

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6131983号  
(P6131983)

(45) 発行日 平成29年5月24日(2017.5.24)

(24) 登録日 平成29年4月28日(2017.4.28)

(51) Int.Cl.

G06F 17/30 (2006.01)

F 1

G06F 17/30 210A  
G06F 17/30 170A  
G06F 17/30 330C

請求項の数 7 (全 16 頁)

(21) 出願番号	特願2015-95131 (P2015-95131)	(73) 特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番 1号
(22) 出願日	平成27年5月7日(2015.5.7)	(74) 代理人	100099759 弁理士 青木 篤
(62) 分割の表示	特願2013-100838 (P2013-100838) の分割	(74) 代理人	100119987 弁理士 伊坪 公一
原出願日	平成19年3月30日(2007.3.30)	(74) 代理人	100133835 弁理士 河野 努
(65) 公開番号	特開2015-164066 (P2015-164066A)	(74) 代理人	100135976 弁理士 宮本 哲夫
(43) 公開日	平成27年9月10日(2015.9.10)	(72) 発明者	坂本 慶彦 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	平成27年5月7日(2015.5.7)		
前置審査			

最終頁に続く

(54) 【発明の名称】情報検索装置、情報検索方法およびそのプログラム

## (57) 【特許請求の範囲】

## 【請求項1】

コンピュータに、

文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうち、いずれの文書から抽出されたキーワードを検索処理に用いるかを特定する情報を検索依頼を行う他のコンピュータから取得すると、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、該特定する情報によって特定される文書から抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを検索処理に用いる。

ことを実行させ、前記検索処理に先立ち、前記特定する情報によって特定される文書から抽出された前記キーワードを表示することを特徴とする情報検索プログラム。

## 【請求項2】

コンピュータが、

文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうち、いずれの文書から抽出されたキーワードを検索処理に用いるかを特定する情報を検索依頼を行う他のコンピュータから取得すると、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、該特定する情報によって特定される文書から抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを検索処理に用いる。

ことを実行し、前記検索処理に先立ち、前記特定する情報によって特定される文書から

抽出された前記キーワードを表示することを特徴とする情報検索方法。

**【請求項 3】**

文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうち、いずれの文書から抽出されたキーワードを検索処理に用いるかを特定する情報を検索依頼を行う他のコンピュータから取得すると、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、該特定する情報によって特定される文書から抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを検索処理に用いる制御部、

を備え、前記検索処理に先立ち、前記特定する情報によって特定される文書から抽出された前記キーワードを表示することを特徴とする情報検索装置。

10

**【請求項 4】**

コンピュータに、

文書の検索処理を行う際に、検索依頼を行う他のコンピュータから取得された、文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうちのいずれかの文書を特定する情報に基づいて特定される文書から、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを用いて、前記検索処理を行う、

ことを実行させ、前記検索処理に先立ち、前記特定する情報によって特定される文書から抽出された前記キーワードを表示することを特徴とする情報検索プログラム。

**【請求項 5】**

20

前記特定される文書は、前記特定される文書から抽出されたキーワードと関連付けて前記他のコンピュータの表示部に表示された文書である

ことを特徴とする請求項 4 記載の情報検索プログラム。

**【請求項 6】**

コンピュータが、

文書の検索処理を行う際に、検索依頼を行う他のコンピュータから取得された、文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうちのいずれかの文書を特定する情報に基づいて特定される文書から、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを用いて、前記検索処理を行い、前記検索処理に先立ち、前記特定する情報によって特定される文書から抽出された前記キーワードを表示する、

30

ことを特徴とする情報検索方法。

**【請求項 7】**

40

文書の検索処理を行う際に、検索依頼を行う他のコンピュータから取得された、文書と文書から抽出されたキーワードとを対応づけて双方表示可能な複数の文書のうちのいずれかの文書を特定する情報に基づいて特定される文書から、該特定する情報によって特定される文書中のいずれかの単語の指定を求めずに、抽出されたキーワードを自動的に検索用キーワードとして生成し、生成した該検索キーワードを用いて、前記検索処理を行う制御部、

