

(19) 日本国特許庁 (JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-55494

(P2010-55494A)

(43) 公開日 平成22年3月11日 (2010.3.11)

(51) Int.Cl.	F I	テーマコード (参考)
G 0 6 F 17/30 (2006.01)	G O 6 F 17/30 3 5 O C	5 B O 7 5
G 0 6 Q 10/00 (2006.01)	G O 6 F 17/60 1 7 4	
	G O 6 F 17/30 3 4 O B	
	G O 6 F 17/30 3 7 O Z	
	G O 6 F 17/30 1 1 O C	
審査請求 未請求 請求項の数 10 O L (全 13 頁)		

(21) 出願番号	特願2008-221606 (P2008-221606)	(71) 出願人	000000295
(22) 出願日	平成20年8月29日 (2008.8.29)		沖電気工業株式会社
			東京都港区西新橋三丁目16番11号
		(74) 代理人	100085198
			弁理士 小林 久夫
		(74) 代理人	100098604
			弁理士 安島 清
		(74) 代理人	100125494
			弁理士 山東 元希
		(74) 代理人	100061273
			弁理士 佐々木 宗治
		(74) 代理人	100070563
			弁理士 大村 昇
		(74) 代理人	100087620
			弁理士 高梨 範夫

最終頁に続く

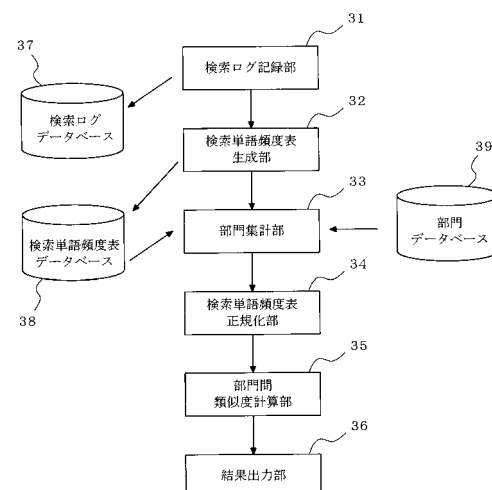
(54) 【発明の名称】 検索分析サーバ装置及び検索分析方法

(57) 【要約】

【課題】ある部門間で、まだコミュニケーションがなされていなくても、その部門間の業務上のつながりを手間をかけずに自動的に発見することが可能な技術を提供する。

【解決手段】組織内の複数の各部門に設置された複数の端末10からそれぞれ送信された検索要求の履歴を、検索単語とその検索要求を行った個人の個人IDと対応付けて記録する検索ログ記録部31と、検索ログを個人ごとに集計し、検索単語とその検索回数とを対応付けた個人別検索単語頻度表を生成する検索単語頻度表生成部32と、個人別検索単語頻度表を部門ごとに集計し、部門別検索単語頻度表を生成する部門集計部33と、部門別検索単語頻度表の検索回数を正規化する検索単語頻度表正規化部34と、正規化して得られた部門ごとの頻度データを基に、部門間の検索単語の類似度を計算する類似度計算部35と、類似度結果を出力する結果出力部36とを備えた。

【選択図】 図3



【特許請求の範囲】

【請求項 1】

組織内の複数の各部門に設置された複数の端末にネットワークを介して接続される検索分析サーバ装置であって、

前記複数の端末からそれぞれ送信された検索要求の履歴を、該検索要求による検索単語とその検索要求を行った個人の個人 ID と対応付けて記録する検索ログ記録部と、

前記検索ログ記録部に記録された検索ログを個人ごとに集計し、検索単語とその検索回数とを対応付けた個人別検索単語頻度表を生成する検索単語頻度表生成部と、

前記個人ごとの個人別検索単語頻度表を部門ごとに集計し、部門別検索単語頻度表を生成する部門集計部と、

前記部門別検索単語頻度表の検索回数を正規化する検索単語頻度表正規化部と、

該検索単語頻度表正規化部で正規化して得られた部門ごとの頻度データを基に、部門間の検索単語の類似度を計算する類似度計算部と、

該類似度計算部で計算された類似度結果を出力する結果出力部とを備えたことを特徴とする検索分析サーバ装置。

【請求項 2】

前記検索単語頻度表正規化部は、前記個人別検索単語頻度表の検索回数を正規化し、前記類似度計算部は、前記検索単語頻度表正規化部で正規化して得られた個人ごとの頻度データ及び前記部門ごとの頻度データを基に、個人間または個人と部門間の検索単語の類似度を計算することを特徴とする請求項 1 記載の検索分析サーバ装置。

