(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
30. Mai 2013 (30.05.2013)

WIPO | PCT

(10) Internationale Veröffentlichungsnummer
WO 2013/075745 A1

(51) Internationale Patentklassifikation:
G06Q 30/00 (2012.01) *G06F 17/30* (2006.01)

(21) Internationales Aktenzeichen: PCT/EP2011/070873

(22) Internationales Anmeldedatum:
23. November 2011 (23.11.2011)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(71) Anmelder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **GENZMEHR, Marcel** [DE/DE]; Denhardtstrasse 16, 39106 Magdeburg (DE). **LANGER, Stefan** [DE/DE]; Pappelallee 13, 39106 Magdeburg (DE).

(72) Erfinder; und
(71) Anmelder : **BEEL, Jöran** [DE/DE]; Zur Salzhaube 3a, 31832 Springe (DE).

(74) Anwalt: **BETTINGER SCHNEIDER SCHRAMM**; Postfach 86 02 67, 81629 München (DE).

(81) Bestimmungsstaaten *(soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Bestimmungsstaaten *(soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ,

[Fortsetzung auf der nächsten Seite]

(54) **Title**: METHOD AND SYSTEM FOR CREATING USER MODELS

(54) **Bezeichnung** : VERFAHREN UND SYSTEM ZUM ERSTELLEN VON NUTZERMODELLEN

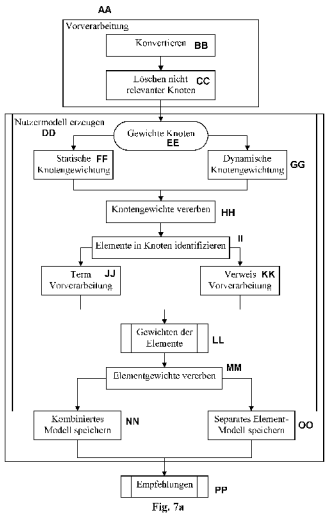

Fig. 7a

Fig 7a:
AA Preprocessing
BB Convert
CC Delete irrelevant nodes
DD Produce user model
EE Weight nodes
FF Static node weighting
GG Dynamic node weighting
HH Transmit node weights
II Identify elements in nodes
JJ Term preprocessing
KK Reference preprocessing
LL Weighting of the elements
MM Transmit element weights
NN Store combined model
OO Store separate element model
PP Recommendations

(57) **Abstract**: Disclosed are a method and a system for creating a user model, in particular for a recommender system, from at least one tree data structure, wherein the user model comprises information about a user, wherein the at least one tree data structure can be associated with the user. The tree data structure comprises a root node and a number of child nodes which are connected to the root node or to a child node via edges, and at least one element is associated with at least one node. The elements associated with the nodes and representing a content of the corresponding node are assessed, the assessed elements are weighted and an element weight is assigned to each element, and a user model is generated, the generated user model comprising the assessed elements and the element weight assigned to the corresponding element.

(57) **Zusammenfassung**: Bereitgestellt werden ein Verfahren und ein System zum Erzeugen eines Nutzermodells, insbesondere für einen Empfehlungsdienst, aus zumindest einer baumförmigen Datenstruktur, wobei das Nutzermodell Informationen über einen Nutzer umfasst, wobei die zumindest eine baumförmige Datenstruktur dem Nutzer zuordenbar ist, wobei die baumförmige Datenstruktur einen Wurzelknoten und eine Anzahl von Kinderknoten umfasst, welche über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind, wobei zumindest einem Knoten zumindest ein Element zugeordnet ist, und wobei die den Knoten zugeordneten Elemente ermittelt werden, wobei die Elemente einen Inhalt des jeweiligen Knoten repräsentieren, die ermittelten Elemente gewichtet werden und jedem Element eine Elementgewichtung zugeordnet wird, und ein Nutzermodell generiert wird, wobei das generierte Nutzermodell die ermittelten Elemente und die dem jeweiligen Element zugeordnete Elementgewichtung umfasst.

WO 2013/075745 A1

MD, RU, TJ, TM), europäisches (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— *mit internationalem Recherchenbericht (Artikel 21 Absatz 3)*

Verfahren und System zum Erstellen von Nutzermodellen

Gebiet der Erfindung

Die Erfindung betrifft ein Verfahren und ein System zum Erstellen von Nutzermodellen und darauf basierenden Empfehlungen, vorzugsweise durch Analyse von baumförmigen Datenstrukturen.

Hintergrund der Erfindung und Stand der Technik

Nutzer von Computersystemen unterscheiden sich in vielerlei Hinsicht, unter anderem hinsichtlich ihrer Interessen, ihres Wissens und ihrer demographischen Daten. Viele Computersysteme versuchen diesen Unterschieden gerecht zu werden, indem sie abhängig etwa vom Wissen und den Interessen des Nutzers, individuelle Informationen oder Benutzeroberflächen zur Anzeige bringen. Um Softwaresysteme, etwa Computerprogramme oder Internet-basierte Anwendungen individuell an seine Nutzer anpassen zu können, benötigen die Softwaresysteme Zugriff etwa auf die Interessen der Nutzer. Diese Daten können entweder manuell vom Nutzer angegeben, oder automatisch vom System erzeugt werden. In jedem Fall werden die Informationen über die Nutzer in so genannten Nutzermodellen gespeichert.

Häufig werden solche Nutzermodelle von Empfehlungsdiensten verwendet. Abhängig von den Interessen eines Nutzers zeigen diese Empfehlungsdienste individuelle Empfehlungen beispielsweise für Filme, Bücher, Musik, oder auch auf den Nutzer abgestimmte Werbung an. Bei einem Empfehlungsdienst handelt es sich immer um ein sogenanntes „User-Item Matching Problem“: Die Frage bei diesem

Problem ist, welche kleine Auswahl an relevanten Items (z.B. Musikstücke, Webseiten, Bücher, etc.) aus einer großen Menge von verfügbaren Items einem Nutzer empfohlen werden soll. Dieser aus dem Stand der Technik bekannte Ansatz ist in **Fig. 1** dargestellt. Gezeigt ist in Fig. 1 eine Menge von Nutzern (User 1 bis User 3) und eine Menge von Items (Item 1 bis Item 3). Mit entsprechenden Verfahren wird die Relevanz von Nutzern und Items zueinander berechnet. Danach können alle Items die einen bestimmten Schwellenwert bzgl. der Relevanz überschreiten den entsprechenden Nutzern empfohlen werden.

Aus dem Stand der Technik bekannte Empfehlungsdienste nutzen zwei grundsätzliche Verfahren, um Nutzermodelle zu erzeugen bzw. Empfehlungen zu geben. Diese Verfahren sind bekannt als "Content Based Filtering" bzw. "Collaborative Filtering".

Beim Content Based Filtering (CBF) nimmt das Computersystem an, dass der Inhalt (Content) der Items mit denen ein Nutzer in Verbindung steht, die Interessen und/oder das Wissen des Nutzers wiederspiegelt. Dieser aus dem Stand der Technik bekannte Ansatz ist in **Fig.2** gezeigt.

Ein Nutzer (User 1) steht mit einer Anzahl von Items (Item 1 bis Item j) in Verbindung. Ein Item kann jedes mögliche Objekt sein. Ein Item kann etwa ein Dokument (Bücher, Webseiten, Emails, etc.), ein Multimediaobjekt (Filme, Musik, Fotos), eine Personen oder ein Ort sein. Items können aber auch Menüeinträge einer Computeranwendung oder Komponenten grafischer Benutzeroberflächen sein.

In Verbindung steht ein Nutzer mit einem Item, wenn irgendein Bezug zwischen ihnen besteht. Das heißt, wenn der Nutzer beispielsweise ein Buch gelesen, gekauft oder auch nur kurz betrachtet hat, eine Person kennt oder einen Film geschaut, heruntergeladen oder auf einem Filmportal bewertet hat, steht der Nutzer mit dem Buch, der Person, bzw. dem Film in Verbindung. Die Verbindung kann

dabei unterschiedlich stark gewichtet werden, je nach Art der Verbindung. Beispielsweise könnte das Kaufen eines Buches stärker gewichtet werden als das bloße Betrachten des Buchcovers. Oder eine Verbindung zu einem Item kann umso stärker gewichtet werden, je öfter das Item genutzt wurde.

Beim Content Based Filtering wird der Inhalt der verbundenen Objekte genutzt um ein Nutzermodell zu erstellen. In der Regel wird dieses Verfahren bei textuellen Items, also Dokumenten, angewandt, da der Inhalt von Dokumenten (also der Text) gut von Computern verarbeitet werden kann im Gegensatz, z. B. zu Bildern. Um den Inhalt der Items zu nutzen, wird für jedes Item ein Modell erstellt. Bei Dokumenten wird häufig das sogenannte "Vector Space Model" genutzt, ein Modell, welches Dokumente als Vektor ihrer Terme darstellt. Jeder Vektor drückt durch seine Länge aus, wie gut der entsprechende Term das eigentliche Dokument beschreibt. Diese Gewichtung kann mit verschiedenen Verfahren errechnet werden. Ein gängiges Verfahren ist das sogenannte TF-IDF Verfahren. Hierbei ist das Gewicht eines Terms für ein Dokument umso größer je öfter der Term in dem Dokument vorkommt und je weniger Dokumente in der gesamten Kollektion es mit diesem Term gibt.

Das Nutzermodell wird dann aus den Modellen der verschiedenen verbundenen Items erzeugt. Dies bedeutet, wenn ein Nutzer viele Bücher besitzt die den Term „Recommender“ mit hohem Gewicht enthalten, dann bekommt auch das Nutzermodell diesen Term mit einem hohem Gewicht zugeordnet. Die verschiedenen Item-Modelle können dabei mit unterschiedlicher Gewichtung in das Nutzermodell einfließen. Üblich ist es beispielsweise, Items die vor kurzem genutzt wurden, stärker zu gewichten als Items deren Nutzung bereits längere Zeit zurückliegt. Üblicherweise wird das Nutzermodell in dem gleichen Format gespeichert, wie die Item-Modelle – also beispielsweise als Vector Space Model.

Von den Items, die später dem Nutzer gegebenenfalls empfohlen werden sollen, wird ebenfalls ein Modell erstellt, beispielsweise wieder mit TF-IDF-Verfahren

und dem Vector Space Model. Diese Items müssen nicht notwendigerweise die gleichen Items sein, die mit Nutzern in Verbindung stehen. Beispielsweise wäre es möglich, ein Nutzermodell zu erstellen basierend auf Webseiten die ein Nutzer besucht hat, und ihm basierend auf diesem Modell Bücher zu empfehlen, oder auch personalisierte Werbung anzuzeigen. Das Matching der Nutzermodelle mit den zu empfehlenden Items basiert vorzugsweise auf Ähnlichkeitsvergleichen zwischen den Nutzer- und den Item-Modellen. Bei textbasierten Items bedeutet das, wenn das Nutzermodell die gleichen Terme mit hohem Gewicht enthält wie die Item-Modelle der zu empfehlenden Items, dann ist die Ähnlichkeit groß und das Item wird dem Nutzer empfohlen. Ein übliches Ähnlichkeitsmaß im Vector Space Model ist etwa die Cosine Similarity.

Beim sogenannten Collaborative Filtering (CF), welches in **Fig. 3** gezeigt ist, spielt der Inhalt von Items keine Rolle. Es wird lediglich die Information verwendet, welche Items mit welchen Nutzern (wie stark) in Verbindung stehen. Die Gewichtung wird entweder direkt vom Nutzer angegeben, indem der Nutzer ein Item bewertet, oder indirekt, indem das System die Nutzung des Items überwacht.

