

## Wissensmanagement- ein neuer Hype der Informatiker?

H. Maurer, Technische Universität Graz / Österreich

hmaurer@iicm.edu

www.iicm.edu/maurer

### Kurzfassung

In dieser Arbeit wird an Hand von zwei konkreten Beispielen erläutert, was Wissensmanagement (WM) ist, und warum WM eine so wichtige Rolle in der Zukunft spielen wird. Dies gilt freilich nur wenn es sich bei WM um mehr handelt, als die Verwendung von Informationssystemen oder verteilten Datenbanken für anspruchsvolle Aufgaben. Durch die Vorstellung des Maurer-Tochtermann Modells für WM und die Erläuterung der Grundzüge neuer Techniken, die weit über traditionelle Ansätze hinausgehen wird anschließend belegt, dass WM tatsächliche etwas grundlegend Neues ist. Das in Europa entwickelte System Hyperwave wird als vielleicht führendes Beispiel dieser neuen Art von Software, d.h. als Beispiel für ein WM-System abschließend kurz erwähnt.

### 1. Einleitung

Seit einiger Zeit wird zunehmend das Schlagwort „Wissensmanagement“ (kurz WM, engl. Knowledge Management) im Zusammenhang mit der effizienten Vermittlung strukturierter Information in großen Organisationen und der Verwendung heterogener Wissensquellen verwendet. Da Informationssysteme, Datenbanken, und vernetzte Systeme wie Internet oder Intranets ähnliche Zielrichtungen haben, stellt sich die Frage, ob man WM wirklich als etwas Neues betrachten kann, oder ob es sich hier nur um „alten Wein in neuen Flaschen handelt.“ In dieser Arbeit wird argumentiert, dass WM im Kern neue Ziele verfolgt, und bereits erste bemerkenswerte Konzepte und Techniken entstanden sind. Der Begriff WM ist vage, weil selbst der Begriff „Wissen“, obwohl wir alle wissen (!), was er bedeutet, schwierig zu fassen ist. Ein Beispiel auch für die sprachliche Verwirrung zwischen Begriffen wie Daten, Information und Wissen möge genügen.

Auf die Frage: „Wann wurde Erzherzog Johann von Österreich deutscher Reichsverweser?“ ist die Antwort „1848“, also ein ganz einfaches Datenstück. Und doch, wenn man dieses in Relation setzt zu „Erzherzog Johann“ und „deutscher Reichsverweser“ dann wissen (!) wir plötzlich etwas, was wir vorher nicht gewusst haben. Man könnte also

argumentieren, dass aus Daten, Informationen und bestehendem Wissen durch geeignete Strukturierung und Verknüpfung neues Wissen entsteht, Wissen also „strukturierte Information“ ist: obwohl auch diese „Definition“ nicht ganz befriedigen kann (wissen (!) wir doch, dass „Wissen“ noch mehr ist), soll sie im Weiteren als Arbeitshypothese verwendet werden. Wenn wir also schon beim umgangssprachlichen Erfassen des Begriffs „Wissens“ große Probleme haben, wie können wir dann erwarten, dass diese bei „Wissensmanagement“ oder gar „Wissensmanagement-Systemen“ einfacher sind?

Insofern ist es nicht verwunderlich, dass es zur Zeit eine ganze Reihe von verschiedenen Definitionen des Begriffs WM gibt, z.B. [1], [4], [8], [19].

Um definitorische Schwierigkeiten zu vermeiden (die wir bei Begriffen wie „Intelligenz“, „Leben“, „Bewusstsein“ usw. genauso treffen), erscheint es sinnvoll, auf genaue Definitionen zu verzichten und pragmatischer vorzugehen. Ein Ansatz in diese Richtung wird im Folgenden vorgestellt und es wird gezeigt, dass damit die Bedeutung von WM als neuer Wissens(!)bereich, der auch wesentlich in die Informatik hineinspielt, außer Frage gestellt wird.

Vollständigkeitshalber muss darauf hingewiesen werden, dass die Abgrenzung zwischen WM und anderen Disziplinen heute noch sehr fließend ist, und von verschiedenen Gruppen verschieden gesehen wird. Beispielsweise werden in dem Buch [3] viele Bereiche als Teilbereiche von WM gesehen, die dieser Autor eher getrennt behandeln würde. Aber auch andere Extrempositionen, die WM weiter entfernt von der Informatik ansiedeln und in die Nähe der Organisationslehre rücken, sind zu finden [4]. Der Autor dieses Beitrags übersieht keinesfalls die organisatorischen Aspekte, die in gutes WM hineinspielen, konzentriert sich aber bewusst in dieser Arbeit auf jene Teilbereiche, die mit der Informatik in Verbindung stehen.

## 2. Pragmatische Ausgangspunkte für Wissensmanagement

Der Stossseufzer, „Wenn unsere Mitarbeiter nur wüssten, was unsere Mitarbeiter wissen, dann wären wir die beste Organisation der Welt“, beleuchtet vielleicht am besten den wesentlichsten Aspekt von WM: eine Gruppe von Menschen hat stets mehr Wissen, als jeder einzelne in dieser Gruppe, ja selbst das selbe Wissen wird von verschiedenen Personen aus verschiedenen Sichten gesehen.

