

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4810469号
(P4810469)

(45) 発行日 平成23年11月9日(2011.11.9)

(24) 登録日 平成23年8月26日(2011.8.26)

(51) Int.Cl. F 1
G 0 6 F 17/30 (2006.01)
 G 0 6 F 17/30 3 3 0 Z
 G 0 6 F 17/30 3 4 0 Z

請求項の数 11 (全 33 頁)

(21) 出願番号	特願2007-52843 (P2007-52843)	(73) 特許権者	000003078 株式会社東芝
(22) 出願日	平成19年3月2日(2007.3.2)		東京都港区芝浦一丁目1番1号
(65) 公開番号	特開2008-217333 (P2008-217333A)	(74) 代理人	100089118 弁理士 酒井 宏明
(43) 公開日	平成20年9月18日(2008.9.18)	(72) 発明者	布目 光生 東京都港区芝浦一丁目1番1号 株式会社東芝内
審査請求日	平成21年3月26日(2009.3.26)	(72) 発明者	石谷 康人 東京都港区芝浦一丁目1番1号 株式会社東芝内
		(72) 発明者	鈴木 優 東京都港区芝浦一丁目1番1号 株式会社東芝内

最終頁に続く

(54) 【発明の名称】 検索支援装置、プログラム及び検索支援システム

(57) 【特許請求の範囲】

【請求項1】

文書及び当該文書中の特定の文字列が指示されたキーワードを取得する文書取得手段と、

前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成手段と、

各形態素と、当該形態素の属性とを対応付けた語句属性テーブルを記憶する第1の記憶手段と、

前記語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別手段と、

前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成手段と、

各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルを記憶する第2の記憶手段と、

前記各分野を表すテーマと、当該テーマに関係する検索条件とを対応付けた検索条件テーブルを記憶する第3の記憶手段と、

前記文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定手段と、

前記推定された文書のテーマに関係する検索条件を、前記検索条件テーブルから選定する検索条件選定手段と、

を備えたことを特徴とする検索支援装置。

【請求項 2】

所定の文字列と、当該文字列が意味する事象とを対応付けた意味推定テーブルを記憶する第 4 の記憶手段と、

前記意味推定テーブルに基づいて、前記文書中に含まれた文字列が意味する事象を判別する意味判別手段と、

Web サイトのサイト種別と、当該サイト種別に関連する前記テーマ及び事象とを対応付けたサイト種別推定テーブルを記憶する第 5 の記憶手段と、

前記 Web サイトのサイト種別と、当該サイト種別に関する検索条件とを対応付けた第 2 の検索条件テーブルを記憶する第 6 の記憶手段と、

前記サイト種別推定テーブルに基づいて、前記文書のテーマ及び当該文書中の文字列が意味する事象から、当該文書を提供した Web サイトのサイト種別を推定するサイト種別推定手段と、

を更に備え、

前記検索条件選定手段は、前記推定されたサイト種別に関する検索条件を、前記第 2 の検索条件テーブルから選定することを特徴とする請求項 1 に記載の検索支援装置。

【請求項 3】

コンテンツの種別を示したコンテンツタイプと、当該コンテンツタイプに関連する前記テーマ及び事象とを対応付けたコンテンツタイプ推定テーブルを記憶する第 7 の記憶手段と、

前記コンテンツタイプと、当該コンテンツタイプに関する検索条件を対応付けた第 3 の検索方法テーブルを記憶する第 8 の記憶手段と、

前記コンテンツタイプ推定テーブルに基づいて、前記文書のテーマ及び当該文書中の文字列が意味する事象から、当該文書のコンテンツタイプを推定するコンテンツタイプ推定手段と、

を更に備え、

前記検索条件選定手段は、前記推定されたコンテンツタイプに関する検索条件を、前記第 3 の検索方法テーブルから選定することを特徴とする請求項 2 に記載の検索支援装置。

【請求項 4】

前記キーワードの前後に出現する前記文書中の形態素に基づいて、当該キーワードの文字列の範囲を形態素単位で拡張し、キーワード候補として取得するキーワード候補取得手段を更に備え、

前記キーワード情報生成手段は、前記キーワード及びキーワード候補に含まれた形態素と、当該形態素の属性とを対応付け、前記キーワード情報として生成することを特徴とする請求項 1 に記載の検索支援装置。

【請求項 5】

前記キーワード候補毎に、当該キーワード候補に関する文書の検索を外部の検索エンジンに依頼する検索依頼手段と、

前記検索エンジンの検索結果に基づき、前記キーワード候補から検索件数が所定値以上のキーワード候補を選定するキーワード候補選定手段と、

を更に備え、

前記キーワード情報生成手段は、前記選定されたキーワード候補に基づいて、前記キーワード情報を生成することを特徴とする請求項 4 に記載の検索支援装置。

【請求項 6】

前記検索条件選定手段により選定された検索条件を提示する提示手段を更に備えたことを特徴とする請求項 1 ~ 3 の何れか一項に記載の検索支援装置。

【請求項 7】

前記選定された検索条件に該当する文書の検索を外部の検索エンジンに依頼する検索依頼手段を更に備え、

10

20

30

40

50

前記提示手段は、前記検索エンジンの検索結果に基づき、検索件数に応じた順序で、前記選定された検索条件を提示することを特徴とする請求項 6 に記載の検索支援装置。

【請求項 8】

前記文書に含まれた各語句を形態素毎に切り出す語句切り出し手段と、
前記形態素毎の語句のうち、特定の語句を指定可能な語句指定手段と、
を更に備え、

前記文書取得手段は、前記語句指定手段により指定された語句を前記キーワードとして取得することを特徴とする請求項 1 に記載の検索支援装置。

【請求項 9】

前記キーワードの入力を受け付ける入力手段と、

前記入力手段で受け付けたキーワードに関する文書の検索を、外部の検索エンジンに依頼する検索依頼手段と、

を更に備え、

前記文書取得手段は、前記検索エンジンにより検索された文書を取得することを特徴とする請求項 1 に記載の検索支援装置。

【請求項 10】

コンピュータに、

文書及び当該文書中の特定の文字列が指示されたキーワードを取得する文書取得機能と、

前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成機能と、

各形態素と、当該形態素の属性とを対応付けた語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別機能と、

前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成機能と、

各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定機能と、

前記推定された文書のテーマに関係する検索条件を、前記各分野を表すテーマと、当該テーマに関係する検索条件とを対応付けた検索条件テーブルから選定する検索条件選定機能と、

を実現させることを特徴とするプログラム。

【請求項 11】

情報通信端末と、検索支援装置とが通信可能に接続された検索支援システムであって、

前記情報通信端末は、

外部装置が提供する文書を取得するコンテンツ閲覧手段と、

前記文書中の特定の文字列をキーワードとして指示可能なキーワード指定手段と、

前記文書及びキーワードを前記検索支援装置に送信する閲覧リソース送信手段と、

表示手段と、

を備え、

前記検索支援装置は、

前記文書及びキーワードを取得する文書取得手段と、

前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成手段と、

各形態素と、当該形態素の属性とを対応付けた語句属性テーブルを記憶する第1の記憶手段と、

前記語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別手段と、

前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成手段と、

10

20

30

40

50

各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルを記憶する第2の記憶手段と、

前記各分野を表すテーマと、当該テーマに係る検索条件とを対応付けた検索条件テーブルを記憶する第3の記憶手段と、

前記文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定手段と、

前記推定された文書のテーマに係る検索条件を、前記検索条件テーブルから選定する検索条件選定手段と、

を備えたことを特徴とする検索支援システム。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、Webサイトの検索を支援する検索支援装置、プログラム及び検索支援システムに関する。

【背景技術】

【0002】

近年、携帯電話の著しい普及と技術革新に伴う性能向上によって、その用途は音声通話に留まらず、電子メールの送受信や携帯電話向けサイトの閲覧に利用されることが一般的となっている。また、機種によっては、Webブラウザの搭載により、一般のWebコンテンツの閲覧も可能となっており、携帯電話はPC(Personal Computer)やPDA(Personal Digital Assistance)に類する高機能な情報端末としての役割を担ってきている。

20

【0003】

上記のように携帯電話が情報端末としての役割を拡大し、PCと同様に膨大なコンテンツにアクセスが可能となるに伴い、携帯電話のような限られた入力インターフェースと画面の上でも、膨大な情報の中から所望する情報を簡単に検索でき、目的の情報へ容易にたどり着くことが要求されている。

【0004】

こうした背景の下で、従来、携帯電話を用いて検索を行い、目的のサイトや情報にたどり着く手段には、大きく分けて次の二つの手法がある。一つは、コンテンツサービスで用意されている階層メニューをたどることにより目的のサイトを探すディレクトリ型の探索方法であり、もう一つは、検索エンジンに対してキーワードを入力することで、検索結果の要約リストを獲得し、それを参照して目的の情報にたどり着くキーワード検索型の探索方法である。

30

【0005】

また、従来、検索方法に関し、種々の技術が提案されている。例えば、特許文献1には、検索対象の利用目的を記述したフレーズを用いて検索対象実態へのポイント情報を検索することで、ユーザの希望する情報の検索の効率化を図った技術が開示されている。また、特許文献2には、指定されたハイパーテキストファイルからキーワードを抽出し、抽出された前記キーワードを用いてコンテンツの登録場所の検索を行うことで、類似したサイトを容易に見つけることが可能な技術が開示されている。また、特許文献3には、検索入力文と予め用意された事例文とをマッチングし、マッチングした事例文に関連付けられた検索エンジンに、検索入力文から抽出したキーワードの検索を依頼し、検索結果をユーザに提示することで、検索エンジン選択の負担を軽減した技術が開示されている。

40

【0006】

【特許文献1】特許第3759700号明細書

【特許文献2】特開2003-208434号公報

【特許文献3】特開2003-132060号公報

【発明の開示】

【発明が解決しようとする課題】

50

【0007】

しかしながら、上記したディレクトリ型の検索方法では、事前に定義されている分類（階層カテゴリ）の粒度が細かすぎると、目的のサイトに到達するまでの操作量が増大し、操作が煩雑になるという問題がある。また、階層カテゴリの粒度が粗い場合には、特定のカテゴリに属するサイトの数が多くなるため、上記同様目的のサイトを探索するまでの操作量が増大し、操作が煩雑になるという問題がある。

【0008】

一方、キーワード検索型の検索方法では、検索の結果、大量の結果がリスト表示された場合、携帯電話の限られた表示領域では、多くの画面スクロール操作等が必要となるため、操作性が悪く、また、検索結果を絞り込むために検索用語を追加するには、携帯電話の限られた入力手段で多数の操作が必要となるため、操作が煩雑になるという問題がある。

10

【0009】

また、特許文献1に記載の技術では、検索精度を上げるため、多種多様なユーザの嗜好を表したフレーズを、網羅的に登録しておく必要があるため、現実的でなく、実用性に欠けるという問題がある。さらに、検索条件となるフレーズを作成する際の操作が煩雑となるという問題がある。また、特許文献2に記載の技術では、ユーザからの操作は特段必要とされていないが、指定されたハイパーテキストファイルのサイトに類似したサイトしか検索することができないため、ユーザが所望するサイトを提示できない可能性がある。また、特許文献3に記載の技術では、予め用意された事例文に基づいて、特定の分野に一義的に分別してしまうため、自由度が低く、ユーザの所望する分野に分別されない可能性がある。

20

【0010】

本発明は、上記に鑑みてなされたものであって、検索する際の操作にかかる負荷を減少させるとともに、ユーザが所望する情報に容易に到達させることが可能な検索支援装置、プログラム及び検索支援システムを提供することである。

【課題を解決するための手段】

【0011】

上述した課題を解決し、目的を達成するために、本発明は、文書及び当該文書中の特定の文字列が指示されたキーワードを取得する文書取得手段と、前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成手段と、各形態素と、当該形態素の属性とを対応付けた語句属性テーブルを記憶する第1の記憶手段と、前記語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別手段と、前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成手段と、各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルを記憶する第2の記憶手段と、前記各分野を表すテーマと、当該テーマに係る検索条件とを対応付けた検索条件テーブルを記憶する第3の記憶手段と、前記文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定手段と、前記推定された文書のテーマに係る検索条件を、前記検索条件テーブルから選定する検索条件選定手段と、を備える。

