

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7091685号
(P7091685)

(45)発行日 令和4年6月28日(2022.6.28)

(24)登録日 令和4年6月20日(2022.6.20)

(51)国際特許分類		F I	
G 0 6 F	16/9032(2019.01)	G 0 6 F	16/9032
G 0 6 F	16/903(2019.01)	G 0 6 F	16/903

請求項の数 9 (全19頁)

(21)出願番号	特願2018-20734(P2018-20734)	(73)特許権者	000005223 富士通株式会社
(22)出願日	平成30年2月8日(2018.2.8)		神奈川県川崎市中原区上小田中4丁目1 番1号
(65)公開番号	特開2019-139413(P2019-139413 A)	(74)代理人	100087480 弁理士 片山 修平
(43)公開日	令和1年8月22日(2019.8.22)	(72)発明者	寺本 和真 神奈川県川崎市中原区上小田中4丁目1 番1号 富士通株式会社内
審査請求日	令和2年11月10日(2020.11.10)	審査官	松尾 真人

最終頁に続く

(54)【発明の名称】 検索処理プログラム、検索処理方法及び検索処理装置

(57)【特許請求の範囲】

【請求項1】

複数のユーザのそれぞれに対応する検索条件を受け付け、
前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付け、
受け付けた前記評価を演算処理した結果から前記評価が高いユーザを特定し、
前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する、
処理をコンピュータに実行させることを特徴とする検索処理プログラム。

【請求項2】

前記評価が高いユーザとは異なるユーザは、前記評価が高いユーザと同一の業務に従事するユーザであることを特徴とする請求項1に記載の検索処理プログラム。

【請求項3】

前記検索条件は、検索ワードを含み、
前記検索条件を変換する場合、前記検索ワードに類似する類似ワードを特定し、前記類似ワードのうち前記評価が高いユーザが検索に利用した頻度が高いものを優先して、前記検索ワードと変換する、
ことを特徴とする請求項1又は2に記載の検索処理プログラム。

【請求項4】

前記検索条件は、複数の検索ワードを組み合わせた検索ワード群を含み、

前記検索条件を変換する場合、前記検索ワード群から変換を行わない検索ワードと変換対象の検索ワードとを特定し、特定した前記変換対象の検索ワードを、前記評価が高いユーザの検索履歴に基づいて変換する、
ことを特徴とする請求項 1 ~ 3 のいずれか一項に記載の検索処理プログラム。

【請求項 5】

前記検索を実行する処理は、複数の文書の中から前記検索ワード群を含む文書を検索する処理であり、

前記検索ワード群に含まれる複数の検索ワードのうち、前記複数の文書における IDF (Inverse Document Frequency) 値が大きいものを優先して、前記変換を行わない検索ワードとして特定する、

10

ことを特徴とする請求項 4 に記載の検索処理プログラム。

【請求項 6】

前記変換対象の検索ワードを変換する場合、前記変換を行わない検索ワードを含む検索ワード群を前記評価が高いユーザの検索履歴から抽出し、抽出した前記検索ワード群に基づいて、前記変換対象の検索ワードを変換する、

ことを特徴とする請求項 5 に記載の検索処理プログラム。

【請求項 7】

前記変換対象の検索ワードを変換する場合、抽出した前記検索ワード群に含まれる前記変換を行わない検索ワード以外の検索ワードのうち、前記複数の文書それぞれにおける TF (Term Frequency) - IDF 値の分散が大きいものを優先して、前記変換対象の検索ワードと変換する、

20

ことを特徴とする請求項 6 に記載の検索処理プログラム。

【請求項 8】

複数のユーザのそれぞれに対応する検索条件を受け付け、

前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付け、

受け付けた前記評価を演算処理した結果から前記評価が高いユーザを特定し、

前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する、

処理をコンピュータが実行することを特徴とする検索処理方法。

30

【請求項 9】

複数のユーザのそれぞれに対応する検索条件を受け付ける検索条件受付部と、

前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付ける評価受付部と、

受け付けた前記評価を演算処理した結果から前記評価が高いユーザを特定する特定部と、
前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する検索部と、

を備える検索処理装置。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、検索処理プログラム、検索処理方法及び検索処理装置に関する。

【背景技術】

【0002】

従来、社内のデータベースに蓄積されたレポートなどの文書の中から必要とする文書を探し出す場合、ユーザは、探し出したい文書に含まれている可能性の高い検索ワードを用いてデータベース内を検索するのが一般的である。この場合、業務経験が豊富なユーザであれば、経験に基づいて適切な検索ワードを設定することができる。

【先行技術文献】

50

【特許文献】

【0003】

【文献】特開2005-301856号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、業務経験が乏しいユーザの場合には、適切な検索ワードを設定することができないことがあるため、自分が必要としている文書を探し出せないおそれがある。

【0005】

1つの側面では、本発明は、検索精度を向上させることが可能な検索処理プログラム、検索処理方法及び検索処理装置を提供することを目的とする。

10

【課題を解決するための手段】

【0006】

一つの態様では、検索処理プログラムは、複数のユーザのそれぞれに対応する検索条件を受け付け、前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付け、受け付けた前記評価を演算処理した結果から前記評価が高いユーザを特定し、前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する、処理をコンピュータに実行させる検索処理プログラムである。

【発明の効果】

20

【0007】

検索精度を向上させることができる。

【図面の簡単な説明】

【0008】

【図1】一実施形態に係る検索処理システムの構成を概略的に示す図である。

【図2】図2(a)は、サーバのハードウェア構成を示す図であり、図2(b)は、ユーザ端末のハードウェア構成を示す図である。

【図3】サーバの機能ブロック図である。

【図4】図4(a)は、ユーザテーブルを示す図であり、図4(b)は、検索履歴テーブルを示す図である。

30

【図5】図5(a)は、頻度テーブルを示す図であり、図5(b)は、TF-IDFテーブルを示す図である。

【図6】ユーザが検索画面を利用して検索を行う場合のサーバの処理を示すフローチャートである。

【図7】図7(a)～図7(c)は、図6の処理が行われる場合にユーザ端末に表示される画面例を示す図である。

【図8】ユーザが検索画面(初心者用)を利用して検索を行う場合のサーバの処理を示すフローチャートである。

【図9】図9(a)～図9(c)は、図8の処理が行われる場合にユーザ端末に表示される画面例を示す図である。

40

【図10】図10(a)～図10(c)は、画面例の別例を示す図である。

【発明を実施するための形態】

【0009】

以下、検索処理システムの一実施形態について、図1～図10に基づいて詳細に説明する。本実施形態の検索処理システムは、会社で管理する文書(電子データ)をユーザが検索する場合に、業務経験が豊富なユーザ(有識者と呼ぶ)でないユーザであっても、有識者と同様に検索ができるように検索を支援し、ユーザに対して検索結果及び検索された文書のデータを提供するシステムである。

【0010】

図1には、本実施形態に係る検索処理システム100の構成が概略的に示されている。図

50

1に示すように、検索処理システム100は、検索処理装置としてのサーバ10と、ユーザ端末70と、を備える。サーバ10及びユーザ端末70は、インターネットやLAN (Local Area Network)などのネットワーク80に接続されている。

【0011】

(サーバ10)