Die Herausforderung, die sich aus obigem Stossseufzer ergibt, ist offenbar, das Wissen (oder Teile des Wissens) aus den Köpfen von Menschen irgendwie computermässig zu erfassen und andern Menschen jederzeit leicht zugänglich zu machen. Die Vorteile für eine Organisation, der das gelingt sind offenbar gewaltig: man würde Arbeitsduplizierung vermeiden; die Kooperation zwischen Mitarbeitern wäre erleichtert; das Wissen ginge nicht verloren, wenn Mitarbeiter ausscheiden; die Einschulung neuer Mitarbeiter wäre wesentlich einfacher; usw. Freilich müsste die Wissenserfassung ohne wesentliche Belastung der Mitarbeiter erfolgen können, und müsste das Wissen vom System fast automatisch geliefert werden, wenn es benötigt wird, auch dann, wenn sich Mitarbeiter gar nicht bewusst sind, gerade jetzt zusätzliches Wissen zu benötigen.

Es wird später gezeigt werden, dass es tatsächlich Verfahren gibt, die WM im obigen Sinn bis zu einem gewissen Grade erlauben. Ein solches WM soll in der weiteren Diskussion „WM für Organisationen“ genannt werden. Man beachte aber auch, dass bereits einfachere Lösungen wesentlich helfen können: beispielsweise ist es vielleicht nicht notwendig, dass alle Mitarbeiter wissen, was alle anderen wissen. Aber Mitarbeiter, die an einem Thema x interessiert sind, sollten zumindest wissen, ob es andere Mitarbeiter gibt, die beim Thema x helfen können. Ein solches Management von Wissen ist viel weniger ambitioniert. Es ist in manchen Organisationen bereits verwirklicht, wobei dafür oft Ausdrücke wie „Knowledge Domains“ [5] oder „Yellow Pages“ verwendet werden.

Umgekehrt kann man sich auch noch viel ambitionierteres WM vorstellen als „WM für Organisationen“, nämlich ein WM, das nicht nur innerhalb von Organisationen, sondern für die gesamte Gesellschaft von Bedeutung ist, indem z.B. über eine moderne Version des Internets und zukünftige omnipräsente Computer allen Menschen große Teile des gesammelten Wissens der Menschheit jederzeit zur Verfügung stehen, fast als Erweiterung des menschlichen Gehirns. Überlegungen, wie und warum so etwas in Zukunft nicht ganz unrealistisch ist, finden sich z.B. in [6].

Es gibt noch einen weiteren pragmatischen Ansatz zum WM, der aus einer anderen Richtung kommt, und der einem zweiten Stossseufzer entspringt. Einem Stossseufzer in Organisationen, die mit grossen Datenmengen arbeiten, welche aus heterogenen Quellen kommen: „Wenn es uns nur gelänge, neue Daten mit bestehenden so zu verknüpfen, dass zusammengehörige als solche erkannt und klassifiziert werden, und verwandte Informationen miteinander verknüpft werden, dann hätten wir

das Informationsarchivierungsproblem endlich gelöst. Dieses Problem soll kurz „WM für Archive“ genannt werden.

Zwei Beispiele mögen diese Situation erläutern: im Journal for Universal Computer Science, J.UCS, siehe [13] und [14], sind Beiträge nicht nur nach mehreren Merkmalen klassifiziert und daher nach verschiedensten Kriterien auffindbar, sondern ist der Einbau von „Links in die Zukunft“ vorgesehen. Damit ist Folgendes gemeint: ist ein Beitrag A beispielsweise 1995 verfasst, und verweist ein neuer Beitrag B im Jahre 2002 auf A, dann ist es in digitalen Bibliotheken [15] üblich, dass B ein Link auf A enthält, also ein „Link in die Vergangenheit“. In J.UCS wird aber auch der Beitrag A aus dem Jahre 1995 ein Link auf den Beitrag B im Jahre 2002 besitzen, also ein „Link in die Zukunft“. Solche Links in die Zukunft lassen sich innerhalb von digitalen Bibliotheken relativ leicht erstellen, wenn die Literaturangaben ein fixes Format haben; sie sind aber auch über die Grenzen von digitalen Bibliotheken hinaus durch verschiedene Maßnahmen generierbar: die einfachste ist die Verwendung sogenannter „Citation Indizes“; andere könnten durch Verfahren entstehen, die „Ähnlichkeiten“ oder „Zusammenhänge“ zwischen Dokumenten erkennen: auf diese wird im Abschnitt 3 näher eingegangen.