Wie beim Content Based Filtering kann die Gewichtung beispielsweise umso stärker sein, je öfter ein Item genutzt wird. Oder ein Item, welches gekauft wurde, wird stärker gewichtet als ein Item welches kostenlos heruntergeladen wurde. Berechnet wird dann auch nicht die Ähnlichkeit zwischen Nutzer-Modellen und Item-Modellen sondern ausschließlich die Ähnlichkeit von Nutzermodellen zueinander (User-User Matching). Hier gibt es wieder viele bekannte Verfahren. Im Wesentlichen wird bei allen Verfahren geprüft, welche Nutzermodelle möglichst viele Items in ähnlicher Gewichtung gemeinsam haben. Wurden nun ähnliche Nutzermodelle identifiziert, werden dem Nutzer 1 die Items empfohlen die mit dem ähnlichen Nutzer 2 in starker Verbindung stehen (und die Nutzer 1 gegebenenfalls noch nicht kennt).

Es ist auch möglich, die Nutzermodelle wie beim Content Based Filtering zu erzeugen und basierend auf diesen Nutzermodellen gleiche Nutzer zu identifizieren. Dieser Ansatz ist in **Fig. 4** gezeigt.

Zumindest beim Content Based Filtering (CBF) ist das Erstellen der Item-Modelle ein zentraler Bestandteil, da alles andere, d.h., das Erstellen der Nutzermodelle und das Matching von Nutzern und Items, hierauf basiert. Wie vorstehend erwähnt, ist ein häufig verwendetes Modell das Vector Space Model, welches ein textuelles Item als Vektor seiner Terme speichert, wobei die Länge des Vektors die Gewichtung des jeweiligen Terms in Bezug auf das Item repräsentiert. Um das Gewicht der Terme zu bestimmen, gibt es zahlreiche Verfahren. Nachteilig ist hierbei jedoch, dass mit diesen aus dem Stand der Technik bekannten Verfahren nur die Modellierung von "normalen" textuellen Items, also Dokumenten, wie Emails, Webseiten, Büchern, News Artikel, wissenschaftliche Artikel, etc., möglich ist.

Aufgabe der Erfindung

Aufgabe der Erfindung ist es daher, ein Verfahren und ein System bereitzustellen, welche es auf einfache Weise erlauben Item-Modelle bzw. Nutzermodelle auch für hierarchische, d.h. baumförmige Strukturen zu erzeugen, um basierend hierauf Nutzermodelle und Empfehlungen zu erstellen.

Erfindungsgemäße Lösung

Diese Aufgabe wird durch ein Verfahren und ein System gemäß den unabhängigen Ansprüchen gelöst. Vorteilhafte Ausgestaltungen der Erfindung sind in den jeweiligen abhängigen Ansprüchen angegeben.

Bereit gestellt wird demnach ein Verfahren zum Erzeugen eines Nutzermodells, insbesondere für einen Empfehlungsdienst, aus zumindest einer baumförmigen

Datenstruktur, wobei das Nutzermodell Informationen über einen Nutzer umfasst, wobei die zumindest eine baumförmige Datenstruktur dem Nutzer zuordenbar ist, wobei die baumförmige Datenstruktur einen Wurzelknoten und eine Anzahl von Kinderknoten umfasst, welche über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind, wobei zumindest einem Knoten zumindest ein Element zugeordnet ist, und wobei

- die den Knoten zugeordneten Elemente ermittelt werden, wobei die Elemente einen Inhalt des jeweiligen Knoten repräsentieren,
- die ermittelten Elemente gewichtet werden und jedem Element eine Elementgewichtung zugeordnet wird, und
- ein Nutzermodell generiert wird, wobei das generierte Nutzermodell die ermittelten Elemente und die dem jeweiligen Element zugeordnete Elementgewichtung umfasst.

Die Knoten der baumförmigen Datenstruktur können gewichtet werden und jedem Knoten kann eine Knotengewichtung zugeordnet werden.

In einem Initialisierungsschritt kann jedem Element eine vorbestimmte Elementgewichtung oder die Knotengewichtung des zugeordneten Knotens zugeordnet werden.

Das Verfahren kann ferner einen Vorverarbeitungsschritt umfassen, bei dem

- Knoten, denen keine Elemente zugeordnet sind, gelöscht werden, und/oder
- Knoten gelöscht werden, denen ein vorbestimmtes Element zugeordnet oder nicht zugeordnet ist, und/oder
- Knoten gelöscht werden, welche vorbestimmte Attribute aufweisen oder nicht aufweisen, und/oder
- Knoten und/oder Elemente der Knoten gelöscht werden, welche nicht direkt dem Nutzer zugeordnet sind.

Das Gewichten der Knoten kann eine statische Knotengewichtung und/oder eine dynamische Knotengewichtung umfassen, wobei

- bei der statischen Knotengewichtung die Anzahl der dem jeweiligen Knoten zugeordneten Kinderknoten, die Anzahl der jeweiligen Geschwisterknoten, die Tiefe des jeweiligen Knotens in der baumförmigen Datenstruktur, die Sichtbarkeit des Knotens, oder eine Kombination hiervon berücksichtigt werden, und
- bei der dynamischen Knotengewichtung für jeden Knoten das Alter, der Zeitpunkt der letzen Änderung, die Anzahl der Änderungen, die Anzahl der Verschiebungen innerhalb der baumförmigen Datenstruktur, die Anzahl der Markierungen, die Sichtbarkeit des Knotens, ein Dämpfungsfaktor, oder eine Kombination hiervon berücksichtigt werden.

Das Ermitteln der den Knoten zugeordneten Elemente kann ein Vorverarbeiten der ermittelten Elemente umfassen, wobei beim Vorverarbeiten der Elemente Text in Token und/oder Terme zerlegt wird, sofern das Element ein Textelement ist, und/oder Verweise verarbeitet werden, sofern das Element ein Verweiselement ist.

Die in dem Initialisierungsschritt den Elementen zugeordneten Elementgewichtungen können angepasst werden, wobei beim Anpassen der jeweiligen Elementgewichtung der Elementtyp, Attributsausprägungen der von dem Element zugeordneten Attribute, eine Häufigkeit des Elements innerhalb der baumförmigen Datenstruktur, die Anzahl der baumförmigen Datenstrukturen in einer Kollektion von baumförmigen Datenstrukturen in denen das Element vorkommt, eine Häufigkeit des Elements innerhalb einer Kollektion von baumförmigen Datenstrukturen, die Größe der baumförmigen Datenstruktur im Verhältnis zu anderen baumförmigen Datenstrukturen in einer Kollektion von baumförmigen Datenstrukturen, die Position des Elementes innerhalb des Knotens, die Sprache des Elementes, die Anzahl der Elemente innerhalb des Knotens, der Abstand des Elementes zu gleichartigen Elementen anderer Knoten, Häufigkeit des Elementes in dem Pfad

zwischen dem Knoten und dem Wurzelknoten, das Alter des Elements, der Zeitpunkt der letzen Änderung, die Anzahl der Änderungen, die Anzahl der Markierungen, die Sichtbarkeit des Elements, ein Dämpfungsfaktor, oder eine Kombination hiervon berücksichtigt werden.

Bei der statischen Knotengewichtung und/oder bei der dynamischen Knotengewichtung oder nach der statischen Knotengewichtung und/oder nach der dynamischen Knotengewichtung und/oder bei oder nach der Elementgewichtung kann eine Vererbung des Knotengewichts bzw. des Elementgewichts berücksichtigt werden.

Das generierte Nutzermodell kann in einer Speichereinrichtung gespeichert werden, um dem Empfehlungsdienst zur Verfügung gestellt zu werden.

Alle Elemente können zusammen mit den jeweiligen Elementgewichtungen als Nutzermodell gespeichert werden, oder für jeden Elementtyp kann ein eigenes Nutzermodell gespeichert werden, wobei die Nutzermodelle der verschiedenen Elementtypen ein Gesamtnutzermodell bilden.

Bei mehreren dem Nutzer zuordenbaren baumförmigen Datenstrukturen kann für jede baumförmige Datenstruktur eine Anzahl von Nutzermodellen generiert werden, welche zusammen ein dem Nutzer zugeordnetes Gesamtnutzermodell bilden.

Jeder baumförmigen Datenstruktur kann eine Baumgewichtung zugeordnet werden.

Für eine neue dem Nutzer zuordenbare baumförmige Datenstruktur kann das dem Nutzer zugeordnete Nutzermodell angepasst werden.

Von der baumförmigen Datenstruktur referenzierte Elemente können in das Nutzermodell eingefügt werden und wie Elemente der baumförmigen Datenstruktur behandelt werden.

Einem generierten Nutzermodell kann eine Information über den Nutzermodelltyp zugeordnet werden.

Das Verfahren kann ferner ein Auswählen von Objekten anhand vorbestimmter Auswahlkriterien umfassen, wobei ein Objekt ein Nutzermodell oder ein Itemmodell umfasst.

Die Auswahlkriterien können umfassen:

- Objekte eines vorbestimmten Typs, und/oder
- Objekte die eine vorbestimmte Ähnlichkeit zu dem Nutzermodell und/oder Itemmodell aufweisen, wobei vor dem Auswählen Ähnlichkeitswerte zwischen dem generierten Nutzermodell und/oder Itemmodell und den Objekten ermittelt werden.

Durch ein Itemmodell kann ein Förderprogramm repräsentiert werden, wobei das das Förderprogramm repräsentierende Itemmodell ausgewählt wird, wenn das Nutzermodell eine vorbestimmte Ähnlichkeit zu dem Itemmodell aufweist.

Bereit gestellt wird ferner ein System zum Erzeugen eines Nutzermodells, insbesondere für einen Empfehlungsdienst, aus zumindest einer baumförmigen Datenstruktur, wobei das Nutzermodell Informationen über einen Nutzer umfasst, wobei die baumförmige Datenstruktur dem Nutzer zuordenbar ist, wobei die baumförmige Datenstruktur einen Wurzelknoten und eine Anzahl von Kinderknoten umfasst, welche über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind, und wobei zumindest einem Knoten zumindest ein Element zugeordnet ist, wobei das System aufweist:

- wenigstens eine Speichereinrichtung zum Speichern wenigstens einer baumförmigen Datenstruktur, und
- eine Verarbeitungseinrichtung, welche mit der Speichereinrichtung gekoppelt ist und welche angepasst ist ein Verfahren nach einem der vorhergehenden Ansprüche auszuführen, um ein Nutzermodell zu generieren und das generierte Nutzermodell in der Speichereinrichtung abzuspeichern und einem Empfehlungsdienst zur Verfügung zu stellen.

Des Weiteren wird ein Datenträgerprodukt bereit gestellt, mit einem darauf gespeicherten Programmcode, welcher in einen Computer und / oder in ein Computernetzwerk ladbar ist und angepasst ist, ein erfindungsgemäßes Verfahren auszuführen.

Kurzbeschreibung der Figuren

Die Erfindung wird anhand eines Ausführungsbeispiels und der Zeichnung näher erläutert. In der Zeichnung zeigt:

Fig. 1	einen aus dem Stand der Technik bekannten Ansatz für ein sogenanntes "User-Item Matching";
Fig. 2	ein aus dem Stand der Technik bekanntes "Content Based Filtering" Verfahren;
Fig. 3	ein sogenanntes "Collaborative Filtering" Verfahren, wie es aus dem Stand der Technik bekannt ist;
Fig. 4	eine Abwandlung des aus dem Stand der Technik bekannten "Collaborative Filtering" Verfahrens;
Fig. 5	eine baumförmige (hierarchische) Datenstruktur;
Fig. 6a, 6b	zwei baumförmige Datenstrukturen, welche im Sinne der Erfindung die gleiche Aussage haben; und
Fig. 7a, 7b	ein Ablaufdiagramm eines erfindungsgemäßen Verfahrens.

Detaillierte Beschreibung der Erfindung

Definitionen

- Baumförmige Datenstruktur - Als baumförmige Datenstruktur (im Folgenden BD) wird eine Datenstruktur bezeichnet mit der sich eine Monohierarchie abbilden lässt. Dabei sind in der Datenstruktur Knoten mittels Kanten baumförmig verbunden. Es gibt genau einen Wurzelknoten, der beliebig viele Kinderknoten haben kann. Jeder Kinderknoten kann wiederum beliebig viele Kinderknoten haben.