【請求項 3】

前記検索単語頻度表正規化部は、TF-IDF (Term Frequency-Inverse Document Frequency) に基づき検索単語の検索回数を正規化することを特徴とする請求項 1 または請求項 2 記載の検索分析サーバ装置。

【請求項 4】

前記検索単語頻度表正規化部は、検索回数を正規化して頻度ベクトルを求め、前記類似度計算部は、前記類似度を、前記頻度ベクトルのコサイン距離として計算することを特徴とする請求項 1 乃至請求項 3 の何れかに記載の検索分析サーバ装置。

【請求項 5】

前記結果出力部は、部門ごとに、その部門で検索回数が多い検索単語と、その部門と類似度が高い部門名を類似度が高い順に示したリストとを類似度結果として作成して出力することを特徴とする請求項 1 乃至請求項 4 の何れかに記載の検索分析サーバ装置。

【請求項 6】

組織内の複数の各部門に設置された複数の端末からそれぞれ送信された検索要求の履歴を、該検索要求による検索単語とその検索要求を行った個人の個人 ID とを対応づけて検索ログ記録部に検索ログとして記録する記録工程と、

前記検索ログ記録部に記録された検索ログを個人ごとに集計し、検索単語とその検索回数とを対応付けた個人別検索単語頻度表を生成する検索単語頻度表生成工程と、

前記個人ごとの個人別検索単語頻度表を部門ごとに集計し、部門別検索単語頻度表を生成する部門集計工程と、

前記部門別検索単語頻度表の検索回数を正規化する検索単語頻度表正規化部と、

前記正規化して得られた部門ごとの頻度データを基に、部門間の検索単語の類似度を計算する類似度計算工程と、

前記計算された類似度結果を出力する結果出力工程とを備えることを特徴とする検索分析方法。

【請求項 7】

前記検索単語頻度表正規化工程は、前記個人別検索単語頻度表の検索回数を正規化しし、前記類似度計算工程は、前記正規化して得られた個人ごとの頻度データ及び前記部門ごとの頻度データを基に、個人間または個人と部門間の検索単語の類似度を計算することを特徴とする請求項 6 記載の検索分析方法。

【請求項 8】

前記検索単語頻度表正規化工程では、T F - I D F (Term Frequency-Inverse Document Frequency) に基づき検索単語の頻度を正規化することを特徴とする請求項 6 または請求項 7 記載の検索分析方法。

【請求項 9】

前記検索単語頻度表正規化工程では、検索回数を正規化して頻度ベクトルを求め、前記類似度計算部は、前記類似度を、前記頻度ベクトルのコサイン距離として計算することを特徴とする請求項 6 乃至請求項 8 の何れかに記載の検索分析方法。

【請求項 10】

前記結果出力工程では、部門ごとに、その部門で検索回数が多い検索単語と、その部門と類似度が高い部門名を類似度が高い順に示したリストとを類似度結果として作成して出力することを特徴とする請求項 6 乃至請求項 9 の何れかに記載の検索分析方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、個人や部門間での業務の関連性を取得するための検索分析サーバ装置及び検索分析方法に関する。

【背景技術】

【0002】

近年、情報化社会の進展に伴い、企業内での知的生産性を高める必要が生じている。組織の知的生産性を高めるためには、情報の共有や部門間の協力関係の構築が重要なのは言うまでもない。企業内での異なる部門であっても業務上の関連性が高ければ、いち早くそのような部門を発見して、協力関係を構築することで組織全体としての知的な生産性が向上する。そのために社内のネットワークで部門ごとにホームページを作成して情報発信を行ったり、社内でのブログや SNS などの取り組みもなされている。

20

【0003】

ただ、このような方法では情報発信に手間がかかったり、他の部門の情報をチェックするなどの手間がかかるという問題がある。そこで、人手によらずにコンピューターを用いて機械的に個人や部門間の関係性を把握しようとする試みがなされている。例えば、組織内でのメールなどによるコミュニケーションを計測して、部門間のコミュニケーションの生産性を評価しようとする試みが開示されている（例えば、特許文献 1 参照）。

