(19) **Bundesrepublik Deutschland**
Deutsches Patent- und Markenamt

(10) **DE 699 33 187 T2** 2007.09.13

(12) # Übersetzung der europäischen Patentschrift

(97) **EP 1 003 111 B1**
(21) Deutsches Aktenzeichen: **699 33 187.0**
(96) Europäisches Aktenzeichen: **99 122 074.0**
(96) Europäischer Anmeldetag: **16.11.1999**
(97) Erstveröffentlichung durch das EPA: **24.05.2000**
(97) Veröffentlichungstag
der Patenterteilung beim EPA: **13.09.2006**
(47) Veröffentlichungstag im Patentblatt: **13.09.2007**

(51) Int Cl.[8]: ***G06F 17/30*** (2006.01)

(30) Unionspriorität:
32894098 19.11.1998 JP

(84) Benannte Vertragsstaaten:
DE, FR, GB, NL

(73) Patentinhaber:
Hitachi, Ltd., Tokyo, JP

(74) Vertreter:
BEETZ & PARTNER Patentanwälte, 80538 München

(72) Erfinder:
Iwayama, Ltd.,Intell. Prop. Group, Makoto Hitachi, Chiyoda-ku, Tokyo 100-8220, JP; Nishioka, Hitachi, Shingo, Chiyoda-ku, Tokyo 100-8220, JP; Niwa, Hitachi, Yoshiki, Chiyoda-ku, Tokyo 100-8220, JP; Takano, Hitachi, Akihiko, Chiyoda-ku, Tokyo 100-8220, JP

(54) Bezeichnung: **Dokumentensuchverfahren und Dienst**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

Hintergrund der Erfindung

[0001] Diese Erfindung bezieht sich auf ein Dokumentensuchverfahren, welches zwischen mehreren Dokumentdatenbanken wechselt und Beziehungen zwischen mehreren Dokumentdatenbanken bildet.

[0002] Da immer mehr Informationen über Dokumente in elektronische Formate konvertiert werden, erwächst ein größerer Bedarf, verschiedene Arten von Dokumentdatenbanken gleichzeitig zu durchsuchen. Beispielsweise möchten Nutzer oft Verzeichnisse einsehen, die sich auf Zeitungsartikel beziehen, für die sich interessieren können.

[0003] In der Vergangenheit war es möglich, eine unabhängige Suche durch den Wechsel zwischen mehreren Dokumentdatenbanken durchzuführen, aber es gab keine Möglichkeit, die Bedeutung eines Dokumentensatzes in anderen Datenbanken zu einem Dokumentensatz in einer bestimmten Datenbank zu überprüfen.

[0004] In Voorhees et al.: „Learning Collection Fusion Strategies", SIGIR '95, Seattle, WA, 1995, Seiten 172-179, ISBN 0-89791-714-6 wird ein Verfahren dargestellt, in dem die Suche auf dieselbe Datenbank beschränkt wird, so dass es möglich ist, andere Dokumentensätze innerhalb dieser Datenbank zu suchen. In diesem Fall wird eine ausreichende Suchgeschwindigkeit oft dadurch erreicht, dass die Bedeutung zwischen Dokumenten vor der Suche berechnet wird. Selbst mit verschiedenen Datenbanken ist es möglich, mehrere Dokumentdatenbanken gleichzeitig zu durchsuchen, falls solch eine Berechnung im Voraus ausgeführt wird. Da jedoch der Berechnungsaufwand mit der Anzahl an Datenbanken aufgrund der wachsenden Anzahl an Kombinationen wächst, ist dieses Verfahren nicht realistisch.

[0005] Es ist ebenso möglich, erst den Satz Schlüsseldokumente auf Seite des Nutzers zu analysieren, um eine Suchanfrage zu erstellen und dann in anderen Dokumentdatenbanken unter Verwendung dieser Anfrage zu suchen. In diesem Fall muss jedoch die Nutzerseite alle Informationen über den Satz Schlüsseldokumente erhalten, wobei die Datenmenge riesig wird, wenn sich die Dokumentdatenbanken in einem Netzwerk befinden.

[0006] In Velez B. et al.: „Fast and effective query refinement", Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, 27. Juli 1997, Seiten 6-15, wird eine Abfrageverfeinerungstechnik für ein Dokumentensuchverfahren gezeigt, welches automatisch Begriffe vorschlägt, die semantisch mit dem Informationsbedarf des Nutzers in Beziehung stehen. Diese Literaturstelle nennt jedoch nicht die Verwendung mehrerer Dokumentdatenbanken und zeigt somit nicht, dass die Suchergebnisse von einer Dokumentdatenbank als Anfrage zur Suche in einer anderen Datenbank verwendet werden.