  Beispiele für BD im Sinne der Erfindung sind vor allem, aber nicht ausschließlich, Verzeichnisstrukturen und/oder Dateisysteme auf einer Festplatte (Ordner und Dateien) oder sogenannte Mind Maps. Handelt es sich bei der BD um ein Dateisystem entsprechen die "Blätter" (das sind die jeweils letzten Knoten eines Pfades in einer BD) Dateien oder Dateiverknüpfungen und alle anderen Knoten entsprechen Verzeichnissen bzw. Ordnern.

  Knoten einer BD enthalten in der Regel ein oder mehrere Elemente. Diese Elemente können unterschiedlichen Typs sein. Übliche Elemente bzw. Elementtypen sind: Text (im Falle eines Dateisystems wäre der Knotentext der Datei- oder Verzeichnisname), zusätzliche Notizen, Tabellen, Termine, Multimediaobjekte (Musik, Film, Bild), Icons, Formeln, Verweise (in der Regel auf externe Items), Zahlen, und / oder Binärcode (insbesondere falls es sich bei der BD um eine Verzeichnisstruktur handelt und der Knoten eine Datei ist). Ein Verweis kann eine eindeutige URI (Uniform Resource Identifier) sein, z.B. Hyperlink, lokaler Link/Verknüpfung auf eine Datei auf einem Speichermedium (z.B. Festplatte). Ein Verweis kann aber auch eine nicht eindeutige Beschreibung sein, die ein Item identifiziert (z.B. Titel eines Dokumentes, Autorenname, Foto, BibTeX Key, Name eines Ortes oder Produktes).

Jedes der Elemente kann eine Anzahl von Attributen besitzen. So kann Text unterschiedlich formatiert sein, also hinsichtlich z.B. Größe und Farbe unterschiedliche Werte annehmen. Auch Knoten selbst können Attribute besitzen, insbesondere um die Anzeige der Knoten zu formatieren oder dem Knoten bestimmte Funktionen zuzuordnen. Beispielsweise können Knoten mittels Attributen als "eingeklappt" oder "ausgeklappt" dargestellt werden, das heißt für den Nutzer sichtbar oder unsichtbar sein. Genauso wie Knoten können einzelne Elemente eines Knotens sichtbar oder unsichtbar für den Nutzer sein.

Kanten in einer BD sind in der Regel ungerichtet und enthalten üblicherweise keine textuellen Informationen. Kanten können aber auch gerichtet sein.

- Item - Items sind beliebige Objekte, d.h., zum Beispiel Dokumente (Bücher, Webseiten, wissenschaftliche Artikel), Dateien, Werbeanzeigen (in Bild, Text, Ton), Personen, Musikstücke oder Musikalben, Filme, Produkte, geographische Orte, etc. oder deren digitale Repräsentation (d.h. nicht zwangsweise ein physisches Buch, sondern z.B. die digitale Kopie/Repräsentation des Buches in verschiedensten Formaten).

- Nutzer - Ein Nutzer ist eine Person die das erfindungsgemäße System anwendet bzw. nutzt. Ein Nutzer kann auch ein sogenannter Agent, eine Art elektronische Person bzw. ein System, welches das Verhalten einer realen Person simuliert.

- Nutzermodell - Ein Nutzermodell umfasst die Interessen, das Wissen oder andere Informationen über die Person, üblicherweise in maschinenlesbarer Form. Im Folgenden werden die Begriffe Interessen bzw. Wissen eines Nutzers bzw. Information über den Nutzer synonym verwendet.

- Verbindung zwischen BD und Nutzer - Eine BD steht mit einem Nutzer in Verbindung bzw. ist dem Nutzer zuordenbar, wenn dieser Nutzer z.B. die BD erstellt, editiert, heruntergeladen, oder geöffnet hat oder sich die BD im Besitz des Nutzers befindet oder befand (z.B. auf der Festplatte des Nutzers gespeichert ist bzw. war).

- Kollektion - Eine Kollektion ist die Menge aller BD auf die das erfindungsgemäße System Zugriff hat.

Erfindungsgemäß werden baumförmige Datenstrukturen BD analysiert, die mit dem Nutzer in Verbindung stehen bzw. einem Nutzer zuordenbar sind, um ein Modell des Nutzers, d.h. ein Nutzermodell, zu erstellen. Ein Nutzermodell umfasst insbesondere, aber nicht ausschließlich, Informationen über die Interessen und das Wissen des Nutzers.

Fig. 5 zeigt eine erfindungsgemäße baumförmige Datenstruktur BD. Eine baumförmige Datenstruktur BD umfasst eine Anzahl von Knoten, wobei ein spezieller Knoten den Root-Knoten bzw. den Wurzelknoten repräsentiert. Die anderen Knoten werden als Kinderknoten bezeichnet, wobei die Kinderknoten über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind. Knoten, welche keine Kinderknoten enthalten, werden als "Blätter" bezeichnet. Jeder Knoten kann einen oder mehrere Verweise auf externe Items enthalten. In Fig. 5 besitzt der Knoten 2.i einen solchen Verweis auf ein Item.

Erfindungsgemäß beschreiben der Inhalt einer baumförmige Datenstruktur und gegebenenfalls die Items die aus einer baumförmige Datenstruktur verlinkt werden bzw. die Inhalte der verlinkten Items die Interessen des Nutzers und können zur Generierung eines Nutzermodells verwendet werden. Vereinfacht gesagt bedeutet dies, wenn die Knoten einer baumförmigen Datenstruktur häufig das Wort "Patent" enthalten, kann daraus geschlossen werden, dass der Nutzer der baumförmige Datenstruktur bzw. der Nutzer, dem die baumförmige Datenstruktur zu-

ordenbar ist sich für Patente interessiert oder Kenntnisse auf diesem Gebiet hat. Der gleiche Schluss kann auch gezogen, wenn das Wort nicht in der baumförmige Datenstruktur selbst vorkommt, aber viele Dokumente (z.B. Patentschriften oder Webseiten) in der baumförmige Datenstruktur verlinkt sind, die das Wort "Patent" enthalten.

Fig. 7a und **Fig. 7b** zeigen ein Ablaufdiagramm eines erfindungsgemäßen Verfahrens zum Erzeugen eines Nutzermodells aus zumindest einer baumförmigen Datenstruktur.

In einem ersten Schritt findet eine Vorverarbeitung statt, bei der die baumförmigen Datenstrukturen für eine weitere Verarbeitung angepasst bzw. aufbereitet werden. Der Schritt der Vorverarbeitung ist ein optionaler Schritt und muss nicht notwendigerweise durchgeführt werden, etwa wenn die baumförmigen Datenstrukturen bereits das für die weitere Verarbeitung notwendige Format aufweisen.

Die Vorverarbeitung kann ein Konvertieren der baumförmigen Datenstrukturen in ein für das System lesbares Format umfassen. Ferner kann die Vorverarbeitung ein Löschen von Knoten aus den baumförmigen Datenstrukturen umfassen, d.h., es können bestimmte Knoten gelöscht werden, die für das Erzeugen eines Nutzermodells nicht relevant sind. Löschen eines Knoten heißt, dass dieser aus der baumförmige Datenstruktur entfernt wird und die Kinderknoten des zu löschenden Knotens entweder ebenfalls entfernt werden oder die Kinderknoten dem Elternknoten des zu löschenden Knotens zugeordnet werden. Ein Knoten kann etwa gelöscht werden, wenn der Knoten eines oder mehrere der folgenden Kriterien erfüllt:

- Der Knoten ist leer;
- Der Knoten enthält ein bestimmtes Element (nicht), wie z.B. Text oder Verweis;
- Der Knoten besitzt ein bestimmtes Attribut (nicht);

- Der Knoten oder Elemente des Knotens stehen nicht direkt mit dem Nutzer in Verbindung. Dies kann der Fall sein, wenn ein Knoten nicht vom Nutzer selbst erzeugt wurde (z.B. bei Knoten einer Mind Map die auf eine Datei verlinken und wo der Text des Knotens gleich dem Dateinamen der verlinkten Datei ist kann angenommen werden, dass der Knoten automatisch, z.B. durch "Drag & Drop" entstanden ist, also der Text des Knotens nicht vom Nutzer erzeugt wurde, und deshalb keine oder nur eine geringe Aussagekraft besitzt).

Selbstverständlich können beim Löschen eines Knotens auch noch weitere Kriterien berücksichtigt werden.

Nach der (optionalen) Vorverarbeitung wird in einem nächsten Schritt basierend auf der baumförmigen Datenstruktur ein Nutzermodell erzeugt. Hierbei werden die Knoten einer baumförmigen Datenstruktur bzw. die Elemente der Knoten analysiert, um die Interessen etc. des Nutzers zu identifizieren und in einem Nutzermodell zu speichern, welches dem Nutzer zugeordnet wird. Dies geschieht in den folgenden Schritten:

A) Gewichten der Knoten

Dem Gewichten der Knoten liegt die Annahme zugrunde, dass einige Knoten bzw. ihre Elemente aussagekräftiger sind um die Interessen des Nutzers zu beschreiben als andere Knoten bzw. ihre Elemente. Bei der Gewichtung der Knoten können zwei Teilgewichte berechnet werden. Es ist aber auch möglich nur eines der beiden Teilgewichte zu berechnen und dieses eine Teilgewicht als Knotengewicht eines Knotens zu betrachten. Die beiden Teilgewichte umfassen die statische Knotengewichtung und die dynamische Knotengewichtung. Selbstverständlich können auch noch weitere hier nicht genannte Teilgewichte berechnet werden. Die Kombination der berechneten Teilgewichte ergibt das Knotengewicht eines Knotens.

Bei der statischen Knotengewichtung können folgende Kriterien berücksichtigt werden:

- Anzahl der Kinderknoten: Ein Knoten wird abhängig von der Anzahl der Kinderknoten gewichtet, z.B. je mehr Kinder der Knoten hat desto mehr Gewicht erhält dieser Knoten.
- Anzahl der Geschwisterknoten: Geschwisterknoten eines Knotens sind jene Knoten, die denselben Elternknoten wie der betrachtete Knoten haben. Hier wird der Knoten abhängig von der Anzahl der Geschwisterknoten des Knotens gewichtet, z.B. je mehr Geschwister ein Knoten hat, desto weniger Gewicht bekommt er.
- Knoten-Tiefe: Je weiter oben (also je näher zum Wurzelknoten) in einer baumförmige Datenstruktur ein Knoten ist, desto mehr Gewicht erhält er. Der Wurzelknoten bekommt also viel Gewicht, die Blattknoten weniger.
- Sichtbarkeit des Knotens: Sichtbare Knoten erhalten mehr Gewicht als unsichtbare Knoten.
- Attribute: Ist der Knoten durch ein bestimmtes Attribut hervorgehoben, z.B. durch farbige Markierung oder Unterstreichung, so erhält er mehr Gewicht. Ist er durch bestimmte Attribute abgeschwächt, z.B. indem er ausgegraut oder durchgestrichen wurde, erhält er weniger Gewicht.