30

40

【0012】

また、本発明は、コンピュータに、文書及び当該文書中の特定の文字列が指示されたキーワードを取得する文書取得機能と、前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成機能と、各形態素と、当該形態素の属性とを対応付けた語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別機能と、前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成機能と、各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定機能と、前記推定された文書のテーマに関

50

係する検索条件を、前記各分野を表すテーマと、当該テーマに関係する検索条件とを対応付けた検索条件テーブルから選定する検索条件選定機能と、を実現させる。

【0013】

また、本発明は、情報通信端末と、検索支援装置とが通信可能に接続された検索支援システムであって、前記情報通信端末は、外部装置が提供する文書を取得するコンテンツ閲覧手段と、前記文書中の特定の文字列をキーワードとして指示可能なキーワード指定手段と、前記文書及びキーワードを前記検索支援装置に送信する閲覧リソース送信手段と、表示手段と、を備え、前記検索支援装置は、前記文書及びキーワードを取得する文書取得手段と、前記文書に含まれた形態素のうち、特定の属性を有した形態素を抽出し、当該文書中の形態素の出現状況に関する統計情報を生成する統計情報生成手段と、各形態素と、当該形態素の属性とを対応付けた語句属性テーブルを記憶する第1の記憶手段と、前記語句属性テーブルに基づいて、前記文書中に含まれた各形態素の属性を判別する属性判別手段と、前記キーワードに含まれた形態素毎に当該形態素の属性を対応付けたキーワード情報を生成するキーワード情報生成手段と、各分野を表すテーマと、当該テーマに関連する用語とを対応付けた文書テーマ推定テーブルを記憶する第2の記憶手段と、前記各分野を表すテーマと、当該テーマに関係する検索条件とを対応付けた検索条件テーブルを記憶する第3の記憶手段と、前記文書テーマ推定テーブルに基づいて、前記統計情報及びキーワード情報に含まれる形態素から、前記文書のテーマを推定する文書テーマ推定手段と、前記推定された文書のテーマに関係する検索条件を、前記検索条件テーブルから選定する検索条件選定手段と、を備える。

10

20

【発明の効果】

【0014】

本発明によれば、文書及び当該文書に含まれた特定の文字列から、当該文書のテーマを推定し、このテーマに基づいて当該文書の特性に応じた検索条件を選定することで、ユーザは、関心のある文書中から特定の文字列を指示するのみで、当該文書のテーマに関連した検索条件を享受することができるため、検索の操作にかかる負荷を減少させるとともに、ユーザが所望する情報に容易に到達させることができる。

【発明を実施するための最良の形態】

【0015】

以下に添付図面を参照して、検索支援装置、プログラム及び検索支援システムの最良な実施形態を詳細に説明する。

30

【0016】

[第1の実施形態]

図1は、本実施形態の検索支援システム1の構成を示した概略図である。図1に示すように、検索支援システム1は、情報通信端末10と、検索支援サーバ(検索支援装置)20と、検索エンジン30、Webサーバ40とを有している。各機器はインターネット等のネットワークを介して相互に通信可能に接続されている。

【0017】

まず、情報通信端末10について説明する。情報通信端末10は、携帯電話等の情報通信端末装置であって、後述する通信部16を介してネットワークに接続することで、検索支援サーバ20、検索エンジン30及びWebサーバ40との間で情報の授受を行う。

40

【0018】

図2は、情報通信端末10のハードウェア構成を示した図である。図2に示すとおり、情報通信端末10は、CPU(Central Processing Unit)11、操作部12、表示部13、ROM(Read Only Memory)14、RAM(Random Access Memory)15、通信部16等を備える。

【0019】

CPU11は、RAM15の所定領域を作業領域として、ROM14に予め記憶された各種制御プログラムとの協働により各種処理を実行し、情報通信端末10を構成する各部の動作を統括的に制御する。

50

【 0 0 2 0 】

また、CPU 11は、ROM 14に予め記憶された所定のプログラムとの協働により、図1に示したコンテンツ閲覧部101、キーワード指定部102、閲覧リソース送信部103、検索方法提示選択部104、検索クエリ送信部105及び検索結果表示部106の各機能部を実現させる。なお、各機能部の動作については後述する。

【 0 0 2 1 】

操作部12は、各種入力キー等を備え、ユーザから操作入力された情報を入力信号として受け付け、その入力信号をCPU 11に出力する。

【 0 0 2 2 】

表示部13は、LCD(Liquid Crystal Display)等により構成され、CPU 11からの表示信号に基づいて、各種情報を表示する。なお、表示部13は、操作部12と一体的にタッチパネルを構成する態様としてもよい。

10

【 0 0 2 3 】

ROM 14は、情報通信端末10の制御にかかるプログラムや各種設定情報等を書き換え不可能に記憶する。

【 0 0 2 4 】

RAM 15は、SDRAM(Synchronous DRAM)等の記憶手段であって、各種データを書き換え可能に記憶する性質を有していることから、CPU 11の作業エリアとして機能し、バッファ等の役割を果たす。

【 0 0 2 5 】

通信部16は、ネットワークを通じ外部の機器(検索支援サーバ20、検索エンジン30、Webサーバ40)との間で通信を行うインターフェースであって、外部機器から送信された各種情報をCPU 11に出力し、また、CPU 11から出力される各種情報を外部機器に送信する。

20

【 0 0 2 6 】

次に、検索支援サーバ20について説明する。検索支援サーバ20は、情報通信端末10から送信された後述する閲覧リソースに基づき、当該閲覧リソースに応じた検索方法の候補を情報通信端末10に提供する検索支援装置である。図3は、検索支援サーバ20のハードウェア構成を示した図である。図3に示すとおり、情報通信端末10は、CPU 21、操作部22、表示部23、ROM 24、RAM 25、記憶部26、通信部27等を備える。

30

【 0 0 2 7 】

CPU 21は、RAM 25の所定領域を作業領域として、ROM 24又は記憶部26に予め記憶された各種制御プログラムとの協働により各種処理を実行し、検索支援サーバ20を構成する各部の動作を統括的に制御する。

【 0 0 2 8 】

また、CPU 21は、ROM 24又は記憶部26に予め記憶された所定のプログラムとの協働により、図1に示した入力文書リソース取得部201、キーワード候補選定部202、文書語句属性抽出部203、キーワード意味推定部204、入力文書テーマ推定部205、サイト種別推定部206、コンテンツタイプ推定部207、検索方法候補選定部208、検索方法候補検証部209及び検索方法候補メニュー生成部210の各機能部を実現させる。なお、各機能部の動作については後述する。

40

【 0 0 2 9 】

操作部22は、各種入力キー等を備え、ユーザから操作入力された情報を入力信号として受け付け、その入力信号をCPU 21に出力する。

【 0 0 3 0 】

表示部23は、LCD等により構成され、CPU 21からの表示信号に基づいて、各種情報を表示する。なお、表示部23は、操作部22と一体的にタッチパネルを構成する態様としてもよい。

【 0 0 3 1 】

50

R O M 2 4 は、検索支援サーバ 2 0 の制御にかかるプログラムや各種設定情報等を書き換え不可能に記憶する。

【 0 0 3 2 】

R A M 2 5 は、S D R A M 等の記憶手段であって、各種データを書換え可能に記憶する性質を有していることから、C P U 2 1 の作業エリアとして機能し、バッファ等の役割を果たす。

【 0 0 3 3 】

記憶部 2 6 は、等の記憶手段であって、磁氣的又は光学的に記録可能な記録媒体を有し、検索支援サーバ 2 0 の制御にかかるプログラムや各種設定情報等を記憶する。

【 0 0 3 4 】

また、記憶部 2 6 は、図 1 に示した形態素解析辞書 2 6 1、キーワード意味推定ルール 2 6 2、語句属性データベース 2 6 3、文書テーマ推定ルール 2 6 4、サイト種別推定ルール 2 6 5、コンテンツタイプ推定ルール 2 6 6、検索方法モデル 2 6 7 を所定の領域に記憶する。これら記憶部 2 6 に記憶された各情報については後述する。なお、本実施形態では、各情報を記憶部 2 6 に記憶する態様としたが、これに限らず、複数の記憶装置に分散して記憶する態様としてもよいし、検索支援サーバ 2 0 外の外部記憶装置に記憶する態様としてもよい。

【 0 0 3 5 】

通信部 2 7 は、ネットワークを通じ外部の機器（情報通信端末 1 0、検索エンジン 3 0、W e b サーバ 4 0）との間で通信を行うインターフェースであって、外部機器から送信された各種情報を C P U 2 1 に出力し、また、C P U 2 1 から出力される各種情報を外部機器に送信する。

【 0 0 3 6 】

検索エンジン 3 0 は、図示しないネットワーク上に存在するコンテンツを検索する機能を、情報通信端末 1 0 や検索支援サーバ 2 0 等に提供するサーバ装置又はシステムである。検索エンジン 3 0 は、情報通信端末 1 0 や検索支援サーバ 2 0 等から後述する検索条件（一又は複数の語句を“ O R ”、“ A N D ”等の論理条件で組み合わせた文字列）を受信すると、この検索条件に対応する W e b サーバ 4 0（W e b サイト）を検索し、この検索条件に合致した W e b サイトの総数（検索件数）や、検索条件に合致した W e b サイトの一覧、各 W e b サイトから提供されるコンテンツの要約等を返信する。

【 0 0 3 7 】

W e b サーバ 4 0 は、H T M L 文書や画像等のコンテンツを H D D（Hard Disk Drive）等の記憶装置に格納し、これらコンテンツを情報通信端末 1 0 等が閲覧（参照）可能に提供する W e b サイトである。なお、ネットワークに接続される W e b サーバ 4 0 の数量は、特に問わないものとする。

【 0 0 3 8 】

以下、本実施形態の検索支援システム 1 で行われる動作を説明する。図 4 は、情報通信端末 1 0 と検索支援サーバ 2 0 との間で行われる連動型検索処理の動作を示したラダーチャートである。なお、ステップ S 1 1 ~ S 1 8 は、情報通信端末 1 0 で行われる処理を示しており、ステップ S 2 1 ~ S 3 1 は、検索支援サーバ 2 0 で行われる処理を示している。

【 0 0 3 9 】

まず、情報通信端末 1 0 では、W e b サーバ 4 0 の宛先を示した U R L 等のアドレス情報が操作部 1 2 等を介して入力されると、コンテンツ閲覧部 1 0 1 は、このアドレス情報に対応した W e b サーバ 4 0 からコンテンツを取得し、表示部 1 3 に表示させる（ステップ S 1 1）。ステップ S 1 1 で、表示部 1 3 に表示されたコンテンツの一例を図 5 - 1 に示す。

【 0 0 4 0 】

次いで、本連動型検索処理を実行する旨の指示信号が操作部 1 2 等を介して入力されると、キーワード指定部 1 0 2 は、表示部 1 3 に表示されたコンテンツ中の特定の語句（文

10

20

30

40

50

字列)をキーワードとして選択させるための入力支援手段を提供し(ステップS 1 2)、当該キーワードが入力されるまで待機する(ステップS 1 3; No)。

【0041】

図5 - 2から図5 - 5は、ステップS 1 2からステップS 1 3にかけて表示部1 3に表示される画面の一例を示した図である。まず、図5 - 1に示した「ブラウザメニュー」から、図5 - 2に示すように、連動型検索処理の実行を指示する「連動型検索」が選択され、この指示信号がCPU 2 1に入力されると、キーワード指定部1 0 2は、ユーザに対して、現在閲覧中のコンテンツに含まれる文字列から、任意の文字列を選択可能なインターフェース(入力支援手段)を提供する。ここで、ユーザが選択可能なキーワードとしては、特に制限はなく、例えば、語句中の部分的な文字列等であってもよい。

10

【0042】

図5 - 3から図5 - 4では、上記インターフェースとして、選択したいキーワードを、その始点及び終点をカーソルにより指定した場合の態様を示しており、このインターフェースにより、「田中」という文字列がキーワードとして選択されている(図5 - 5参照)。なお、キーワード指定部1 0 2が提供する入力支援手段は、この態様に限らないものとする。

【0043】

図4に戻り、表示部1 3に表示されたコンテンツに基づき、ユーザから操作部1 2を介して特定のキーワード(文字列)が入力されると(ステップS 1 3; Yes)、閲覧リソース送信部1 0 3は、ステップS 1 3で入力されたキーワードと、現在閲覧中のコンテンツのURL又はコンテンツのHTML文書を、閲覧リソースとして検索支援サーバ2 0に送信する(ステップS 1 4)。