Zusammenfassung der Erfindung

[0007] Es ist daher eine Aufgabe der Erfindung, die der bestehenden Technologie innewohnenden Probleme dadurch zu lösen, dass einem Benutzer ermöglicht wird, einen beliebigen Satz von Dokumenten in einer beliebigen Dokumentendatenbank anzugeben und effizient Dokumentensätze innerhalb irgendeiner bestimmten Datenbank zu suchen, die zu diesem Dokumentensatz in Bezug stehen, während der Netzwerkverkehr reduziert wird.

[0008] Die oben beschriebene Aufgabe wird durch ein Dokumentensuchverfahren entsprechend Anspruch 1 gelöst. Die abhängigen Ansprüche beziehen sich auf bevorzugte Ausführungsformen.

[0009] Im Falle einer großen Suchanfrage, wie z. B. in dem Fall eines Satzes von Schlüsseldokumenten, ist es schneller, anstelle der Verwendung aller Informationen in der Suchanfrage, eine Suche nur mit Schlagwörtern der Sucheingabe als Zusammenfassung durchzuführen, wobei auch die Netzwerklast reduziert wird. Im Sinne dieser Beschreibung bedeutet „Zusammenfassung" „Satz von Schlagwörtern für einen Dokumentensatz".

[0010] Die Dokumentdatenbanken befinden sich auf Servern in einem Netzwerk, die ein Modul zum Erstellen einer Zusammenfassung durch Auswählen von Schlagwörtern für einen Dokumentensatz innerhalb der Dokumentdatenbank, und ein Modul zum Durchführen einer Suche mittels einer beliebigen Zusammenfassung enthalten. Ein Nutzer, der eine Suche durchführt, gibt einen Dokumentensatz über einen Client an einen Server, in dem eine Quelldatenbank gespeichert ist, und erhält eine Zusammenfassung.

[0011] Als Nächstes wird die Zusammenfassung an einen Server gesendet, auf dem sich die zu durchsuchende Zieldatenbank befindet, und ein Suchergebnis wird empfangen.

[0012] Als Suchschnittstelle des Clients wird als Erstes ein Anzeigebereich für einen Dokumentensatz angegeben, in dem der benötigte Satz von Schlüsseldokumenten angegeben werden kann und auch die zu durchsuchende Datenbank ausgewählt werden kann. In dem Client wählt der Nutzer dann einen interessanten Dokumentensatz aus den in dem Anzeigebereich angezeigten Dokumentensätzen, und falls notwendig, wechselt er die zu durchsuchende Dokumentdatenbank.

[0013] Diese und weitere Aufgaben, Merkmale und Vorteile der vorliegenden Erfindung werden in Anbetracht der folgenden detaillierten Beschreibung der bevorzugten Ausführungsformen in Verbindung mit den begleitenden Figuren deutlicher.

Kurze Beschreibung der Zeichnungen:

[0014] Fig. 1 stellt ein Diagramm dar, das ein Beispiel der Gesamtkonstruktion eines Systems zeigt, das das Verfahren zur Suche in mehreren Dokumentdatenbanken umsetzt.

[0015] Fig. 2 stellt ein Diagramm dar, das ein Beispiel des Aufbaus einer Suchassistenten-Schnittstelle in einem Client zeigt.

Beschreibung der bevorzugten Ausführungsformen

[0016] Fig. 1 zeigt eine typische, allgemeine Anordnung, wobei ein Client **11** einen beliebigen Schlüsseldokumentensatz in einer Dokumentdatenbank **131** eines Servers **13** spezifiziert, und einen Dokumentensatz mit einer hohen Relevanz (Ähnlichkeit) mit dem spezifizierten Schlüsseldokumentensatz von einer Dokumentdatenbank **141** eines anderen Servers **14** erhält. Hierbei befinden sich die Quell- und Zieldokumentdatenbanken **131**, **141** auf Servern an verschiedenen Orten auf die entsprechend über ein Netzwerk **12** zugegriffen werden kann.

[0017] Als Erstes spezifiziert der Client **11** einen Schlüsseldokumentensatz in der Quelldokumentendatenbank **131** entsprechend der Nutzeranforderung und sendet diese Information als eine Satz der einen Dokumentenbezeichner, z. B. eine ID, die jedem Dokument zugeordnet ist, die der Server **13** verstehen kann, über das Netzwerk **12** an den Server **13**. Der Dokumentensatz wird in einem Fenster zum Anzeigen der Suchergebnisse P1, das später beschrieben wird, spezifiziert.

[0018] Der Server **13** identifiziert einen Dokumentensatz, der von dem Client gesendet wurde. Eine Zusammenfassung des Dokumentensatzes wird dann für den gesuchten Dokumentensatz durch ein Zusammenfassungserzeugungsmodul **132** angefertigt, und wird dann zurück an den Client **11** über das Netzwerk **12** gesendet. Hierbei bedeutet die Bezeichnung „Zusammenfassung" ein Satz Schlagwörter, der sich auf einen Dokumentensatz bezieht. Das Zusammenfassungserzeugungsmodul kann durch irgendeine der bekannten Verfahren wie z. B. in der japanischen Offenlegungsschrift Nr. Hei 9-62693, „Method of Document Classification by Probability Model" beschrieben, konstruiert werden.