を備え、前記検索処理に先立ち、前記特定する情報によって特定される文書から抽出された前記キーワードを表示することを特徴とする情報検索装置。

**【発明の詳細な説明】**

**【技術分野】**

**【0001】**

本発明は、情報検索装置、情報検索方法およびそのプログラムに関し、特に、電子化された教材を用いて学習を行うW E Bベースの学習システムにおいて、学習者の質問内容からその教材の学習に最適な情報、すなわち質問に対する回答を得るための情報を検索可能にする情報検索装置、情報検索方法およびそのプログラムに関する。

**【背景技術】**

50

**【0002】**

今日、コンピュータネットワークの拡大とその進歩に伴い、情報を入手したい時、インターネットで公開されている情報の中から所望の情報をキーワードなどを使って検索可能にするサーチエンジンが広く利用されている。サーチエンジンは、WWWなどで公開されている情報の全文を貯えておき、キーワードによって検索する全文検索型と、カテゴリー別に分類されているディレクトリ型の2種類に大別できる。最近では、ディレクトリ型のサーチエンジンにも検索機能が搭載されるようになってきている。

**【0003】**

電子化された教材を用いて学習を行うWEBベースの学習システムにおいても、情報の入手にキーワードによる検索が必須である。WEBベースの学習システムは、学習者からの質問を受け付ける機能を持つものが一般的であるが、質問に対する回答は講師が手作業により作成する必要がある。このため、コストが掛かり、またその回答を得るまでに時間が掛かる。講師が回答を作成する代わりに、学習者自身がWEB上のサーチエンジンを用いて検索し回答を得ることはできるものの、所望の情報を得るためにサーチエンジンに入力する、キーワードから構成される検索式を適切に作成して回答を得ることは困難である。

10

**【0004】**

また、WEBベースの学習システムにおいて、学習者の意欲を保つためには、疑問点に対して即座のレスポンスが必要であり、メールや掲示版では遅い。チャットを利用する上で即時性は保てるが、チャット利用者に即座に対応するため常に講師が待機する必要があり、コストが掛かる。学習者は、疑問点の解決のため、サーチエンジンを用いて、ファイルやWEBページを検索することはできるが、一般的にサーチエンジンの利用者にとって、所望の情報を得るためにサーチエンジンに入力するのに適したキーワードから構成される検索式を作成することは難しく、1つの検索キーワードだけ入力して検索する場合が多い。この場合、一度に大量のWEBページが検索され、有効な関連WEBページを得るのは困難である。

20

**【0005】**

また、従来のWEBベースの学習システムにおいて、利用者の入力したテキスト文（文書）から、カテゴリーを判別し、検索結果を絞り込む技術はあるが、事前の準備が必要であり、コストが掛かる。

**【0006】**

30

しかるに、特許文献1に記載の情報検索システムは、情報ネットワーク上の情報を検索する検索装置と、原文と該原文に関するユーザからの質問とに基づいて検索装置に対する検索要求を行う検索要求装置と、を含む情報検索システムにおいて、検索装置は、原文を指定する原文指定部と、学習者の質問文から検索キーワードを抽出するキーワード抽出部と、その検索キーワードに基づいて質問の質問属性を指定する質問属性指定部と、指定された質問属性に基づいて検索装置による検索対象となる範囲を絞り込むことを利用できる単語の単語属性を指定し、原文からその単語属性に属する単語を抽出し、抽出した単語を検索キーワードへの追加キーとする追加キー抽出部と、追加キーが抽出された場合にその追加キーを検索キーワードに追加し、検索装置に対し質問に基づいた検索要求を作成し、検索要求を原文から抽出された追加キーに関連付ける関連付け部とを有する。

40

**【0007】**

端的に言えば、特許文献1は、質問文から抽出された検索キーワードに基づいて指定された質問属性に基づき検索範囲を絞り込むための単語属性を指定し、原文から該単語属性に属する単語を抽出し、抽出した単語を検索キーワードに加えて検索を行う技術を開示したものである。

**【先行技術文献】****【特許文献】****【0008】**

【特許文献1】特開2005-11149号公報（明細書の[請求項1]、段落番号[0002]、[0004]、[0007]～[0009]、[0014]、[00015]