20

【0044】

一方、検索支援サーバ2 0では、情報通信端末1 0から送信された閲覧リソースを受信すると、入力文書リソース取得部2 0 1は、当該閲覧リソースに含まれたキーワードとHTML文書とを取得し(ステップS 2 1)、後段のキーワード候補選定部2 0 2に出力する。なお、閲覧リソースにHTML文書が含まれておらず、代わりにコンテンツのURLが含まれていた場合には、入力文書リソース取得部2 0 1は、ステップS 2 1において、このURLに対応するWebサーバ4 0からコンテンツを取得し、このコンテンツからHTML文書を取得する。

30

【0045】

続いてキーワード候補選定部2 0 2は、入力文書リソース取得部2 0 1からキーワード及びHTML文書を受け取ると、ステップS 2 2のキーワード候補選定処理を実行する。以下、図6を参照して、ステップS 2 2のキーワード候補選定処理について説明する。

【0046】

図6は、キーワード候補選定部2 0 2により実行されるキーワード候補選定処理の手順を示したフローチャートである。まず、キーワード候補選定部2 0 2は、入力されたHTML文書を入力文書とし、この入力文書に含まれた文字列を、形態素解析辞書2 6 1に基づいて形態素単位で分割する(ステップS 2 2 1)。

【0047】

ここで、形態素解析辞書2 6 1は、種々の形態素と、当該形態素の属性とが予め対応付けられたテーブルであって、形態素解析辞書2 6 1に含まれた形態素と、入力文書に含まれた文字列とを比較し、一致した文字列毎に分割することで、形態素単位で分割が行われる。なお、ここで形態素の属性とは、当該形態素の品詞や活用型、付加情報(例えば、国名、人名等)、読み、発音等であって、形態素毎に予め定義されている。

40

【0048】

次に、キーワード候補選定部2 0 2は、入力されたキーワードが、入力文書中の何文字目から何文字目に出現するかを特定し(ステップS 2 2 2)、この出現位置の始点をi、終点をjとした後、これら位置情報をRAM 2 5等に保持する(ステップS 2 2 3)。

【0049】

50

続いて、キーワード候補選定部202は、キーワードの始点*i*から、当該キーワードに先行する形態素に対応する語の文字列長分だけ減算、又は、キーワードの終点*j*から、当該キーワードに後続する形態素に対応する語の文字列長分だけ加算することで、キーワードの選択範囲を形態素単位で拡張し(ステップS224)、選択範囲が拡張されたキーワードをキーワード候補としてRAM25等に保持する(ステップS225)。

【0050】

例えば、図5-5を例にとると、キーワード「田中」を基準とし、上記選択範囲の拡張処理(ステップS224)を文書の前方向(キーワードに先行する方向)へ1回適用した場合には、キーワード候補は「MF田中」となる。また、同様にキーワード「田中」を基準として、上記選択範囲の拡張処理(ステップS224)を文書の後方向(キーワードに後続する方向)に1回適用した場合には、キーワード候補は「田中寿英」となる。なお、本実施形態では、キーワードに先行又は後続する形態素について行うこととしたが、キーワードに先行及び後続する形態素について行うこととしてもよい。

10

【0051】

次いで、キーワード候補選定部202は、キーワード候補となった文字列の長さが、キーワードが存在する一文の長さと同様か、又は、選択範囲の拡張を予め定められた回数分実行したか否かを判定し、何れにも該当しない場合には(ステップS226; No)、ステップS224に再び戻り、先の選択範囲の拡張処理で求められたキーワード候補を基準に、キーワードの選択範囲をさらに拡張する。

【0052】

一方、ステップS226において、キーワード候補となった文字列の長さが、キーワードが存在する一文の長さと同様であると判定した場合、又は、選択範囲の拡張を予め定められた回数分実行したと判定した場合には(ステップS226; Yes)、キーワード候補選定部202は、これまでにキーワード候補として得られた各文字列を、検索エンジン30に夫々送信し(ステップS227)、各文字列に応じた検索結果を、検索エンジン30から夫々取得する(ステップS228)。

20

【0053】

そして、キーワード候補選定部202は、取得した検索結果のうち、検案件数が多く得られた上位のキーワード候補を、入力されたキーワードとともに、キーワード拡張結果として、文書語句属性抽出部203に出力し(ステップS229)、図4のステップS23へと移行する。なお、文書語句属性抽出部203に出力するキーワード候補の個数は特に問わず、例えば、最上位のものをキーワード候補としてもよいし、上位3つをキーワード候補としてもよい。また、ステップS227、228を行わず、ステップS229では、ステップS226までに得られたキーワード候補とキーワードとをキーワード拡張結果として文書語句属性抽出部203に出力する態様としてもよい。

30

【0054】

図4に戻り、文書語句属性抽出部203は、キーワード候補選定部202からキーワード拡張結果を受け取ると、ステップS23の語句属性解析処理を実行する。以下、図7、図8を参照して、ステップS23の語句属性解析処理について説明する。

【0055】

図7は、文書語句属性抽出部203により実行される語句属性解析処理の手順を示したフローチャートである。まず、文書語句属性抽出部203は、形態素解析辞書261に基づいて、入力文書に対し形態素解析を適用することで、文書を形態素毎に分割し(ステップS231)、各形態素の属性を判別することで、各形態素とその属性とを対応付けたテーブルを生成し、形態素解析結果としてRAM25等に保持する(ステップS232)。

40

【0056】

次いで、文書語句属性抽出部203は、形態素解析結果から、入力文書全体に含まれた属性が名詞の形態素を抽出し、その出現頻度を対応付けた統計情報を生成する(ステップS233)。また、文書語句属性抽出部203は、形態素解析辞書261に基づいて、キーワード拡張結果に対し形態素解析を適用することで、キーワード及び当該キーワード候

50

補に含まれた文字列を形態素毎に分割し、各形態素の属性を判別することで、各形態素とその属性とを対応付けたキーワード情報を生成する（ステップS 2 3 4）。そして、文書語句属性抽出部2 0 3は、統計情報とキーワード情報とを語句属性データベース2 6 3として、記憶部2 6に保持する（ステップS 2 3 5）。

【0 0 5 7】

なお、本実施形態では、ステップS 2 3 1において、入力文書の形態素解析を行うこととしたが、上述したキーワード候補選定処理（図6参照）のステップS 2 2 1で、ステップS 2 3 1と同様の形態素解析を行った場合には、ステップS 2 2 1の形態素解析結果を用いる態様としてもよい。

【0 0 5 8】

図8は、語句属性データベース2 6 3を説明するための図である。なお、入力文書として、A 1 1に示した文字列が入力されたものとする。

【0 0 5 9】

ステップS 2 3 1において、文書語句属性抽出部2 0 3は、入力文書に対し形態素解析を適用することで、文書を形態素毎に分割し、各形態素に属性を対応付けた形態素解析結果を生成する。A 1 2には、A 1 1の形態素解析結果を示しており、A 1 1の文字列に含まれた形態素（ワールドカップ等）毎に、当該形態素の属性（名詞 - 固有名詞 - 一般等）が対応付けられている。

【0 0 6 0】

A 1 3は、A 1 2の形態素解析結果に基づいて導出された統計情報の一例を示している。このA 1 3の例では、まず入力文書全体の形態素解析結果から固有名詞（名詞 - 固有名詞）が抽出された例が示されている。

【0 0 6 1】

具体的に、文書語句属性抽出部2 0 3は、形態素解析結果（A 1 2参照）から、属性が一般固有名詞（名詞 - 固有名詞 - 一般）に該当する形態素として、まず「ワールドカップ」を選択し、この品詞の入力文書中における出現位置を入力文書から特定する。そして、文書語句属性抽出部2 0 3は、品詞種別と、当該品詞種別に該当する品詞と、出現位置とを対応付ける。なお、A 1 3では、入力文書の1文字目から7文字目（1、7）に「ワールドカップ」が出現することを意味している。

【0 0 6 2】

また、文書語句属性抽出部2 0 3は、上記と同様形態素解析結果から、属性が固有名詞に形態素「W杯（名詞 - 固有名詞 - 一般）」、「ドイツ（名詞 - 固有名詞 - 国名）」、「田中寿英（名詞 - 固有名詞 - 人名 - 性、名詞 - 固有名詞 - 人名 - 名）」を選択し、各形態素について品詞種別と、出現位置とを対応付ける。そして、文書語句属性抽出部2 0 3は、属性の種別（一般固有名詞、国名、人名等）毎に選定した形態素の個数を算出し、属性の種別毎に対応する形態素と出現位置とを関係付け、統計情報として生成する（A 1 3参照）。なお、A 1 3において、属性の種別に付加された数値（例えば“2”）は、その属性に属する形態素の個数を示している。

【0 0 6 3】

A 1 4は、キーワード情報の一例を示している。このA 1 4の例では、キーワードとしてA 1 1に含まれた「田中」がキーワードとして指定された場合の例を示しており、また、キーワード候補として「たMF田中寿英選手」等が入力された場合の例を示している。なお、A 1 4において、基準となるキーワード（基準キーワード）「田中」には、その旨を示す出現位置情報「0（該当文字）」が付加されている。

【0 0 6 4】

具体的に、文書語句属性抽出部2 0 3は、形態素解析結果（A 1 2参照）に基づいて、キーワード及びキーワード候補に含まれた文字列を、A 1 4に示したように、「た」、「MF」、「田中」、「寿英」、「選手」の形態素毎に分割する。そして、文書語句属性抽出部2 0 3は、分割した形態素毎に、その属性を対応付けたキーワード情報を生成する。

【0 0 6 5】

10

20

30

40

50

なお、キーワード情報に記録される各形態素には、基準キーワードに対する出現位置に応じた出現位置情報が付与されるものとする。本実施形態では、A 1 4 に示したように、基準キーワードの前方に出現する品詞については、その形態素分の個数を 0 から減算した出現位置情報（- 1（1 品詞前）等）が付与され、また、基準キーワードの後方に出現する形態素については、その形態素分の個数を 0 に加算した出現位置情報（+ 1（1 品詞後）等）が付与されるものとする。

【 0 0 6 6 】

図 4 に戻り、続くステップ S 2 4 では、キーワード意味推定部 2 0 4 により、キーワード意味推定処理が実行される。以下、図 9、図 1 0 を参照して、ステップ S 2 4 のキーワード意味推定処理について説明する。

10

【 0 0 6 7 】

図 9 は、キーワード意味推定部 2 0 4 により実行されるキーワード意味推定処理の手順を示したフローチャートである。まず、キーワード意味推定部 2 0 4 は、文書語句属性抽出部 2 0 3 で取得された形態素解析結果を、R A M 2 5 等から読み出す（ステップ S 2 4 1）。

【 0 0 6 8 】

次いで、キーワード意味推定部 2 0 4 は、形態素解析結果に基づいて、入力文書に含まれた各形態素から、国名や人名等の付加情報を有する形態素を特定する（ステップ S 2 4 2）。

【 0 0 6 9 】

続いてキーワード意味推定部 2 0 4 は、記憶部 2 6 に記憶されたキーワード意味推定ルール 2 6 2 を読み出し（ステップ S 2 4 3）、このキーワード意味推定ルールに基づいて、入力文書に含まれた文字列（属性が名詞又は記号の形態素等）が意味する事象を判別する（ステップ S 2 4 4）。

20

【 0 0 7 0 】

ここで、キーワード意味推定ルール 2 6 2 とは、特定の文字列と、当該文字列意味する事象とを予め対応付けたテーブルであって、キーワード意味推定部 2 0 4 は、このキーワード意味推定ルール 2 6 2 に定義された文字列と、入力文書中の文字列とを比較することで、一致した文字列が意味する事象を判別する。なお、ここで「意味する事象」とは、所定の文字列から想起可能な事柄を意味し、例えば、文字列「ワールドカップ」、「W 杯」

30

【 0 0 7 1 】

次いで、キーワード意味推定部 2 0 4 は、入力文書に含まれた文字列（形態素）のうち、ステップ S 2 4 2 で特定した形態素及びステップ S 2 4 3 で判別した文字列について、対応する付加情報及び事象（以下、意味情報という）を対応付け、キーワード意味推定結果として R M A 2 5 等に保持した後（ステップ S 2 4 5）、図 4 のステップ S 2 5 に移行する。

