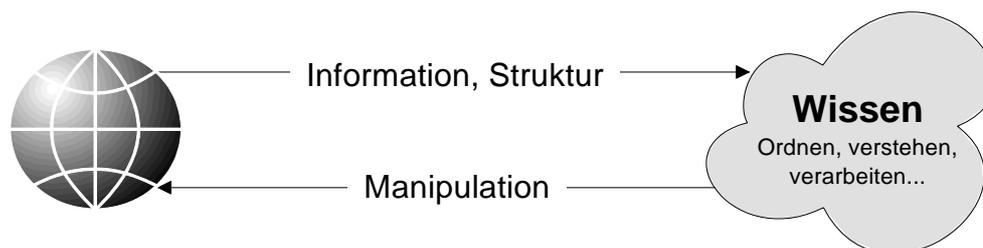


Bild 1

Uns begegnen gegenwärtig zwei Formen des kollektiven Wissens (Bild 2):

- Das mit einer natürlichen Sprache, d.h. mit Literatur kommunizierte Kollektivwissen der Menschen,
- das mit technisch-semantischen Sprachen und in Datenbanken kommunizierte Wissen in Computern.



Bild 2

In beiden Fällen ist das Wissen semantisch gespeichert. Das bedeutet, dass eine Sprache strukturell verknüpfte Objekte und Prozesse der abgebildeten Welt in entsprechend strukturell verknüpfte Begriffe und Syntax der Abbildungssprache umsetzt. Nur natürliche oder technisch-semantische Sprachen können Aussagen über die Welt machen.

¹ Im Sinne des inhaltlichen und strukturellen "Wesens" der Welt.

3 Drei Voraussetzungen für ein technisches Wissensnetzwerk

Ein technisches Wissensnetzwerk muss die folgenden Voraussetzungen erfüllen, wenn es effektiv sein will (Bild 3).

1. Das Netzwerk muss über eine möglichst vollständige Wissensbasis verfügen.
2. Zur Repräsentation und Kommunikation muss eine einzige, durchgängige semantische Sprache verwendet werden.
3. Jeder beteiligte Computer muss eine mit der semantischen Sprache korrespondierende geeignete Datenorganisation haben.

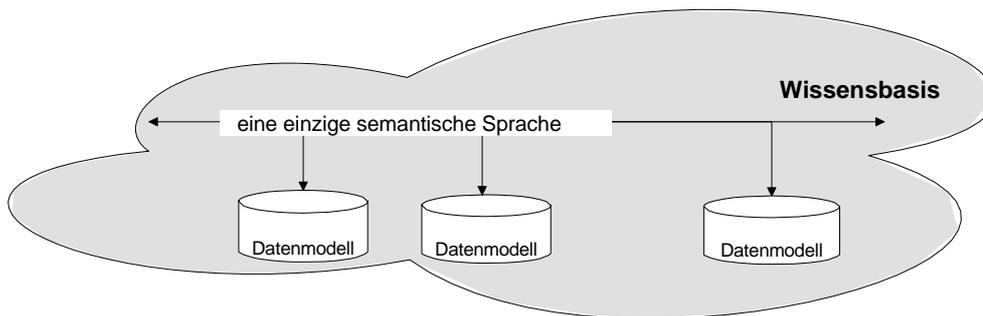


Bild 3

Im Jahr 2001 forderte T. Berners-Lee im viel zitierten Artikel "Semantic Web" (Scientific American [BLHL01]) sinngemäß dasselbe:

Das Web ist die Wissensbasis. RDF (Resource Description Framework) und OWL (Web Ontology Language) sind zwei semantische Sprachen. Softwarehersteller schaffen eine heterogene Datenorganisation.