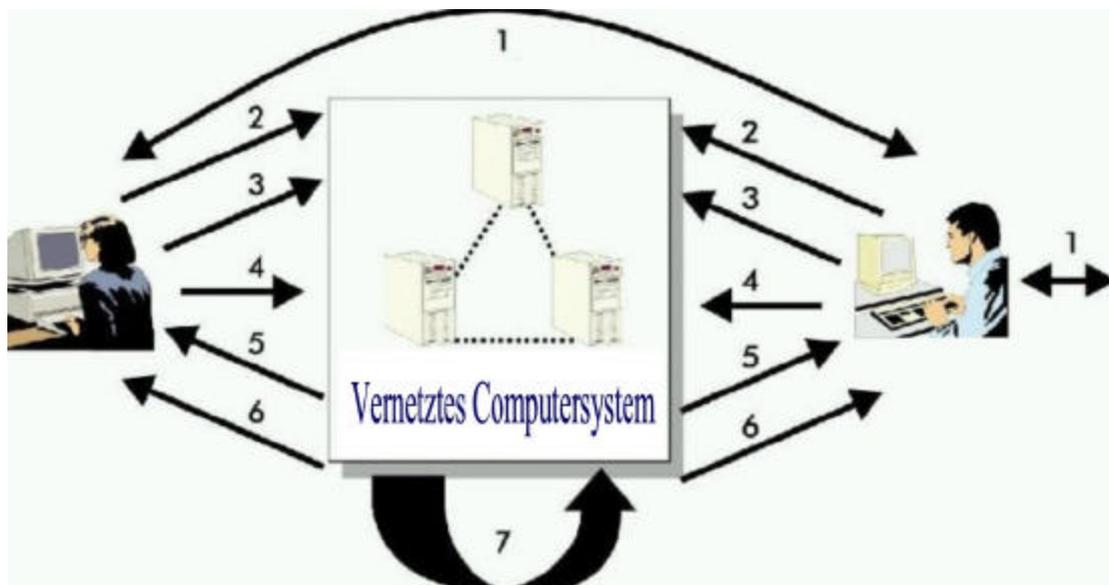
Ein weiteres Beispiel ist die elektronische Version des Brockhaus Lexikons [16]. Bei jedem Beitrag gibt es ein „Wissensnetz“, das grafisch Beiträge, die mit dem gegenwärtigen einen Zusammenhang haben, aufzeigt. Die Erkennung des Zusammenhangs geschieht beim Brockhaus einerseits über den schon einmal angesprochenen Ansatz „Ähnlichkeiten“ zu erkennen, unterstützt durch entsprechende Metadaten [17]: jeder Beitrag ist einer oder mehreren Kategorien zugeordnet. Damit sind z.B. Querverweise vom Politiker Kohl zum Gemüse Kohl oder anderen Gemüsen, und umgekehrt, eliminierbar. Im Jahr 2003 soll erstmals das Wissensnetz des elektronischen Brockhaus auf einer vollständigen Ontologie [12] der deutschen Sprache aufgebaut werden!

Es wird sich im Folgenden herausstellen, dass „WM für Archive“ ein wichtiger Teilbereich von „WM für Organisationen“ ist, weshalb wir uns im weiteren auf den Bereich WM für Organisationen beschränken wollen.

### 3. Ein Modell für das Wissensmanagement in Organisationen.

Neben einigen anderen schon zitierten Ansätzen, die die Kernpunkte eines WM-Systems aufzeigen, hat sich das sogenannte „Kommunikationsmodell“ von Maurer-Tochtermann in Abbildung 1

besonders bewährt, da es deutlich zeigt, wo sich WM von (verteilten) Informationssystemen und Datenbanken unterscheidet.



**Abbildung 1**

Abbildung 1 zeigt eine Gruppe von Menschen, die gegenseitig Wissen austauschen, wobei ein grosser Teil des Wissensaustausches über ein vernetztes Computersystem (und damit auch asynchron, also zeitversetzt) vor sich geht. Den mit 1 bis 7 markierten Pfeilen kommt dabei naturgemäß eine besondere Bedeutung zu:

Pfeil 1 zeigt, dass Menschen direkt (in der Kaffeepause, beim Betriebsausflug, übers Telefon,...) Wissen austauschen: es ist dieser Pfeil, mit dem sich die organisatorische Seite des Wissensmanagements hauptsächlich beschäftigt, die in dieser Arbeit nicht weiter behandelt wird, und nur durch eben diesen einen Pfeil angedeutet sein soll.