Bei der statischen Knotengewichtung wird die baumförmige Datenstruktur nur zu einem bestimmten Zeitpunkt betrachtet. Erfindungsgemäß können die Knoten aber auch dynamisch gewichtet, d.h., Veränderungen und Nutzungsintensivität der baumförmigen Datenstruktur über die Zeit können in die Knotengewichtung einfließen. Wenn beispielsweise ein Knoten intensiver genutzt wird, bzw. in der Vergangenheit intensiver genutzt wurde als andere Knoten, so kann dieser Knoten ein höheres Gewicht erhalten als die anderen Knoten. Die Gewichtung kann sich unter anderem ergeben aus:

- Alter des Knotens: Ältere Knoten können mehr oder weniger Gewicht erhalten als jüngere Knoten. Vorzugsweise erhalten jüngere Knoten ein höheres Gewicht. Es kann auch ein Schwellenwert vorgesehen sein, beispielsweise der Art "Knoten die mindestens 12 Stunden und maximal 5 Tage alt sind". Knoten, die das Schwellenwertkriterium erfüllen erhalten ein höheres Gewicht.
- Zeitpunkt der letzten Bearbeitung (z.B. Editierung): Knoten die kürzlich editiert wurden erhalten ein höheres Gewicht.
- Anzahl der Bearbeitungen: Vorzugsweise kann ein Knoten, der öfter editiert wurde als andere Knoten ein höheres Gewicht erhalten.
- Anzahl der Verschiebungen: Je öfter ein Knoten verschoben (ausgeschnitten und wieder eingefügt) wurde, desto mehr Gewicht erhält er.
- Anzahl der Markierungen: Je öfter ein Knoten ausgewählt/markiert wurde desto mehr Gewicht erhält er.
- Sichtbarkeitsdauer: Je länger ein Knoten sichtbar war, desto mehr Gewicht erhält er.
- Anzahl der Sichtbarkeiten: Je öfter ein Knoten unsichtbar und wieder sichtbar gemacht wurde, das heißt ein- und ausgeklappt wurde, desto stärker ist sein Gewicht.
- Anzahl der verfolgten Verweise: Je öfter ein Verweis eines Knoten geöffnet wurde, desto größer ist das Gewicht des Knoten.

Jede der vorstehend genannten (dynamischen) Gewichtungen kann mittels eines Zeitparameters geschwächt oder verstärkt werden. Hierzu ein Beispiel: Ein Knoten wird doppelt so stark gewichtet wenn er wenigstens zwei Mal editiert wurde. Liegt die letzte Editierung aber schon länger als X Wochen zurück, wird die Gewichtung durch einen dämpfenden Zeitparameter nur 1,5 Mal so stark gewichtet.

Zusätzlich zur vorstehend genannten statischen und/oder dynamischen Knotengewichtung kann eine Vererbung von Gewichten vorgesehen sein. Die Vererbung

wird vorzugsweise dann durchgeführt, nachdem die statische bzw. dynamische Gewichtung durchgeführt worden ist. Bei der Vererbung von Gewichten können Knoten Gewichtungen von ihren umliegenden Knoten "erben". Hat etwa ein Elternknoten ein sehr hohes Gewicht (weil er z.B. oft ausgewählt wurde), kann auch der Kindsknoten ein höheres Gewicht bekommen als wenn er nur für sich betrachtet würde. Bevorzugt bekommen alle Kinderknoten und deren Knoten, alle Geschwisterknoten und alle Elternknoten und deren Eltern bis zur Wurzel ein höheres Gewicht, wobei das zusätzliche Gewicht schwächer wird, je weiter entfernt von dem vererbenden Knoten sich der erbende Knoten befindet. Zudem kann vorgesehen sein, dass nur Knoten, die einen Schwellenwert (z.B. Gewichtung fünf Mal größer als normal) überschreiten das Gewicht an umliegende Knoten vererben können. Gehören mehrere Knoten einer baumförmige Datenstruktur einer bestimmten "Gruppe" an, kann die Gewichtung der einzelnen Knoten der Gruppe aneinander angeglichen bzw. vererbt werden. Gruppen können visuell in der baumförmige Datenstruktur erkennbar sein oder sich durch bestimmte Attribute bzw. Elementtypen auszeichnen. Beispiel: Eine baumförmige Datenstruktur enthält einige Knoten, die Verweise haben. Alle diese Knoten sind der Gruppe "Verweis-Knoten" zugeordnet. Obwohl nur 95% dieser Knoten eine sehr hohe Gewichtung haben, vergibt das System an alle Knoten (also auch an die restlichen 5%) eine sehr hohe Gewichtung. Die schwachen Knoten einer Gruppe erben quasi von ihren anderen Gruppenknoten.

B) Identifizieren der Elemente in den Knoten

In einem weiteren Schritt werden die Elemente in den Knoten identifiziert und ggf. einer Vorverarbeitung zugeführt. Hierbei werden zunächst die in jedem Knoten enthaltenden Elemente und deren Attribute identifiziert bzw. ermittelt.

Handelt es sich bei dem ermittelten Element um ein Textelement, so wird dieser Text weiter zerlegt und zwar in Token bzw. Terme (Begriffe). Häufig kann als Term ein einzelnes Wort gelten, manchmal aber auch zusammengesetzte Wörter

wie "Mind Map". Im Folgenden gilt jeder Term als eigenständiges Element vom Typ Text.

Die Terme können weiter verarbeitet werden. Hierfür können aus dem Stand der Technik bekannte Verfahren, etwa aus dem Bereich Information Retrieval herangezogen werden. Beispiele für solche Verfahren sind etwa

- Stemming: Wörter werden auf ihre Stämme reduziert. Beispielsweise würde das Wort "Stämme" auf "Stamm" gestemmt bzw. reduziert.
- Stop Word Removal: Sehr häufig vorkommende Wörter mit wenig Aussagekraft (beispielsweise der, die, das, wo, wer, weshalb, schon, so, darum, ...) werden entfernt.
- Latent Semantic Indexing (LSI): Mit Latent Semantic Indexing werden Synonyme von Wörtern zusammengefasst bzw. berücksichtigt.
- Translation: die Worte werden in eine Referenzsprache, z.B. Englisch, übersetzt.
- Spelling Correction: Rechtschreibfehler werden erkannt und korrigiert oder gelöscht.

Bei dem ermittelten Element kann es sich auch um einen Verweis handeln. Verweise können ebenfalls vorverarbeitet werden, indem beispielsweise für jeden Verweis die URI (Uniform Ressource Identifier) und/oder die Sonderzeichen auf ein einheitliches Format konvertiert werden oder falls es sich um keinen eindeutigen Verweis handelt (z.B. lediglich der Titel eines Dokumentes), versucht wird einen eindeutigen Identifikator zu finden (im Falle eines Dokumentes beispielsweise die ISBN).

C) Gewichten der Elemente

Ähnlich wie die Knoten kann auch jedes Element eines Knotens gewichtet werden. Besonders Text- und Verweiselemente sind wichtig für die Erstellung des erfindungsgemäßen Nutzermodells. Vorzugsweise erhält jedes Element zunächst eine vorbestimmte Gewichtung (Initialgewichtung), etwa die Gewichtung 1 oder die Gewichtung seines zugehörigen Knotens. Dies kann etwa in einem Initialisierungsschritt erfolgen, bei dem alle Elemente mit einer Initialgewichtung versehen werden.

Die Initialgewichtung eines Elementes kann verstärkt bzw. geschwächt werden, vorzugsweise basierend auf einen oder mehreren der folgenden Faktoren:

- Element-Typ: Elemente bestimmter Typen können unterschiedliche Basisgewichtungen erhalten. Beispielsweise kann ein Text-Element welches den allgemeinen Knotentext darstellt eine höhere Gewichtung erhalten als ein Text-Element welches eine zusätzliche Notiz darstellt.
- Attribute: Abhängig von den Attributen können Elemente eine stärkere oder schwächere Gewichtung bekommen. Beispielsweise kann ein Text-Element, welches fett formatiert ist stärker gewichtet werden als ein Text-Element ohne Formatierung.
- Element Frequenz: Je öfter ein Element in der baumförmigen Datenstruktur vorkommt, desto stärker kann seine Gewichtung sein.
- BD Frequenz: In je weniger baumförmigen Datenstrukturen der gesamten Kollektion ein Element vorkommt, desto stärker wird es gewichtet. Dies beruht auf der Annahme, dass ein Element, welches nur wenige Male in allen baumförmigen Datenstrukturen vorkommt, aussagekräftiger ist als ein Element, das in fast jeder baumförmigen Datenstruktur vorkommt. Enthält beispielsweise in einer Kollektion von 100 baumförmigen Datenstrukturen nur eine einzige baumförmige Datenstruktur den Term "Baum", dann würde dieser Term stärker gewichtet in Bezug auf die baumförmige Datenstruktur als wenn 90 weitere baumförmige Datenstrukturen diesen Term ebenfalls enthalten.

- Kollektionsfrequenz: Je seltener das Element in der Gesamtmenge aller Elemente der gesamten Kollektion vorkommt, desto stärker wird es gewichtet. Dies ist sehr ähnlich zur BD Frequenz, mit dem Unterschied, dass bei der BD Frequenz die Anzahl der baumförmigen Datenstrukturen gezählt wird in denen das Element vorkommt und bei er Kollektionsfrequenz die Gesamtanzahl der Elemente selbst.

- BD Größe: Je größer die baumförmige Datenstruktur, desto weniger stark wird das Element gewichtet. Dies unterliegt der Annahme, dass große baumförmige Datenstrukturen tendenziell mehr Elemente enthalten aber nicht gegenüber kleinen baumförmige Datenstrukturen bevorzugt werden sollen. Die Größe einer baumförmigen Datenstruktur kann angegeben werden durch die Anzahl der Knoten einer baumförmigen Datenstruktur oder durch die Anzahl der Elemente in einer baumförmigen Datenstruktur.

- Position im Knoten: Elemente die vorne im Knoten stehen werden anders gewichtet als Elemente weiter hinten im Knoten. Enthält ein Knoten beispielsweise 100 Terme, dann kann vorgesehen sein, dass nur die ersten 10 Terme berücksichtigt werden. Ferner kann vorgesehen sein, dass die weiteren Terme (z.B. die nächsten 10 Terme) mit weniger Gewicht berücksichtigt werden.

- Sprache (falls der Knoten ein Text-Elemente enthält): Im Gegensatz zu Dokumenten, wie Webseiten, kommt es bei baumförmigen Datenstrukturen häufig vor, dass Terme in verschiedenen Sprachen enthalten sind. Die Elemente eines Knotens können abhängig von der Sprache unterschiedlich stark gewichtet werden. Das heißt auch, dass wenn z.B. der Text eines Knotens in einer bestimmten Sprache ist, die anderen Elemente des Knotens (zum Beispiel ein Verweis) weniger oder mehr gewichtet werden.

- Knotenlänge: Elemente werden abhängig von der Knotenlänge gewichtet. Je weniger Elemente ein Knoten enthält, desto stärker können seine Elemente gewichtet werden.

- Abstand zu gleichartigen Elementen: Je weniger gleichartige Elemente es in der Nähe eines Knotens gibt zu dem das Element gehört, desto mehr Gewicht bekommt das Element. Zum Beispiel: Hat ein Knoten einen Verweis auf ein Item und die umliegenden Knoten (z.B. alle Kinder, Geschwister und Elternknoten) enthalten keine Verweise, dann könnte dieser Verweis ein besonders hohes Gewicht bekommen, da die Vermutung nahe liegt, dass sich der Verweis auch auf die umliegenden Knoten bezieht. Haben hingegen Geschwisterknoten ebenfalls Verweise, bekommt dieser Verweis kein besonders hohes Gewicht.

- Element-Wiederholung: Baumförmige Datenstrukturen können sehr benutzerspezifisch erstellt werden. Beispielsweise kann es vorkommen, dass ein Nutzer Elemente in den Knoten oft wiederholt, ein anderer Nutzer aber nicht. Hier kann es vorteilhaft sein, das Gewicht eines Elementes zu verringern, je öfter einer der Elternknoten (bis hoch zum Wurzelknoten) oder Geschwisterknoten dieses Element bereits enthält. Fig. 6a und Fig. 6b verdeutlichen diesen Fall. **Fig. 6a** und **Fig. 6b** zeigen jeweils eine baumförmige Datenstruktur mit der gleichen Aussage von zwei Nutzern, wobei die baumförmigen Datenstrukturen dennoch unterschiedlich aussehen. In Fig. 6a wiederholt sich der Term "Recommender" mehrfach, in Fig. 6b hingegen nicht. Trotzdem wäre der Term "Recommender" für beide baumförmigen Datenstrukturen bzw. Nutzer gleichermaßen zutreffend und sollte gleichermaßen gewichtet werden.