【 0 0 7 2 】

図 1 0 は、キーワード意味推定処理を説明するための図である。図 1 0 において、上段の A 2 1 は、入力文書に含まれた各形態素から、国名や人名等の付加情報を有する形態素を特定した状態を模式的に示している。この例では、形態素「ドイツ」が国名（C O U N T R Y）の付加情報を有しており、同様に形態素「田中」が人名 - 姓（P E R S O N _ s e i）が、形態素「寿英」が人名 - 名（P E R S O N _ m e i）が、形態素「ポトルン」が都市名（C I T Y、T E A M）の付加情報を有している。

40

【 0 0 7 3 】

また、中段の A 2 2 は、入力文書にキーワード意味推定ルール 2 6 2 を適用し、特定の文字列が意味する事象を判別した結果を模式的に示している。なお、キーワード意味推定ルール 2 6 2 には、文字列「ワールドカップ」及び「W 杯」について、事象「イベント（E V E N T）」が対応付けられているものとする。また、属性が人名（P E R S O N）の文字列 + 文字列「選手」の組み合わせについて、事象「運動選手（A T H L E T E）」が

50

対応付けられているものとする。

【 0 0 7 4 】

また、キーワード意味推定ルール 2 6 2 には、属性が人名 (P E R S O N) の文字列 + 任意の文字列「 * (ワイルドカード) 」 + 用語「 () + 属性が数 (n u m e r a l) の文字列 + 文字列「 」の組み合わせについては、属性が数 (n u m e r a l) の文字列に対し、事象「年齢 (A G E) 」が対応付けられているものとする。さらに、属性が数 (n u m e r a l) の文字列 + 文字列「日」の組み合わせについては、事象「日付 (D A T E) 」が対応付けられているものとする。

【 0 0 7 5 】

下段の A 2 3 は、意味解析結果の一例を示している。この例では、A 2 3 に示したように、該当する文字列 (形態素) の前後に、タグ (「 < 」と「 > 」で囲まれた文字列) 形式で表した意味情報を埋め込むことで、対応する文字列に対応付けられている。なお、本実施形態では、意味情報をタブ形式で対応付ける態様としたが、これに限らず、他の方法を用いることとしてもよい。

10

【 0 0 7 6 】

図 4 に戻り、続くステップ S 2 5 では、入力文書テーマ推定部 2 0 5 により、文書テーマ推定処理が実行される。以下、図 1 1、図 1 2 を参照して、ステップ S 2 5 の文書テーマ推定処理について説明する。

【 0 0 7 7 】

図 1 1 は、入力文書テーマ推定部 2 0 5 により実行される文書テーマ推定処理の手順を示したフローチャートである。まず、入力文書テーマ推定部 2 0 5 は、記憶部 2 6 に保持された語句属性データベース 2 6 3 を読み出すとともに (ステップ S 2 5 1)、R A M 2 5 等に記憶された意味解析結果を読み出す (ステップ S 2 5 2) 。

20

【 0 0 7 8 】

次いで、入力文書テーマ推定部 2 0 5 は、記憶部 2 6 に記憶された文書テーマ推定ルール 2 6 4 を読み出し (ステップ S 2 5 3)、語句属性データベース 2 6 3 に含まれた語 (形態素) と、文書テーマ推定ルール 2 6 4 に定義された用語とを比較する (ステップ S 2 5 4) 。

【 0 0 7 9 】

ここで、図 1 2 の A 3 3 を参照して、文書テーマ推定ルール 2 6 4 について説明する。文書テーマ推定ルール 2 6 4 には、各分野を表すテーマと、当該テーマに関連する用語とが対応付けて定義されている。A 3 3 の例においては、文書テーマ推定ルール 2 6 4 は階層構造で表され、各分野のテーマを表す上位概念語と、各上位概念語に関連する用語として下位概念語とが関連付けて定義されている。A 3 3 に示すように、例えば、上位概念語「スポーツ」の下位概念語として、「ゴルフ、サッカー、野球」が関連付けられている。

30

【 0 0 8 0 】

また、各概念語には、当該概念語に関連する語が、語彙リストとして定義されている。例えば、上位概念語「スポーツ」に対しては、スポーツに関連する語彙リストとして「選手、競技、試合、・・・」が定義されている。また、「ゴルフ」に関連する語彙リストとしては「ホール、打差、パー」が定義されており、同様に、「サッカー」に対しては「得点、W杯、キーパー」が、「野球」に対しては「球場、ホームラン、・・・」が定義されている。ステップ S 2 5 4 において、入力文書テーマ推定部 2 0 5 は、この語彙リストに含まれた用語と適合する語が、語句属性データベース 2 6 3 に存在するか否かを判定する。

40

【 0 0 8 1 】

図 1 1 に戻り、ステップ S 2 5 4 の比較の結果、語句属性データベース 2 6 3 中に、文書テーマ推定ルール 2 6 4 に定義された語と適合する語がないと判定した場合には (ステップ S 2 5 5 ; N o)、図 4 のステップ S 2 6 に直ちに移行する。

【 0 0 8 2 】

また、ステップ S 2 5 4 の比較の結果、語句属性データベース 2 6 3 中に、文書テーマ

50

推定ルール 264 に定義された用語と適合する語があると判定した場合には (ステップ S255 ; Yes)、入力文書テーマ推定部 205 は、その適合した用語に係るテーマに対し、所定のスコアを付与する (ステップ S256)。ここでスコアとは、統計情報及びキーワード情報に含まれた特定の形態素に対して予め付与された数値であって、特定のテーマに特有の用語 (形態素) であるほど、高い値が付与されているものとする。

【0083】

続いて、入力文書テーマ推定部 205 は、適合した用語に係るテーマと、そのスコアとを対応付け、入力文書推定結果として RAM 25 等に格納した後 (ステップ S257)、図 4 のステップ S26 に直ちに移行する。

【0084】

図 12 は、文書テーマ推定処理を説明するための図である。図 12 において、A31 は、語句属性データベース 263 に含まれた統計情報を示しており、A32 は、キーワード情報を示している。また、A33 は、文書テーマ推定ルール 264 の一例を示している。

【0085】

以下、A31 ~ A34 を参照して、文書テーマ推定処理の一例を説明する。なお、統計情報に含まれた「W杯」に対しスコア“1”が、キーワード情報に含まれた「選手」に対しスコア“10”が付与されているものとする。

【0086】

まず、入力文書テーマ推定部 205 は、A31、A32 で示した語句属性データベース 263 に対して、A33 の文書テーマ推定ルール 264 を適用することで、キーワード情報の「+2 (2品詞語)」に定義された「選手」が、文書テーマ推定ルール 264 の上位概念語「スポーツ」に適合すると判断する。この場合、入力文書テーマ推定部 205 は、上位概念語「スポーツ」について、「選手」に割り振られたスコア“10”を加算する。

【0087】

次に、入力文書テーマ推定部 205 は、統計情報の「W杯」が、文書テーマ推定ルール 264 の下位概念語「スポーツ - サッカー」に適合すると判断する。この場合、入力文書テーマ推定部 205 は、下位概念語「スポーツ - サッカー」について「W杯」に割り振られたスコア“1”を加算する。

【0088】

このように、入力文書テーマ推定部 205 は、A34 に示したように、「スポーツ」がスコア“10”、「スポーツ - サッカー」がスコア“1”と最終的に導出し、これらを入力文書推定結果として RAM 25 等に格納する。

【0089】

図 4 に戻り、続くステップ S26 では、サイト種別推定部 206 により、サイト種別推定処理が実行される。以下、図 13、図 14 を参照して、ステップ S26 のサイト種別推定処理について説明する。

【0090】

図 13 は、サイト種別推定部 206 により実行されるサイト種別推定処理の手順を示したフローチャートである。まず、サイト種別推定部 206 は、RAM 25 に格納されたキーワード意味推定結果を読み出すとともに (ステップ S261)、入力文書推定結果を読み出す (ステップ S262)。

【0091】

次に、サイト種別推定部 206 は、記憶部 26 に予め記憶されたサイト種別推定ルール 265 を読み出し (ステップ S263)、このサイト種別推定ルール 265 をキーワード意味推定結果及び入力文書推定結果に対して適用することで、後述するサイト種別とサイト関連度とを取得する (ステップ S264)。そして、サイト種別推定部 206 は、取得したサイト種別とサイト関連度とを対応付け、サイト種別推定結果として RAM 25 等に格納した後 (ステップ S265)、図 4 のステップ S27 に移行する。

【0092】

図 14 は、サイト種別推定処理を説明するための図である。図 14 において、A41 は

10

20

30

40

50

、入力文書推定結果の一例を示しており、A 4 2 は、キーワード意味推定結果の一例を示している。また、A 4 3 は、サイト種別推定ルール 2 6 5 の一例を示しており、A 4 4 は、サイト種別推定結果の一例を示している。なお、A 4 2 のキーワード意味推定結果は、上述した図 1 0 の A 2 3 とは表現が異なるが、同様の意味を有するものである。

【 0 0 9 3 】

まず、図 1 4 の A 4 3 を参照して、サイト種別推定ルール 2 6 5 について説明する。サイト種別推定ルール 2 6 5 には、各分野を表すテーマと、意味情報と、サイト種別とが対応付けて定義されている。A 4 3 の例では、まずサイト種別として「ショッピングサイト」、「Blog」、「ニュースサイト」、「トラベルサイト」等の名称が定義されている。そして、これらのサイト種別に対して、各分野を表す「テーマ」及び意味情報に対応する「キーワード」が夫々定義されている。また、「テーマ」及び「キーワード」に含まれた各語については、夫々の語の重み付けとして所定の数値が定義されている。

10

【 0 0 9 4 】

例えば、サイト種別が「ショッピングサイト」であれば、関連するテーマとして「コンピュータ、書籍、チケット...」が定義されており、夫々の語の重み付けとして、“ 0 . 3 ”、“ 0 . 1 ”、“ 0 . 1 ”が定義されていることを示している。また、「ショッピングサイト」に関連するキーワードも同様に定義されている。ここでは、「MONEY」等が関連付けられており、その重み付けは“ 0 . 8 ”と定義されている。

【 0 0 9 5 】

同様に、サイト種別「Blog」では、関連するテーマとして「スポーツ、芸能、コンピュータ」等が夫々“ 0 . 1 ”の重み付けで定義されており、キーワードとして「PERSON」、「EVENT」等が夫々“ 0 . 3 ”の重み付けで定義されている。

20

【 0 0 9 6 】

また、サイト種別「ニュースサイト」では、関連するテーマとして「政治、経済、スポーツ、芸能...」等が夫々“ 0 . 8 ”の重み付けで定義されており、キーワードとして「PERSON」が“ 0 . 3 ”、「COUNTRY」が“ 0 . 1 ”、「ATHLETE」が“ 0 . 3 ”の重み付けで定義されている。

【 0 0 9 7 】

さらに、サイト種別「トラベルサイト」では、関連するテーマとして「旅行」が“ 0 . 8 ”、「乗り物」が“ 0 . 4 ”で定義されており、キーワードとして「COMPANY」、「CITY」が夫々“ 0 . 7 ”、「MONEY」が“ 0 . 2 ”の重み付けで定義されている。

30

【 0 0 9 8 】

サイト種別推定部 2 0 6 は、上記したサイト種別推定ルール 2 6 5 を、前段の処理までに得られた入力文書推定結果と、キーワード意味推定結果とに対して適用することで、まず入力文書推定結果に含まれた「スポーツ」が、サイト種別「Blog」並びに「ニュースサイト」のテーマ「スポーツ」に適合すると判断する。

【 0 0 9 9 】

ここで、「スポーツ」について割り当てられた重み付けは「Blog」で“ 0 . 1 ”、「ニュースサイト」で“ 0 . 8 ”と定義されているため、サイト種別推定部 2 0 6 は、これらの値を読み出し、夫々の重み付け値と入力文書推定結果の「スポーツ」に付与されたスコア“ 1 0 ”とを乗算する。具体的には、サイト種別「Blog」について、 $0 . 1 \times 1 0 = 1 . 0$ 、サイト種別「ニュースサイト」について、 $0 . 8 \times 1 0 = 8 . 0$ と算出する。