図2(a)には、サーバ10のハードウェア構成が示されている。図2(a)に示すように、サーバ10は、CPU (Central Processing Unit) 90、ROM (Read Only Memory) 92、RAM (Random Access Memory) 94、記憶部(ここではHDD (Hard Disk Drive)) 96、ネットワークインタフェース97、及び可搬型記憶媒体用ドライブ99等を備えている。これらサーバ10の構成各部分は、バス98に接続されている。サーバ10では、ROM 92あるいはHDD 96に格納されているプログラム(検索処理プログラムを含む)、或いは可搬型記憶媒体用ドライブ99が可搬型記憶媒体91から読み取ったプログラムをCPU 90が実行することにより、図3に示す、各部の機能が実現される。なお、図3には、サーバ10のHDD 96等に格納されているデータベース(database: DB)やテーブルも図示されている。

10

【0012】

図3には、サーバ10の機能ブロック図が示されている。図3に示すように、サーバ10においては、CPU 90がプログラムを実行することにより、入力受付部20、検索ワード変換部22、検索実行部24、提供部26、特定部としての評価部28、更新部30、としての機能が実現されている。なお、図3の各部の機能は、例えば、ASIC (Application Specific Integrated Circuit)やFPGA (Field Programmable Gate Array)等の集積回路により実現されてもよい。

20

【0013】

入力受付部20は、ユーザが検索画面(図7(a)参照)又は検索画面(初心者用)(図9(a)参照)に入力した検索ワードを取得し、検索実行部24又は検索ワード変換部22に受け渡す。また、入力受付部20は、ユーザにより画面上のボタンが押されたという情報を取得し、各部に通知する。なお、ユーザは、文書の検索を行う際には、図7(a)の検索画面又は図9(a)の検索画面(初心者用)のいずれかを選択して用いるものとする。

【0014】

検索ワード変換部22は、必要に応じて、ユーザが入力した検索ワードを変換し、変換後の検索ワードを検索実行部24に受け渡す。なお、検索ワード変換部22は、類義語辞書DB 45、ユーザテーブル42、頻度テーブル46、TF-IDFテーブル48を用いた処理を実行する。類義語辞書DB 45は、類義語を関連付けて格納するデータベースであり、例えば、「性能試験」と「性能テスト」と、「有用性試験」が類義語であること等の情報が格納されている。なお、ユーザテーブル42、頻度テーブル46及びTF-IDFテーブル48の詳細や、検索ワード変換部22の処理の詳細については後述する。

30

【0015】

検索実行部24は、ユーザが入力した検索ワードや検索ワード変換部22が変換した後の検索ワードを用いて、ドキュメントDB 40に格納されている膨大な量の文書の中から文書を検索する。また、検索実行部24は、検索された文書の情報を検索結果としてユーザ端末70に送信する。更に、検索実行部24は、有識者が図7(a)の検索画面を用いて検索を行った場合に、検索に用いた検索ワードを検索履歴テーブル44に格納する。

40

【0016】

ここで、検索実行部24は、図7(a)の検索画面を用いて検索をしたユーザが「有識者」であるか「一般」であるかを、ユーザテーブル42に基づいて判断する。図4(a)には、ユーザテーブル42のテーブル構造が示されている。ユーザテーブル42には、「ユーザ」、「評価平均値」、「区分」、「部門」の各情報が関連付けて格納されている。「ユーザ」の欄には、ユーザの氏名や識別番号等が格納される。「評価平均値」の欄には、ユーザが自らの検索について評価した点数の平均値が格納される。「区分」の欄には、各

50

ユーザの評価平均値から定まる各ユーザの区分が格納される。区分には、「有識者」と「一般」がある。「有識者」は、例えば評価平均値が閾値以上であり、業務経験が豊富なユーザを意味し、「一般」は、例えば評価平均値が閾値未満であり、業務経験が然程豊富でないユーザを意味するものとする。なお、これに限らず、全ユーザのうち、評価平均値が高い方から上位所定割合のユーザを有識者として設定することとしてもよい。「部門」の欄には、各ユーザが所属する会社内の部門の情報であり、経理部、事業部、A部門、B部門などの情報が格納される。

【0017】

図4(b)には、検索履歴テーブル44のテーブル構造が示されている。図4(b)に示すように、検索履歴テーブル44には、「有識者」と、「ワード1」、「ワード2」、...の情報が格納されている。「有識者」の欄には、有識者であるユーザの氏名や識別番号等が格納され、「ワード1」、「ワード2」、...の欄には、有識者が検索に用いた検索ワードが格納される。なお、検索履歴テーブル44において列方向に並ぶ1回の検索で用いられたワードを、「検索ワード群」とも呼ぶものとする。

10

【0018】

提供部26は、検索実行部24の検索結果の中からユーザが1つの文書を選択すると、選択された文書の概要を示す情報をユーザ端末70に対して送信する。また、提供部26は、ユーザから文書をダウンロードする旨の要求があると、ダウンロード要求のあった文書をドキュメントDB40から読み出し、ユーザ端末70に対して送信(ダウンロード)する。

20

【0019】

評価部28は、ユーザがユーザ端末70において文書をダウンロードした場合に、ユーザに対し、今回の検索についての評価を問い合わせ、評価結果に基づいてユーザテーブル42を更新する。なお、図3では、破線がテーブルの「更新」を意味している。

【0020】

更新部30は、ユーザテーブル42及び検索履歴テーブル44のデータに基づいて頻度テーブル46を更新するとともに、検索履歴テーブル44のデータ及びドキュメントDB40に格納されている文書に基づいて、TF-IDFテーブル48を更新する。なお、更新部30は、各テーブルの更新処理を1日1回や1週間に1回などの頻度でバッチ処理により実行する。

30

【0021】

図5(a)には、頻度テーブル46のテーブル構造が示されている。図5(a)に示すように、頻度テーブル46には、各検索ワード(「溶媒」、「アセトン」等)を、各部門の有識者がどのくらいの頻度で検索に用いたかの情報(出現頻度)が格納されている。例えば、A部門におけるワードの出現頻度 F_A は、A部門の有識者がワードを用いた検索を行った回数を M_A 、A部門の有識者が行った全検索回数を N_A とすると、次式(1)で表すことができる。

$$F_A = M_A / N_A \quad \dots (1)$$

【0022】

更新部30は、ユーザテーブル42に基づいて各部門の有識者を特定し、検索履歴テーブル44を参照して、部門ごとに各ワードの出現頻度を算出し、頻度テーブル46を更新する。

40

【0023】

図5(b)には、TF-IDFテーブル48のテーブル構造が示されている。図5(b)に示すように、TF-IDFテーブル48には、各検索ワード(「溶媒」、「アセトン」等)についての、「TF(Term Frequency)値」、「IDF(Inverse Document Frequency)値」、「TF-IDF値」、「TF-IDF値の分散」の各値が格納されている。ここで、「TF値」は、各検索ワードが、ドキュメントDB40に格納されている全ての文書において出現する頻度を意味する。また、「IDF値」は、各検索ワードがいくつの文書内で共通して使われているかを表す値であり、値が小さいほど文書にあまり出て

50

こないことを意味する。「TF-IDF値」は、上述したTF値と、上述したIDF値の積である。TF-IDF値の分散は、ある検索ワードについて文書毎に求められるTF値と上記IDF値の積である文書毎のTF-IDF値が、どれだけばらついているかを示す値である。TF-IDF値の分散が小さい検索ワードは、どの文書でもそのワードが同程度に出現していることを示し、TF-IDF値の分散が大きい検索ワードほど、文書の分類に貢献しているワードであることを意味している。