Die Gewichtung der Elemente kann auch mit den gleichen Verfahren stattfinden mit dem die Knoten gewichtet werden. Beispielsweise können ältere Elemente schwächer gewichtet werden als neuere und auch bei den Elementgewichtungen kann eine Vererbung stattfinden.

D) Speichern des Nutzermodells

In einer vorteilhaften Ausgestaltung der Erfindung kann es vorteilhaft sein, das generierte Nutzermodell zu speichern, um es etwa einem Empfehlungsdienst zur

Verfügung zu stellen. Alternativ kann ein Nutzermodell aber auch auf Anforderung erstellt werden, ohne es zu speichern.

Erfindungsgemäß können mindestens zwei verschiedene Ansätze genutzt werden, um ein Nutzermodell zu speichern. Die beiden hier gezeigten Ansätze sind das Typ-Neutrale Speichern und das Typ-Abhängige Speichern eines Nutzermodells.

Bei der Typ-Neutralen Speicherung werden alle Elemente mit ihrer Gewichtung gespeichert. Das heißt, Terme, Links/Verweise, Bilder, etc. werden alle gemeinsam in dem Model gespeichert. In einer konkreten Ausgestaltung des erfindungsgemäßen Verfahrens kann hierfür das eingangs beschriebene Vector Space Model genutzt werden, welches durch die Erfindung so erweitert wird, dass nicht nur Terme mit einer Gewichtung gespeichert werden können, sondern vielmehr beliebige Elemente verschiedenen Typs mit ihrer Gewichtung und ihrem Typ.

Bei der Typ-Abhängigen Speicherung kann für jeden Elementtyp ein separates Nutzermodell erzeugt werden, welche zusammen ein Gesamtnutzermodell bilden. Ein Nutzermodell umfasst dann beispielsweise ein Text-Modell und ein Verweis-Modell. Hierfür können Standardverfahren aus dem Information Retrieval Bereich bzw. User Modelling Bereich genutzt werden. Ein Standardmodell für ein Text-basiertes Modell wäre beispielsweise wieder das genannte Vector Space Model in dem die einzelnen Terme entsprechend des oben beschriebenen Verfahrens gewichtet sind. Verweise können auch in anderen Modellen gespeichert werden, die beispielsweise auch die Reihenfolge der Elemente in der baumförmigen Datenstruktur berücksichtigen.

E) Weitere Schritte

Die nachfolgenden Schritte können optional zu den vorstehend genannten Schritten ausgeführt werden.

Hat ein Nutzer Beziehungen zu mehreren baumförmigen Datenstrukturen, kann für jede baumförmige Datenstruktur ein (bzw. mehrere) Modelle erzeugt werden, wie vorstehend beschrieben, und die verschiedenen Modelle am Ende zu einem Gesamtmodell zusammen gefügt werden. Hierbei können unterschiedliche baumförmige Datenstrukturen mit unterschiedlicher Gewichtung versehen werden. Die Gewichtung erfolgt nach ähnlichen Prinzipien wie die Gewichtung der Knoten oder der Elemente. Beispielsweise kann eine neuere baumförmige Datenstruktur oder baumförmige Datenstrukturen, die häufiger geöffnet oder editiert wurden, stärker gewichtet werden.

Entsteht eine neue Beziehung zwischen einer baumförmigen Datenstruktur und einem Nutzer zu dem bereits ein Nutzermodell existiert, kann das bestehende Nutzermodell um die Elemente der neuen baumförmigen Datenstruktur erweitert werden.

Enthält eine baumförmige Datenstruktur Verweise auf Items, können diese Items ebenfalls für die Generierung eines Nutzermodells genutzt werden. Das heißt, die Elemente in dem verlinkten Item werden in das Nutzermodell eingefügt und zwar auf vergleichbare Weise wie Elemente der baumförmige Datenstruktur selbst. Diese Items können mit einer niedrigeren Gewichtung versehen werden. Ist das verlinkte Item eine baumförmige Datenstruktur, werden deren Elemente mit dem vorstehend beschriebenen Verfahren gewichtet. Ist das verlinkte Element z.B. eine Webseite, dann kann die Gewichtung mit Standardverfahren, wie dem TF-IDF durchgeführt werden.

Die vorstehend genannten Typ-Neutralen und Typ-Abhängigen Modelle können in Untermodelle unterteilt sein, beispielsweise in

- Modelle für Kurzzeitinteressen: Dieses Modell würde beispielsweise nur Daten aus einer Session bzw. der zuletzt editierten baumförmigen Datenstruktur enthalten (oder Daten der baumförmigen Datenstruktur, die in einem bestimmten Zeitraum editiert wurden).

- Modelle für Langzeitinteressen: Dieses Modell würde Interessen basierend auf allen oder zumindest mehreren baumförmigen Datenstrukturen enthalten.

- Modelle für verschiedene Interessen: Es ist denkbar, dass Nutzer verschiedene baumförmige Datenstrukturen erstellen für z.B. verschiedene Projekte. Das heißt, eine baumförmige Datenstruktur (oder auch mehrere) werden genutzt für Projekt A und eine andere baumförmige Datenstruktur (oder auch mehrere) für ein anderes Projekt B. Erfindungsgemäß können baumförmige Datenstrukturen, die sehr unterschiedlich sind für die Erstellung unterschiedlicher Modelle genutzt werden (die ggf. auch wieder unterteilt werden in Langzeit und Kurzzeitinteressen). Die Identifizierung von zusammengehörigen baumförmigen Datenstrukturen kann folgendermaßen erfolgen:

  - Inhaltliche Analyse: Hier werden mit den Verfahren zur Gewichtung von Termen oder Verweisen ähnliche baumförmige Datenstrukturen ermittelt. Unterschreiten die baumförmigen Datenstrukturen einen bestimmten Ähnlichkeits-Schwellenwert, werden sie für unterschiedliche Nutzermodelle genutzt.

  - Zeitliche Analyse: Baumförmige Datenstrukturen, die selten oder nie zur gleichen Zeit genutzt, geöffnet, etc. werden, werden für unterschiedliche Nutzermodelle genutzt.

Baumförmige Datenstrukturen können für unterschiedliche Arten von Anwendungen genutzt werden, zum Beispiel Dateiverwaltung, Brainstorming, Dokumentenmanagement, Projektplanung etc. Die Art der Anwendung wird im Nutzermodell vermerkt. Wenn ein Nutzer verschiedene baumförmige Datenstrukturen für verschiedene Arten von Anwendungen erstellt, werden wieder jeweils verschiedene Nutzermodelle erstellt. Die Art der Anwendung kann wie folgt festgestellt werden:

- Über die Anwendung, mit der die baumförmige Datenstruktur erstellt wurde. Beispielsweise kann pauschal angenommen werden, dass eine baumförmige

Datenstruktur, mit dem Windows Explorer erstellt worden ist zur Dateiverwaltung dient.

- Durch manuelle Angabe des Nutzers: Der Nutzer kann in der Anwendung zum Erstellen der baumförmigen Datenstruktur angeben für welchen Zweck er die baumförmige Datenstruktur erstellen will (z.B. Brainstorming, Projektplanung, etc.).
- Automatische Analyse: Das System analysiert die baumförmige Datenstruktur und schließt, z.B. an Hand ihres Aufbaus, ihrer Nutzung oder ihres Quellformates, automatisch auf die Art der Anwendung.

Bei der automatischen Analyse können folgende Regeln angewandt werden:

- Enthalten baumförmige Datenstrukturen viele Verweise, ist die primäre Anwendung Dateiverwaltung, Webseitenverwaltung bzw. Dokumentenverwaltung.
- Werden erst sehr schnell sehr viele Knoten in der baumförmigen Datenstruktur erstellt, diese dann verschoben und editiert und die baumförmige Datenstruktur danach nie oder nur selten geöffnet, wurde sie für Brainstorming erstellt.
- Wächst die baumförmige Datenstruktur langsam und kontinuierlich ist es keine baumförmige Datenstruktur für Brainstorming.

Bei der automatischen Analyse können Faktoren wie Wachstumsrate, Größe, Nutzungsdauer, Art der Nutzung, Anwendung zum Erstellen der baumförmigen Datenstruktur und/oder weitere Faktoren eine Rolle spielen.

Mit dem vorstehend beschriebenen erfindungsgemäßen Verfahren sind für eine Anzahl von Nutzer jeweils ein oder auch mehrere Nutzermodelle erzeugt worden. Diese Nutzermodelle werden erfindungsgemäß genutzt, um einem Nutzer Empfehlungen für Items zu geben. Das heißt, basierend auf den Nutzermodellen können Items identifiziert werden, die der Nutzer wahrscheinlich als interes-

sant/relevant empfindet. Um Items zu empfehlen, wird gemäß der Erfindung ein Verfahren vorgeschlagen, welches sowohl Item-Modelle als auch Nutzermodelle als gleich ansieht. Im Folgenden werden Items und Nutzer zusammengefasst als Objekt bezeichnet, wobei jedes Objekt einen Typ haben kann (Typ = Nutzer; Typ = Webseite; Typ = Email; Typ = Wissenschaftlicher Artikel; etc.).

Das erfindungsgemäße Verfahren zum Vorschlagen von Objekten kann zumindest die nachfolgend genannten Schritte umfassen:

Die Nutzermodelle wurden bereits mit vorstehen genannten erfindungsgemäßen Verfahren erstellt und gespeichert. Nun wird von allen Items die potentiell empfohlen werden können, ein Item-Modell mit aus dem Stand der Technik bekannten Verfahren erzeugt. Handelt es sich bei dem Item beispielsweise um eine Webseite könnten Terme mittels TF-IDF gewichtet und als Vector Space Model gespeichert werden. Handelt es sich bei dem Item um eine wissenschaftliche Arbeit könnte ebenfalls TF-IDF genutzt werden, aber auch andere Verfahren, wie etwa Citation Proximity Analysis um das Item abzubilden.

Es kann auch das vorstehend genannte Typ-Neutrale Verfahren genutzt werden, um ein entsprechendes Modell von Items zu erstellen die, wie baumförmige Datenstrukturen, mehrere Elementtypen enthalten. Dies trifft zum Beispiel auf wissenschaftliche Artikel zu. Eine wissenschaftliche Arbeit enthält in der Regel Text und Verweise (Referenzen auf andere wissenschaftlichen Arbeiten), vergleichbar zu einer baumförmigen Datenstruktur, die ebenfalls Text und Verweise (z.B. auf Dateien) enthält. Darum können beide relativ leicht mit kompatiblen Modelltypen abgebildet werden. Wichtig ist, dass erfindungsgemäß alle Objekte in dem gleichen bzw. einem kompatiblen Modell abgebildet werden, um diese später miteinander vergleichen zu können.

Nachdem von allen Objekten Modelle erstellt und gespeichert wurden, werden diese auf Ähnlichkeit verglichen. Enthalten die Objektmodelle mehrere Untermo-

delle (beispielsweise für Kurz- und Langzeitinteressen oder für verschiedene Elementtypen), wird jedes dieser Untermodelle mit den anderen Objektmodellen verglichen. Der Vergleich kann mit Standardverfahren stattfinden, wie Cosine für Vergleiche im Vector Space Model oder Ähnlichkeitsmaßen wie Greedy Citation Tiling für verweisbasierte Modelle. Das heißt letztlich, dass durch das Objekt-Objekt Matching die Ähnlichkeit der Nutzer zueinander, die Ähnlichkeit der Items zueinander und die Ähnlichkeit der Items und Nutzer zueinander in einem Schritt berechnet werden. Dies hat den Vorteil, dass einem Nutzer später sehr flexible Empfehlungen gegeben werden können.