40

【 0 1 0 0 】

次に、サイト種別推定部 2 0 6 は、キーワード意味推定結果を参照し、当該キーワード意味推定結果に含まれた「PERSON」が、サイト種別「Blog」の関連キーワードに重み付け値“ 0 . 3 ”で定義された「PERSON」に適合すると判定するとともに、サイト種別「ニュースサイト」の関連キーワードに重み付け値“ 0 . 3 ”で定義された「PERSON」に適合すると判定する。また、サイト種別推定部 2 0 6 は、キーワード意

50

味推定結果に含まれた「ATHLETE」が、サイト種別「ニュースサイト」の関連キーワードに重み付け値“0.3”で定義された「ATHLETE」に適合すると判断する。

【0101】

サイト種別推定部206は、これらの重み付け値を、先に導出したサイト種別の重み付け値とスコアとの乗算結果に加算することで、各サイト種別との関連度（サイト関連度）を導出する。具体的には、サイト種別「Blog」とのサイト関連度が、 $1.0 + 0.3 = 1.3$ 、サイト種別「ニュースサイト」とのサイト関連度が、 $8.0 + 0.3 + 0.3 = 8.6$ と算出する。そして、サイト種別推定部206は、A44に示したように、サイト種別と関連度との組をサイト種別推定結果として取得する。ここで、サイト種別推定結果に含まれたサイト種別は、情報通信端末10のユーザが閲覧した可能性のあるWebサイトのサイト種別を推定したものとなっており、サイト関連度は、その推定の確かさ（可能性）を数値の大小により表している。

10

【0102】

図4に戻り、続くステップS27では、コンテンツタイプ推定部207により、コンテンツタイプ推定処理が実行される。以下、図15、図16を参照して、ステップS27のコンテンツタイプ推定処理について説明する。

【0103】

図15は、コンテンツタイプ推定部207により実行されるコンテンツタイプ推定処理の手順を示したフローチャートである。まず、コンテンツタイプ推定部207は、RAM25に格納されたキーワード意味推定結果を読み出すとともに（ステップS271）、入力文書推定結果を読み出す（ステップS272）。

20

【0104】

次に、コンテンツタイプ推定部207は、記憶部26に予め記憶されたコンテンツタイプ推定ルール266を読み出し（ステップS273）、このコンテンツタイプ推定ルール266をキーワード意味推定結果及び入力文書推定結果に対して適用することで、後述するコンテンツタイプとコンテンツ関連度とを取得する（ステップS274）。そして、コンテンツタイプ推定部207は、取得したコンテンツタイプとコンテンツ関連度とを対応付け、コンテンツタイプ推定結果としてRAM25等に格納した後（ステップS275）、図4のステップS28に移行する。

【0105】

図16は、コンテンツタイプ推定処理を説明するための図である。図16において、A51は、入力文書推定結果の一例を示しており、A52は、キーワード意味推定結果の一例を示している。また、A53は、コンテンツタイプ推定ルール266の一例を示しており、A54は、コンテンツタイプ推定結果の一例を示している。

30

【0106】

まず、図16のA53を参照して、コンテンツタイプ推定ルール266について説明する。コンテンツタイプ推定ルール266には、各分野を表すテーマと、意味情報と、コンテンツタイプとが対応付けて定義されている。A53の例では、まずコンテンツタイプとして「カタログ」、「レビュー」、「記事」、「プレスリリース」等の名称が定義されている。そして、これらのコンテンツタイプに対して、分野を表すテーマ「テーマ」及び意味情報に対応する「キーワード」が夫々定義されている。また、「テーマ」及び「キーワード」に含まれた各語については、夫々の語の重み付けとして所定の数値が定義されている。

40

【0107】

例えば、コンテンツタイプが「カタログ」であれば、関連するテーマとして「コンピュータ...」が定義されており、夫々の語の重み付けとして、“0.3”が定義されていることを示している。また、「カタログ」に関連するキーワードも同様に定義されている。ここでは、「COMPANY」が関連付けられており、その重み付けは“0.6”と定義されている。

【0108】

50

同様に、コンテンツタイプ「レビュー」では、関連するテーマとして「レストラン、旅行、書籍、コンピュータ」等が夫々“0.2”の重み付けで定義されており、キーワードとして「EVENT」、「CITY」等が夫々“0.2”、“0.1”の重み付けで定義されている。

【0109】

また、コンテンツタイプ「記事」では、関連するテーマとして「政治、経済、スポーツ、芸能…」等が夫々“0.8”の重み付けで定義されており、キーワードとして「PERSON」が“0.3”、「COUNTRY」が“0.1”、「ATHLETE」が“0.3”の重み付けで定義されている。

【0110】

さらに、コンテンツタイプ「プレスリリース」では、関連するテーマとして「コンピュータ、芸能」が夫々“0.3”で定義されており、キーワードとして「COMPANY」、「DATE」が夫々“0.7”、また「PERSON」が“0.1”の重み付けで定義されている。

【0111】

コンテンツタイプ推定部207は、上記したコンテンツタイプ推定ルール266を、前段の処理までに得られた入力文書推定結果と、キーワード意味推定結果とに対して適用することで、まず入力文書推定結果に含まれた「スポーツ」が、コンテンツタイプ「記事」に適合すると判定する。

【0112】

ここで、「スポーツ」について割り当てられた重み付けは“0.8”と定義されているため、コンテンツタイプ推定部207は、この値を読み出し、この重み付け値と入力文書推定結果の「スポーツ」に付与されたスコア“10”とを乗算する。具体的には、コンテンツタイプ「記事」について、 $0.8 \times 10 = 8.0$ と算出する。

【0113】

次に、コンテンツタイプ推定部207は、キーワード意味推定結果を参照し、当該キーワード意味推定結果に含まれた「PERSON」が、コンテンツタイプ「記事」の関連キーワードに重み付け値“0.3”で定義された「PERSON」に適合すると判定するとともに、コンテンツタイプ「プレスリリース」の関連キーワードに重み付け値“0.1”で定義された「PERSON」に適合すると判定する。

【0114】

コンテンツタイプ推定部207は、これらの重み付け値を、先に導出したコンテンツタイプの重み付け値とスコアとの乗算結果に加算することで、各コンテンツタイプとの関連度（コンテンツ関連度）を導出する。具体的には、コンテンツタイプ「記事」とのコンテンツ関連度が、 $8.0 + 0.3 + 0.3 = 8.6$ 、コンテンツタイプ「プレスリリース」とのコンテンツ関連度が、 $0.0 + 0.1 = 0.1$ と算出する。そして、コンテンツタイプ推定部207は、A54に示したように、コンテンツとコンテンツ関連度との組をコンテンツタイプ推定結果として取得する。ここで、コンテンツタイプ推定結果に含まれたコンテンツタイプは、情報通信端末10のユーザが閲覧した可能性のあるコンテンツの種類を推定したものとなっており、コンテンツ関連度は、その推定の確かさ（可能性）を数値の大小により表している。

【0115】

図4に戻り、続くステップS28では、検索方法候補選定部208により、検索方法候補選定処理が実行される。以下、図17～図21を参照して、ステップS28の検索方法候補選定処理について説明する。

【0116】

図17は、検索方法候補選定部208により実行される検索方法候補選定処理の手順を示したフローチャートである。まず、検索方法候補選定部208は、前段の処理までに取得された入力文書推定結果、サイト種別推定結果、コンテンツタイプ推定結果を、RAM25から読み出す（ステップS281）。

10

20

30

40

50

【 0 1 1 7 】

次いで、検索方法候補選定部 208 は、記憶部 26 に予め記憶された検索方法モデル 267 を読み出す (ステップ S282)。以下、図 18 を参照して、検索方法モデル 267 について説明する。

【 0 1 1 8 】

図 18 は、検索方法モデル 267 の例を示した図である。検索方法モデル 267 には、入力文書推定結果、サイト種別推定結果及びコンテンツタイプ推定結果の夫々に含まれる各項目として「テーマ」、「サイト種別」、「コンテンツタイプ」と、検索対象を表す「ラベル」とが対応付けて定義されている。

【 0 1 1 9 】

図 18 - 1 に示した検索方法モデル 267 には、「テーマ」に対応付けられた検索の「ラベル」が定義されている。例えば、「コンピュータ」であれば、「スペックを調べる、発売時期を調べる、評価を見る」といったラベルが定義されている。同様に「スポーツ」であれば、「試合結果を見る、プロフィールを見る」といったラベルが定義されており、「政治」であれば、「プロフィールを見る、意味を調べる、組織を調べる」といったラベルが定義されており、「旅行」であれば、「地図を見る、価値を調べる、ツアーを探す」といったラベルが定義されている。なお、「ID」は、各テーマ名を一意に識別するための ID 番号である。

【 0 1 2 0 】

また、図 18 - 2 に示した検索方法モデル 267 には、「サイト種別」に対応付けられた検索の「ラベル」が定義されている。例えば、「ショッピングサイト」であれば、「価格を調べる、評判を調べる、販売店を調べる」といったラベルが定義されている。同様に「ニュースサイト」であれば、「ニュースを見る、プロフィールを見る、写真を見る」といったラベルが定義されており、「Blog」であれば、「評価を見る、プロフィールを見る」といったラベルが定義されており、「トラベルサイト」であれば、「ツアーを探す、地図を見る、地名を調べる、感想を見る」といったラベルが定義されている。なお、「ID」は、各サイト種別名を一意に識別するための ID 番号である。

【 0 1 2 1 】

さらに、図 18 - 3 に示した検索方法モデル 267 には、「コンテンツタイプ」に対応付けられた検索の「ラベル」が定義されている。例えば、「カタログ」であれば、「用語の意味を調べる、価格を調べる」といったラベルが定義されている。同様に「レビュー」であれば、「スペックを調べる、評価を見る」といったラベルが定義されており、「記事」であれば、「ニュースを見る、プロフィールを見る、地名を調べる」といったラベルが定義されており、「プレスリリース」であれば、「用語の意味を調べる、会社を調べる、価格を調べる」といったラベルが定義されている。なお、「ID」は、各コンテンツタイプ名を一意に識別するための ID 番号である。

【 0 1 2 2 】

また、各ラベルには、その検索対象に応じた検索キーワード群が、「検索クエリ」として予め対応付けて定義されている。図 19 は、ラベルに対応付けられた「検索クエリ」の例を示した図である。

【 0 1 2 3 】

例えば、ラベル名が「価格を調べる」であれば、「価格、値段、実売価格、プライス、特価」といった語群が、“OR”で連結された検索クエリ(検索条件)として定義されている。同様に、ラベル名が「試合結果を見る」であれば、「得点、試合結果、勝利、点差」といった語群が、ラベル名が「ニュースを見る」であれば、「ニュース、記事」といった語が、ラベル名「プロフィールを見る」であれば、「profile、プロフィール、自己紹介」といった語群が、ラベル名が「写真を見る」であれば、「画像、イメージ、jpg、jpeg」といった語群が、夫々“OR”で連結された検索クエリとして定義されている。

【 0 1 2 4 】

10

20

30

40

50

図17に戻り、検索方法候補選定部208は、検索方法モデル267を、入力文書推定結果、サイト種別推定結果及びコンテンツタイプ推定結果の各項目に適用することで、各項目に応じたラベルを夫々決定する(ステップS283)。そして、検索方法候補選定部208は、決定したラベルを検索方法候補としてRAM25等に保持し(ステップS284)、図4のステップS29へと移行する。

【0125】

ここで、図20、図21を参照して、ステップS283の処理について説明する。入力文書推定結果(図12、A34参照)、サイト種別推定結果(図14、A44参照)及びコンテンツタイプ推定結果(図16、A54参照)に対し、図18で示した検索方法モデルを適用すると、検索方法候補選定部208は、各項目(テーマ、サイト種別、コンテンツタイプ)について、図20に示した検索方法モデル(ID200、ID020、ID003)が適合すると判断することになる。

10

【0126】

上各検索方法モデルの判断は、例えば、以下の方法で決定することができる。まず、検索方法候補選定部208は、図21に示したように、テーマ、サイト種別、コンテンツタイプ毎に、その内容をノードとして表記し、これら異なる3つの項目から、各々一つのノードを選択する。その際選択されるのは、スコア及び関連度に、より大きな値が付与されたテーマ名、サイト種別名及びコンテンツタイプ名に対応するノードが選択されるものとする。この結果、検索方法候補選定部208は、テーマ名「スポーツ」、サイト種別名「ニュース」、コンテンツタイプ名「記事」の各ノードを選択する。