【0024】

更新部30は、検索履歴テーブル44に基づいて、検索に用いられた検索ワードを特定し、特定された検索ワードごとに、ドキュメントDB40を用いてTF-IDFテーブル48の各値を算出し、更新する。

10

【0025】

(ユーザ端末70)

図1に戻り、ユーザ端末70は、会社内のユーザが利用する端末である。ユーザは、ユーザ端末70に検索ワードを入力して所定の操作を行うことで、サーバ10により検索された文書を閲覧したり、利用したりすることができる。

【0026】

図2(b)には、ユーザ端末70のハードウェア構成が示されている。図2(b)に示すように、ユーザ端末70は、CPU190、ROM192、RAM194、記憶部(HDD)196、ネットワークインタフェース197、表示部193、入力部195、及び可搬型記憶媒体191に記憶された情報を読み取り可能な可搬型記憶媒体用ドライブ199等を備えている。表示部193は、液晶ディスプレイ等を含み、入力部195は、マウスやキーボード、タッチパネル等を含む。これらユーザ端末70の構成各部は、バス198に接続されている。

20

【0027】

(サーバ10の処理について)

次に、サーバ10の処理について、図6及び図8のフローチャートに沿って、その他図面を適宜参照しつつ詳細に説明する。

【0028】

図6には、ユーザが図7(a)の検索画面を利用して検索を行う場合のサーバ10の処理が示され、図8には、ユーザが図9(a)の検索画面(初心者用)を利用して検索を行う場合のサーバ10の処理が示されている。

30

【0029】

(検索画面を利用して検索を行う場合の処理)

まず、図6に基づいて、ユーザが検索画面を利用して検索を行う場合のサーバ10の処理について説明する。なお、図6の処理の前提として、ユーザからの要求に応じて、入力受付部20は、図7(a)に示すような検索画面をユーザ端末70に送信しているものとする。したがって、ユーザ端末70の表示部193には、図7(a)の検索画面が表示されているものとする。図7(a)の検索画面には、検索ワードを入力する欄と、「検索」ボタンと、検索結果を表示する欄と、が設けられている。

【0030】

図6の処理では、まず、ステップS10において、入力受付部20が、検索画面に検索ワードが入力され「検索」ボタンが押されるまで待機する。図7(a)に示すように、例えばユーザが検索ワード群「溶媒 植物性 有用性」を入力して、「検索」ボタンを押すと、ステップS12に移行する。なお、「検索」ボタンが押されると、ユーザ端末70から入力された検索ワード(検索ワード群)が送信されてくるので、入力受付部20は、受信した検索ワードを検索実行部24に対して受け渡す。

40

【0031】

ステップS12に移行すると、検索実行部24は、入力された検索ワードを用いてドキュメントDB40内の文書の検索を実行し、検索結果をユーザ端末70に送信する。これにより、ユーザ端末70の表示部193には、検索結果が表示されるようになっている。な

50

お、図7(a)には、検索結果として、3つの文書の情報が表示された状態が示されている。

【0032】

次いで、ステップS14では、入力受付部20が、検索結果の1つがユーザ端末70において選択されるまで待機する。ユーザ端末70において、ユーザが検索結果の中から1つの文書を選択すると、入力受付部20は、その旨を提供部26に通知し、ステップS16に移行する。

【0033】

ステップS16では、提供部26が、選択された文書の概要を示す画面をユーザ端末70に対して送信する。これにより、ユーザ端末70の表示部193には、図7(b)に示すような文書の概要を表示する画面が表示されることになる。なお、図7(b)の画面には、「ダウンロード」ボタンも表示されているものとする。

10

【0034】

次いで、ステップS18では、入力受付部20が、ユーザ端末70において、ユーザにより「ダウンロード」ボタンが押されるまで待機する。ユーザ端末70においてユーザにより「ダウンロード」ボタンが押されると、入力受付部20は、その旨を提供部26に通知し、ステップS20に移行する。

【0035】

ステップS20に移行すると、提供部26は、ドキュメントDB40からユーザによって選択された文書のデータを読み出し、ユーザ端末70に送信する(ダウンロードを行う)。

20

【0036】

ステップS20の後は、ステップS22に移行し、評価部28が、検索評価画面をユーザ端末70に送信する。これにより、ユーザ端末70の表示部193には、図7(c)に示すような検索評価画面が表示されることになる。なお、検索評価画面は、ユーザが実施した検索結果に対する満足度を例えば11段階(0点~10点)で評価する画面である。ユーザは、点数を選択するとともに、「送信」ボタンを押すことで、評価結果をサーバ10(入力受付部20)に対して送信することができる。

【0037】

次いで、ステップS24では、入力受付部20が、ユーザ端末70において「送信」ボタンが押されるまで待機する。ユーザによって「送信」ボタンが押されると、入力受付部20は、ユーザ端末70から送信されてくる点数を評価部28に通知し、ステップS26に移行する。

30

【0038】

ステップS26に移行すると、評価部28は、ユーザによって選択された(入力された)評価点と、これまでに得られた評価点との平均値を算出して、ユーザテーブル42の「評価平均値」の欄を更新する。なお、評価部28は、評価平均値の更新により、ユーザの区分を変更する必要がある場合には、区分の変更も行う。例えば、ユーザが「有識者」となるために、「評価平均値が所定の閾値以上であること」という基準を満たす必要がある場合には、評価部28は、更新後の評価平均値が閾値以上であるか否かを判断し、判断結果に基づいて区分を更新する。

40

【0039】

次いで、ステップS28では、検索実行部24が、検索を行ったユーザが「有識者」である場合に、入力された検索ワードで検索履歴テーブル44を更新する。より具体的には、検索実行部24は、ユーザが有識者であるか否かをユーザテーブル42を参照して判断し、ユーザが有識者であれば、図7(a)の画面において入力された検索ワード群「溶媒植物性 有用性」と、ユーザの識別情報を検索履歴テーブル44に追加する。

【0040】

以上により、図6の全処理が終了する。なお、図6の処理は、ユーザから図7(a)の検索画面を用いた検索要求が出されるたびに、繰り返し実行される。

【0041】

50

(検索画面 (初心者用) を利用して検索を行う場合の処理)

次に、図 8 に基づいて、ユーザが検索画面 (初心者用) を利用して検索を行う場合のサーバ 10 の処理について説明する。なお、図 8 の処理の前提として、ユーザからの要求に応じて、入力受付部 20 は、図 9 (a) に示すような検索画面 (初心者用) をユーザ端末 70 に送信しているものとする。したがって、ユーザ端末 70 の表示部 193 には、図 9 (a) の検索画面 (初心者用) が表示されているものとする。図 9 (a) の画面には、「検索」ボタンの他、「類義語変換」ボタンと「履歴から変換」ボタンが設けられている。なお、「類義語変換」ボタンが押されるまでは「履歴から変換」ボタンは押せない (選択できない) 状態になっているものとする。

【 0042 】

図 8 のステップ S 102 では、入力受付部 20 が、検索画面 (初心者用) に検索ワードが入力されるまで待機する。図 9 (a) に示すように、例えばユーザが検索ワード群「アセトン 海外 性能試験」を入力した場合、ステップ S 104 に移行する。

【 0043 】

ステップ S 104 に移行すると、入力受付部 20 が、既に「類義語変換」ボタンが押されたか否かを判断する。このステップ S 104 の判断が否定されると、ステップ S 106 に移行する。