Die Pfeile 2 bis 4 symbolisieren verschiedene Wege, wie Informationen in das WM-System gelangen, während 5 bis 6 andeuten sollen, dass es mindest zwei Wege gibt, wie Wissen aus dem WM-System zu den Benutzern gelangt. Konkreter, Pfeil 1 symbolisiert die explizite Eingabe von Informationen in das System und Pfeil 2 die implizite Eingabe: darunter ist die Erfassung von Informationen und Wissen als Abfallprodukt anderer, ohnehin durchgeführter Aktivitäten zu verstehen, d.h. hier entsteht im WM-System neues Wissen ohne zusätzliche Arbeitsbelastung. Pfeil 4 deutet an, dass ein WM-System auch in der Lage ist, „systemisch“ Informationen zu generieren, nämlich durch

Beobachtung der Benutzer. Pfeil 5 symbolisiert den traditionellen Informationsabruf aus einem Informationssystem oder einer Datenbank: Benutzer stellen explizite Anfragen und bekommen darauf vom WM-System entsprechende Antworten geliefert: diese können kleine Informationsbruchstücke sein, oder große kohärente Dokumente, wie z.B. Manuals, Bücher oder Unterrichtsmodule. „Interessanter“ ist wieder Pfeil 6: er soll andeuten, dass das WM-System auch von sich aus aktiv werden kann, und ohne Aufforderung des Benutzers Informationen bzw. Wissen liefern kann. Schließlich zeigt Pfeil 7, dass ein WM-System selbst aus bestehendem Wissen neues generieren kann.

Die Abbildung 1 verdeutlicht sehr anschaulich den Unterschied zwischen klassischen Informationssystemen (Datenbanken) und WM-Systemen: lässt man nämlich die Pfeile 3, 4, 6 und 7 in der Abbildung weg, dann hat man genau ein klassisches Informationssystem. Ob WM-Systeme sich also wirklich von klassischen Informationssystemen unterscheiden, hängt davon ab, ob die Pfeile 3, 4, 6 und 7 tatsächlich realisierbar sind. Dies wird im nächsten Abschnitt durch konkrete Beispiele positiv beantwortet.

Abbildung 1, für „WM für Organisationen“ konzipiert, gilt aber auch für „WM für Archive“, wenn man dabei die Pfeile 1 und 3 weitgehend ignoriert: in diesem Sinne ist die frühere Aussage gerechtfertigt, dass man sich auf „WM für Organisationen“ beschränken darf, weil man damit „WM für Archive“ eigentlich mitbehandelt.

#### 4. Einige Techniken des Wissensmanagements

In diesem Abschnitt wird angedeutet, wie die durch die Pfeile 3, 4, 6 und 7 in Abbildung 1 angedeuteten Aktionen in modernen WM-Systemen realisiert werden können.

Pfeil 3 in Abbildung 1 symbolisiert die implizite Eingabe von Informationen, d.h. die Vermehrung des Wissensbestandes im WM-System als „Abfallprodukt“ von Handlungen von Menschen, die diese ohnehin für andere Zwecke ausführen. Dazu gehören so einfache, wie z.B. die Ankündigung einer Veranstaltung, die gedruckt und als eMail an einen gewissen Verteiler versandt wird, wobei dieser und das WM-System so konzipiert sind, dass die Veranstaltung auch in den Veranstaltungskalender des WM-Systems aufgenommen wird, nach dem Datum der Veranstaltung diese automatisch zum Ordner „Vergangene Veranstaltungen dieses Jahres“ verschoben, und nach Ablauf des Jahres der Ordner „Veranstaltungen seit xxxx“ ergänzt wird. Alle Informationen, die ohnehin für einen Webserver eingegeben werden

(vom Telefonverzeichnis, bis zur Struktur und den Aufgaben des Unternehmens und seiner Abteilungen) wandern selbstverständlich in das WM-System. Da kaum eine große Organisation heute ohne ISO 9000 Zertifizierung konkurrenzfähig arbeiten kann, stellen die umfangreichen Informationen, die zu diesem Zweck laufend mitgeführt werden müssen, eine mächtige und natürliche Quelle für das WM-System dar: dadurch werden laufende Projekte mit ihren Zielen, Personen, Werkzeugen, Meilensteinen, Problemen, Protokollen, Dokumentationen, usw. erfasst und speichern damit umfangreiches Wissen über das Geschehen in der Organisation. Auch fertiggestellte Projekte erweisen sich plötzlich als große Hilfe, weil man aus positiven und negativen Erfahrungen mit diesen Projekten viel für laufende und zukünftige Vorhaben lernen kann. Der Workflow einer Organisation ist eine weitere Komponente, die heute sinnvoller Weise ohnehin elektronisch vorliegt, und automatisch in das WM-System integriert werden muss. Ähnliches gilt für Jahresberichte, Produktlisten mit Beschreibungen und Preisen, Manuals und andere interne Berichte, und geht natürlich hin bis zur Integration bestehender Informationssysteme, wofür sogenannte „Wissensportale“ [20] benötigt werden. Aber auch Email sollte (mit entsprechenden Autorisierungsmaßnahmen) zentral im WM-System, und nicht getrennt von jedem Benutzer verwaltet werden. Warum die Zusammenführung aller dieser Informationen so wichtig ist, wird im Folgenden klar werden: nur wenn ein ausreichender Vorrat von Informationen vorhanden ist, können die Informationen vom WM-System automatisch in Verbindung zueinander gebracht werden, entsteht daraus „strukturierte Information“, also Wissen.