[0019] In einem Beispiel werden als Erstes die Wortdichten durch Aufteilen in Wörter all der Dokumente in dem Dokumentensatz, für die eine Zusammenfassung erstellt werden soll, zusammen gezählt. Im Allgemeinen, da der Grad, mit dem ein gegebener Dokumentensatz durch bestimmte Wörter beschrieben wird, höher ist für die Wörter, die öfter innerhalb desselben auftreten, werden Wörter in der Zusammenfassung umso häufiger mit aufgenommen, je höher ihr Auftreten in dem Dokumentensatz ist. Allgemeine Wörter jedoch, die oft in allen Dokumenten vorkommen, wie z. B. „do", usw. sind nicht für die Zusammenfassung geeignet. Daher werden gewöhnlich Wörter zur Einbeziehung in die Zusammenfassung unter Berücksichtigung ihres Vorkommens in der Dokumentdatenbank ausgewählt, zu der der Dokumentensatz gehört. Im Besonderen sind Schlagwörter wünschenswert die eine hohe Vorkommensfrequenz in einem bestimmten Dokumentensatz, aber eine niedrige Vorkommensfrequenz in der gesamten Dokumentdatenbank haben, d. h. sie sind geeignet für eine Zusammenfassung, die den Dokumentensatz charakterisiert. Somit werden Wörter für die Zusammenfassung durch Berechnen ihrer Gewichtung aus geeigneten Parametern unter Verwendung der Aufkommensfrequenz in dem Dokumentensatz und der Aufkommensfrequenz in der Dokumentdatenbank, und der Annahme von Wörtern mit einer Gewichtung gleich oder größer als ein bestimmter Schwellenwert berechnet.

[0020] Je höher die ist die Gewichtung ist, umso höher ist die Relevanz des Wortes zu einem gegebenen Dokument, und je geringer die Gewichtung ist, umso geringer ist die Relevanz des Wortes zu dem Dokument.

[0021] Der Server **13** gibt an den Client über das Netzwerk **12** einen Satz Wörter zurück, die eine Gewichtung haben, die durch das oben beschriebene Verfahren berechnet wurde. Diese Wörter werden als „Schlagwörter" in Fig. 2 angezeigt.

[0022] Als Nächstes evaluieren oder verarbeiten Nutzer an dem Client **11** die Zusammenfassung (Zusammenfassung des Schlüsseldokumentensatzes) von dem Server **13**, und der Client **11** überträgt diese an den Zielserver **14** über das Netzwerk **12**.

[0023] In der Evaluierung oder Verarbeitung, die durch Nutzer an dem Client durchgeführt wird, entfernen Nutzer beispielsweise Wörter von der Zusammenfassung, die nicht als relevant angesehen werden, oder ersetzen Nutzer beispielsweise Wörter in der Zusammenfassung.

[0024] Unter Verwendung des Suchmoduls **143** berechnet der Server **14** die Relevanz der Zusammenfassung des Schlüsseldokumentensatzes, die von dem Client an die Zieldokumentdatenbank **141** gesendet wurde, und liefert hoch relevante Dokumentenbezeichner mit hoher Relevanz mit einer Relevanzgewichtung an den Client **11** zurück. Das Such-

modul hier kann durch eine aus dem Stand der Technik bekannte Schlüsselwortsuche implementiert werden. Insbesondere, da die Zusammenfassung des Dokumentensatzes als Satz bestehend aus Wörtern mit Gewichtung eingegeben wird, können diese Wörter als gewichtete Eingabeschlüsselwörter betrachtet werden und eine ODER-Schlüsselwortsuche kann durchgeführt werden. In diesem Fall kann die Gewichtung (Relevanz) des Dokuments, das das Suchergebnis darstellt, berechnet werden. Dies wird dadurch erreicht, dass die Wörter, die in der Zusammenfassung und im zu suchenden Dokument auftreten, herangezogen werden, eine Gesamtgewichtung aus ihrer Gewichtung in der Zusammenfassung und ihrer Gewichtung in dem zu suchenden Dokument berechnet wird (z. B. Produkt aus den zwei Gewichtungen) und dann aufaddieren der Gewichtungen aller Wörter (z. B. Berechnung einer Gesamtsumme), um die Relevanz zu erhalten.