【 0044 】

ステップ S 106 に移行すると、入力受付部 20 が、「類義語変換」ボタンが押されたか否かを判断する。このステップ S 106 の判断が否定された場合には、ステップ S 120 に移行し、入力受付部 20 は、「検索」ボタンが押されたか否かを判断する。このステップ S 120 の判断が否定されると、ステップ S 104 に戻る。すなわち、ステップ S 106 又はステップ S 120 の判断が肯定されるまで、ステップ S 104、S 106、S 120 の判断が繰り返し実行される。なお、ユーザは、図 9 (a) に示す検索ワード群「アセトン 海外 性能試験」で検索を行いたい場合には、「検索」ボタンを押し、検索ワードを適切なワードに変換したい場合には、「履歴から変換」ボタンを押し。

【 0045 】

ユーザが「検索」ボタンを押した場合、ステップ S 120 の判断が肯定されるため、ステップ S 122 に移行する。ステップ S 122 では、検索実行部 24 は、入力された検索ワード群「アセトン 海外 性能試験」を用いてドキュメント DB 40 を検索し、検索結果を検索画面 (初心者用) に表示する。

【 0046 】

一方、ユーザが「類義語変換」ボタンを押した場合、ステップ S 106 の判断が肯定されるため、ステップ S 108 に移行する。ステップ S 108 に移行すると、検索ワード変換部 22 が、各検索ワードの類義語を類義語辞書 DB 45 から抽出する。例えば、入力された検索ワードが検索ワード群「アセトン 海外 性能試験」である場合において、類義語辞書 DB 45 に「性能試験」と「性能テスト」と、「有用性試験」が類義語であるという情報が含まれていたとする。この場合、検索ワード変換部 22 は、「性能試験」の類義語として「性能テスト」と、「有用性試験」を抽出する。なお、「アセトン」や「海外」に類義語が存在する場合には、同様に抽出するが、説明の簡素化のため、類義語が存在しなかったものとする。

【 0047 】

次いで、ステップ S 110 では、検索ワード変換部 22 が、同じ部門の有識者の検索における各類義語 (検索ワードを含む) の出現頻度を特定する。具体的には、検索ワード変換部 22 は、ユーザテーブル 42 を参照して、検索を行っているユーザが例えば A 部門であることを特定すると、頻度テーブル 46 を参照して、A 部門のユーザの検索履歴における、「性能試験」と「性能テスト」と、「有用性試験」の出現頻度を特定する。

【 0048 】

次いで、ステップ S 112 では、検索ワード変換部 22 が、類義語 (検索ワードを含む) のうち出現頻度が最も高いものを検索ワードとする。上記例において、「性能試験」の出

10

20

30

40

50

現頻度が0.1、「性能テスト」の出現頻度が0.2、「有用性試験」の出現頻度が0.3であったとすると、出現頻度が最も高い「有用性試験」が検索ワードとなる。すなわち、検索ワード変換部22は、図9(b)に示すように、検索ワード群を図9(a)の「アセトン 海外 性能試験」から、「アセトン 海外 有用性試験」に変換する。これにより、ユーザが検索したい内容を変えることなく、入力した検索ワードを部門内でよく使われる検索に適した検索ワードに変換することができる。なお、複数の検索ワードに類義語が存在していた場合には、それぞれについて、出現頻度に基づく変換を行う。

【0049】

次いで、ステップS114では、入力受付部20が、「履歴から変換」ボタンが押されたか否かを判断する。このステップS114の判断が否定された場合には、ステップS120に移行し、入力受付部20は、「検索」ボタンが押されたか否かを判断する。ステップS120の判断が否定されると、ステップS104に戻るが、ステップS104の判断は肯定されるため、ステップS114に再度移行することになる。したがって、入力受付部20は、ステップS114、S120のいずれかの判断が肯定されるまで待機する。なお、ユーザは、図9(b)に示す検索ワード群「アセトン 海外 有用性試験」で検索を行いたい場合には、「検索」ボタンを押し、検索ワードを更に変換したい場合には、「履歴から変換」ボタンを押し。

10

【0050】

ユーザが「検索」ボタンを押した場合、ステップS120の判断が肯定されるため、ステップS122に移行する。ステップS122に移行すると、検索実行部24は、図9(b)に示す検索ワード群「アセトン 海外 有用性試験」を用いてドキュメントDB40を検索し、検索結果を表示する。

20

【0051】

一方、ユーザが「履歴から変換」ボタンを押した場合には、ステップS114の判断が肯定されるため、ステップS116に移行する。ステップS116に移行すると、検索ワード変換部22が、1つの検索ワード以外の検索ワードを固定する。この場合、検索ワード変換部22は、1つの検索ワード以外の検索ワードを検索の軸として固定し、固定しない1つの検索ワードを変換対象として設定する。例えば、「アセトン」と「有用性試験」を比較すると、「アセトン」の方が特徴的であり、文書検索において重要かつ有用な検索ワードであるため、検索ワードから除外しないようにする必要がある。したがって、本実施形態では、検索ワード変換部22は、有用な検索ワードを固定するため、TF-IDFテーブル48に格納されている各ワードのIDF値を用いることとしている。

30

【0052】

例えば、検索ワード群が「アセトン 海外 有用性試験」である場合、検索ワード変換部22は、TF-IDFテーブル48を参照して、3つの検索ワードのうちIDF値が最も小さいワード以外を固定する。このようにすることで、普遍的に用いない(多くの文書で汎用されない)固有名詞のようなワード(IDF値が大きくなる傾向にあるワード)を検索ワードとして固定することができる。なお、上記例では、「アセトン 海外 有用性試験」のうち、「アセトン」と「海外」が固定されたものとする。

【0053】

次いで、ステップS118では、検索ワード変換部22が、固定されていない検索ワードを、ユーザと同部門の有識者の検索履歴に基づいて変換する。具体的には、検索ワード変換部22は、検索履歴テーブル44を参照し、ユーザと同部門の有識者の検索履歴から、ステップS116で固定した検索ワードが全て含まれる検索履歴(検索ワード群)を抽出する。ここでは、例えば、「アセトン 海外 使用制限」や「アセトン 海外 使用例」や「アセトン 海外 使用例 中国」が抽出されたとする。この場合、「アセトン 海外」以外のワード「使用制限」、「使用例」、「使用例 中国」が、「有用性試験」と変換する候補のワード(変換候補のワード)となる。

40

【0054】

次いで、検索ワード変換部22は、TF-IDFテーブル48を参照して、固定した検索

50

ワード以外のワード（変換候補のワード）のTF-IDF値の分散を抽出する。変換候補のワードが「使用制限」であれば「使用制限」のTF-IDF値の分散をTF-IDFテーブル48から抽出する。また、変換候補のワードが「使用例」であれば「使用例」のTF-IDF値の分散をTF-IDFテーブル48から抽出する。また、変換候補のワードが「使用例 中国」であれば、「使用例」のTF-IDF値の分散と「中国」のTF-IDF値の分散をTF-IDFテーブル48から抽出して、その平均値を変換候補のワード「使用例 中国」のTF-IDF値の分散とする。

【0055】

そして、検索ワード変換部22は、「使用制限」、「使用例」、「使用例 中国」のうち、TF-IDF値の分散が最も大きいものを固定していない検索ワードと代えて用いることとする。このように、TF-IDF値の分散が大きいワードを検索ワードとして用いる理由は、TF-IDF値の分散が大きいワードほど、文書の分類に貢献しているワードであると考えられるからである。上記例において、例えば「使用制限」のTF-IDF値の分散が最も大きければ、検索ワード変換部22は、図9(c)に示すように、図9(b)の検索ワード群「アセトン 海外 有用性試験」を「アセトン 海外 使用制限」に変換する。その後は、ステップS120に移行する。