Im folgenden werden diese Voraussetzungen diskutiert:

3.1 Wissensbasis

Wissen ist umfassend. Es geht dabei um alles Wissen einer Person oder einer Organisation (Bild 4). Deshalb müssen in einem Wissensnetzwerk Menschen und Computer Zugang zu allem relevanten Wissen haben, damit neues Wissen generierbar wird. Synergien und Zufallstreffer entstehen dann dort, wo Wissensbereiche unterschiedlicher Herkunft sich kreuzen.

Wissen muss die Information aller Fachbereiche und Datenbanken in der Gegenwart, für die Zukunftsplanung und im Rückgriff auf die Historie erfassen. Ganz praktisch gehören z.B. Inhalte der folgenden Bereiche dazu:

- Warenwirtschaft, Vertriebs- und Entwicklungsanwendungen, etc.
- Intranet, Firmendatenbanken, Projekt- und Planungsdaten,
- Internetausschnitte und Groupware

Das Wissensmanagement in einem Unternehmen hat die generelle Gestaltungsaufgabe, alles verfügbare organisierte und nicht organisierte Wissen logisch vernetzt bereitzustellen und auf verständliche Art zugänglich zu machen.

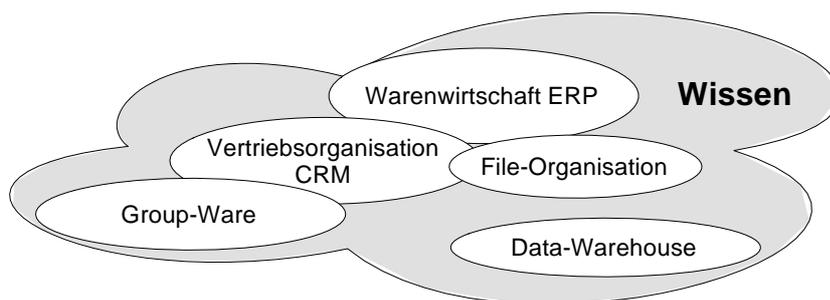


Bild 4

Eine große Wissensbasis ist nach dem Stand der Technik zwar elektronisch verfügbar, aber sie liegt in unterschiedlichen Formaten in viele Wissensausschnitte zerteilt vor, d.h. verteilt auf die Warenwirtschaftssoftware, auf Vertriebs- und Entwicklungssysteme oder z.B. eine Wissensdatenbank. Dadurch ist die Wissensbasis nicht durchgängig ansprechbar.

3.2 Eine einzige semantische Sprache

Wissen können wir nur teilen, wenn wir eine gemeinsame Sprache sprechen. Deshalb muss alles Wissen logisch organisiert sein und - durch eine einzige Sprache verbunden - abrufbar sein. Eine einzige semantische Sprache ist aber heute nicht verfügbar. Dafür wird Wissen in einer Unzahl von Softwareanwendungen, Sprachen und Protokollen kommuniziert.

Menschen kommunizieren Wissen mit der natürlichen Sprache. Das Erfassen und Verarbeiten von Wissen hat damit in jeder Wissenschaft mit viel Erfolg stattgefunden. Dabei dienten die semantischen Natursprachen aller Welt als gutes Speicherungs- und Kommunikationssystem. Zur Wissensrepräsentation und -kommunikation kommen hier Grammatik² und Wortschatz zur Anwendung.

Auch eine technisch-semantische Sprache muss ähnlich aufgebaut sein. Eine technische Sprache darf man nicht einfach erfinden. Sie ergibt sich vielmehr aus der Natur der abzubildenden Welt. Zuerst muss man die Struktur der realen Welt analysieren. Danach wird ein Abbildungssystem in Form einer Sprache - basierend auf diesen Strukturen - erschaffen. Nur wenn eine durchgängige, strukturelle Äquivalenz von abgebildeter Welt und semantischer Sprache erreicht wird, wird diese schnell und effektiv funktionieren (Bild 5).