[0025] Unter Verwendung des obigen Verfahrens kann der Client **11** einen Dokumentensatz aus der Dokumentendatenbank **141** erhalten, der sich auf einen beliebigen Schlüsseldokumentensatz in der Dokumentdatenbank **131** bezieht. Das charakterisierende Merkmal dieses Verfahrens ist, dass der Netzwerkverkehr durch Überlassen der Verarbeitung (Erzeugung der Zusammenfassung) des originale, zu suchenden Dokumentensatzes der Serverseite überlassen wird. Es wird erkannt werden, dass die Verkehrsmenge wesentlich geringer ist als in dem Fall, in dem der Client alle Informationen der zu suchenden Dokumente empfangen und verarbeiten muss. Das Suchassistentenmodul **112** des Clients muss dann im Grunde genommen nur die Zusammenfassung des Dokumentensatzes von dem Quellserver an den Zielserver senden und nahezu die gesamte in der Suche enthaltene Verarbeitung kann den beiden Servern überlassen werden. Vielmehr muss die Serverseite nur ein Zusammenfassungserzeugungsmodul und ein Suchmodul für die in Frage kommende Dokumentendatenbank, aufweisen und es ist daher vollkommen unnötig, Informationen in anderen Dokumentdatenbanken zu berücksichtigen.

[0026] In der vorangegangenen Beschreibung wurde ein Verfahren beschrieben, in dem die Dokumentdatenbank **131** die Quelldatenbank und die Dokumentdatenbank **141** die Zieldatenbank war, aber das gleiche Verfahren kann angewendet werden, wenn die Dokumentdatenbank **141** die Quelldatenbank und die Dokumentdatenbank **131** die Zieldatenbank ist. In diesem Fall erhält der Client eine Zusammenfassung des Schlüsseldokumentensatzes von einem Zusammenfassungserzeugungsmodul **142** des Servers **14**, überträgt dieses an den zu durchsuchen Server **13**, und erhält relevante Dokumente in der Dokumentdatenbank **131** von dem Suchmodul **133** des Servers **13**. Wenn das Obige verallgemeinert wird, und ein Server mit einem Zusammenfassungserzeugungsmodul und einem Suchmodul für eine neue Datenbank bereitgestellt wird, kann diese Dokumentdatenbank als Quelldatenbank oder Zieldatenbank für alle Dokumentdatenbanken, die zu dem Netzwerk verbunden sind, allein durch Verbinden des Servers zu dem Netzwerk funktionieren.

[0027] In Fig. 1 befinden sich das Zusammenfassungserzeugungsmodul und das Suchmodul (**132**, **133** und **142**, **143**) entsprechend auf verschiedenen Servern (**13**, **14**), aber die Ausführungsform dieser Erfindung ist nicht auf diese bestimmte Anordnung beschränkt. Beispielsweise können das Zusammenfassungserzeugungsmodul und das Suchmodul in einem Server verschieden von der Dokumentdatenbank installiert sein, und ein Zusammenfassungserzeugungsmodul und ein Suchmodul können durch diesen Server für mehrere Datenbanken zur Verfügung gestellt werden.

[0028] Schließlich zeigt Fig. 2 eine Ausführungsform, die den Client betrifft. **111** ist ein Beispiel einer Suchassistentenschnittstelle, die in dem Client installiert ist. Dies ist grundsätzlich die gleiche Schnittstelle wie die Schnittstelle, die durch den Erfinder der vorliegenden Erfindung in der japanischen Offenlegungsschrift Hei 11-85786 „Document search support method and document search support service" (entspricht US Patentanmeldung S.N. 09/145,155, angemeldet am 09/01/98 durch Nishioka et al.) oder der japanischen Offenlegungsschrift Hei 10-74210, „Document search method and device, and document search service using the saure" (entspricht US Patentanmeldung S.N. 08/888,017, eingereicht 07/03/97 by Niwa et al.) vorgeschlagen wurde. E1 ist ein Fenster zur Eingabe einer Suchanfrage, wobei der Nutzer eine Suchanfrage durch die Eingabe einer Abfolge von Schlüsselwörtern oder in Form eines Satzes eingeben kann. M1 ist ein Fenster zur Auswahl einer Dokumentdatenbank, wobei der Nutzer einen bestimmten Teil auf der rechten Seite mit einer Maus herunterziehen kann, um so eine Liste von Dokumentdatenbanken anzuzeigen und einige gewünschte Dokumentdatenbank auszuwählen. B1 ist ein Suchknopf, der eine Suche startet. Demnach gibt der Nutzer eine beliebige Suchanfrage in das Fenster E1 ein, wählt eine zu durchsuchende Dokumentdatenbank in dem Fenster M1 aus, und führt eine gewöhnliche Suche durch die Eingabe von Schlüsselwörtern in das Fenster E1, die sich auf die in Fenster M1 gewählte Dokumentdatenbank beziehen, durch Drücken des Knopfes B1 aus. Diese Suche wird durch Unterstützung durch das in Fig. 1 gezeigte Suchassistentenmodul **112** durchgeführt, aber da Details des Suchverfahrens in der vorangegangenen Anmeldung beschrieben wurden, werden sie hier nicht wiederholt.