30

【0004】

また、会議等のスケジュール情報から出席者の共起関係を調べ、人脈情報の生成、表示を行うようにした技術が開示されている（例えば、特許文献 2 参照）。

【特許文献 1】特開 2007-280000 号公報

【特許文献 2】特開 2005-108123 号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかし、特許文献 1 では、既にコミュニケーションがなされている場合に、関連した指標を計測するものの、まだコミュニケーションがなされていない部門間の関係については処理の対象としていない。従って、まだコミュニケーションがなされていない部門間であっても、業務上の関連がある場合にそれを発見して部門間の情報共有、提携を促す目的には向かない。

40

【0006】

また、特許文献 2 でも、やはり会議等の既にコミュニケーションがなされている場合を主な処理の対象としているため、まだ交流がない部門間で業務の関連性を発見する目的には向かないと考えられる。

【0007】

本発明はこのような点に鑑みなされたもので、ある部門間で、まだコミュニケーション

50

がなされていなくても、その部門間の業務上のつながりを手間をかけずに自動的に発見することが可能な検索分析サーバ装置及び検索分析方法を得ることを目的とする。

【0008】

また、個人間や個人と部門間においても、業務上のつながりを手間をかけずに自動的に発見することが可能な検索分析サーバ装置及び検索分析方法を得ることを目的とする。

【課題を解決するための手段】

【0009】

本発明に係る検索分析サーバ装置は、組織内の複数の各部門に設置された複数の端末にネットワークを介して接続される検索分析サーバ装置であって、複数の端末からそれぞれ送信された検索要求の履歴を、検索要求による検索単語とその検索要求を行った個人の個人IDと対応付けて記録する検索ログ記録部と、検索ログ記録部に記録された検索ログを個人ごとに集計し、検索単語とその検索回数とを対応付けた個人別検索単語頻度表を生成する検索単語頻度表生成部と、個人ごとの個人別検索単語頻度表を部門ごとに集計し、部門別検索単語頻度表を生成する部門集計部と、部門別検索単語頻度表の検索回数を正規化する検索単語頻度表正規化部と、検索単語頻度表正規化部で正規化して得られた部門ごとの頻度データを基に、部門間の検索単語の類似度を計算する類似度計算部と、類似度計算部で計算された類似度結果を出力する結果出力部とを備えたものである。

【0010】

また、本発明に係る検索分析サーバ装置は、検索単語頻度表正規化部が、個人別検索単語頻度表の検索回数を正規化し、類似度計算部が、検索単語頻度表正規化部で正規化して得られた個人ごとの頻度データ及び部門ごとの頻度データを基に、個人間または個人と部門間の検索単語の類似度を計算するものである。

【発明の効果】

【0011】

本発明によれば、社内における部門または個人が検索に用いた単語を分析することで、まだ互いにコミュニケーションがなされていない部門間、または個人間の業務の関係性を自動的に発見することが可能となる。

【発明を実施するための最良の形態】

【0012】

以下、図面を参照して本発明の実施形態を説明する。本発明は組織内の個人が行っている情報検索に注目する。なぜならば、そのような検索作業は普段の業務の必要から生じるものであり、従って検索の際にどのような単語を用いたかを調べることで、その個人や部門の業務内容を相当程度把握することができると考えられるからである。本発明の本質は、このような検索に用いた単語を個人または部門ごとに集計し、個人や部門等の類似度を調べることで個人または部門間の関連性を発見することにある。

【0013】

図1は、本発明の一実施の形態のネットワーク構成を示す図である。

図1に示す検索分析システムは、PC等の各端末10（図1中の端末1～端末N）と、各端末10から送信された検索要求を分析する検索分析サーバ装置20とがネットワーク30を介して接続されている。各端末10は、組織内の複数の各部門に分散して設けられており、社内LAN等のネットワーク30及び検索分析サーバ装置20を介してインターネットなどのネットワーク40に接続されている。

【0014】

各端末10は、組織内の個人によって操作されるもので、この端末10を使用して各個人が情報検索を行う。各端末10から送信される検索要求は、ネットワーク30を通して検索分析サーバ装置20へと送信される。検索分析サーバ装置20は、受信した検索要求を処理して検索結果を返信するとともに、以下に詳述する検索分析処理を行う。このような検索は社内の文書の検索であってもよいし、インターネット検索でもよい。後者の場合には検索分析サーバ装置20は例えばウェブ・プロキシであってもよい。

【0015】

図 2 は、本発明の検索分析サーバ装置のハードウェア構成を示すブロック図である。

検索分析サーバ装置 20 は、具体的には例えばパソコンなどのコンピュータで構成され、CPU 等の演算部 21 と、ROM や RAM で構成される主記憶装置 22 と、ハードディスクなどの補助記憶装置 23 と、端末 10 等の外部装置とデータ送信を行うための通信インタフェース (I/F) 24 と、ディスプレイなどの表示装置 25 とを備えている。