10

【0056】

ステップS116、S118を経てステップS120に移行した後、ユーザによって「検索」ボタンが押されると、ステップS122に移行し、検索実行部24は、図9(c)に示す検索ワード群「アセトン 海外 使用制限」を用いてドキュメントDB40を検索し、検索結果を表示する。

20

【0057】

ステップS122の後は、ステップS124に移行し、入力受付部20は、図6のステップS14と同様、検索結果の1つがユーザ端末70において選択されるまで待機する。ユーザ端末70において、ユーザが検索結果の中から1つの文書を選択すると、入力受付部20は、その旨を入力受付部20が提供部26に通知し、ステップS126に移行する。ステップS126では、提供部26は、図6のステップS16と同様、選択された文書の概要を示す画面をユーザ端末70に対して送信する(図7(b)参照)。

【0058】

次いで、ステップS128では、入力受付部20が、図6のステップS18と同様、ユーザ端末70において、ユーザにより図7(b)の画面において「ダウンロード」ボタンが押されるまで待機する。ユーザ端末70においてユーザにより「ダウンロード」ボタンが押されると、入力受付部20は、その旨を提供部26に通知し、ステップS130に移行する。ステップS130に移行すると、提供部26は、図6のステップS20と同様、ドキュメントDB40から選択された文書のデータを読み出し、ユーザ端末70に送信することで、ダウンロードを実行する。

30

【0059】

以上により、図8の全処理が終了する。なお、図8の処理は、ユーザから図9(a)の検索画面(初心者用)を用いた検索要求が出されるたびに、繰り返し実行される。

【0060】

なお、本実施形態では、上述したように、「履歴から変換」ボタンが押された場合に、検索ワード変換部22は、固定していない検索ワードを変換した結果を検索画面(初心者用)に表示する場合について説明した。しかしながら、これに限られるものではない。例えば、図10(a)に示すように「アセトン 海外 有用性試験」が検索ワード群として設定されている状態で、「履歴から変換」ボタンが押されたとする。このとき、固定されていない検索ワード「有用性試験」の変換候補ワードが複数抽出された場合には、図10(b)に示すように、各変換候補のワードで変換した後の検索ワード群を選択可能に表示するようにしてもよい。ユーザは、表示された検索ワード群のいずれか(例えば「アセトン 海外 使用制限」)を選択することで、図10(c)に示すように、検索に用いる検索ワード群を変更することができる。なお、図10(b)においては、例えば、変換候補のワ

40

50

ードのTF-IDF値の分散が大きいものから順に、上から表示することができる。これにより、検索ワード変換部22が適切と判断したな検索ワード群ほど上位に（ユーザの目につきやすく）表示することができる。

【0061】

なお、本実施形態では、検索ワードが1つであっても「履歴から変換」ボタンを押すことができるものとする。例えば、検索ワードとして「アセトン」が設定されている状態で、ユーザが「履歴から変換」ボタンを押したとする。この場合、検索ワード変換部22は、検索ワード「アセトン」を固定して、ユーザと同部門の有識者の検索履歴から、固定した検索ワード「アセトン」が含まれる検索履歴を抽出する。そして、検索ワード変換部22は、TF-IDFテーブル48を参照して、固定した検索ワード「アセトン」以外のワードのTF-IDF値の分散を抽出し、TF-IDF値の分散が最も大きいものを検索ワードに追加する。このようにすることで、ユーザが入力した検索ワードが1つの場合に、適切な検索ワードを追加した上で検索を行うことが可能となる。

10

【0062】

なお、これまでの説明から明らかなように、本実施形態では、入力受付部20により、複数のユーザのそれぞれに対応する検索条件を受け付ける検索条件受付部、及び複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付ける評価受付部、としての機能が実現されている。また、検索ワード変換部22と検索実行部24とにより、評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、評価が高いユーザの検索履歴に基づいて検索条件を変換し、変換して得られた検索条件による検索を実行する検索部としての機能が実現されている。

20

【0063】

以上、詳細に説明したように、本実施形態によると、入力受付部20が、検索画面において複数のユーザが入力した検索ワードを受け付けるとともに（S10：肯定）、検索評価画面において複数のユーザが入力した検索結果に関する評価を受け付け（S24：肯定）、評価部28が、受け付けた評価が高いユーザを「有識者」として特定する（S26）。そして、入力受付部20が、有識者とは異なるユーザが検索画面（初心者用）において入力した検索ワードを受け付けた場合に（S102：肯定）、検索ワード変換部22が、有識者の検索履歴に基づいて検索ワードを変換し（S106～S118）、検索実行部24が、変換後の検索ワードを用いた検索を実行する（S122）。これにより、本実施形態では、有識者の検索履歴に基づいて、有識者以外のユーザが入力した検索ワードをより適切な検索ワードに変換した上で検索を実行することができるため、検索精度を向上することが可能となる。また、ユーザは、必要な文書を探し出すまで、検索ワードの一部を変更して検索を実行するという作業を繰り返す必要がなくなるため、検索に要する労力を低減し、検索に要する時間を削減することができる。また、検索ワードが自動的に変換されることで、ユーザは、有識者であればどのような検索ワードを設定するかを容易に知ることができる。これにより、ユーザの業務能力の向上を図ることが可能となる。

30

【0064】

また、本実施形態では、検索ワード変換部22は、ステップS110やS118において検索を行うユーザと同一部門の有識者の検索履歴を用いることとしている。これにより、ユーザが参考にすべき知識を有する有識者の検索履歴に基づいて、ユーザが設定した検索ワードを適切な検索ワードに変換することができる。

40

【0065】

また、本実施形態では、検索ワード変換部22は、ユーザによって「類義語変換」ボタンが押された場合に、各検索ワードの類義語（類似ワード）を類義語辞書DB45から抽出し、検索ワードとその類義語のうち、同部門の有識者の検索履歴における出現頻度が最も高いものを検索ワードとする。これにより、ユーザが検索したい内容を変えずに、入力した検索ワードを部門内でよく使われる検索に適した検索ワードに変換することができる。

【0066】

50

また、本実施形態では、検索ワード変換部 22 は、ユーザによって「履歴から変換」ボタンが押されると、1つの検索ワード以外の検索ワードを IDF 値に基づいて固定し、検索の軸として設定する (S 116)。これにより、検索の方向性を固定したうえで、固定しない検索ワードを適切な検索ワードに変換することが可能となる。

【0067】

また、本実施形態では、検索ワード変換部 22 は、固定された検索ワードに基づいて、検索履歴テーブル 44 から固定されていない検索ワードの変換候補のワードを特定し、特定された変換候補のワードのうち TF - IDF 値の分散が最大のもので、固定されていない検索ワードを変換する (S 118)。これにより、変換候補のワードのうち、文書の分類に貢献している適切なワードで検索ワードを変換することができる。