[0029] P1 ist ein Fenster zur Anzeige eines Suchergebnisses. In dem oberen Teil ist ein Feld angezeigt,

das die gesamte Anzahl der Dokumente anzeigt, die als Ergebnis des Suchprozesses erhalten wurden und eine Anzahl der Dokumente, die durch den Nutzer wie nachfolgend beschrieben ausgewählt wurden. Darunter ist ein Feld für den Nutzer vorgesehen, um ausgewählt/nicht ausgewählt (P13) anzugeben, und ein Dokumententitelteil, der die Relevanz (P12) der Suchanfrage und Titel (P11) der Dokumente, in Form einer Liste dargestellt. Dieses Anzeigefenster hat eine Scrollfunktion, so dass durch Scrollen der User einen Teil sehen kann, der zur einer bestimmten Zeit nicht in der Anzeige dargestellt werden kann. In einem Ausgewählt-/Nicht-ausgewählt-Feld werden Dokumente bei jedem Mausklick entweder ausgewählt oder abgewählt. Wenn Dokumente durch Klicken angewählt werden, wird eine Zusammenfassung der korrespondierenden Dokumente als grafische Darstellung eines Satzes von Wörtern mit Gewichtung in einem Zusammenfassungsanzeigefenster P2 dargestellt. Das Zusammenfassungsanzeigefenster P2 hat auch ein Feld in seinem oberen Teil, in dem die Gesamtzahl der Schlagwörter und Anzahl der Schlagwörter, die durch den Nutzer angewählt wurden, dargestellt wird. Dokumententitel werden gewöhnlich entsprechend ihrer Relevanz sortiert.

[0030] Das Fenster P1 zum Anzeigen der Suchergebnisse in dem Diagramm zeigt, dass eine Gesamtzahl von 22 Dokumenten als Ergebnis der Suche erhalten wurde, und dass 3 Dokumente durch den Nutzer aufgrund ihres Titels als interessante Dokumente ausgewählt wurden. Den ausgewählten Dokumenten wird ein Haken durch Klicken zugeordnet. In dem Zusammenfassungsanzeigefenster P2 werden 5 Schlagwörter entsprechend den ausgewählten Dokumenten angezeigt.

[0031] Obwohl in dieser Ausführung ausgelassen, können umgekehrt auch Dokumente für die die Schlagwörter die in dem Zusammenfassungsanzeigefenster P2 als repräsentativ ausgewählt wurden in dem Fenster P1 angezeigt werden. Daher kann der Nutzer eine fortschrittlichere Suche durch Erzeugen einer Zusammenfassung, die seinen Präferenzen entspricht, durchführen. Dies wird en detail in der oben aufgeführten japanischen Patentoffenlegungsschrift Hei 11-085786 beschrieben.

[0032] Somit kann der Nutzer Dokumente anwählen/abwählen, während er sich auf die Titel und die Schlagwörter der ausgewählten Dokumente bezieht, und er kann mehrere ihn interessierende Dokumente auswählen.

[0033] Dementsprechend kann der Nutzer, wenn er in der Nutzung anderer Dokumentdaten für den Dokumentensatz entsprechend des Suchergebnisses interessiert ist, die Dokumentdatenbank im Fenster M1 wechseln und den Knopf B1 drücken, um so eine neue Suche zu beginnen.

[0034] Somit sendet der Client einen Bezeichner der mehreren, ausgewählten Dokumente an den Server, wo die Quelldokumentendatenbank gespeichert ist (z. B. der Server **13**), erhält eine Zusammenfassung dieser mehreren Dokumente, sendet diese Zusammenfassung an den Server, auf dem die Zieldokumentendatenbank gespeichert ist (z. B. Server **14**) und erhält ein Suchergebnis von dem Zielserver (z. B. Server **14**). Das neue Suchergebnis wird im Fenster P1 angezeigt. Mit anderen Worten wird in diesem Fall P1 mit dem neuerlich gesuchten Dokumentensatz aktualisiert.

[0035] Um ein erneutes Suchergebnis mit einem vorherigen Suchergebnis zu vergleichen, kann der Nutzer den Zurückknopf B2 drücken, um das vorherige Suchergebnis in dem Fenster P1 wieder anzuzeigen, und um zu der Anzeige des Fensters P1 in dem Zustand zurückzukehren, in dem es sich vor Durchführung der Suche befand. Ebenso kann das Fenster P1 zu dem neuen Suchergebnis durch Drücken des Vorwärtsknopfes B3 vorgeblättert werden.

[0036] Da der Nutzer andere Dokumentdatenbanken entsprechend einem Suchergebnis zu jedem Zeitpunkt der Suche durchsuchen kann, kann der Nutzer durch Wiederholen des Suchzykluses frei von einer Datenbank zu einer anderen Datenbank wechseln. Natürlich ist es auch möglich, diesen Zyklus innerhalb der gleichen Dokumentdatenbank, d. h. ohne Wechsel der Dokumentdatenbank zu wiederholen.