【0016】

図 3 は、本発明の検索分析サーバ装置の機能的な構成を示すブロック図である。

検索分析サーバ装置 20 は、検索ログ記録部 31 と、検索単語頻度表生成部 32 と、部門集計部 33 と、検索単語頻度表正規化部 34 と、部門間類似度計算部 35 と、結果出力部 36 と、検索ログデータベース 37 と、検索単語頻度表データベース 38 と、部門データベース 39 とを備えている。各処理部 31 ~ 36 は、CPU と主記憶装置 22 内の制御プログラムとにより機能的に実現されている。

【0017】

検索ログ記録部 31 は、各端末 10 からそれぞれ送信された検索要求の履歴を記録するもので、個人が検索に用いた単語と、その検索を行った個人の個人 ID と対応させて検索ログデータベース 37 に検索ログとして記録する。なお、検索分析サーバ装置 20 は、各端末 10 から検索要求を受け付ける際、その端末 10 を操作する個人に対し、個別の個人 ID の入力を求めるようにしており、検索ログ記録部 31 は、その際に入力された個人 ID を記録するものである。また、検索分析サーバ装置 20 がインターネット検索のウェブ・プロキシである場合には、個人が特定の検索サイト (Google など) を用いて検索している際にその検索単語を個人 ID と共にログとして記録するようにすればよい。

【0018】

検索単語頻度表生成部 32 は、検索ログデータベース 37 に記録された検索ログを各個人 ID ごとに集計し、当該個人が検索に用いた文字列と、当該個人によるその文字列の検索回数とからなる個人別検索単語頻度表を生成して検索単語頻度表データベース 38 に記録する。図 4 は、検索単語頻度表を示したものであり、検索に用いた文字列 (検索単語) とその検索回数 (頻度) とを示している。

【0019】

また、部門データベース 39 には、部門 ID とその部門に所属する全個人の個人 ID リストとが対応付けて記憶されている。部門集計部 33 は、この部門データベース 39 を用いて、各部門ごとに、その部門に所属する全ての個人の個人別検索単語頻度表を検索単語頻度表データベース 38 から取得して合成し、部門ごとの部門別検索単語頻度表を生成する。

【0020】

検索単語頻度表正規化部 34 は、部門集計部 33 で得られた部門ごとの部門別検索単語頻度表の頻度を、TF-IDF (Term Frequency-Inverse Document Frequency) 等によって正規化する。部門間類似度計算部 35 は、検索単語頻度表正規化部 34 で正規化された各部門それぞれの正規化結果を用いて、部門間の類似度を計算する。

【0021】

結果出力部 36 は、部門間類似度計算部 35 で得られた部門間の類似度を表示装置 25 に出力して表示させる。

【0022】

以下、検索分析システムの動作の流れを説明する。

個人は、自己の端末 10 を操作して検索画面を表示し、自己の個人 ID と検索文字列を入力して検索を実行する。この検索の際には、文字列を複数入力するものとする。ここで入力される検索語句は、個人の普段の業務に関連する単語であったり、関心のある単語であったりする。また、個人が、ある特定の分野に関して活動を行う場合には、その活動に関連する語句も検索語句として入力されることが想定される。従って、個人から入力される検索語句は、ある特定の分野に関して活動を行おうとする意志を表したものと言える。このように個人から端末 10 に入力された検索文字列は、個人 ID と共に検索要求として

検索分析サーバ装置 20 へと送信される。

【0023】

検索分析サーバ装置 20 は、端末 10 からの検索要求を受信すると、検索要求に含まれる検索文字列を、個人 ID と対応させて検索ログデータベース 37 に検索ログとして記録する。なお、検索要求に含まれる検索文字列を、そのまま検索ログデータベース 37 に記録してもよいが、形態素分析を使って一度単語に分解し、各単語を個人 ID と対応させて検索ログデータベース 37 に記録するようにしてもよい。検索単語の検索ログデータベース 37 への記録は、個人が検索を行う度に常時行う。