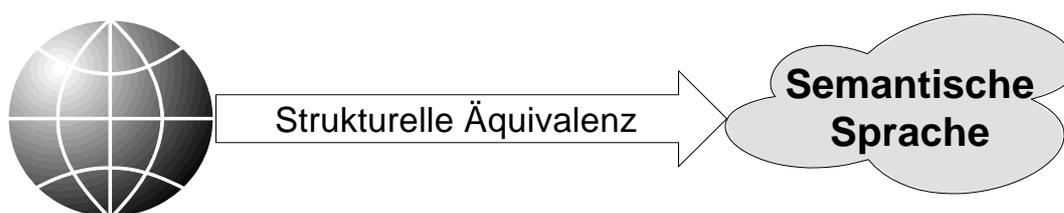


Bild 5

Betrachtet man z.B. Sterne, Produkte, oder Lebewesen, so haben wir es immer mit substantieller Materie zu tun. Dies sind Dinge, Stoffe, Menschen, Tiere und Pflanzen. Im sprachlichen Raum werden sie durch Worte der Wortart Substantiv charakterisiert.

Darüber hinaus kann man Zustandsänderungen oder Prozesse beobachten. Auch Sie haben ein sprachliches Äquivalent: das Verb. In jeden Prozess sind typischerweise Objekte mit hinein verwickelt. Im Ereignis: "Ein Mensch landete 1969 mit einer Rakete auf dem Mond" sind dies die beteilig-

² Die Grammatiken der natürlichen Sprachen gleichen sich in den Grundstrukturen.

ten Substantive Mensch, Rakete, Mond. Sie müssen in der strukturellen Sprachabbildung dieses Prozesses durch einen Satzbauplan mit dem Verb "landen" verknüpft werden. Aus diesem Beispiel soll klar werden, dass Sprachstrukturen sich direkt aus realen Prozessstrukturen ergeben und nicht frei gewählt werden dürfen.

Die Tabelle zeigt unvollständig, ohne ins Detail zu gehen, weitere Entsprechungen aus Welt und Sprache (Bild 6).

beobachtete Objekte der Welt	korrespondierende Wortarten der semantischen Sprache
Dinge, Stoffe, Menschen, Tiere, Pflanzen..	Substantiv
Eigenschaften	Adjektiv
Zustandsänderungen / Prozesse	Verben und dazugehörige Satzbaupläne
Raumangaben	Adverbiale des Raums
Zeitangaben	Adverbiale der Zeit
Angaben der Art und Weise	Modal Adverb

Bild 6

Diesen Strukturen werden die Grammatiken der natürlichen Sprachen gut gerecht. So weisen alle natürlichen Sprachen mindestens die in voriger Tabelle aufgeführten Wortarten Substantiv, Verb, Adjektiv und Adverb auf [HP02], welche somit auch für eine technisch-semantische Sprache erforderlich sind.

Darüber hinaus bieten die natürlichen Sprachen zur Verkürzung eines Textes Vereinfachungen und Abstrahierungen an, bei welchen die Äquivalenz zu den abgebildeten Vorgängen nicht mehr direkt sichtbar ist. Eine technische Sprache darf diese Vereinfachungen nicht unverstanden übernehmen.

Die technisch-semantische Sprache muss neben Aussagen in Gegenwart, Vergangenheit und Zukunft auch Fragen und Befehle ausdrücken können, d.h. sie muss technisch gesehen auch als Abfragesprache und Programmiersprache taugen.

Die Terminologie und Struktur einer technisch-semantischen Sprache muss einem Abiturient (nicht nur einem Programmierer) verständlich sein. Sonst wird der Anwender die Sprache nicht verwenden. Die grammatische Terminologie der Natursprachen wie z.B. Substantiv, Verb, Adjektiv, etc. müssen deshalb auch in einer technisch-semantischen Sprache Anwendung finden.