20

【0127】

次に、検索方法候補選定部208は、これまでに選択した各ノードに対応する「ラベル」を枚挙することで、検索方法候補として「試合結果を見る、プロフィールを見る、ニュースを見る、写真を見る、地名を調べる…」を取得し、後段の検索方法候補検証部209に出力する。

【0128】

図4に戻り、続くステップS29では、検索方法候補検証部209により、検索方法検証処理(ステップS29)が実行される。以下、図22を参照して、ステップS29の検索方法検証処理の動作について説明する。

【0129】

図22は、検索方法候補検証部209により実行される検索方法検証処理の手順を示したフローチャートである。まず、検索方法候補検証部209は、検索方法候補に含まれた各ラベルを参照し、各ラベルに対応する「検索クエリ」を、記憶部26に記憶された検索方法モデル267から読み出す(ステップS291)。

30

【0130】

次いで、検索方法候補検証部209は、読み出した「検索クエリ」毎に、当該「検索クエリ」を「キーワード」とともに検索エンジン30に送信することで、当該検索エンジン30に「検索クエリ」及び「キーワード」に関する文書の検索を依頼する(ステップS292)。ここで「キーワード」とは、閲覧リソースに含まれたキーワード、即ち情報通信端末10のユーザが指定したキーワードであるものとする。なお、「検索クエリ」と「キーワード」とは“AND”で連結する態様としてもよいし、“OR”で連結する態様としてもよい。

40

【0131】

続いて検索方法候補検証部209は、各検索クエリについて、検索エンジン30から得られた検索件数の多い順に、各ラベルをソートし(ステップS293)、ソートしたラベルと、検索クエリとを対応付けて後段の検索方法候補メニュー生成部210に出力した後(ステップS294)、図4のステップS30に移行する。

【0132】

図4に戻り、検索方法候補メニュー生成部210は、検索方法検証処理でソートされたラベル名を、情報通信端末10の表示部13で表示可能なメニューとして生成し(ステッ

50

プ S 3 0)、このメニューを各ラベルに対応する「検索クエリ」とともに、情報通信端末 1 0 に送信した後 (ステップ S 3 1)、検索支援サーバ 2 0 側の処理を終了する。

【 0 1 3 3 】

なお、生成される「メニュー」には、ソートされた全てのラベル名が表示される態様としてもよいし、例えば、ソートされたラベルのうち、上位に位置するラベル (例えば、上位 4 つのラベル名) のみが表示される態様としてもよい。

【 0 1 3 4 】

一方、情報通信端末 1 0 では、検索支援サーバ 2 0 から「メニュー」及び「検索クエリ」を受け取ると、検索方法提示選択部 1 0 4 は、この「メニュー」により表される画面 (メニュー画面) を表示部 1 3 に表示させた後 (ステップ S 1 5)、操作部 1 2 を介してユーザから特定のラベルが指示されるまで待機を行う (ステップ S 1 6 ; N o)。

10

【 0 1 3 5 】

図 2 3 は、ステップ S 1 5 で表示部 1 3 に表示されたメニュー画面の一例を示した図である。情報通信端末 1 0 のユーザは、このメニュー画面に提示された各項目 (ラベル) から、特定のラベルを、操作部 1 2 を介して選択することが可能となっている。ユーザから操作部 1 2 を介して特定のラベル (例えば、「試合結果を見る」) が指定されると、当該ラベル名に応じた指示信号が検索方法提示選択部 1 0 4 に入力される。

【 0 1 3 6 】

ステップ S 1 6 において、検索方法提示選択部 1 0 4 により、特定のラベルを指示する指示信号の入力が確認されると (ステップ S 1 6 ; Y e s)、検索クエリ送信部 1 0 5 は、この指示されたラベルに対応する「検索クエリ」と、ステップ S 1 3 で入力された「キーワード」とを、検索エンジン 3 0 に送信することで、当該検索エンジン 3 0 に「検索クエリ」及び「キーワード」に関する文書の検索を依頼する (ステップ S 1 7)。なお、「検索クエリ」と「キーワード」とは “ A N D ” で連結して検索させる態様としてもよいし、“ O R ” で連結して検索させる態様としてもよい。

20

【 0 1 3 7 】

例えば、図 2 3 に示したメニュー画面に基づき、ラベル「試合結果を見る」が選択されたとすると、対応する「検索クエリ」は、「得点、試合結果、勝利、点差」 (図 1 9 参照) を “ O R ” で連結した文字列であるため、検索クエリ送信部 1 0 5 は、この「検索クエリ」をユーザが指定した「キーワード」 (例えば、「田中」) とともに検索エンジン 3 0 に送信することになる。

30

【 0 1 3 8 】

続いて検索結果表示部 1 0 6 は、検索エンジン 3 0 から送信された検索結果を受け取ると、この検索結果を表示部 1 3 に表示させ (ステップ S 1 8)、情報通信端末 1 0 での処理を終了する。

【 0 1 3 9 】

図 2 4 は、ステップ S 1 8 で表示部 1 3 に表示された検索結果の一例を示した図である。情報通信端末 1 0 のユーザは、この検索結果に基づき、所望する W e b サーバ 4 0 を、操作部 1 2 を介して選択することで、当該 W e b サーバ 4 0 のコンテンツを閲覧することができる。なお、ここで閲覧したコンテンツから再度キーワードを選択し、上述したブラウザ連動型検索処理を行うこととしてもよい。

40

【 0 1 4 0 】

以上のように、本実施形態によれば、文書及び当該文書に含まれた特定の語句 (キーワード) から、当該文書のテーマ、当該文書を提供した W e b サーバ 4 0 のサイト種別、当該文書のコンテンツタイプを推定し、これらテーマ、サイト種別及びコンテンツタイプに応じた検索条件を選定することで、ユーザは、関心のある文書中から特定の語句を指示するのみで、当該文書のテーマ、当該文書を提供した W e b サーバ 4 0 のサイト種別、当該文書のコンテンツタイプに応じた検索条件を享受することができるため、検索の操作にかかる負荷を減少させるとともに、ユーザが所望する情報に容易に到達させることができる。

50

【 0 1 4 1 】

なお、本実施形態の検索支援サーバ20で実行されるプログラムは、ROM24等に予め組み込まれて提供するものとするが、これに限らず、インストール可能な形式又は実行可能な形式のファイルでCD-ROM、フレキシブルディスク(FD)、CD-R、DVD(Digital Versatile Disk)等のコンピュータで読み取り可能な記録媒体に記録して提供するように構成してもよい。また、このプログラムを、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成してもよく、インターネット等のネットワーク経由で提供または配布するように構成してもよい。

【 0 1 4 2 】

また、本実施形態では、文書のテーマ、サイト種別及びコンテンツタイプに基づいて、検索条件を選定することとしたが、これに限らず、文書のテーマ、サイト種別及びコンテンツタイプのうち、何れかの特性に基づいて、検索条件を選定することとしてもよい。

【 0 1 4 3 】

[第2の実施形態]

次に、検索支援装置、プログラム及び検索支援システムの第2の実施形態について説明する。なお、上述した第1の実施形態と同様の要素については、同一の符号を用いて示し、その説明は適宜省略する。

【 0 1 4 4 】

図25は、本実施形態の検索支援システム2の構成を示した概略図である。図25に示すように、検索支援システム2は、情報通信端末50と、検索支援サーバ60と、検索エンジン30、Webサーバ40とを有し、各機器はインターネット等のネットワークを介して相互に通信可能に接続されている。

【 0 1 4 5 】

図25に示すとおり、本実施形態の情報通信端末50は、第1の実施形態で説明した情報通信端末10の機能構成(図1参照)から、キーワード指定部102を取り除き、キーワードリスト提示選択部501を新たに加えた構成となっている。即ち、本実施形態では、コンテンツ中の特定のキーワードが検索支援サーバ60に送信されるのではなく、コンテンツの閲覧を実施中に、ブラウザ連動型検索処理を実行する旨の指示信号が入力されると、この現在閲覧中のコンテンツのURL又はコンテンツのHTML文書が、閲覧リソースとして検索支援サーバ60に送信されるようになっている。

【 0 1 4 6 】

ここで、図26-1~図26-3を参照して、本実施形態の情報通信端末10から、閲覧リソースが送信されるまでの手順を説明する。図26-1に示すように、コンテンツが表示部13に表示された状態で、「ブラウザメニュー」から、図5-2に示すように、ブラウザ連動型検索処理の実行を指示する「連動検索」が選択され、この指示信号が入力されると、閲覧リソース送信部103は、図5-3で示したように、表示部13に表示されたコンテンツのURL又はコンテンツのHTML文書を、閲覧リソースとして検索支援サーバ60に送信する。

【 0 1 4 7 】

また、本実施形態の検索支援サーバ60は、第1の実施形態で説明した検索支援サーバ20の機能構成(図1参照)に、キーワード切り出し部601、キーワード候補選定部602を加えた構成となっている。

【 0 1 4 8 】

ここで、キーワード切り出し部601は、入力文書リソース取得部201から入力されたHTML文書を入力文書とし、この入力文書に形態素解析を適用し、当該入力文書を形態素毎に分割し、RAM25等に保持する。

【 0 1 4 9 】

また、本実施形態のキーワード候補選定部602は、キーワード切り出し部601により分割された形態素毎に、第1の実施形態と同様、各品詞(キーワード)の選択範囲を拡

10

20

30

40

50

張する。そして、602は、キーワード候補として得られた各文字列を、検索エンジン30に夫々送信することで取得した検索結果のうち、検索結果が多く得られた上位のキーワードを選定し、選定したキーワードをキーワード候補として情報通信端末50に送信する。なお、選定するキーワードの個数は特に問わないものとする。

【0150】

また、キーワード候補選定部602は、情報通信端末50のキーワードリスト提示選択部501からキーワードを受信すると、このキーワード及び入力文書に基づいて、キーワード候補選定処理(図6参照)を実行する。なお、移行の処理は、上述した第1の実施形態と同様であるため、説明は省略する。

【0151】

一方、情報通信端末50では、検索支援サーバ60から送信されたキーワード候補が受信されると、キーワードリスト提示選択部501は、このキーワード候補を表示部13に表示させることでユーザに提示する。

【0152】

図27-1は、キーワードリスト提示選択部501により表示部13に表示されたキーワード候補の一例を示した図である。情報通信端末10のユーザは、図27-2に示すように、表示部13に提示されたキーワード候補から、特定のキーワード名を、操作部12を介して選択することが可能となっている。そして、特定のキーワード名(例えば、「田中寿英」)が選択されると、当該キーワード名に応じた指示信号がキーワードリスト提示選択部501に入力される。

【0153】

キーワードリスト提示選択部501は、特定のキーワード名を指示する指示信号の入力を確認すると、この指示信号で指示されたキーワード名に対応するキーワードを、検索支援サーバ60のキーワード候補選定部602に送信する。

【0154】

以上のように、本実施形態によれば、ユーザは、閲覧中の文書から選出された複数のキーワード候補から、特定のキーワードを選択するのみで、当該文書のテーマ、当該文書を提供したWebサーバ40のサイト種別及び当該文書のコンテンツタイプに応じた検索条件を享受することができるため、検索の操作にかかる負荷を減少させるとともに、ユーザが所望する情報に容易に到達させることができる。

【0155】

[第3の実施形態]

次に、検索支援装置、プログラム及び検索支援システムの第3の実施形態について説明する。なお、上述した第1の実施形態と同様の要素については、同一の符号を用いて示し、その説明は適宜省略する。

【0156】

図28は、本実施形態の検索支援システム3の構成を示した概略図である。図28に示すように、検索支援システム3は、情報通信端末70と、検索支援サーバ80と、検索エンジン30、Webサーバ40とを有し、各機器はインターネット等のネットワークを介して相互に通信可能に接続されている。

【0157】

図28に示すとおり、本実施形態の情報通信端末70は、第1の実施形態で説明した情報通信端末10の機能構成(図1参照)から、コンテンツ閲覧部101、キーワード指定部102及び閲覧リソース送信部103を取り除き、キーワード入力部701及びキーワード送信部702を新たに加えた構成となっている。

【0158】

キーワード入力部701は、連携検索の対象となるキーワードを、操作部12を介して直接入力することが可能な入力支援手段を情報通信端末10のユーザに提供する。ここで、入力支援手段とは、例えば、図29に示したような、表示部13に表示されるインターフェース等であって、ユーザはこの入力支援手段を介して連携検索の対象とするキーワー