【0024】

＜検索分析処理の方法＞

以下、検索ログデータベース 37 に記録された検索ログを利用した検索分析処理の流れを説明する。

図 5 は、検索分析サーバ装置における検索分析処理の流れを示したフローチャート図である。

検索単語頻度表生成部 32 は、検索ログデータベース 37 から検索ログを取得する (S1)。そして、検索ログを各個人 ID ごとに集計し、当該個人が検索に用いた各文字列と、当該個人によるその文字列の検索回数とからなる個人別検索単語頻度表を生成し (S2)、検索単語頻度表データベース 38 に記録する。個人別検索単語頻度表が一旦生成された後に、その個人別検索単語頻度表に対応する個人からの検索要求に基づき検索ログデータベース 37 が更新された場合には、該当の個人別検索単語頻度表も更新する。すなわち、個人が検索に用いた文字列（または、形態素解析の出力である各単語について）が個人別検索単語頻度表に既に存在するかどうかを調べる。もし既に存在すれば、その文字列の頻度の数値を 1 だけ増加させる。もし存在しなければ、その文字列を検索単語とし、頻度を 1 とする新たな項目をその個人別検索単語頻度表に追加する。

なお、検索単語頻度表生成部 32 における個人別検索単語頻度表の生成・更新のタイミングは、検索ログ記録部 31 で検索ログデータベース 37 が更新される度に行っても良いし、一定期間ごとに定期的に行うようにしても良い。

【0025】

部門集計部 33 は、検索単語頻度表生成部 32 で作成された個人別検索単語頻度表を、部門データベース 39 に記憶された部門別の個人 ID リストを用いて各部門ごとに合成し、部門ごとの検索単語頻度表（以下、部門別検索単語頻度表という）を作成する (S3)。この処理および以降の処理は個人から検索要求が送信される度に行うよりも、ある一定時間ごと、例えば一日ごとに行うほうがよいだろう。

【0026】

以下、部門集計部 33 における合成処理について具体的に説明する。

図 6 は、個人別検索単語頻度表の合成処理の説明図である。

まず、個人 ID リストのうちの最初の個人 ID を取り出し、その個人 ID に対応する検索単語頻度表 T1 を検索単語頻度表データベース 38 から取得する (S11)。この検索単語頻度表 T1 をそのまま部門別検索単語頻度表 X0 とする (S12)。

【0027】

次に、部門集計部 33 は、個人 ID リストから次の個人 ID を取り出し、その個人 ID に対応する検索単語頻度表 T2 を検索単語頻度表データベース 38 から得る (S13)。そして、検索単語頻度表 T2 の各要素について、その検索単語が部門別検索単語頻度表 X0 に存在するかどうか判断し、存在する場合には、対応する頻度を検索単語頻度表 T2 の頻度だけ増加させる。もし部門別検索単語頻度表 X0 にその検索単語が存在しない場合には、新たな要素として、検索単語とその頻度とを検索単語頻度表 T2 から抽出して部門別検索単語頻度表 X0 に追加する (S14)。そして、部門別検索単語頻度表 X を完成する (S15)。以上の処理をその部門の全ての個人 ID について繰り返すことで、部門別検索単語頻度表 X を完成する。

以上の処理を全ての部門について行い、各部門ごとに、それぞれ部門別検索単語頻度表

10

20

30

40

50

Xを作成する。

【0028】

検索単語頻度表正規化部34は、部門集計部33で作成された各部門ごとの部門別検索単語頻度表の頻度を正規化する(S4)。以下、正規化処理について説明する。なお、後の処理の便宜のために、部門集計部33で作成された部門別検索単語頻度表Xを、検索単語をインデックス化して数値ベクトルとしておくのが都合がよい。すなわち、全ての検索単語の総数をNとした時、各検索単語について、適当に1からNまでの数字で一対一に対応づける。その上で各部門の部門別検索単語頻度表を1からNまでのインデックスで対応づけられたベクトルとし、その各値を、そのインデックスに対応する検索単語の頻度とする。これをその部門の検索単語頻度ベクトルと呼ぶ。

10

【0029】

すなわち、ある部門の検索単語頻度ベクトルとは、数字の列

【0030】

【数1】

$$(n_1, n_2, \dots, n_i, \dots, n_N)$$

であって、 n_i は単語ID=iに対応するその部門での検索単語の頻度である。Nは全検索単語数である。

【0031】

図6の部門別検索単語頻度表Xの場合、

(8, 81, 7, 20, ...)