Pfeil 4, der „systemisch“ Wissen auf Grund von Eingaben symbolisiert, ist zur Zeit noch die größte Schwachstelle von WM-Systemen: die Grundidee ist, dass aus den Eingaben, die aus bestimmten Quellen (z.B. bestimmten Mitarbeitern oder Datenbanken) kommen, allgemeine Regeln abgeleitet werden können, die in ähnlichen Situationen zur Verfügung stehen. Solche Regeln sind häufig mit den Aktivitäten, die Pfeil 5 symbolisiert (explizite Datensuche), verzahnt: beispielsweise wird ein gutes WM-System merken, dass manche Suchwege von einem oder mehreren Benutzern immer wieder verwendet werden, und daher entsprechende Abkürzungen vorsehen, automatische Lesezeichen generieren, oder vormerken, dass fast alle Personen, die z.B. x suchen, auch y suchen. Damit wird dann einem Benutzer, der x abrufen, y automatisch angeboten, im Sinne des Pfeils 6.

Die Mächtigkeit der Pfeile 6 und 7 können besser erläutert werden, wenn zunächst einige Techniken erklärt werden, die für ein gutes WM-System von ausschlaggebender Bedeutung sind.

Zu den wichtigeren Konzepten in einem WM-System gehört die von Maurer propagierte Idee des „aktiven Dokuments“. Überspitzt formuliert kann man diese so darstellen: wann immer man ein Dokument abrufen, kann man dazu eine beliebige Frage stellen (typisch: eintippen, in weiterer Zukunft wohl auch einfach sprechen) ... und das WM-System gibt die Antwort!

Obwohl es natürlich unmöglich ist, dass das WM-System als solches beliebige Fragen beantwortet, gibt es zwei Ansätze, die pragmatisch nahe an das beabsichtigte Resultat herankommen: erstens können gewisse Anfragen in Datenbankabfragen umgesetzt und damit beantwortet werden, andererseits kann bei häufig auftretenden Fragen das erste Mal ein Experte (potentiell auch zeitversetzt) die Frage beantworten, aber eine wiederholte Beantwortung der selben Frage kann mit einem Trick vermieden werden: das WM-System speichert Fragen und die zugehörigen Antworten. Kommt zum selben Dokument eine weitere Frage, so wird überprüft, ob diese mit einer vorhergehenden gleichwertig ist. Wenn dies der Fall ist, kann das System ohne Einschaltung eines Experten die Antwort geben. Dieses Verfahren ist vor allem dann sinnvoll, wenn viele Personen das selbe Dokument (Webseite) abrufen, und dieses sich über längere Zeit hinweg nicht ändert. Freilich gibt es auch dabei das Problem zu erkennen, ob eine Frage  $x$  gleichwertig mit einer Frage  $y$  ist.

Zur Lösung dieses Problems gibt es mehrere Möglichkeiten: die Fragen  $x$  und  $y$  können durch Vergleich von Worten (gegebenenfalls mit Unterstützung von Synonymwörterbüchern, semantischen Netzen oder Ontologien) auf Ähnlichkeit überprüft werden. Besteht der Verdacht, dass die Frage  $x$  mit der Frage  $y$  gleichwertig ist, so wird gefragt, ob die Frage  $y$  den Wünschen entspricht, in welchem Fall das System die Antwort geben kann. Wenn bei keiner vorher gestellten Frage der Verdacht besteht, dass sie gleichwertig mit  $x$  ist, dann und nur dann muss die Frage wieder an einen Experten weiter geleitet werden. Eine Verfeinerung dieser Methode gelingt fallweise durch „case based reasoning“ und andere Ansätze aus der künstlichen Intelligenz, doch sind diese Methoden bisher noch von geringerer praktischer Bedeutung.

Natürlich wäre es eleganter, wenn man überhaupt beweisen könnte ob zwei Textstücke  $x$  und  $y$  gleichwertig sind oder nicht: dies ist aber mit

heutigen und absehbaren Methoden nicht lösbar, außer unter recht eingeschränkten Bedingungen, z.B. wenn man nur gewisse syntaktische Formen zulässt und das Gebiet einschränkt, siehe [9].

Es gibt aber einen weiteren besonders einfachen Ansatz für die Lösung des Problemkreises aktive Dokumente, den man als „lokale FAQs“ bezeichnen könnte. Konkreter, haben Benutzer Fragen, markieren sie das entsprechende Stück des Dokuments und tippen die Frage ein. Diese wird von einem Experten beantwortet: während etwaiger „Sprechstunden“ sofort, sonst zeitversetzt. Vor der markierten Stellen des Dokuments wird ein Icon eingefügt, das späteren Lesern signalisiert: „Hier wurde bereits eine Frage gestellt und von einem Experten beantwortet“. Nachfolgende Benutzer werden sich also den Frage – Antwortdialog ansehen, bevor sie eine etwaige weitere Frage zum selben Dokumentstück stellen. In großen Experimenten hat sich herausgestellt, dass selten mehr als einige wenige Fragen zum selben Dokumentstück gestellt werden (und wenn ja, der Autor des Dokuments gut beraten ist, dieses abzuändern, weil es offenbar Unklarheiten hervorruft!), und dass sich die Liste der Fragen zu einem Dokumentstück bald „stabilisiert“: bei einer Firma mit 150.000 Mitarbeitern, die umfangreiche Informationen für alle Mitarbeiter zum verpflichtendem Durchlesen anbot, wurden nach 600 Mitarbeitern nur noch von 0.03 % der anderen Mitarbeitern neue Fragen gestellt! D.h. für die ersten 600 Mitarbeitern mussten zunächst Experten (im gegenständlichen Fall rund um die Uhr) zur Verfügung stehen, aber bei den anderen 149.000 Mitarbeitern (d.h. bei 99.6%) traten insgesamt nur noch 45 neue Fragen auf, die man ohne große Problematik offenbar auch zeitversetzt beantworten konnte!