10

20

30

40

50

ドを入力することが可能となっている。キーワード入力部 701 は、この入力支援画面に基づいて入力されたキーワードを、キーワード送信部 702 に出力する。

【0159】

キーワード送信部 702 は、キーワード入力部 701 から入力されたキーワードを、検索支援サーバ 80 に送信する。

【0160】

本実施形態の検索支援サーバ 80 は、第 1 の実施形態で説明した検索支援サーバ 20 の機能構成（図 1 参照）から、入力文書リソース取得部 201 を取り除き、予備検索実行部 801 を新たに加えた構成となっている。

【0161】

予備検索実行部 801 は、情報通信端末 70 のキーワード送信部 702 から送信されたキーワードを、検索エンジン 30 に送信することで、当該キーワードに関する文書の検索を依頼する。そして、予備検索実行部 801 は、キーワードについて、検索エンジン 30 から検索結果として得られたコンテンツの HTML 文書を、上位の Web サーバ（Web サイト）から所定数分取得し、この取得した HTML 文書と、情報通信端末 70 から送信されたキーワードとを、キーワード候補選定部 202 に出力する。なお、移行の処理は、上述した第 1 の実施形態と同様であるため、説明は省略する。なお、取得する HTML 文書の個数は、任意に設定可能であるものとする。

【0162】

以上のように、本実施形態によれば、情報通信端末 10 のユーザは、連携検索対象としたキーワードを入力するのみで、このキーワードに関連する文書から、当該文書のテーマ、当該文書を提供した Web サーバ 40 のサイト種別及び当該文書のコンテンツタイプに応じた検索条件を享受することができるため、検索の操作にかかる負荷を減少させるとともに、ユーザが所望する情報に容易に到達させることができる。

【図面の簡単な説明】

【0163】

【図 1】第 1 の実施形態における検索支援システムの構成を示した概略図である。

【図 2】端末のハードウェア構成を示した図である。

【図 3】検索支援サーバのハードウェア構成を示した図である。

【図 4】ブラウザ連動型検索処理の動作を示したラダーチャートである。

【図 5 - 1】端末の表示部に表示された画面の一例を示した図である。

【図 5 - 2】端末の表示部に表示された画面の一例を示した図である。

【図 5 - 3】端末の表示部に表示された画面の一例を示した図である。

【図 5 - 4】端末の表示部に表示された画面の一例を示した図である。

【図 5 - 5】端末の表示部に表示された画面の一例を示した図である。

【図 6】キーワード候補選定処理の流れを示したフローチャートである。

【図 7】語句属性解析処理の流れを示したフローチャートである。

【図 8】語句属性解析処理を説明するための図である。

【図 9】キーワード意味推定処理の流れを示したフローチャートである。

【図 10】キーワード意味推定処理を説明するための図である。

【図 11】文書テーマ推定処理の流れを示したフローチャートである。

【図 12】文書テーマ推定処理を説明するための図である。

【図 13】サイト種別推定処理の流れを示したフローチャートである。

【図 14】サイト種別推定処理を説明するための図である。

【図 15】コンテンツタイプ推定処理の流れを示したフローチャートである。

【図 16】コンテンツタイプ推定処理を説明するための図である。

【図 17】検索方法候補選定処理の流れを示したフローチャートである。

【図 18 - 1】検索方法モデルの一例を示した図である。

【図 18 - 2】検索方法モデルの一例を示した図である。

【図 18 - 3】検索方法モデルの一例を示した図である。

10

20

30

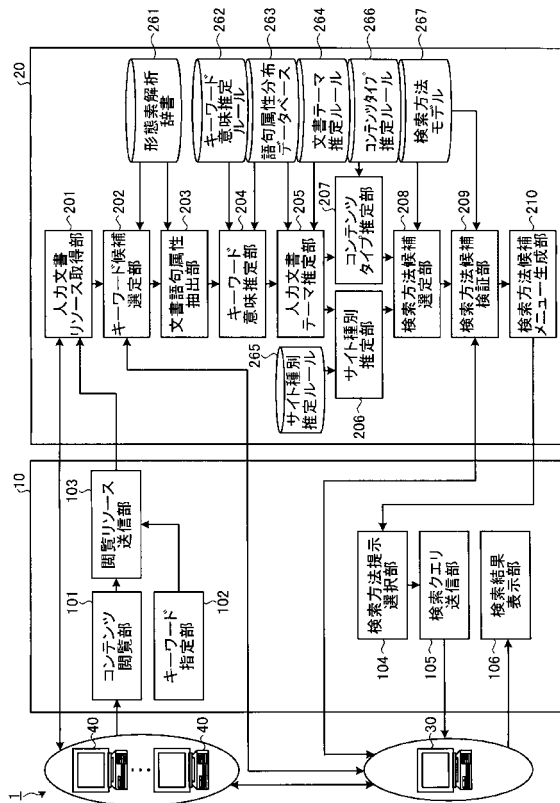
40

50

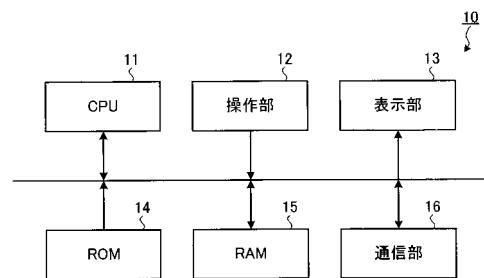
- 【図19】 検索方法モデルの一例を示した図である。
- 【図20】 コンテンツタイプ推定処理を説明するための図である。
- 【図21】 コンテンツタイプ推定処理を説明するための図である。
- 【図22】 検索方法検証処理の流れを示したフローチャートである。
- 【図23】 端末の表示部に表示された画面の一例を示した図である。
- 【図24】 端末の表示部に表示された画面の一例を示した図である。
- 【図25】 第2の実施形態における検索支援システムの構成を示した概略図である。
- 【図26-1】 端末の表示部に表示された画面の一例を示した図である。
- 【図26-2】 端末の表示部に表示された画面の一例を示した図である。
- 【図26-3】 端末の表示部に表示された画面の一例を示した図である。 10
- 【図27-1】 端末の表示部に表示された画面の一例を示した図である。
- 【図27-2】 端末の表示部に表示された画面の一例を示した図である。
- 【図28】 第3の実施形態における検索支援システムの構成を示した概略図である。
- 【図29】 端末の表示部に表示された画面の一例を示した図である。
- 【符号の説明】
- 【0164】
- 1 検索支援システム
- 2 検索支援システム
- 3 検索支援システム
- 10 情報通信端末 20
- 11 CPU
- 12 操作部
- 13 表示部
- 14 ROM
- 15 RAM
- 16 通信部
- 101 コンテンツ閲覧部
- 102 キーワード指定部
- 103 閲覧リソース送信部
- 104 検索方法提示選択部 30
- 105 検索クエリ送信部
- 106 検索結果表示部
- 20 検索支援サーバ
- 21 CPU
- 22 操作部
- 23 表示部
- 24 ROM
- 25 RAM
- 26 記憶部
- 261 形態素解析辞書 40
- 262 キーワード意味推定ルール
- 263 語句属性データベース
- 264 文書テーマ推定ルール
- 265 サイト種別推定ルール
- 266 コンテンツタイプ推定ルール
- 267 検索方法モデル
- 27 通信部
- 201 入力文書リソース取得部
- 202 キーワード候補選定部
- 203 文書語句属性抽出部 50

- 204 キーワード意味推定部
- 205 入力文書テーマ推定部
- 206 サイト種別推定部
- 207 コンテンツタイプ推定部
- 208 検索方法候補選定部
- 209 検索方法候補検証部
- 210 検索方法候補メニュー生成部
- 30 検索エンジン
- 40 Webサーバ
- 50 情報通信端末
- 501 キーワードリスト提示選択部
- 60 検索支援サーバ
- 601 キーワード切り出し部
- 602 キーワード候補選定部
- 70 情報通信端末
- 701 キーワード入力部
- 702 キーワード送信部
- 80 検索支援サーバ
- 801 予備検索実行部