50

および図面の〔図1〕参照)。

【発明の概要】

【発明が解決しようとする課題】

【0009】

しかしながら、上記特許文献1に開示の情報検索システムおよび検索要求装置は、原文から単語属性に属する単語を抽出し、抽出した単語を検索キーワードに加えて検索式を求めており、学習者の質問文および/または原文の記載によっては重要でない検索キーワードが求められてしまい、所望のキーワードを抽出できず、十分な絞込みができないという問題がある。

また、上記特許文献1に記載の発明では、学習者に何等の操作、意識をさせることなく、すなわち自動的に、質問に対する回答を得ることはできないという問題がある。 10

【0010】

本発明は、上記問題を解決するためになされたものであり、情報に関する質問内容からその質問に回答するのに最適な情報を検索する情報検索装置、情報検索方法およびそのプログラムの提供、より具体的には、教材に関する学習者の質問内容からその教材の学習に最適な十分に絞り込んだ情報を自動的に検索する情報検索装置、情報検索方法およびそのプログラムの提供を目的とする。

【課題を解決するための手段】

【0011】

上記目的を達成する本発明による情報検索プログラムは、学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索プログラムであって、コンピュータ装置を複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力手段と、学習者からの質問文の入力を受け付ける質問文受付手段と、前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出手段と、前記出力手段から出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出手段と、前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索手段と、して機能させることを特徴とする。 20

【0012】

上記情報検索プログラムは、さらに、前記コンピュータ装置を、前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成手段、として機能させることを特徴とする。

【0013】

上記目的を達成する本発明による情報検索方法は、学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索方法であって、コンピュータ装置が、複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力ステップと、学習者からの質問文の入力を受け付ける質問文受付ステップと、前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出ステップと、前記出力ステップで出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出ステップと、前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索ステップと、を行うことを特徴とする。 40

【0014】

50

上記情報検索方法は、さらに、前記コンピュータ装置が、前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成ステップ、を行うこと特徴とする。

#### 【0015】

上記目的を達成する本発明による情報検索装置は、学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索装置であって、複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力手段と、学習者からの質問文の入力を受け付ける質問文受付手段と、前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出手段と、前記出力ステップで出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出手段と、前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索手段と、を備えることを特徴とする。10

#### 【0016】

上記情報検索装置は、さらに、前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成手段、を備えること特徴とする。20

#### 【発明の効果】

#### 【0017】

本発明の情報検索装置、情報検索方法およびそのプログラムによれば、学習者は検索式を意識せずに教材の学習内容についての質問文を入力することが可能となり、かつ教材の学習内容に沿った学習者が真に必要とする最適な情報を自動的に検索することが可能となる。

#### 【図面の簡単な説明】

#### 【0018】

【図1】本発明の情報検索装置のブロック構成図である。30

【図2】学習対象の文書に関する質問の回答を得るのに適切な文書を提示する基本手順を示すフローチャートである。

【図3】教材の主文書の記載内容を上段に表示し取出された重要キーワードを下段に表示するディスプレイ画面30の具体例を示す図である。

【図4】教材の主文書の関連文書の記載内容を上段に表示し取出された重要キーワードを下段に表示するディスプレイ画面40の具体例を示す図である。

【図5】キーワードテーブルを示す図であり、(A)は図3に示す教材の主文書におけるキーワードテーブルを示す図であり、(B)は図4に示す教材の主文書の関連文書におけるキーワードテーブルを示す図である。

【図6】学習対象の文書に関する質問の回答を得るのに適切な重要キーワードを加味した文書を提示する手順を示すフローチャートである。40

【図7】学習者が学習対象の文書に関する文書を参照中に質問文を入力し、質問文の形態素解析を行っている所の画面を示す図である。

【図8】質問文から取出されたキーワードを用いて検索した検索結果を示す図である。

【図9】本発明により学習者が学習対象の文書に関する質問内容からその学習に最適な文書を取得している所の画面を示す図である。

【図10】学習者が入力した質問文から取出されたキーワードと質問文の入力時に開いていた教材の文書に対応する重要キーワードとを組合わせたキーワードを用いて検索した検索結果を示す図である。