Nach dem Stand der Technik dienen heute RDF und OWL als semantische Sprachen. Die Sprachlandschaft von RDF und OWL ist jedoch leider zerteilt und hat eine ungeeignete Grammatik:

1. Die Sprache wird zuerst einmal in sechs Sprachen unterteilt:
 - a. für assertionales Wissen oder Wissen zu Aussagen (RDF),
 - b. für terminologisches Wissen (RDFS, OWL FULL, OWL DL oder OWL lite),
 - c. für Abfragen (SPARQL).
2. Der Anwender muss sich mit drei unklaren Wortarten begnügen.

3. RDF kennt keinen "substantivischen Verband". So muss der Anwender zu Hilfskonstruktionen greifen [HKRS08].
4. Die RDF "Grammatik" bietet nur einen einzigen Satzbauplan aus Subjekt, Prädikat und Objekt an, während die Welt Prozesse kennt, die eben mehrere Objekte fordern.

Zusammenfassend kann man schließen: Die Sprachen RDF und OWL sind komplex, nicht für Laien und komplizierter als die natürliche Sprache. Die Sprachen sind nicht durchgängig für alle Wissensgebiete einsetzbar. "Durch diese dezentralisierte Vorgehensweise wird eingestanden, dass es unmöglich ist, ein einzelnes Schema zu entwickeln, das für alle Gebrauchsmöglichkeiten passend wäre" [WIKI01].

Aus grammatischer und semantischer Sicht ist es nicht möglich, mit diesem Angebot Wissen befriedigend darzustellen. Hier finden wir nicht die geforderte "einzige semantische Sprache". Trotz der Mängel wurden inzwischen aber wertvolle Inhalte als "Vokabularien" erstellt. Diese sind nicht verloren und können künftig in andere leistungsstärkere semantische Sprachen importiert werden.

3.3 Dynamische Datenorganisation

Neben der Kommunikation muss in einem EDV System die Informationsspeicherung durch eine geeignete, inhaltsunabhängige Datenorganisation vorgegeben sein.

Die Datenorganisation in herkömmlichen Datenbanken ist normalerweise inhaltsabhängig, d. h. jeder Prozess wird in einem manuell gestalteten Tabellensystem gespeichert. Für die Abspeicherung von Sätzen einer semantischen Sprache wird jedoch ein dynamisches Datenmodell benötigt, welches beliebige Prozesse abbilden kann, ohne dass seine Struktur manuell geändert werden muss.

Nach dem Stand der Technik erlauben unterschiedlich strukturierte Datenbankmodelle mit vielen 100.000 Tabellen keinen durchgängigen Zugriff auf vorhandenes Firmenwissen. In Handarbeit erarbeitete Schnittstellen kosten viel und helfen wenig.

4 Implementierung von logic DATA Base und logic lingua

Bei logic DATA GmbH, Unterägeri wird an einem System aus einer dynamischen Datenbank "logic DATA Base" und der technisch-semantischen Sprache "logic lingua" gearbeitet.

Mit diesem System kann die verfügbare Wissensbasis unterschiedlicher Wissensquellen und Datenbankinhalte unterschiedlicher Software-Anwendungsbereiche konvertiert, eingelesen und in einem Format zur Verfügung gestellt werden. Somit ist die erste Voraussetzung erfüllt.

logic lingua arbeitet mit einer der natürlichen Sprachgrammatik ähnlichen Grammatik. logic lingua ist dabei gleichermaßen Übermittlungsprotokoll, Abfragesprache und Programmiersprache. logic lingua wird vom Menschen und von der Maschine verstanden. Ein Anwender kann damit intuitiv Fragestellungen und transparenten Einblick in komplexe Inhalte unterschiedlicher Datenquellen erlangen. Damit ist die zweite Voraussetzung erfüllt.

Die Wissensrepräsentation im Datenbank-Modell logic DATA Base ist inhaltsunabhängig. Sie lässt sich in die Wissensrepräsentation von logic lingua maschinell konvertieren und umgekehrt. Das Datenbankmodell erlaubt dabei die Zusammenlegung und Teilung von Datenbankinhalten. Damit ist die dritte Voraussetzung erfüllt.