10

【0068】

なお、上記実施形態では、図 9 (b) に示すように「アセトン 海外 有用性試験」が検索ワード群として設定された状態から「履歴から変換」ボタンが押された場合に、検索ワードを固定する方法として、IDF 値を利用する例について説明した。しかしながら、これに限られるものではなく、例えば、検索ワード変換部 22 は、検索ワードを固定する際に、専門用語を格納した専門用語辞典 DB を更に参照するようにしてもよい。具体的には、検索ワード変換部 22 は、検索ワード群「アセトン 海外 有用性試験」のうち、専門用語辞典 DB に含まれているワード (例えばアセトン) を固定する。その後、検索ワード変換部 22 は、検索ワード「海外」、「有用性試験」のうち、IDF 値が最も小さいワード以外を固定する。このようにすることで、専門用語を必ず固定するようにした上で、固定されていない検索ワードをより適切なワードに変換することができるため、検索精度を向上することが可能となる。なお、IDF 値を用いずに、専門用語辞書 DB のみを用いて検索ワードを固定するようにしてもよい。

20

【0069】

なお、上記実施形態では、ステップ S 110 や S 118 において、部門ごとの出現頻度や検索履歴を用いる場合について説明したが、これに限られるものではない。すなわち、部門を考慮せず、会社全体の出現頻度や検索履歴を用いることとしてもよい。

【0070】

なお、上記実施形態では、ステップ S 116 において、IDF 値が最も小さい検索ワード以外の検索ワードを固定する場合について説明したが、これに限らず、例えば、IDF 値が所定の閾値よりも大きい検索ワードを固定するようにしてもよい。また、固定する検索ワードの数の上限を定めるようにしてもよい。

30

【0071】

なお、上記実施形態では、検索ワード変換部 22 が、ステップ S 116 において検索ワードを変換する際に、TF - IDF 値の分散を用いることとしたが、これに限られるものではない。例えば、検索ワード変換部 22 は、ステップ S 116 において、文書毎に求められる TF 値の分散を用いることとしてもよい。

【0072】

なお、上記実施形態では、ユーザテーブル 42 に格納されている評価平均値に基づいて、自動的に「有識者」と「一般」を区別する場合について説明したが、これに限られるものではない。例えば、「有識者」と「一般」を管理者等が手入力により区別することとしてもよい。

40

【0073】

なお、上記実施形態では、扱う業務が共通しており、検索傾向が類似する部門がある場合に、それらの部門を予め関連付けておいてもよい。この場合、異なる部門であっても、ステップ S 110 や S 118 において同一の部門として取り扱うようにしてもよい。

【0074】

なお、上記実施形態では、図 7 (a) の画面と図 9 (a) の画面をユーザが選択できる場合について説明したが、これに限られるものではない。例えば、入力受付部 20 は、検索を行うユーザが「有識者」か「一般」かをユーザテーブル 42 に基づいて判断し、「有識

50

者」であれば、図7(a)の画面をユーザ端末70に送信し、「一般」であれば、図9(a)の画面をユーザ端末70に送信することとしてもよい。また、検索画面としては、図9(a)の画面のみを用意することとしてもよい。この場合、図9(a)の画面において、ユーザが「類義語変換」ボタンや「履歴から変換」ボタンを押さずに、「検索」ボタンを押した場合に、図6の処理を行うようにし、ユーザが「類義語変換」ボタンや「履歴から変換」ボタンを押した場合に、図8のステップS106以降の処理を行うようにしてもよい。

【0075】

なお、上記実施形態では、検索処理システム100が会社内で利用される場合について説明したが、これに限られるものではない。会社内に限らず、学校等で用いることとしてもよいし、予め定めたグループ内で用いることとしてもよい。また、検索処理システム100の利用範囲を会社や学校等に限定しないこととしてもよい。

10

【0076】

なお、上記実施形態では、文書を検索する際の検索ワードを検索結果に関する評価が高いユーザの検索履歴に基づいて変換する場合について説明したが、これに限られるものではない。すなわち、何らかの情報を検索する場合に、入力された検索条件を検索結果に関する評価が高いユーザの検索履歴に基づいて変換することとしてもよい。

【0077】

なお、上記の処理機能は、コンピュータによって実現することができる。その場合、処理装置が有すべき機能の処理内容を記述したプログラムが提供される。そのプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体(ただし、搬送波は除く)に記録しておくことができる。

20

【0078】

プログラムを流通させる場合には、例えば、そのプログラムが記録されたDVD(Digital Versatile Disc)、CD-ROM(Compact Disc Read Only Memory)などの可搬型記録媒体の形態で販売される。また、プログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することもできる。

【0079】

プログラムを実行するコンピュータは、例えば、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、自己の記憶装置に格納する。そして、コンピュータは、自己の記憶装置からプログラムを読み取り、プログラムに従った処理を実行する。なお、コンピュータは、可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することもできる。また、コンピュータは、サーバコンピュータからプログラムが転送されるごとに、逐次、受け取ったプログラムに従った処理を実行することもできる。

30

【0080】

上述した実施形態は本発明の好適な実施の例である。但し、これに限定されるものではなく、本発明の要旨を逸脱しない範囲内において種々変形実施可能である。

40

【0081】

なお、以上の実施形態の説明に関して、更に以下の付記を開示する。

(付記1) 複数のユーザのそれぞれに対応する検索条件を受け付け、

前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付け、

前記評価が高いユーザを特定し、

前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する、

処理をコンピュータに実行させることを特徴とする検索処理プログラム。

(付記2) 前記評価が高いユーザとは異なるユーザは、前記評価が高いユーザと同一の

50

業務に従事するユーザであることを特徴とする付記 1 に記載の検索処理プログラム。

(付記 3) 前記検索条件は、検索ワードを含み、

前記検索条件を変換する場合、前記検索ワードに類似する類似ワードを特定し、前記類似ワードのうち前記評価が高いユーザが検索に利用した頻度が高いものを優先して、前記検索ワードと変換する、

ことを特徴とする付記 1 又は 2 に記載の検索処理プログラム。

(付記 4) 前記検索条件は、複数の検索ワードを組み合わせた検索ワード群を含み、

前記検索条件を変換する場合、前記検索ワード群から変換を行わない検索ワードと変換対象の検索ワードとを特定し、特定した前記変換対象の検索ワードを、前記評価が高いユーザの検索履歴に基づいて変換する、

ことを特徴とする付記 1 ~ 3 のいずれかに記載の検索処理プログラム。

(付記 5) 前記検索を実行する処理は、複数の文書の中から前記検索ワード群を含む文書を検索する処理であり、

前記検索ワード群に含まれる複数の検索ワードのうち、前記複数の文書における IDF (Inverse Document Frequency) 値が小さいものを優先して、前記変換を行わない検索ワードとして特定する、

ことを特徴とする付記 4 に記載の検索処理プログラム。

(付記 6) 前記変換対象の検索ワードを変換する場合、前記変換を行わない検索ワードを含む検索ワード群を前記評価が高いユーザの検索履歴から抽出し、抽出した前記検索ワード群に基づいて、前記変換対象の検索ワードを変換する、

ことを特徴とする付記 5 に記載の検索処理プログラム。

(付記 7) 前記変換対象の検索ワードを変換する場合、抽出した前記検索ワード群に含まれる前記変換を行わない検索ワード以外の検索ワードのうち、前記複数の文書それぞれにおける TF (Term Frequency) - IDF 値の分散が大きいものを優先して、前記変換対象の検索ワードと変換する、