Soll nun einem Nutzer eine Empfehlung gegeben werden, werden basierend auf seinem Objekt bzw. Nutzermodell:

- alle Objekte empfohlen, die einen bestimmten Typ haben (z.B. Webseite bzw. allgemein „Item“) oder ungleich eines bestimmten Typs sind (z.B. Nutzer) und einen gewissen Ähnlichkeitswert überschreiten;
- alle Objekte empfohlen, auf die in einem der im vorherigen Schritt bestimmten Objekte mit hohem Gewicht verwiesen wird. Wurde z.B. im vorherigen Schritt ein zum Nutzermodell ähnliches Objekt vom Typ Webseite bestimmt, werden dem Nutzer die Webseiten empfohlen die häufig auf der ähnlichen Webseite verlinkt sind. Oder, wurde ein ähnliches Objekt vom Typ "Nutzer" bestimmt werden die Items empfohlen, auf die häufig in dem Nutzermodell des ähnlichen Nutzers verwiesen wird bzw. die in enger Verbindung zu dem Nutzer stehen; und/oder
- alle Objekte empfohlen, die ähnlich zu den Objekten sind, auf die im Nutzermodell des Nutzers verwiesen wird. Das heißt: Hat ein Nutzer in seinem Nutzermodell einen Verweis auf eine Webseite X, werden die Objekte empfohlen die ähnlich zu dieser Webseite X sind.

Zusätzlich bzw. alternativ zu dem vorstehend beschriebenen Ansatz kann ein Verfahren basierend auf Machine Learning genutzt werden. Die Verweise eines Nut-

zermodells können mit Machine Learning Verfahren genutzt werden, um die Präferenzen von Nutzern zu lernen. Hierbei gilt jeder Verweis auf ein Item als positive Assoziation, welche das System lernt und darauf basierend Empfehlungen für neue Items gibt.

In einer speziellen Ausführung des erfindungsgemäßen Verfahrens kann dieses verwendet werden, um Empfehlungen für Fördergelder/Förderprogramme zu geben. Hierzu wird zunächst ein aus einer baumförmigen Datenstruktur ein Nutzermodell erzeugt, wie vorstehend beschrieben. Das Förderprogramm selbst wird als Item betrachtet. Das Item wiederum wird repräsentiert durch einen Text welcher das Förderprogramm beschreibt. Dieser Text kann eine Webseite sein, eine Broschüre im PDF Format, Social Tags, etc. Enthält beispielsweise ein Nutzermodell den stark gewichteten Term "Recommender Systems" und gibt es ein Förderprogramm, dessen Webseite ebenfalls diesen Term oft beinhaltet, würde dieses Förderprogramm dem Nutzer empfohlen.

Durch die erfindungsgemäßen Nutzermodelle und Empfehlungen kann der Nutzen vieler Softwareprogramme für den Anwender gesteigert werden, da sie interessante Empfehlungen erhalten.

Ansprüche

1. Computer-implementiertes Verfahren zum Erzeugen eines Nutzermodells, insbesondere für einen Empfehlungsdienst, aus zumindest einer baumförmigen Datenstruktur, wobei das Nutzermodell Informationen über einen Nutzer umfasst, wobei die zumindest eine baumförmige Datenstruktur dem Nutzer zuordenbar ist, wobei die baumförmige Datenstruktur einen Wurzelknoten und eine Anzahl von Kinderknoten umfasst, welche über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind, wobei zumindest einem Knoten zumindest ein Element zugeordnet ist, und wobei

   - die den Knoten zugeordneten Elemente ermittelt werden, wobei die Elemente einen Inhalt des jeweiligen Knoten repräsentieren,
   - die ermittelten Elemente gewichtet werden und jedem Element eine Elementgewichtung zugeordnet wird, und
   - ein Nutzermodell generiert wird, wobei das generierte Nutzermodell die ermittelten Elemente und die dem jeweiligen Element zugeordnete Elementgewichtung umfasst.

2. Verfahren nach Anspruch 1, wobei die Knoten der baumförmigen Datenstruktur gewichtet werden und jedem Knoten eine Knotengewichtung zugeordnet wird.

3. Verfahren nach einem der vorhergehenden Ansprüche, wobei in einem Initialisierungsschritt jedem Element eine vorbestimmte Elementgewichtung oder die Knotengewichtung des zugeordneten Knotens zugeordnet wird.

4. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Verfahren ferner einen Vorverarbeitungsschritt umfasst, bei dem

   - Knoten, denen keine Elemente zugeordnet sind, gelöscht werden, und/oder
   - Knoten gelöscht werden, denen ein vorbestimmtes Element zugeordnet oder nicht zugeordnet ist, und/oder
   - Knoten gelöscht werden, welche vorbestimmte Attribute aufweisen oder nicht aufweisen, und/oder
   - Knoten und/oder Elemente der Knoten gelöscht werden, welche nicht direkt dem Nutzer zugeordnet sind.

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Gewichten der Knoten eine statische Knotengewichtung und/oder eine dynamische Knotengewichtung umfasst, wobei

   - bei der statischen Knotengewichtung die Anzahl der dem jeweiligen Knoten zugeordneten Kinderknoten, die Anzahl der jeweiligen Geschwisterknoten, die Tiefe des jeweiligen Knotens in der baumförmigen Datenstruktur, die Sichtbarkeit des Knotens, oder eine Kombination hiervon berücksichtigt werden, und
   - bei der dynamischen Knotengewichtung für jeden Knoten das Alter, der Zeitpunkt der letzen Änderung, die Anzahl der Änderungen, die Anzahl der Verschiebungen innerhalb der baumförmigen Datenstruktur, die Anzahl der Markierungen, die Sichtbarkeit des Knotens, ein Dämpfungsfaktor, oder eine Kombination hiervon berücksichtigt werden.

6. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Ermitteln der den Knoten zugeordneten Elemente ein Vorverarbeiten der ermittelten Elemente umfasst, wobei beim Vorverarbeiten der Elemente Text in Token und/oder Terme zerlegt wird, sofern das Element ein Textelement ist,

und/oder Verweise verarbeitet werden, sofern das Element ein Verweiselement ist.

7. Verfahren nach einem der vorhergehenden Ansprüche, wobei die in dem Initialisierungsschritt den Elementen zugeordneten Elementgewichtungen angepasst werden, wobei beim Anpassen der jeweiligen Elementgewichtung der Elementtyp, Attributsausprägungen der von dem Element zugeordneten Attribute, eine Häufigkeit des Elements innerhalb der baumförmigen Datenstruktur, die Anzahl der baumförmigen Datenstrukturen in einer Kollektion von baumförmigen Datenstrukturen in denen das Element vorkommt, eine Häufigkeit des Elements innerhalb einer Kollektion von baumförmigen Datenstrukturen, die Größe der baumförmigen Datenstruktur im Verhältnis zu anderen baumförmigen Datenstrukturen in einer Kollektion von baumförmigen Datenstrukturen, die Position des Elementes innerhalb des Knotens, die Sprache des Elementes, die Anzahl der Elemente innerhalb des Knotens, der Abstand des Elementes zu gleichartigen Elementen anderer Knoten, Häufigkeit des Elementes in dem Pfad zwischen dem Knoten und dem Wurzelknoten, das Alter des Elements, der Zeitpunkt der letzen Änderung, die Anzahl der Änderungen, die Anzahl der Markierungen, die Sichtbarkeit des Elements, ein Dämpfungsfaktor, oder eine Kombination hiervon berücksichtigt werden.

8. Verfahren nach einem der Ansprüche 5 bis 7, wobei bei der statischen Knotengewichtung und/oder bei der dynamischen Knotengewichtung oder nach der statischen Knotengewichtung und/oder nach der dynamischen Knotengewichtung und/oder bei oder nach der Elementgewichtung die Vererbung des Knotengewichts bzw. des Elementgewichts berücksichtigt werden.

9. Verfahren nach einem der vorhergehenden Ansprüche, wobei das generierte Nutzermodell in einer Speichereinrichtung gespeichert wird, um dem Empfehlungsdienst zur Verfügung gestellt zu werden.

10. Verfahren nach Anspruch 9, wobei alle Elemente zusammen mit den jeweiligen Elementgewichtungen als Nutzermodell gespeichert werden, oder wobei für jeden Elementtyp ein eigenes Nutzermodell gespeichert wird, wobei die Nutzermodelle der verschiedenen Elementtypen ein Gesamtnutzermodell bilden.

11. Verfahren nach einem der vorhergehenden Ansprüche, wobei bei mehreren dem Nutzer zuordenbaren baumförmigen Datenstrukturen für jede baumförmige Datenstruktur eine Anzahl von Nutzermodellen generiert wird, welche zusammen ein dem Nutzer zugeordnetes Gesamtnutzermodell bilden.

12. Verfahren nach Anspruch 11, wobei jeder baumförmigen Datenstruktur eine Baumgewichtung zugeordnet wird.

13. Verfahren nach Anspruch 11 oder 12, wobei für eine neue dem Nutzer zuordenbare baumförmige Datenstruktur das dem Nutzer zugeordnete Nutzermodell angepasst wird.

14. Verfahren nach einem der vorhergehenden Ansprüche, wobei von der baumförmigen Datenstruktur referenzierte Elemente in das Nutzermodell eingefügt werden und wie Elemente der baumförmigen Datenstruktur behandelt werden.

15. Verfahren nach einem der vorhergehenden Ansprüche, wobei einem generierten Nutzermodell eine Information über den Nutzermodelltyp zugeordnet wird.

16. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Verfahren ferner ein Auswählen von Objekten anhand vorbestimmter Auswahlkriterien umfasst, wobei ein Objekt ein Nutzermodell oder ein Itemmodell umfasst.

17. Verfahren nach Anspruch 16, wobei die Auswahlkriterien umfassen:

    - Objekte eines vorbestimmten Typs, und/oder

    - Objekte die eine vorbestimmte Ähnlichkeit zu dem Nutzermodell und/oder Itemmodell aufweisen, wobei vor dem Auswählen Ähnlichkeitswerte zwischen dem generierten Nutzermodell und/oder Itemmodell und den Objekten ermittelt werden.

18. Verfahren nach einem der vorhergehenden Ansprüche, wobei durch ein Itemmodell ein Förderprogramm repräsentiert wird, und wobei das das Förderprogramm repräsentierende Itemmodell ausgewählt wird, wenn das Nutzermodell eine vorbestimmte Ähnlichkeit zu dem Itemmodell aufweist.

19. System zum Erzeugen eines Nutzermodells, insbesondere für einen Empfehlungsdienst, aus zumindest einer baumförmigen Datenstruktur, wobei das Nutzermodell Informationen über einen Nutzer umfasst, wobei die baumförmige Datenstruktur dem Nutzer zuordenbar ist, wobei die baumförmige Datenstruktur einen Wurzelknoten und eine Anzahl von Kinderknoten umfasst, welche über Kanten mit dem Wurzelknoten oder mit einem Kinderknoten verbunden sind, und wobei zumindest einem Knoten zumindest ein Element zugeordnet ist, aufweisend

    - wenigstens eine Speichereinrichtung zum Speichern wenigstens einer baumförmigen Datenstruktur, und

    - eine Verarbeitungseinrichtung, welche mit der Speichereinrichtung gekoppelt ist und welche angepasst ist ein Verfahren nach einem der vorhergehenden Ansprüche auszuführen, um ein Nutzermodell zu generieren und das generierte Nutzermodell in der Speichereinrichtung abzuspeichern und einem Empfehlungsdienst zur Verfügung zu stellen.

20. Datenträgerprodukt mit einem darauf gespeicherten Programmcode, welcher in einen Computer und / oder in ein Computernetzwerk ladbar ist und angepasst ist, ein Verfahren nach einem der Ansprüche 1 bis 18 auszuführen.