[0037] Entsprechend dieser Erfindung kann der Nutzer eine zu durchsuchende Dokumentdatenbank frei bestimmen und die Suche ohne Berücksichtigung des Ortes oder des Aufbaus jeder Dokumentdatenbank verbessern. Ferner, da ein Server, auf dem eine Dokumentendatenbank sich befindet, modularisiert werden kann, kann der Server als Quelldatenbank oder Zieldatenbank bezüglich allen anderen, zu einem Netzwerk verbundenen Datenbanken einfach durch Verbinden eines Servers mit einem Zusammenfassungserzeugungsmodul und einem Suchmodul mit dem Netzwerk fungieren, falls gewünscht wird eine neue Dokumentdatenbank zu durchsuchen.

Patentansprüche

1. Ein Dokumentensuchverfahren, bei dem Server (**13**, **14**) mit Datenbanken (**131**, **141**) und Programme zum Manipulieren der Datenbanken (**131**, **141**) über ein Netzwerk (**12**) verteilt sind und ein Client (**11**), der mit den Servern (**13**, **14**) verbunden ist, eine Dokumentensuche durchführt, wobei das Verfahren eine Funktion aufweist, um aus einer ausgewählten Dokumentendatenbank Dokumentensätze mit hoher Bedeutung entsprechend der Bedeutung zu einer Sucheingabe zu suchen, wobei die Eingabe aus einem Satz Schlüsselwörter, Teilen eines Dokuments oder jedem gewünschten Satz von Dokumen-

ten besteht, wobei das Suchverfahren eine Funktion zum Wechseln zwischen verschiedenen Dokumentendatenbanken (**131**, **141**) aufweist, und wobei die Suchergebnisse aus einer Quelldokumentendatenbank (**131**) durch einen Nutzer, der eine Suche durchführt als Anfrage zur Suche in einer Zieldokumentendatenbank (**141**) verwendet werden können, wobei der Client (**11**)
– einen Satz von Dokumenten in einer Suchanfrage an den Server (**13**) übermittelt, in dem die Quelldokumentendatenbank (**131**) gespeichert ist,
– von dem Server (**13**) eine Zusammenfassung erhält, die nur Schlagwörter bezüglich des gesendeten Satzes von Dokumenten enthält,
– eine Suchanfrage entsprechend der Zusammenfassung an einen anderen, durch den Nutzer ausgewählten, Server (**14**) sendet, in dem die Zieldokumentendatenbank (**141**) gespeichert ist, und
– von dem anderen Server (**14**) ein Suchergebnis empfängt.

2. Ein Dokumentensuchverfahren wie in Anspruch 1 definiert, wobei eine Schnittstelle vorgesehen ist, in der ein Satz von Dokumenten von dem Suchergebnis einer Dokumentendatenbank (**131**) ausgewählt oder abgewählt werden kann, und eine aus dem Suchergebnis ausgewählter Satz von Dokumenten als Anfrage, um eine Suche in einer anderen Datenbank (**141**) durchzuführen, verwendet werden kann.

3. Ein Dokumentensuchverfahren wie in Anspruch 1 definiert, wobei eine Zusammenfassung, die nur Schlagwörter enthält, als Suchanfrage zur Durchführung einer Suche verwendet wird.

4. Ein Dokumentensuchverfahren wie in Anspruch 1 definiert, wobei der Server eine Zusammenfassung aus Schlagwörtern erzeugt, die für einen Satz von Dokumenten, die von dem Client (**11**) gesendet wurde, relevant sind und diese an den Client (**11**) übermittelt, und einen Satz von Dokumenten mit einer hohen Relevanz zu irgendeiner durch den Client (**11**) gesendeten Zusammenfassung sucht und an den Client (**11**) übermittelt.

5. Ein Dokumentensuchverfahren wie in Anspruch 1 definiert, wobei der Client (**11**) eine Schnittstelle (**111**) hat, um einen Satz von Dokumenten als Suchanfrage und eine zu durchsuchende Dokumentendatenbank anzugeben, der Satz von Dokumenten in der Suchanfrage an einen durch den Nutzer bestimmten Server (**13**) gesendet wird, eine Zusammenfassung des Satzes von Dokumenten von diesem Server (**13**) empfangen wird, die empfangene Zusammenfassung an einen Server (**14**) gesendet wird, der eine andere Dokumentendatenbank (**114**) enthält, und Suchergebnisse von dem letzten Server (**14**) empfangen und angezeigt werden.