Obige Diskussion legt nahe, dass kein gutes WM-System auf das Konzept „aktive Dokumente“ verzichten kann.

Bereits angesprochen im Zusammenhang mit dem Vergleich von zwei Fragen  $x$  und  $y$  wurde das Problem, festzustellen, ob  $x$  und  $y$  „gleichwertig“ oder „ähnlich“ sind. Es gibt eine Reihe von Verfahren (meist verbesserte Versionen des offensichtlichen Ansatzes, die Häufigkeit wichtiger Worte als Maß der Ähnlichkeit zu wählen), denen es gelingt, mit guter Trefferwahrscheinlichkeit festzustellen, ob zwei Textstücke ähnlich sind. Diese Verfahren erlauben damit nicht nur eine automatische Klassifikation von Textstücken (im Sinne des Pfeils 7 in Abbildung 1), sondern auch die aktive Verständigung über den Verdacht der Ähnlichkeit von zwei Dokumenten (im Sinne des Pfeils 6 in Abbildung 1). Einige konkrete Beispiele mögen die Anwendbarkeit solcher Ähnlichkeitsprüfungen erläutern.

Beispiel 1: Man betrachte eine weitverzweigte Organisation. Wird an einem Ort A ein neues Projekt begonnen, so wird dies nach ISO9000 entsprechend dokumentiert. Das WM-System nimmt dieses Dokument, übersetzt es aus der Landessprache ins Englische und vergleicht die so entstehende (schlechte) Übersetzung mit den englischen Beschreibungen aller laufenden Projekte in dieser Organisation. Wird eine Ähnlichkeit mit einem Projekt an einer anderen Stelle B gefunden, werden A und B verständigt, dass hier potentiell eine Arbeitsduplizierung eintreten könnte. Das WM-System mag sich in einigen Fällen „irren“, d.h. A und B auf Ähnlichkeiten hinweisen, die bekannt oder irrelevant sind, aber selbst wenn nur ganz wenige Fälle von Duplizierung vermieden werden können, ist dies finanziell sehr signifikant.

Beispiel 2: In einer Organisation, die über große interne Forschungsabteilungen verfügt, stellt ein Mitarbeiter eine neue Publikation A in die digitale Bibliothek der Organisation. Wenig später kann es sein, dass er auf Grund der Ähnlichkeitsanalyse einen Email-Hinweis erhält, dass in der digitalen Bibliothek bereits zwei ähnliche Arbeiten B und C vorliegen. Gleichzeitig werden auch die Autoren von B und C benachrichtigt, dass eine weitere ähnliche Arbeit A gerade in die digitale Bibliothek eingebracht wurde. Wie im Beispiel 1 können dadurch Duplizierungen vermieden, und kann die Kooperation erhöht werden. Im Idealfall wird die Publikation A übrigens nicht nur mit Dokumenten in der organisationsinternen digitalen Bibliothek, sondern auch mit solchen in anderen digitalen Bibliotheken, die über das Internet zugänglich sind (wie z.B. J.UCS [13], [14]), verglichen.

Beispiel 3: Durch die laufende Überprüfung von nicht-privaten Emails wird festgestellt, dass zwei Gruppen A und B anscheinend ähnliche Probleme bearbeiten: sie werden entsprechend verständigt.

Beispiel 4: In einem Diskussionsforum wird ein vor Monaten ausführlich behandeltes Thema wieder angeschnitten: das WM-System erkennt dies und kürzt eine unnötige neue Diskussion durch den Verweis auf die schon vorliegenden Beiträge ab.

Zusammenfassend muss festgehalten werden, dass die Ähnlichkeitsüberprüfung von Dokumenten zu den mächtigsten Werkzeugen gehört, über die ein WM-System verfügen muss. Errechnete Zusammenhänge lassen sich oft grafisch gut durch ein Wissensnetz darstellen, wie dies z.B. im Multimedialen Brockhaus [16] geschieht. Ein

Beispiel dafür ist Abbildung 2: sie zeigt das für das Wort „Raumsonde“ automatisch berechnete Wissensnetz.