ことを特徴とする付記 6 に記載の検索処理プログラム。

(付記 8) 複数のユーザのそれぞれに対応する検索条件を受け付け、

前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付け、

前記評価が高いユーザを特定し、

前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する、

処理をコンピュータが実行することを特徴とする検索処理方法。

(付記 9) 複数のユーザのそれぞれに対応する検索条件を受け付ける検索条件受付部と、

前記複数のユーザそれぞれに対応する検索結果に関する評価の入力を受け付ける評価受付部と、

前記評価が高いユーザを特定する特定部と、

前記評価が高いユーザとは異なるユーザに対応する検索条件を受け付けた場合に、前記評価が高いユーザの検索履歴に基づいて前記検索条件を変換し、変換して得られた検索条件による検索を実行する検索部と、

を備える検索処理装置。

(付記 10) 前記評価が高いユーザとは異なるユーザは、前記評価が高いユーザと同一の業務に従事するユーザであることを特徴とする付記 9 に記載の検索処理装置。

(付記 11) 前記検索条件は、検索ワードを含み、

前記検索部は、前記検索ワードに類似する類似ワードを特定し、前記類似ワードのうち前記評価が高いユーザが検索に利用した頻度が高いものを優先して、前記検索ワードと変換する、

ことを特徴とする付記 9 又は 10 に記載の検索処理装置。

(付記 12) 前記検索条件は、複数の検索ワードを組み合わせた検索ワード群を含み、

前記検索部は、前記検索ワード群から変換を行わない検索ワードと変換対象の検索ワード

10

20

30

40

50

とを特定し、特定した前記変換対象の検索ワードを、前記評価が高いユーザの検索履歴に基づいて変換する、

ことを特徴とする付記 9 ~ 11 のいずれかに記載の検索処理装置。

(付記 13) 前記検索部は、複数の文書の中から前記検索ワード群を含む文書を検索し、前記検索ワード群に含まれる複数の検索ワードのうち、前記複数の文書における IDF 値が小さいものを優先して、前記変換を行わない検索ワードとして特定する、

ことを特徴とする付記 12 に記載の検索処理装置。

(付記 14) 前記検索部は、前記変換を行わない検索ワードを含む検索ワード群を前記評価が高いユーザの検索履歴から抽出し、抽出した前記検索ワード群に基づいて、前記変換対象の検索ワードを変換する、

10

ことを特徴とする付記 13 に記載の検索処理装置。

(付記 15) 前記検索部は、抽出した前記検索ワード群に含まれる前記変換を行わない検索ワード以外の検索ワードのうち、前記複数の文書それぞれにおける TF - IDF 値の分散が大きいものを優先して、前記変換対象の検索ワードと変換する、

ことを特徴とする付記 14 に記載の検索処理装置。

【符号の説明】

【0082】

- 10 サーバ(検索処理装置)
- 20 入力受付部(検索条件受付部、評価受付部)
- 22 検索ワード変換部(検索部の一部)
- 24 検索実行部(検索部の一部)
- 28 評価部(特定部)

20

30

40

50

【図面】

【図1】

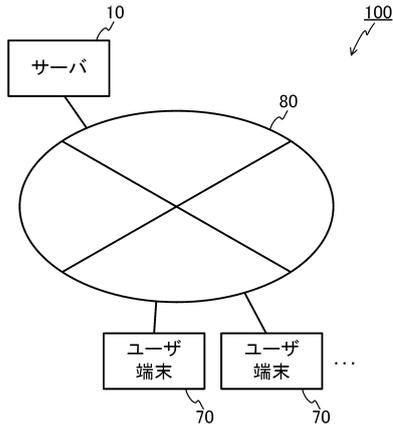
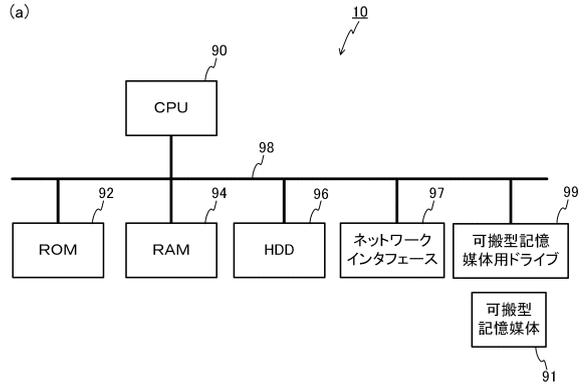
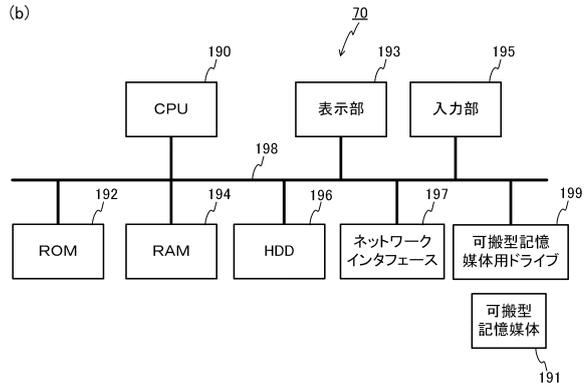


図1

【図2】



10



20

図2

【図3】

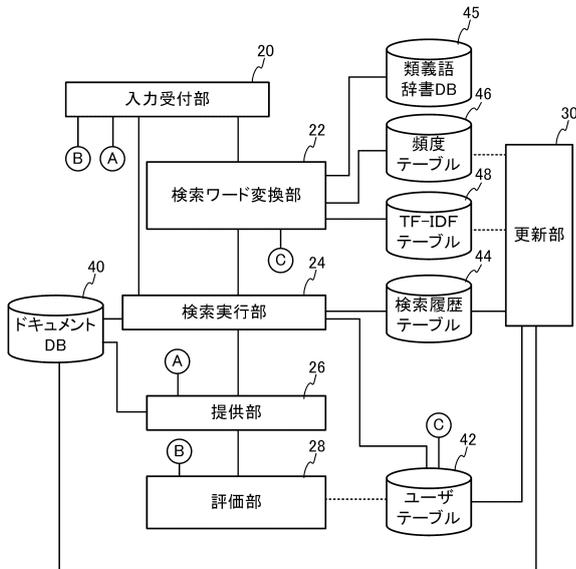


図3

【図4】

(a) <ユーザテーブル42>

ユーザ	評価平均値	区分	部門
A	8.8	有識者	A部門
B	8.5	有識者	A部門
C	7.5	一般	D部門
...

30

(b) <検索履歴テーブル44>

有識者	ワード1	ワード2	ワード3	ワード4	...
A	溶媒	植物性	有用性		...
A	アセトン	海外	制限		...
B	樽エキス	訴求効果	メディア		...
...

40

図4

50

【 図 5 】

(a) <頻度テーブル46>

	A部門での出現頻度	B部門での出現頻度	C部門での出現頻度	...
溶媒	0.1	0.2	0.03	...
アセトン	0.05	0.2	0.01	...
樽エキス	0.001	0.1	0	...
...

(b) <TF-IDFテーブル48>

	溶媒	アセトン	樽エキス	...
TF値	0.3	0.1	0.03	...
IDF値	1.1	1.5	9	...
TF-IDF値	0.33	0.15	0.27	...
TF-IDF値の分散	Va	Vb	Vc	...