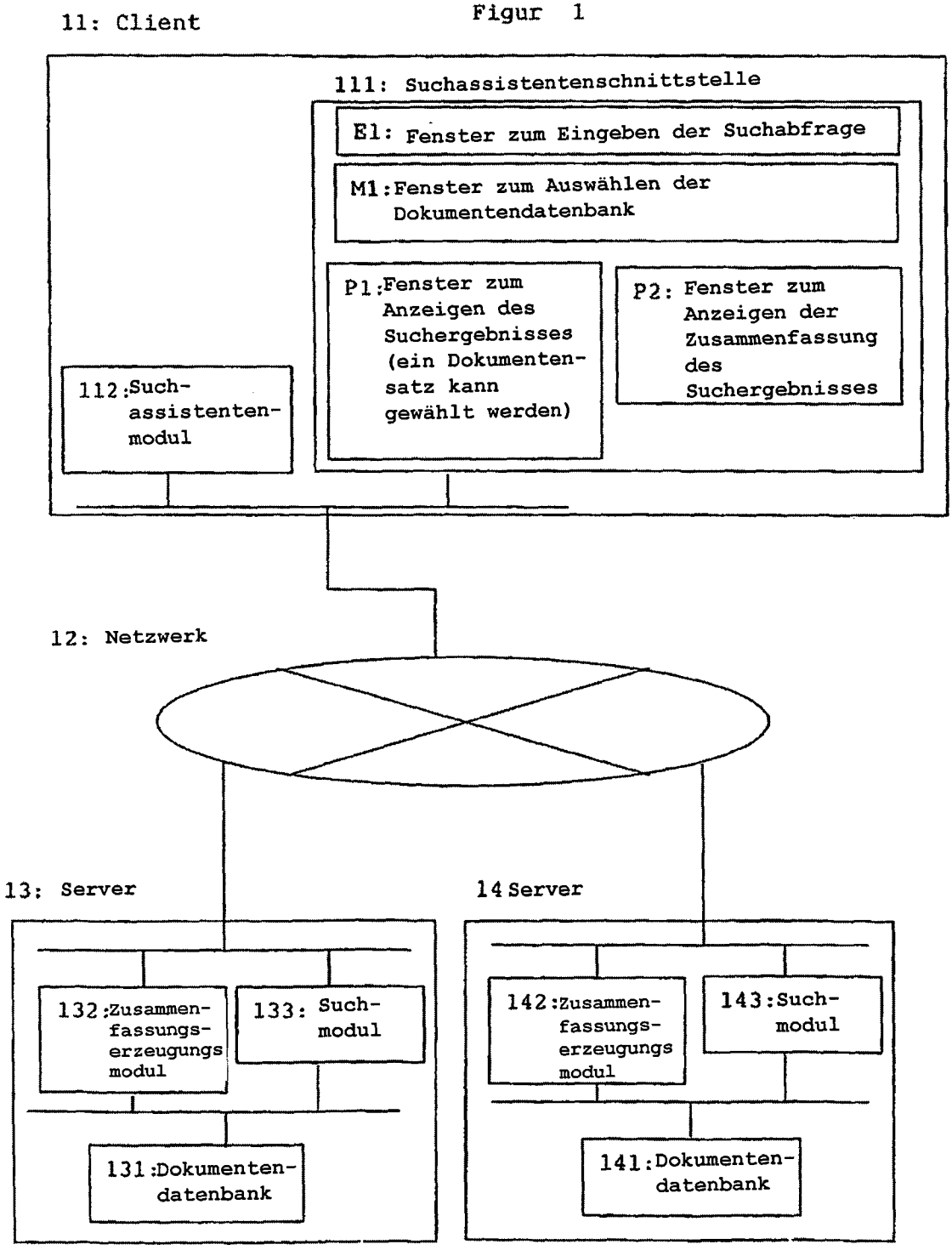
6. Ein Dokumentensuchverfahren wie in Anspruch 2 definiert, wobei eine Zusammenfassung, die nur Schlagwörter enthält, als Suchanfrage zur Durchführung einer Suche verwendet wird.

7. Ein Dokumentensuchverfahren wie in Anspruch 2 definiert, wobei Server (**13**, **14**), die Dokumentendatenbanken (**131**, **141**) und Programme zum Manipulieren der Datenbanken enthalten, über ein Netzwerk (**12**) verteilt sind, wobei ein Client (**11**)
einen Satz von Dokumenten in einer Suchanfrage an einen Server (**13**) übermittelt, in dem eine ausgewählte Dokumentendatenbank (**131**) gespeichert ist,
eine Zusammenfassung empfängt, die nur Schlagwörter in Bezug auf die gesendete Dokumentenanfrage enthält,
eine Suchanfrage entsprechend der Zusammenfassung, die die Evaluierung der Zusammenfassung durch einen Nutzer angibt an einen Server (**14**) sendet, in dem eine andere Dokumentendatenbank (**141**) gespeichert ist, und
ein Suchergebnis empfängt.