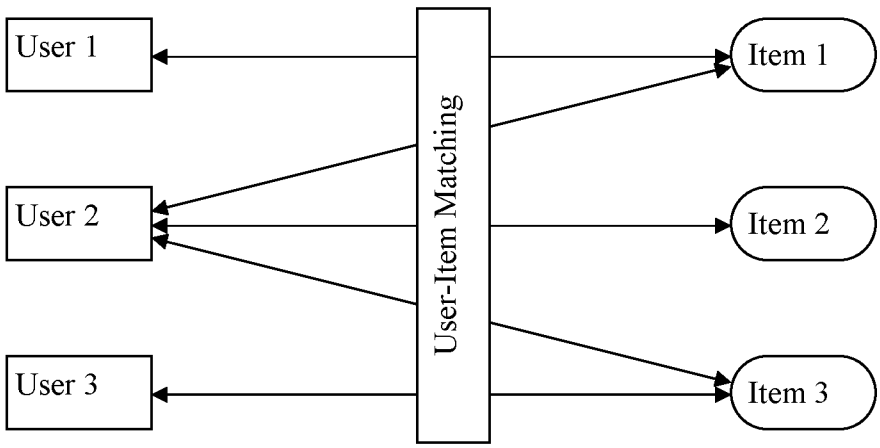

Fig. 1
(Stand der Technik)

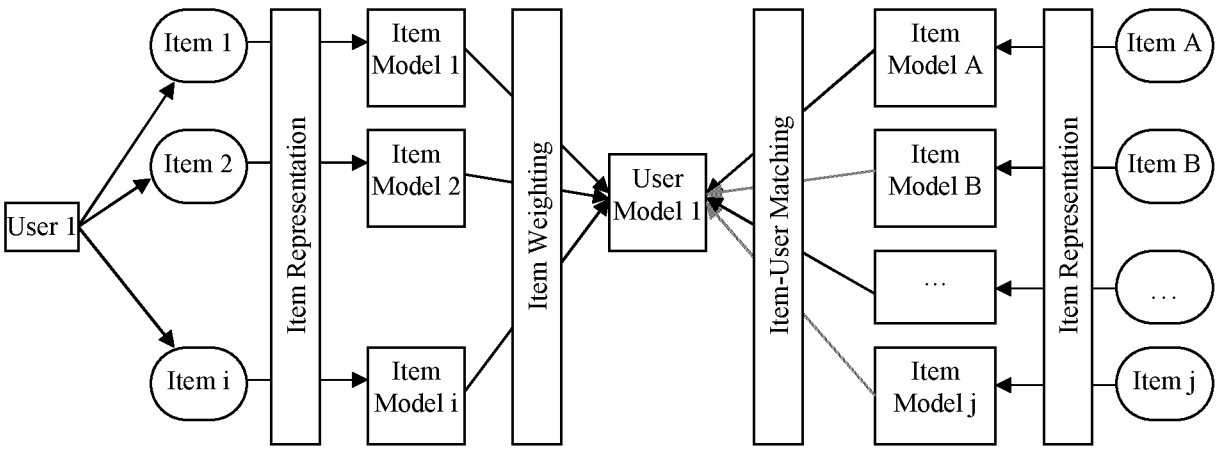

Fig. 2
(Stand der Technik)

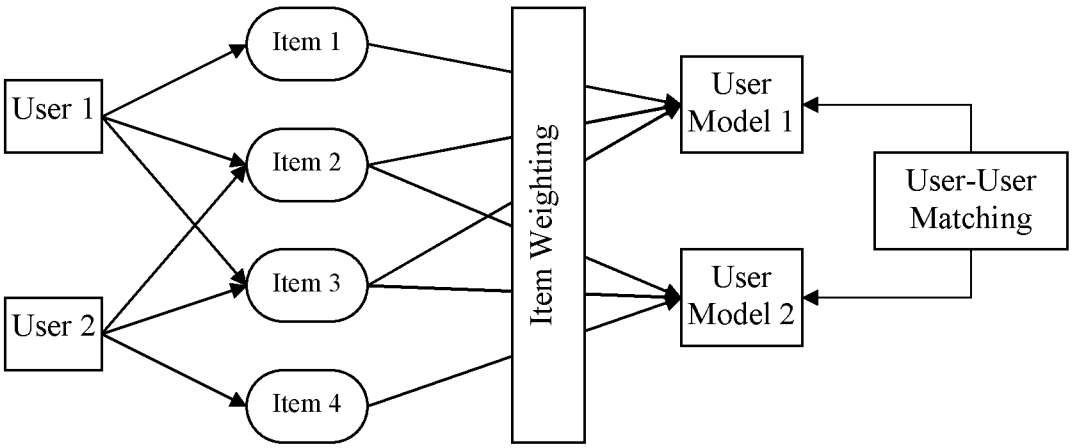

Fig. 3
(Stand der Technik)

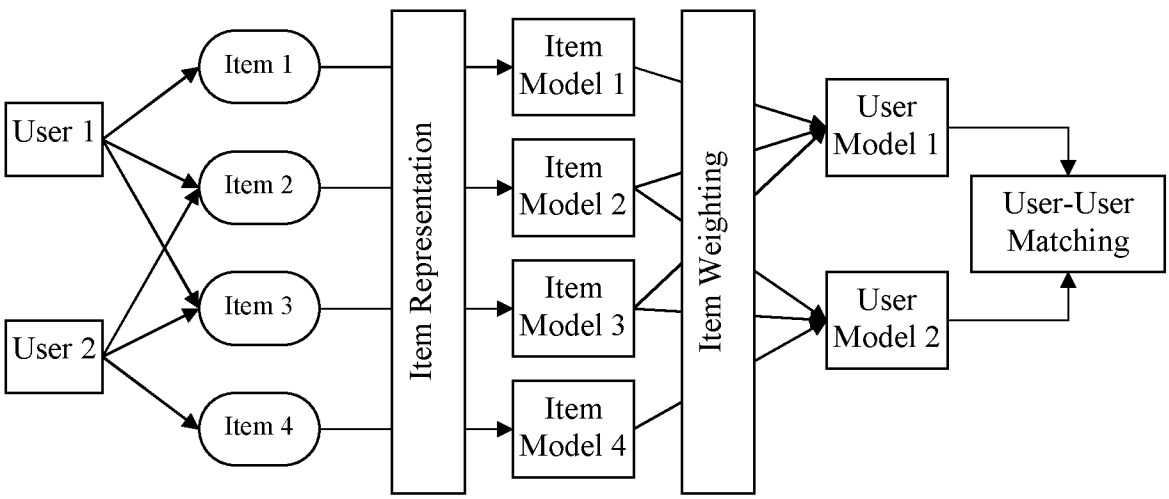

Fig. 4
(Stand der Technik)

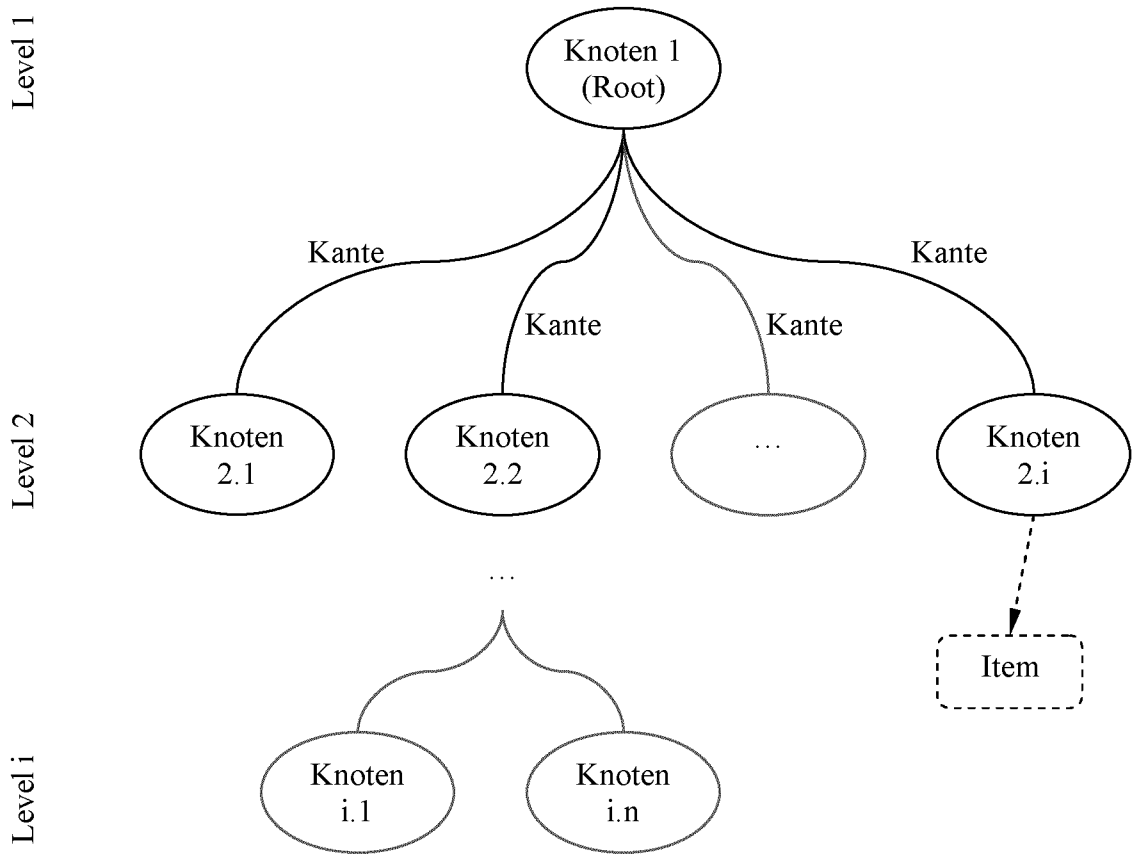
Level 1
Level 2
Level i
Knoten 1
(Root)
Kante
Kante
Kante
Kante
Knoten
2.1
Knoten
2.2
…
Knoten
2.i
…
Item
Knoten
i.1
Knoten
i.n

Fig. 5

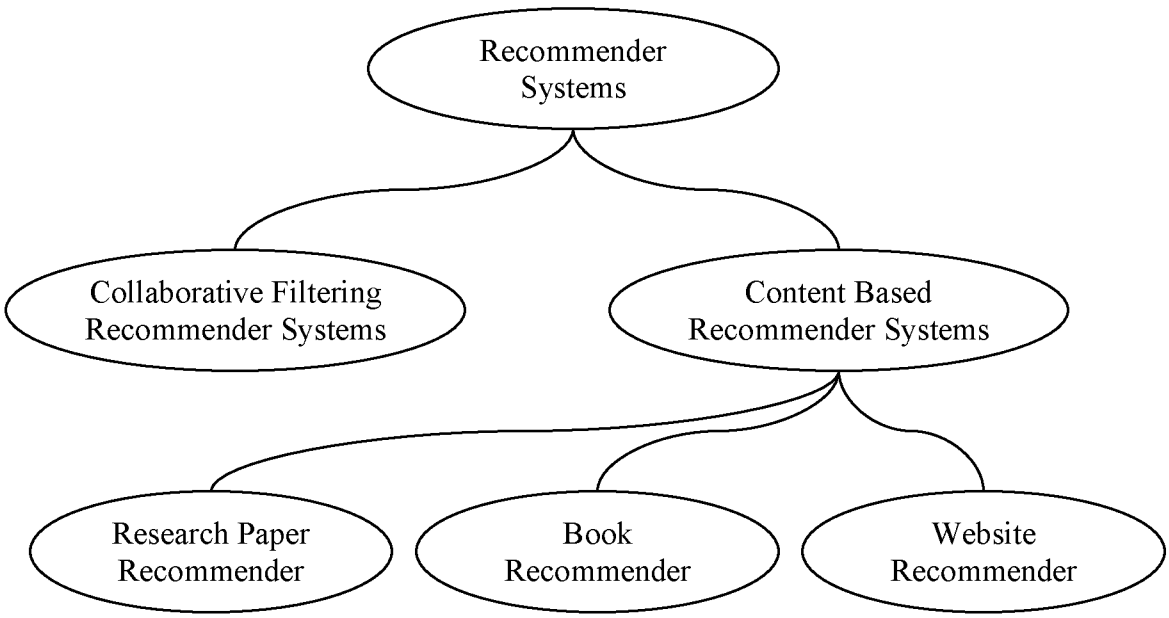
Recommender Systems
Collaborative Filtering Recommender Systems
Content Based Recommender Systems
Research Paper Recommender
Book Recommender
Website Recommender