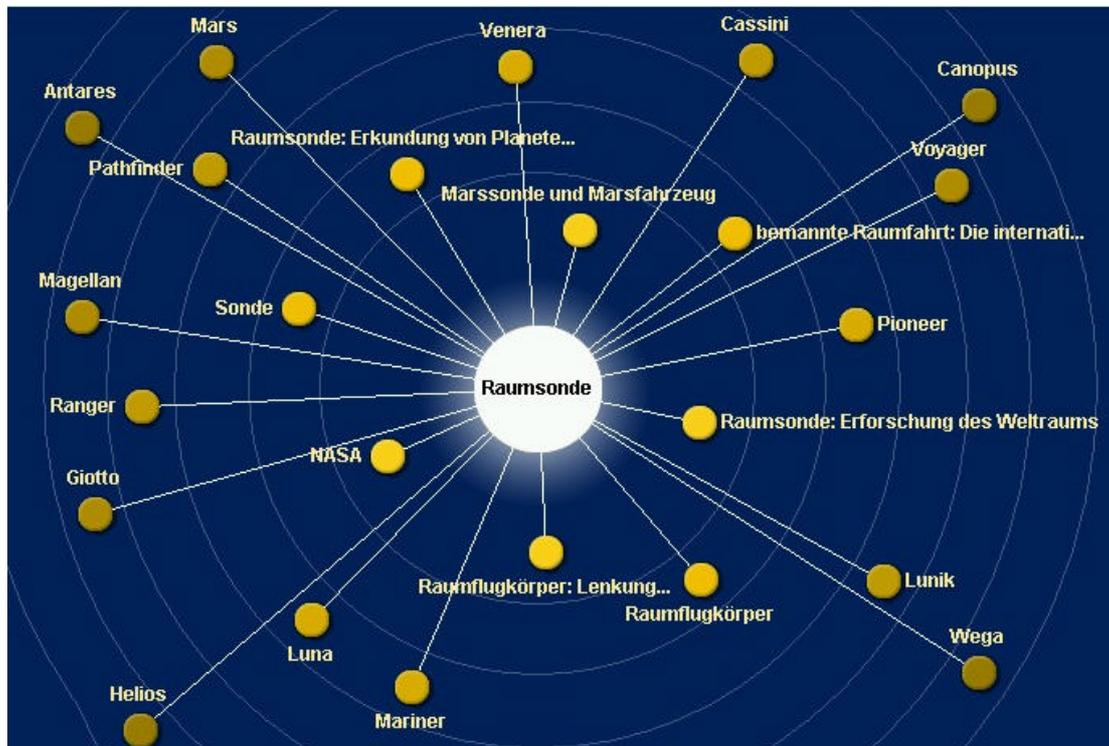


Abbildung 2

Eine komplexere Variante der Ähnlichkeitserkennung ist die „Zusammenhangserkennung“. Die Verfahren dafür sind noch nicht generisch genug, dass sie allgemein einsetzbar sind, sondern müssen auf den Bedarf zugeschnitten werden. Darum möge es genügen, diese Variante an einem Beispiel zu erläutern.

Nehmen wir an, wir haben einen sehr großen Datenbestand, etwa alle öffentlich zugänglichen WWW Server und Datenbanken, inklusive der elektronisch erfassten Berichterstattung in Zeitungen, Nachrichtenagenturen, usw. Aufgabe ist es festzustellen, welche Personen mit einer bestimmten Person X in einem engeren Zusammenhang stehen.

Findet das WM-System irgendwo das Textstück „X hat am 15. Oktober 2000 Nassau besucht“ so wird die Information (15.10.2000, Nassau) in die „X-Datenbank“ aufgenommen. Bei der Analyse einer Eintragung an ganz anderer Stelle „Maurer war vom 10. bis 20. Oktober 2000 auf North Eluthera“ wird (a) „Maurer“ als Eigename durch Syntaxanalyse erkannt und wird (b) der Eintrag mit jedem Eintrag der X-Datenbank verglichen.

Da das WM-System gewisse geografische Beziehungen kennt, „weiß es“ dass North Eluthera zu den Bahamas gehört und Nassau die Hauptstadt der Bahamas ist. Da der 15. Oktober 2000 im Intervall 10. bis 20. Oktober 2000 liegt, wäre es denkbar, dass Maurer die Person X getroffen hat. In der „Zusammenhangsdatenbank“ wird daher das Paar (X, Maurer, 1) gespeichert: es gibt bisher ein Indiz, dass Maurer die Person X getroffen haben könnte. Findet später das WM-System die Informationen „X hat die Person Z getroffen“ und wo anders „Maurer ist mit Z in die Schule gegangen“, wird in der Zusammenhangsdatenbank das Tripel (X,Maurer,1) durch (X,Maurer,2) ersetzt. Erreicht der Zähler eine kritische Größe, d.h. hat man z.B. irgendwann (X, Maurer, 100) dann schlägt das System Alarm: nun ist es höchst wahrscheinlich, dass X und Maurer irgendwie zusammen hängen: das „Alarmschlagen“ ist wieder eine typische Aktion, die dem Pfeil 6 in Abbildung 1 entspricht, während der beschriebene Vorgang der Berechnung der Zusammenhangtripel dem Pfeil 7 zuzuordnen wäre.