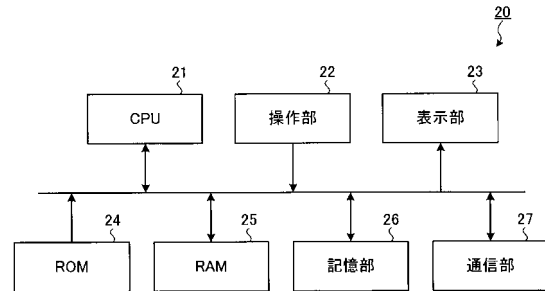
【図1】



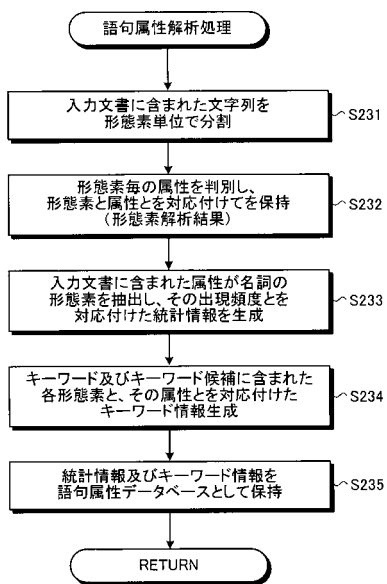
【図2】



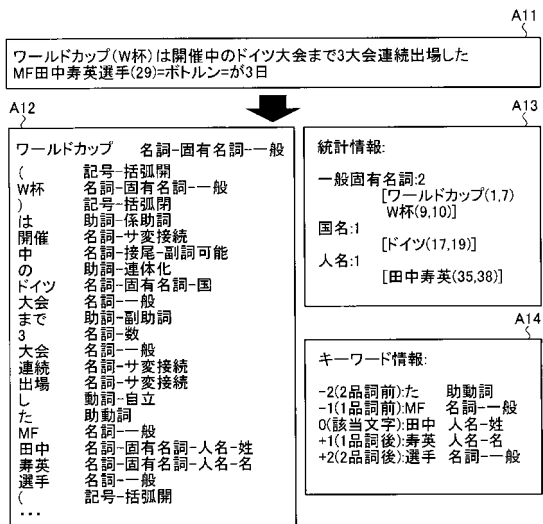
【図3】



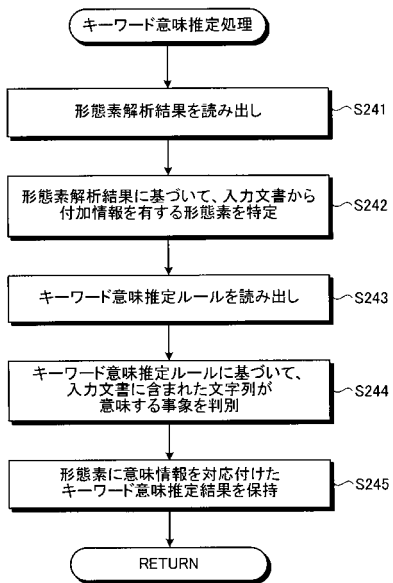
【図7】



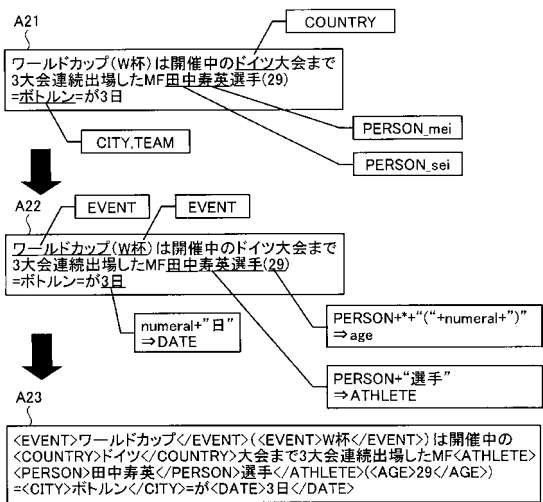
【図8】



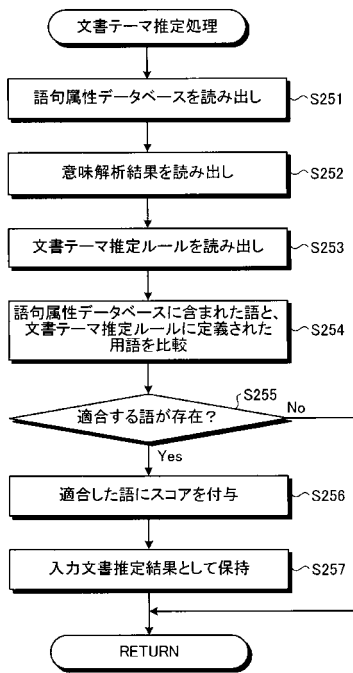
【図9】



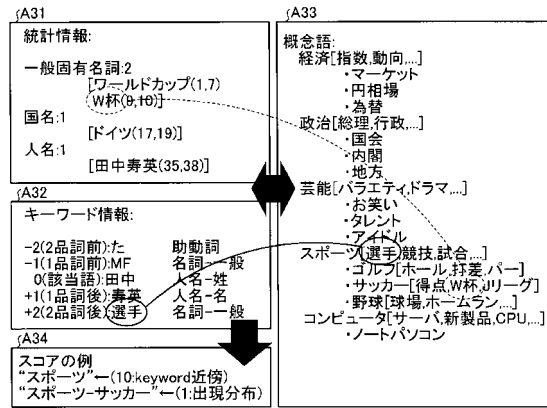
【図10】



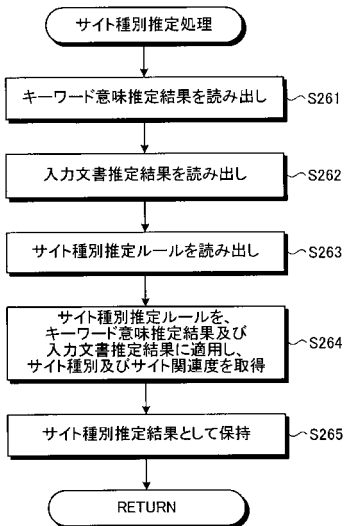
【図 1 1】



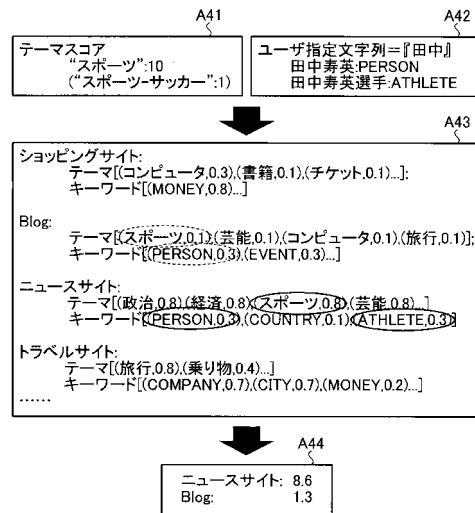
【図 1 2】



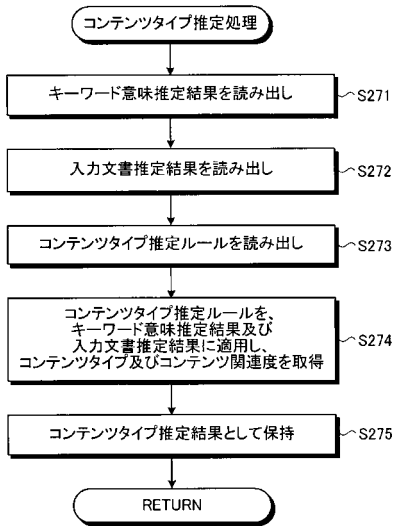
【図 1 3】



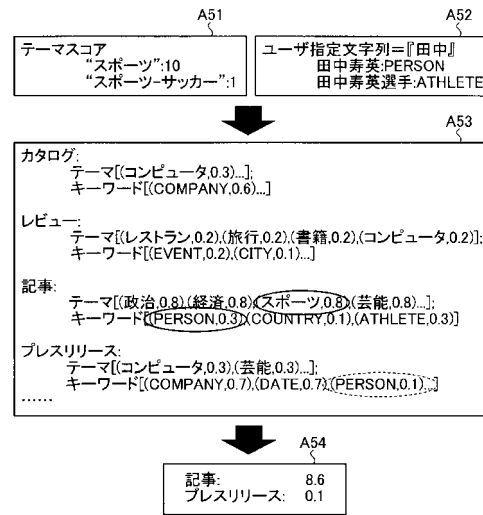
【図 1 4】



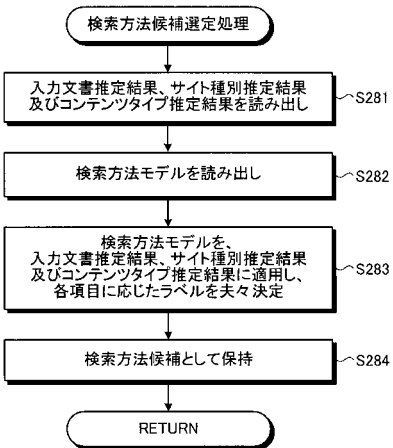
【 図 1 5 】



【 図 1 6 】



【 図 1 7 】



【 図 1 8 - 1 】

ID	テーマ	ラベル
100	コンピュータ	スペックを調べる 発売時期を調べる 評価を見る
200	スポーツ	試合結果を見る プロフィールを見る
300	政治	プロフィールを見る 意味を調べる 組織を調べる
400	旅行	地図を見る 価格を調べる ツアーを探す
...		

【 図 1 8 - 2 】

ID	サイト種別	ラベル
010	ショッピングサイト	価格を調べる 評価を調べる 販売店を調べる
020	ニュースサイト	ニュースを見る プロフィールを見る 写真を見る
030	Blog	評価を見る プロフィールを見る
040	トラベルサイト	ツアーを探す 地図を見る 地名を調べる 感想を見る
...		

【 図 1 8 - 3 】

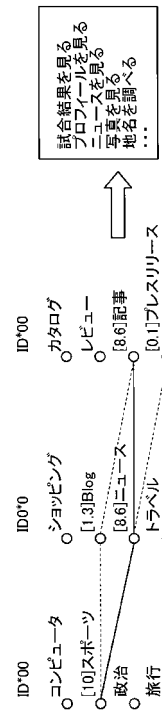
ID	コンテンツタイプ	ラベル
010	カタログ	用語の意味を調べる 価格を調べる
020	レビュー	スペックを調べる 評価を見る
030	記事	ニュースを見る プロフィールを見る 地名を調べる
040	プレスリリース	用語の意味を調べる 会社を調べる 価格を調べる
...		

【図 19】

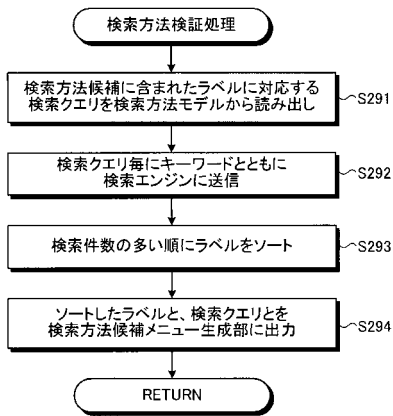
ラベル 価格を調べる 試合結果を探す ニュースを読む プロフィールを見る 写真を見る	付加クエリ 価格 値段 実売価格 プライス 特價 得点 試合結果 勝利 点差 ニュース 記事 profile プロフィール 自己紹介 画像 イメージ jpg jpeg
-----------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------

【図 20】

ID	テーマ	ラベル
200	スポーツ	試合結果を見る プロフィールを見る
ID	サイト種別	ラベル
020	ニュースを見る	ニュースを見る プロフィールを見る 写真を見る
ID	コンテンツタイプ	ラベル
003	記事	ニュースを見る プロフィールを見る 地名を調べる



【図 22】



【図 23】

「田中」について

1. 試合結果を見る 携帯サイト
2. 写真を見る 携帯サイト
3. ニュースを読む 携帯サイト
4. プロフィールを見る 携帯サイト

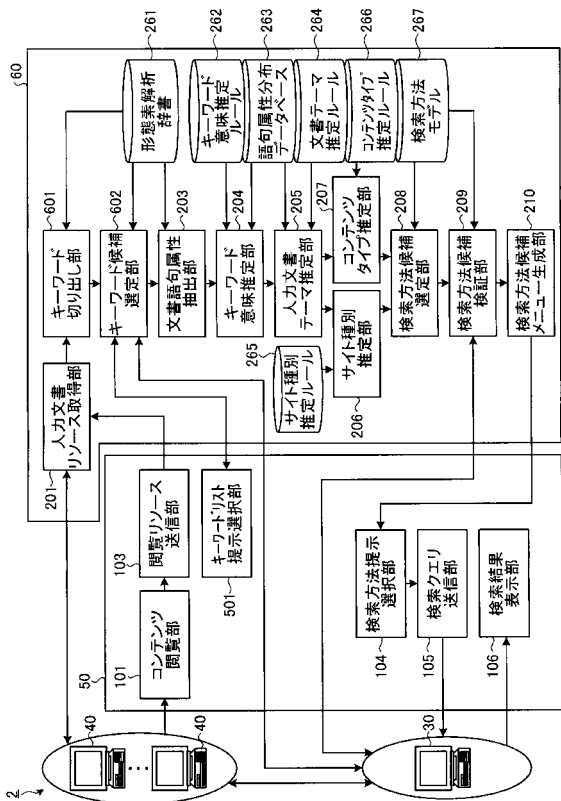
元のページに戻る

【図 24】

Googoo

1. 日-ブラジル1対4で敗れる
中央でボールを持った田中寿英がミドルシュートを放つが、キーパーが正面で
- 2.3-1でオーストラリアの勝利
右サイドでフリーの中村がクロスを上げるが、少し長い。田中寿英はボールを

【図25】



【図26-1】

今日のニュース
 ワールドカップ(W杯)は開催中のドイツ大会まで3大会連続出場したMF田中寿英選手(29)=ポトルン=が3日...

ブラウザメニュー

【図26-2】

今日のニュース
 ワールドカップ(W杯)は開催中のドイツ大会まで3大会連続出場したMF田中寿英選手(29)=ポトルン=が3日...

お気に入り登録
 画面メモ保存
 ...

運動型検索

【図26-3】

今日のニュース
 ワールドカップ(W杯)は開催中のドイツ大会まで3大会連続出場したMF田中寿英選手(29)=ポトルン=が3日...

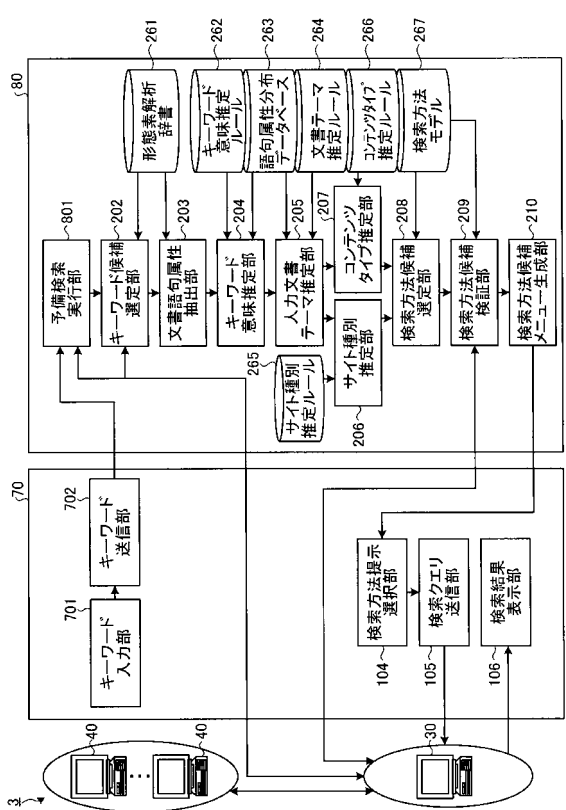
【図27-1】

キーワード
 1.W杯
 2.ドイツ
 3.田中寿英
 ...

【図27-2】

キーワード
 1.W杯
 2.ドイツ
 3.田中寿英
 ...

【図28】



【図 29】

トップメニュー
<input type="text"/> <input type="button" value="検索"/>
最新情報
着ミュージックをプレゼント!

フロントページの続き

審査官 吉田 誠

(56)参考文献 特開2002-222210(JP,A)
特開2001-22788(JP,A)
特開平10-162011(JP,A)

(58)調査した分野(Int.Cl., DB名)
G06F 17/30