Das Design von logic lingua und logic DATA Base entspricht also den oben aufgestellten Forderungen. Damit ist nun doch ein Schema vorhanden, das für alle Gebrauchsmöglichkeiten passt und eine enorme Zugänglichkeit und Datentransparenz gewährt. Weitere Details finden sich in [HR02].

Prototypen wurden implementiert und zeigen vielversprechende Ergebnisse.

Mit logic lingua kann auch neue Software auf der Datenstruktur von logic DATA Base geschrieben werden. Damit entfallen jegliche Konvertierungen und Wissen unterschiedlicher Herkunft kann durchgängig mit logic lingua navigiert werden.

Bild 7 gibt einen Überblick über die Systemarchitektur.

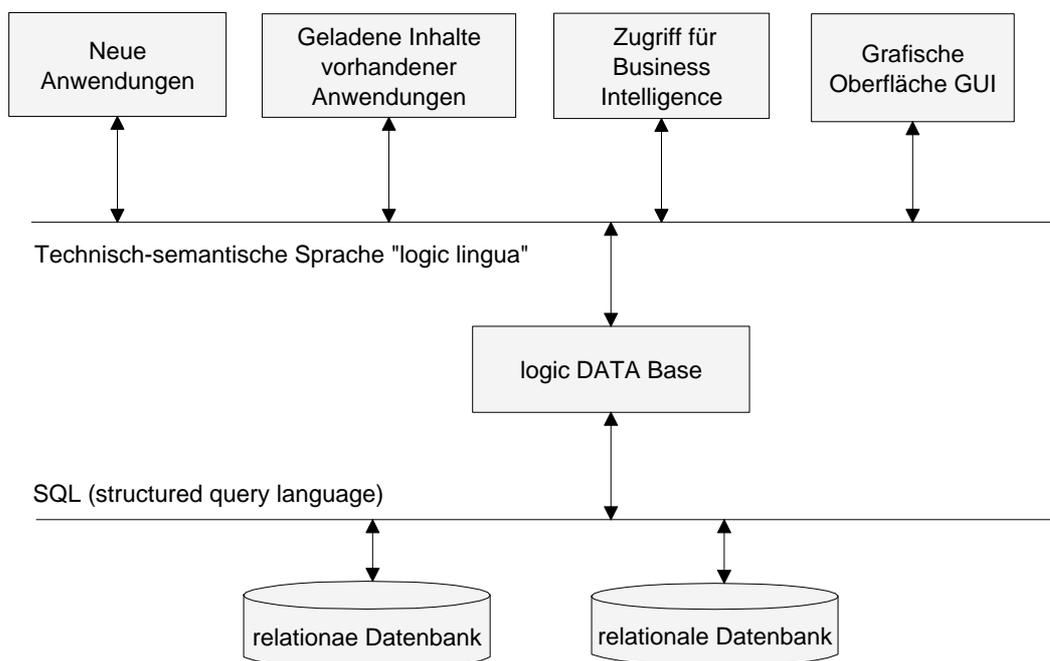


Bild 7

Literatur

[BLHL01] T. Berners-Lee, J. Hendler und O. Lassila. The semantic Web. Scientific American, S. 96-101, May 2001.

[HKRS08] P. Hitzler, M. Krötzsch, S. Rudolph, Y. Sure. Semantic WebB, S.56 Einführung von Hilfsknoten, Berlin, Springer Verlag, 2008.

[HP02] R. Huddleston und G. Pullum. The Cambridge Grammar of the English Language, S. 22 Lexical categories, Cambridge, Cambridge University Press, 2002.

[HR02] H. Rustige. Knowledge Integration Server. In Hanning, U. (publisher): Knowledge Management und Business Intelligence; Berlin, Springer Verlag, 2002.

[WIKI01] Wikipedia, Semantisches Web. Absatz: RDF-Schema, Dez. 2009, http://de.wikipedia.org/wiki/Semantisches_Web