Es sollte klar sein, dass die Herstellung von Beziehungen im beschriebenen Beispiel ein ziemlich komplexes System voraussetzt. Sie hat allerdings z.B. bei Fahndungen in der Vergangenheit bereits sehr wertvolle Dienste geleistet.

Es gibt sehr viel einfachere Anwendungen, wo es auch um Zusammenhänge geht z.B. im Bereich eCommerce, wo man etwa die Ähnlichkeit von zwei Kunden A und B, was Geschmack von Büchern angeht, durch ihre Einkäufe feststellt. Kauft A ein Buch eines bisher nicht von A gelesenen Autors X, und bald darauf ein weiteres Buch von X, dann kann man davon ausgehen, dass A der Autor X gefällt, und man wird B (meist mit Erfolg) auf den Autor X hinweisen.

Die Forschung im Zusammenhang mit Zusammenhangserkennung ist noch voll im Gange: aber schon jetzt ist klar, dass es sich hier um ein weiteres Werkzeug handelt, das für KM-Systeme von größter Bedeutung ist.

## 5. Zusammenfassung

In dieser Arbeit wurde gezeigt, dass und wie sich Wissensmanagement Systeme (WM-Systeme) von klassischen Informationssystemen unterscheiden. Die beschriebenen Funktionen sind auch tatsächlich schon in einigen existierenden Systemen realisiert, wobei Hyperwave [7], [18] als das vermutlich ausgereifteste zu bezeichnen ist, da es aktive Dokumente, automatische Klassifizierung, Ähnlichkeitserkennung und

viele andere wesentliche Komponenten eines WM-Systems zusammen anbietet.

## 6. Literatur

- [1] Woods, E., Sheina, M.: Knowledge Management - Applications, Markets and Technologies, Ovum Report (1998).
- [2] Studer, R., Abecker, A., Decker, S.: Informatik-Methoden für das Wissensmanagement. Angewandte Informatik und Formale Beschreibungsverfahren, Teubner-Texte zur Informatik, Vol 29 (1999).
- [3] Karagiannis, D., Telesko, R.: Wissensmanagement – Konzepte der Künstlichen Intelligenz und des Softcomputing, Lehrbücher Wirtschaftsinformatik, Oldenbourg Verlag (2001).
- [4] Davenport, T., Prusak, L.: Working Knowledge: How Organizations Manage What They Know; Harvard Business School Press, Boston, (1998).
- [5] Helic, D., Maurer, H., Scerbakov, N. : Knowledge Domains: A Global Structuring Mechanism for Learning Resources in WBT Systems, Proceedings of WEBNET 2001, November 2001, AACE, Charlottesville, USA (2001), 509-514.
- [6] Maurer, H.: Die (Informatik) Welt in 100 Jahren, Informatik Spektrum, Springer Verlag (April 2001), 65-70.
- [7] Maurer, H., HyperWave: The Next Generation Web Solution; (Ed.), Addison-Wesley Longman, London (1996).
- [8] Sivan, Y.: The PIE of Knowledge Infrastructure: To Manage Knowledge We Need Key Building Blocks; WebNet Journal 1,1 (1999), 15-17.
- [9] Heinrich, E., Maurer, H.: Active Documents: Concept, Implementation and Applications; Journal of Universal Computer Science 6, 12 (2000), 1197- 1202.
- [10] Tochtermann, K., Maurer, H.: Umweltinformatik und Wissensmanagement. Tagungsband des 14. Internationalen GI-Symposiums Informatik für den Umweltschutz, Bonn, Metropolis Verlag (2000), 462-475.
- [11] Meersman, R., Tari, Z., Stevens, S. (Eds.): Database Semantics; Kluwer Academic Publishers, USA, (1999).
- [12] Uschold, M., Gruninger, M.: Ontologies: principles, methods and applications; The Knowledge Engineering Review 11,2 (1996), 93-136.
- [13] [www.jucs.org](http://www.jucs.org)

- [14] Krottmaier, H., Maurer, H.: Transclusions in the 21st Century, Journal of Universal Computer Science 7,12 (2001), 1125-1136.
- [15] Endres, A., Fellner, D.: Digitale Bibliotheken, Heidelberg, Germany: dpunkt Verlag (2000).
- [16] Der Brockhaus Multimedial 2002 Premium; DVD, Brockhaus Verlag, Mannheim (2001).
- [17] Weibel, S.: "The state of the Dublin core metadata initiative," D-Lib Magazine 5, 2 (1999), 300.
- [18] [www.Hyperwave.de](http://www.Hyperwave.de)
- [19] Petkoff, B.: Wissensmanagement. Addison Wesley (1998).
- [20] Hyperwave: Hyperwave Information Portal - White paper (1999).  
<http://www.hyperwave.de/publish/downloads/Portalwhitepaper.pdf>