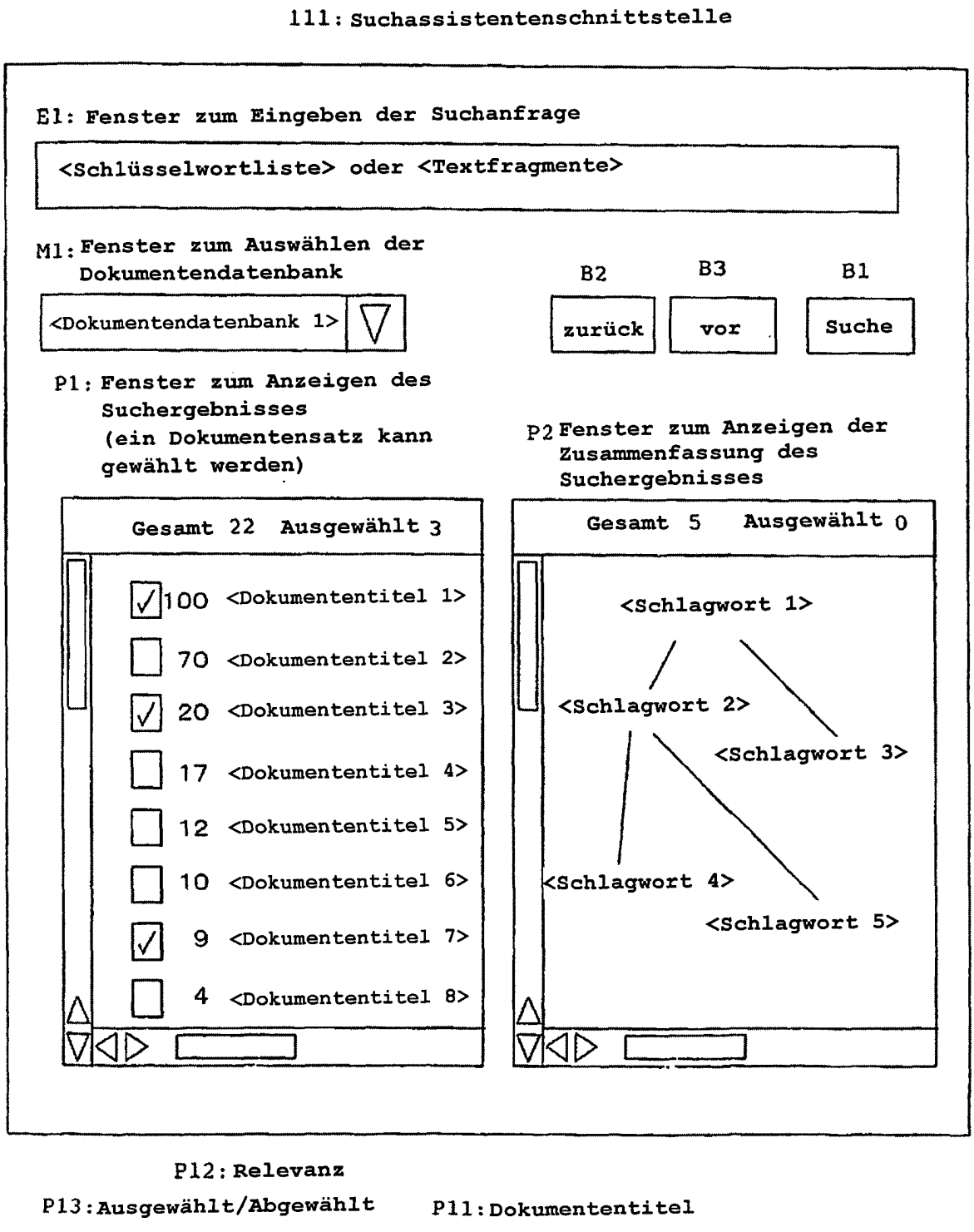
Es folgen 2 Blatt Zeichnungen

Anhängende Zeichnungen

Figur 1



Figur 2


111: Suchassistentenschnittstelle
E1: Fenster zum Eingeben der Suchanfrage
<Schlüsselwortliste> oder <Textfragmente>
M1: Fenster zum Auswählen der Dokumentendatenbank
<Dokumentendatenbank 1>
B2
B3
B1
zurück
vor
Suche
P1: Fenster zum Anzeigen des Suchergebnisses (ein Dokumentensatz kann gewählt werden)
P2 Fenster zum Anzeigen der Zusammenfassung des Suchergebnisses
Gesamt 22 Ausgewählt 3
100 <Dokumententitel 1>
70 <Dokumententitel 2>
20 <Dokumententitel 3>
17 <Dokumententitel 4>
12 <Dokumententitel 5>
10 <Dokumententitel 6>
9 <Dokumententitel 7>
4 <Dokumententitel 8>
Gesamt 5 Ausgewählt 0
<Schlagwort 1>
<Schlagwort 2>
<Schlagwort 3>
<Schlagwort 4>
<Schlagwort 5>
P12: Relevanz
P13: Ausgewählt/Abgewählt
P11: Dokumententitel