となる。

ただし、検索単語

「テレワーク」、「動向」、「臨場感」、「とは」

は、それぞれ単語ID

1、2、3、4

に対応するものとする。

【0032】

こうして作成された部門ごとの各検索単語頻度ベクトルを、後の部門間類似度計算部35で用いるために正規化する。これはTF-IDF (Term Frequency-Inverse Document Frequency) に基づくもので、自然言語処理ではよく知られている手法である。これは、どんな文章でも一般的によく表れる単語とそうではない単語があり、よく表れる単語については検索単語頻度表におけるその頻度のウェイトを下げ、あまり表れない単語については検索単語頻度表におけるその頻度のウェイトを上げることで、部門が検索している単語の特徴をより引き出すことを狙いとしている。

30

【0033】

すなわち、ある部門の正規化された検索単語頻度ベクトルを

【0034】

【数2】

$$(n_1, n_2, \dots, n_i, \dots, n_N)$$

とする。

ここで、 n_i は単語ID=iに対応するその部門での検索単語の頻度である。Nは全検索単語数である。

この時、この部門の正規化された検索単語頻度ベクトル

【0035】

【数3】

$$(f_1, f_2, \dots, f_i, \dots, f_N)$$

をTF-IDF法により以下のように定める。

50

【 0 0 3 6 】

【数 4】

$$f_i = tf_i \cdot idf_i$$

ここに、

【数 5】

$$tf_i = \frac{n_i}{\sum_{k=1}^N n_k}$$

10

$$idf_i = \log \frac{D}{df_i}$$

【 0 0 3 7 】

ここでDは全検索総数であり、検索ログデータベース37に存在する、全ての個人について、検索した回数を合計したものである。df_iは全検索のうち、単語iを含む検索の回数である。なお、Dおよびdf_iについてはここでは検索ログをもとに集計した値としたが、通常は大量のドキュメントをもとに計算するのが普通であり、そのような値を用いてもよい。その場合には大量のドキュメントがあると仮定して、Dは総ドキュメント数、df_iは単語iを含むドキュメント数である。

20

検索単語頻度表正規化部34では、以上のようにして部門ごとに、正規化された検索単語頻度ベクトルを作成する。

【 0 0 3 8 】

続いて、部門間類似度計算部35は検索単語頻度表正規化部34で作成された部門ごとの正規化された検索単語頻度ベクトルを用いて、全ての部門ペアそれぞれについて類似度を算出する(S5)。この類似度は、各部門がそれぞれ検索に用いた単語が類似している度合いを示すものであって、部門の間の業務の近さを表す指標となる。検索に用いた単語は上述したように普段の業務、関心を反映したものであり、2つの部門が似たような単語で検索しているほど、それらの部門は関連の深い業務をしていると推定できる。

30

【 0 0 3 9 】

任意の2つの部門間の検索単語の類似度の計算について以下に説明する。

2つの部門X、Yの正規化された検索単語頻度ベクトルをそれぞれ、

【 0 0 4 0 】

【数 6】

$$X = (x_1, x_2, \dots, x_i, \dots, x_N)$$

$$Y = (y_1, y_2, \dots, y_i, \dots, y_N)$$

とする。ここで、x_iおよびy_iは、それぞれ単語ID=iに対応する部門Xの検索単語頻度ベクトルと部門Yの検索単語頻度ベクトルの値である。Nは、全検索単語の総数である。

40

この時、部門間の類似度はN次元ベクトルのコサイン距離として以下の式から算出される。

【 0 0 4 1 】

【数 7】

$$\cos(X, Y) = \frac{\sum_{i=1}^N x_i \times y_i}{\sqrt{\sum_{i=1}^N x_i^2} \times \sqrt{\sum_{i=1}^N y_i^2}}$$

50

【0042】

類似度は $-1 \leq (\text{類似度}) \leq 1$ の値をとり、部門間の検索単語の特徴が似ているほど類似度は1に近くなる。部門の人間が普段の業務で検索を行う際に使う単語は、上述したように業務の内容や領域を反映したものとなるので、部門間の類似度が1に近いほどその部門間では業務が似ていて関連が深い可能性が高い。

【0043】

部門間類似度計算部35は、以上に説明した類似度の計算処理を全ての部門ペアごとに行い、それぞれの類似度を算出する。

【0044】

結果出力部36は、部門間類似度計算部35で算出された、全ての部門ペアごとの類似度を表示装置25に出力して表示させる(S6)。

【0045】

図7は、類似度計算結果の表示例を図である。なお、図7には部門Aの類似度計算結果の表示例を示している。

図7に示すように、部門Aで検索頻度が高かった検索単語を複数表示するとともに、部門Aと関連がある可能性のある部門名を、類似度が高い順に順位付けしてリスト表示する。この表示により部門Aでは、例えばテレワークの分野などで部門XXXや部門YYY等と業務の関連性が高いことを知ることができる。