Fig. 6a

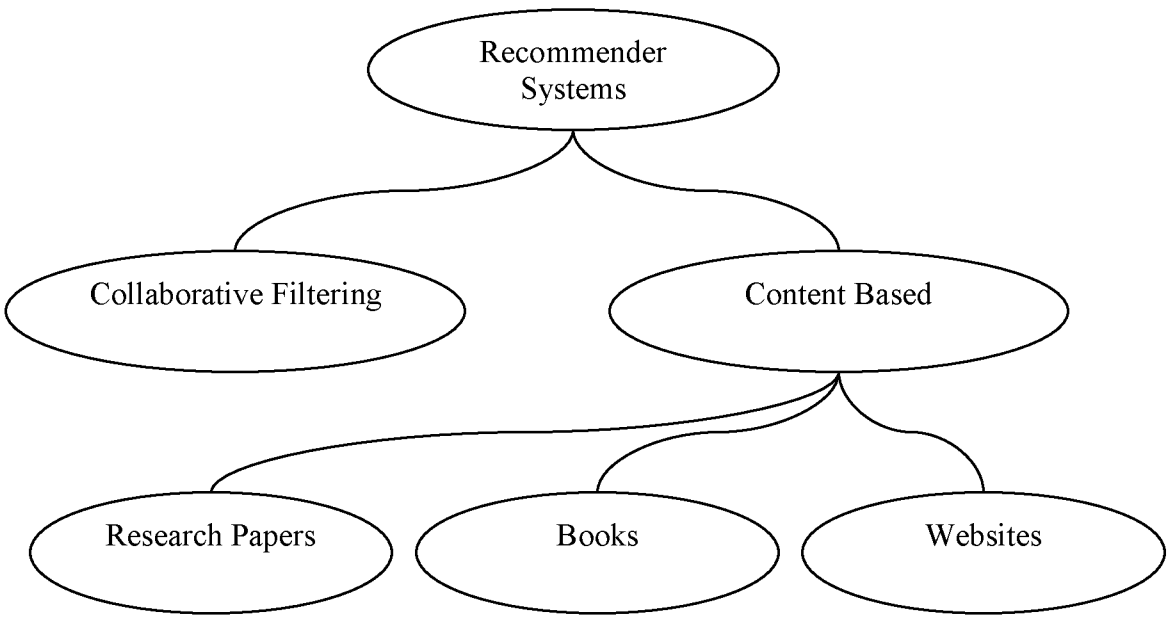
Recommender Systems
Collaborative Filtering
Content Based
Research Papers
Books
Websites

Fig. 6b

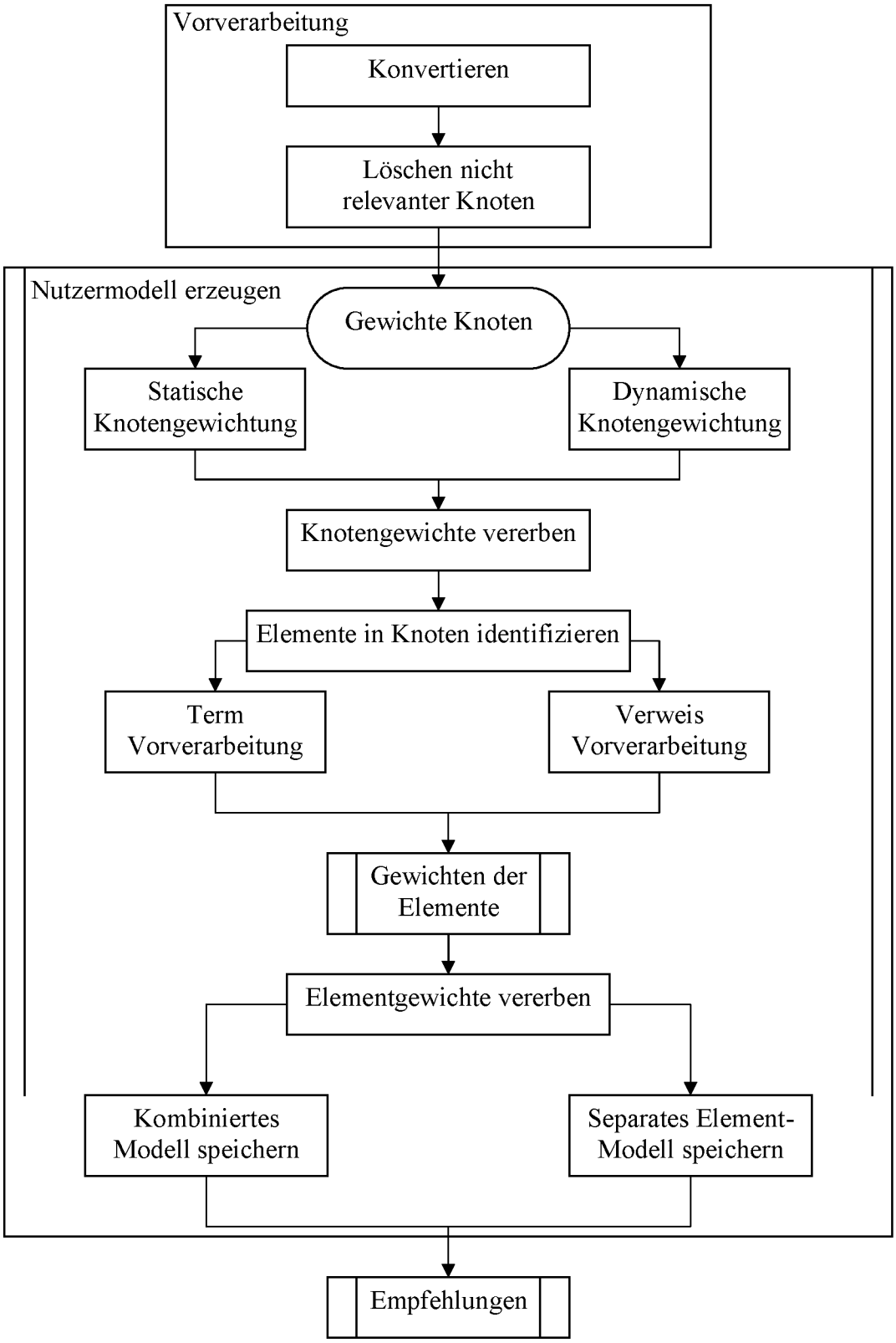
Vorverarbeitung
Konvertieren
Löschen nicht relevanter Knoten
Nutzermodell erzeugen
Gewichte Knoten
Statische Knotengewichtung
Dynamische Knotengewichtung
Knotengewichte vererben
Elemente in Knoten identifizieren
Term Vorverarbeitung
Verweis Vorverarbeitung
Gewichten der Elemente
Elementgewichte vererben
Kombiniertes Modell speichern
Separates Element-Modell speichern
Empfehlungen

Fig. 7a

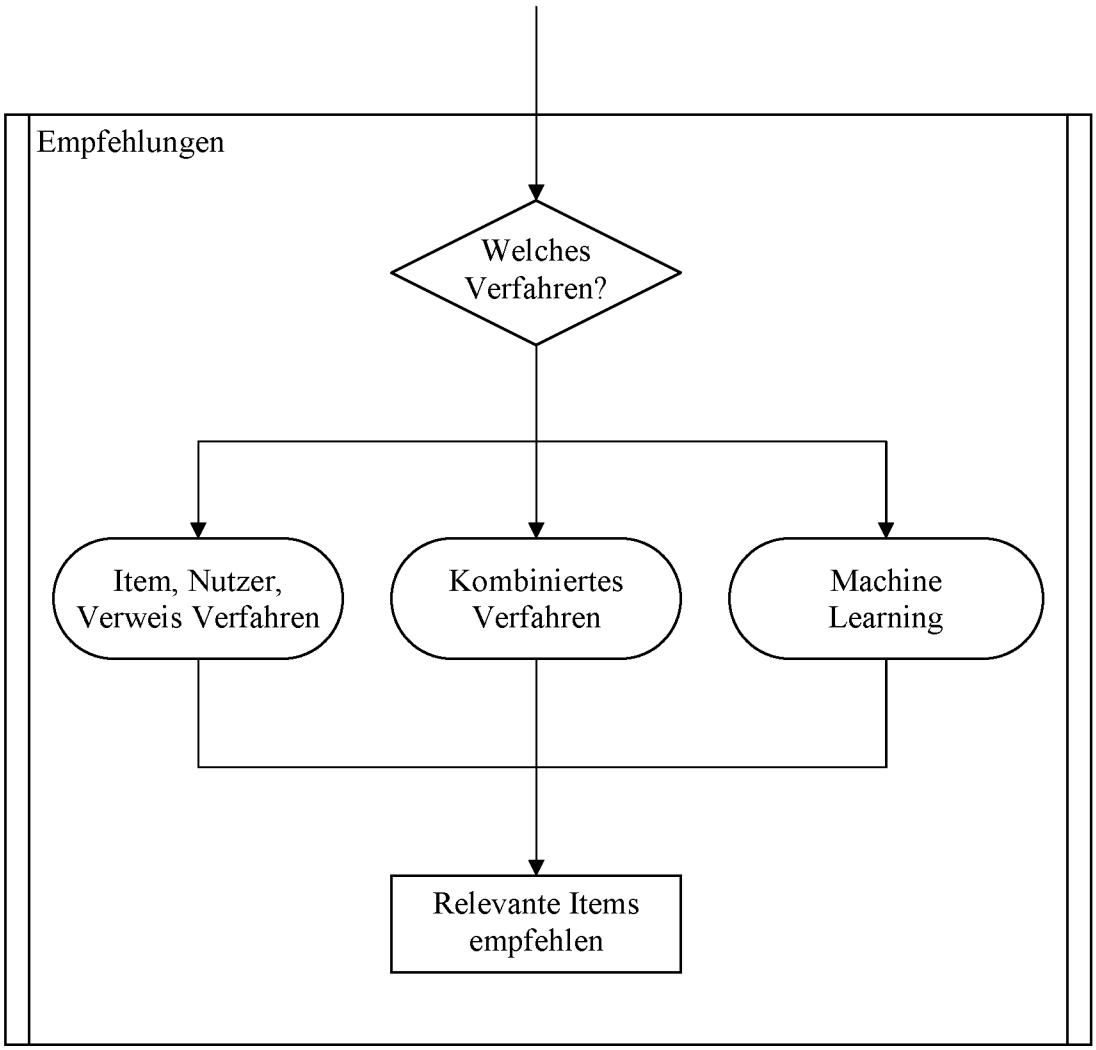
Empfehlungen
Welches Verfahren?
Item, Nutzer, Verweis Verfahren
Kombiniertes Verfahren
Machine Learning
Relevante Items empfehlen

Fig. 7b

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/070873

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06Q30/00 G06F17/30
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
L	Die einzigen technischen Aspekte der beanspruchten Erfindung beziehen sich auf allgemein gebräuchliche Datenverarbeitungstechnologie. Aufgrund seiner weiten Verbreitung kann die allgemeine Bekanntheit eines solchen Standes der Technik nicht bezweifelt werden. Siehe Amtsblatt 11/2007, S. 592. -----	1-20

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 December 2011	02/01/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Laub, Christoph

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/EP2011/070873

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
INV. G06Q30/00 G06F17/30
ADD.

Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
G06Q

Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)
EPO-Internal, WPI Data

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
L	Die einzigen technischen Aspekte der beanspruchten Erfindung beziehen sich auf allgemein gebräuchliche Datenverarbeitungstechnologie. Aufgrund seiner weiten Verbreitung kann die allgemeine Bekanntheit eines solchen Standes der Technik nicht bezweifelt werden. Siehe Amtsblatt 11/2007, S. 592. -----	1-20

☐ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen ☐ Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :
"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist
"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist
"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche	Absendedatum des internationalen Recherchenberichts
19. Dezember 2011	02/01/2012

Name und Postanschrift der Internationalen Recherchenbehörde	Bevollmächtigter Bediensteter
Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Laub, Christoph