図5

【 図 6 】

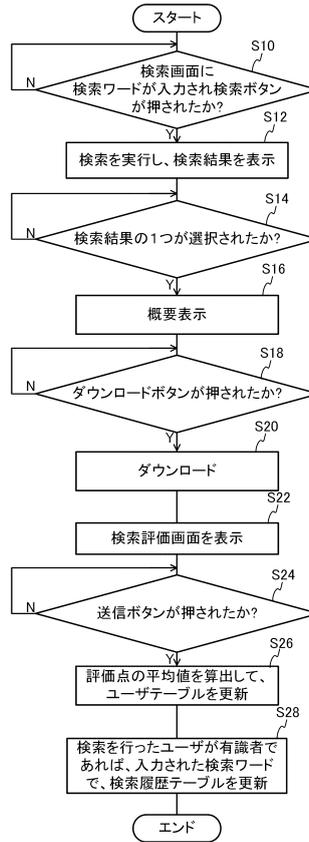


図6

【 図 7 】

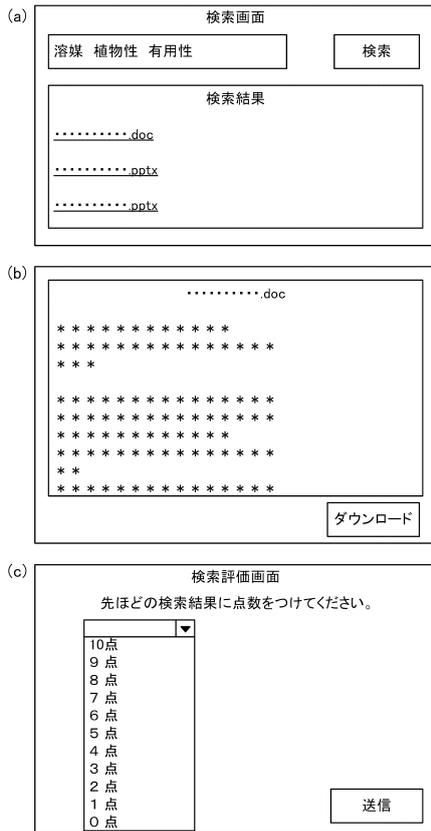


図7

【 図 8 】

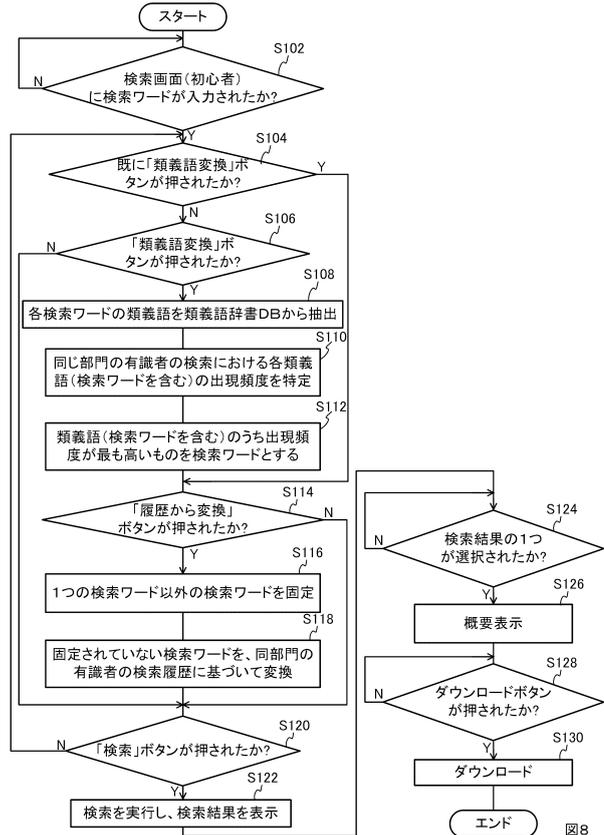


図8

10

20

30

40

50

【 図 9 】

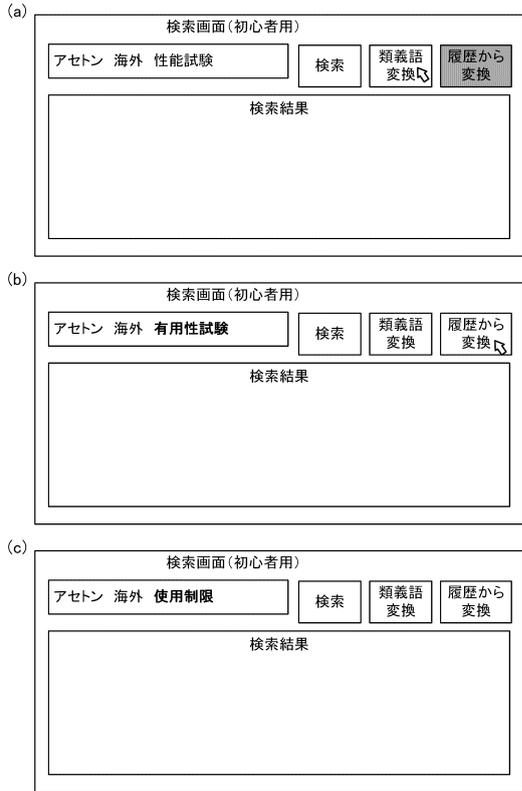


図9

【 図 1 0 】

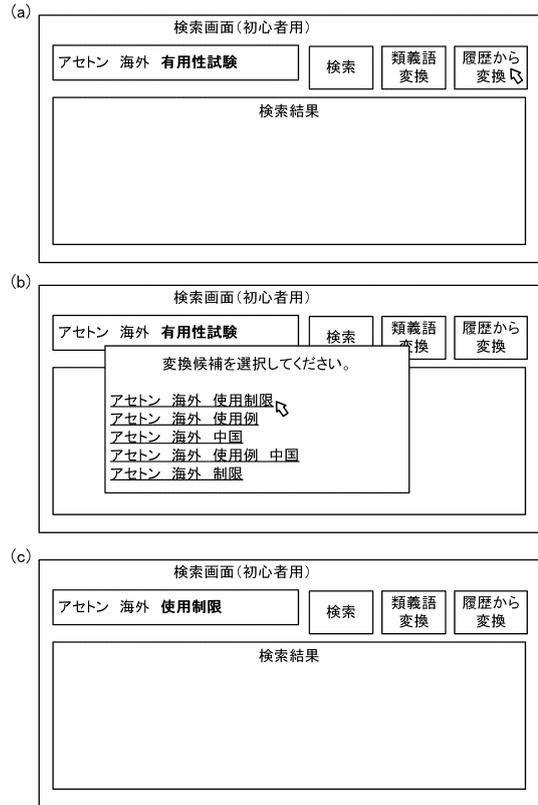


図10

10

20

30

40

50

フロントページの続き

- (56)参考文献 米国特許出願公開第2009/0144262 (US, A1)
特開2005-071002 (JP, A)
特開2003-323205 (JP, A)
特開平06-314296 (JP, A)
米国特許出願公開第2008/0319976 (US, A1)
特開2008-203997 (JP, A)
特開2005-099884 (JP, A)
米国特許出願公開第2017/0364585 (US, A1)
滝川 真弘, ノイズに頑健な分野別単語排他度の提案とTwitterユーザの専門性推定への適用, 第9回データ工学と情報マネジメントに関するフォーラム (第15回日本データベース学会年次大会) [online], 電子情報通信学会データ工学研究専門委員会日本データベース学会 情報処理学会データベースシステム研究会, 2017年02月27日, DEIM Forum 2017 D5-3, Internet URL:<http://db-event.jp.org/deim2017/papers/273.pdf>
- (58)調査した分野 (Int.Cl., DB名)
G06F 16/00 - 16/958