【0046】

類似度計算結果の提示の仕方としては、上述したように検索分析サーバ装置20の表示装置25に表示してもよいし、メールや社内のホームページなどを用いて社内全体に広く提示するようにしてもよい。例えば、全ての部門ペアごとの類似度結果を社内のホームページなどに常時、掲示するようにしてもよい。この場合、各部門はホームページにアクセスすることにより、自分たちと業務上のつながりがありそうな部門を知ることができる。これにより、まだコミュニケーションがなされていないが業務上関連のありそうな部門に注目することができ、必要ならアクセスすることが可能である。そうすることで部門間の情報共有や協力関係の構築を促す。このようにして社内でのコミュニケーション、情報共有、業務の提携などが向上し、組織全体としての生産性が向上することが期待できる。

【0047】

また、類似度結果の提示の仕方としては、他に例えば、端末10からの類似度計算要求に応じて検索分析サーバ装置20で類似度を計算し、類似度計算結果を結果出力部36から通信I/F24を介して端末10に送信し、端末10側で表示するようにしてもよい。

【0048】

このように本実施の形態では、検索で用いられる検索単語が、その検索を行う個人やその個人が属する部門の業務内容に関連する語句であり、また、今後の業務活動に関連する語句でもあることに着目し、各個人の検索ログを利用して部門間の類似度を自動的に計算し、その類似度計算結果を出力するようにした。したがって、この検索分析サーバ装置20を用いることにより、まだ互いにコミュニケーションがなされていない部門間の業務の関係性を自動的に発見することが可能となる。また、算出された類似度に基づいて2つの部門間における業務の関連性の程度を知ることが可能となる。したがって、互いの業務に関連がありそうな部門同士、また、ある分野に関して活動を行おうとする意志を持っている部門同士を早期に結びつけることが可能となる。そして、まだ互いにコミュニケーションがなされていないうちから早期に部門間での情報交換、協力関係の構築等を促すことが可能となり、組織の生産性を高める効果がある。

【0049】

なお、本実施の形態では、部門ごとの検索単語頻度ベクトルを対象に、正規化及び類似度の算出を行う場合を例に説明したが、個人別検索単語頻度表から部門別検索単語頻度表への変換は行わずに、個人別検索単語頻度表から個人の検索単語頻度ベクトルを作成し、それをTF-IDF法により正規化して個人間の類似度を計算するようにしてもよい。この結果は、個人間の情報交換や協力関係の推進に利用することができる。あるいは個人の

10

20

30

40

50

検索単語頻度ベクトルと部門ごとの検索単語頻度ベクトルとを用いて個人と部門間の類似度を計算してもよい。この場合、個人と部門間での協力の可能性の発見が可能である。

【図面の簡単な説明】

【0050】

【図1】本発明の一実施の形態のネットワーク構成を示す図である。

【図2】本発明の検索分析サーバ装置のハードウェア構成を示すブロック図である。

【図3】本発明の検索分析サーバ装置の機能的な構成を示すブロック図である。

【図4】図1の検索分析サーバ装置における検索分析処理の流れを示したフローチャート図である。

【図5】ある個人の検索単語頻度表を示す図である。

10

【図6】検索単語頻度表の合成処理の説明図である。

【図7】類似度計算結果の表示例を示す図である。

【符号の説明】

【0051】

10 端末

20 検索分析サーバ装置

21 演算部

22 主記憶装置

23 補助記憶装置

25 表示装置

20

30 ネットワーク

31 検索ログ記録部

32 検索単語頻度表生成部

33 部門集計部

34 検索単語頻度表正規化部

35 部門間類似度計算部

36 結果出力部

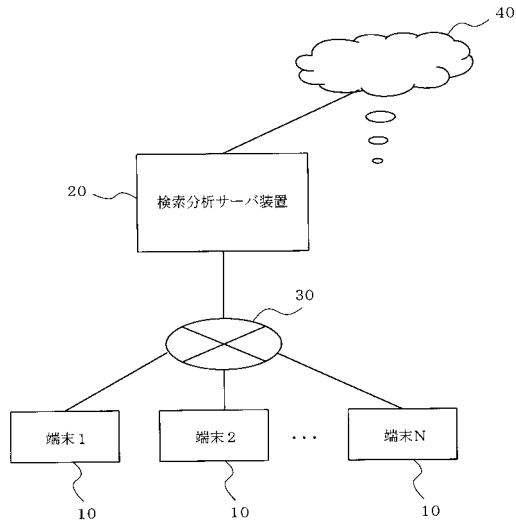
37 検索ログデータベース

38 検索単語頻度表データベース

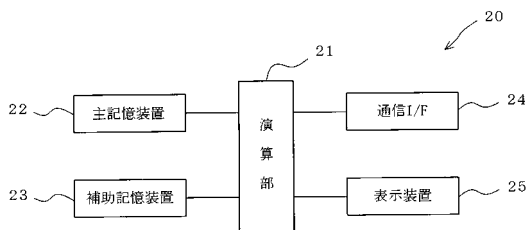
39 部門データベース

30

【図 1】



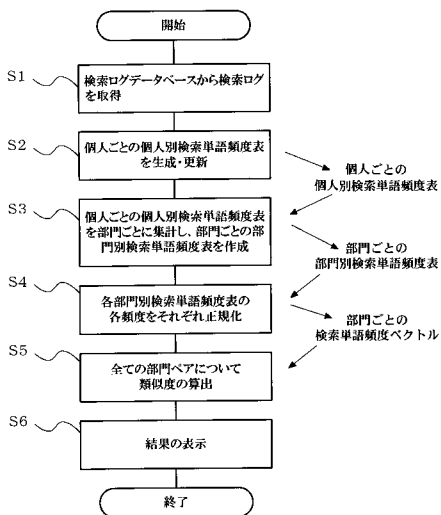
【図 2】



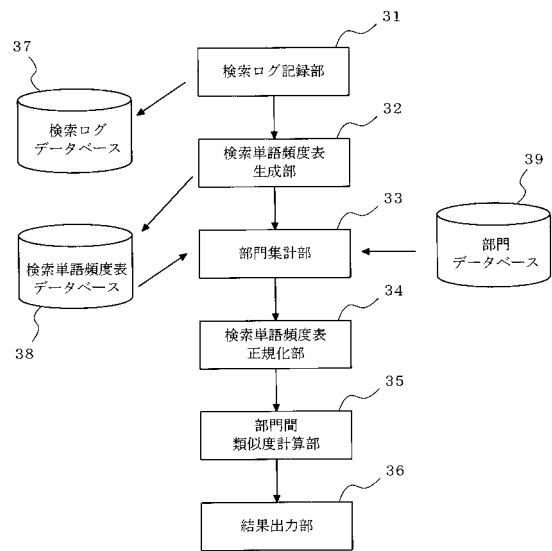
【图 4】

検索単語	頻度
テレワーク	5
臨場感	7
...	...
動向	81
とは	110

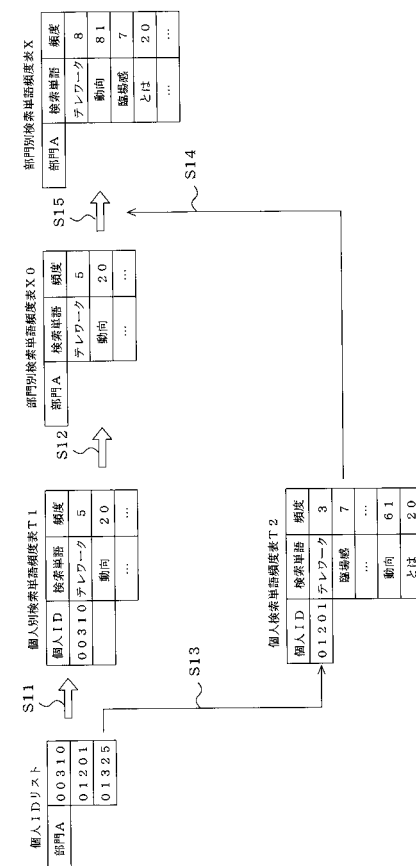
【図 5】



【 図 3 】



【図 6】



【図 7】

部門A		
に関連がある可能性のある部門リスト		
検索頻度が高かった検索用語 テレワーク、臨場感、情報共有		
順位	部門	類似度
1	X X X	0. 8
2	Y Y Y	0. 7 5
3	Z Z Z	0. 7
...	...	...

フロントページの続き

(72)発明者 首藤 和彦

東京都港区西新橋三丁目1番11号 沖電気工業株式会社内

Fターム(参考) 5B075 KK02 PQ02 PQ36 QM05