【図11】図1に示す情報検索装置を用いて、複数の文書における各重要キーワードを取50

得する処理を示すフロー・チャートである。

【発明を実施するための形態】

【0019】

図1は本発明の情報検索装置のプロック構成図である。図1の一点鎖線の上側に示す本発明の情報検索装置1は、サーバであり、図1の一点鎖線の下側に示す、検索エンジンとしてのサーバ18および学習者側のパーソナルコンピュータ(学習者PC)19とネットワーク17を介して通信可能に接続され、学習者と情報検索装置1との間でWEBベースの学習を行う装置である。情報検索装置1は、相互にバスライン10を介して接続されるCPU11、学習者からの質問文の入力を受け付ける質問文受付手段であるマウスやキーボード等の入力装置12、メモリ13、ハードディスク14、ディスプレイ15およびネットワーク17を介してサーバ18および学習者PC19と通信する通信ユニット16を有する。  
10

【0020】

メモリ13には、CPU11を第1キーワード抽出手段131、第2キーワード抽出手段132、検索手段133および重要語定義テーブル生成手段134として機能させるためのプログラムが格納されている。

【0021】

通信ユニット16は、複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末(学習者PC)19に出力する出力手段としての機能も有する。

【0022】

第1キーワード抽出手段は、質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する。

第2キーワード抽出手段は、出力手段から出力された第1学習領域におけるキーワードのうち、質問文に関する情報の検索のために第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して重要キーワードを検索のための第2キーワードとして抽出する。  
20

【0023】

検索手段は、第1キーワードと第2キーワードの双方を用いて質問文に関する情報の検索を行う。

重要語定義テーブル生成手段は、教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを重要キーワードとして重要語定義テーブルを生成する。  
30

【0024】

ハードディスク14には、学習者が質問する学習対象の教材の文書を格納する教材用テキストファイル群141および文書を自立語(名詞、代名詞、動詞、形容詞、形容動詞、副詞、接続詞および感動詞)に分けるために用いる形態素解析辞書ファイル142が格納される。

【0025】

情報検索装置1は、通信ユニット16、ネットワーク17を介して学習対象となる主文書を含む複数の文書をサーバ18から受信し、教材用テキストファイル群141に格納する。サーバ18から受信した複数の文書は、受信後、直接メモリ13に書込んでもよい。本発明の実施形態では、学習対象となる主文書を含む複数の文書はサーバ18に格納されサーバ18から情報検索装置1に送信されるが、他の実施形態として複数の文書はサーバ18以外のコンピュータから送信された情報あるいは自身で作成した情報を情報検索装置1内に格納してもよい。  
40

【0026】

次いで、情報検索装置1は、形態素解析辞書ファイル142を用いて各文書をそれぞれ形態素解析し自立語を取り出し、各文書をそれぞれ解析し自立語の重要度を算出し、各文書における自立語の重要度の中から少なくとも一つの自立語を重要キーワードとして選択する。  
50

**【 0 0 2 7 】**

図2は学習対象の文書に関する質問の回答を得るのに適切な文書を提示する基本手順を示すフローチャートである。

ステップS1で、学習者PC19から通信ユニット16を介して入力された図3に示す学習対象の文書に関する質問文を読取る。

**【 0 0 2 8 】**

ステップS2で、ステップS1で入力した質問文を形態素解析し質問文内のキーワードを取出す。

ステップS3で、学習者が現在参照中の文書を形態素解析し該文書内のキーワードを取出す。現在参照中の文書は、学習者PC19と通信して得るか、質問文の読み取り時に読み取る。  
10

**【 0 0 2 9 】**

ステップS4で、ステップS2で質問文から取出したキーワードとステップS3で参照中の文書から取出したキーワードとを組合させて、所望の情報を得るためにサーチエンジンに入力するキーワードで構成される検索式を生成する。

ステップS5で、ステップS4で生成した検索式で検索を実行し、質問に対する回答を得るための文書情報(WEBページ)を学習者に出力する。

**【 0 0 3 0 】**

図3は教材の主文書の記載内容を上段に表示し取出された重要キーワードを下段に表示するディスプレイ画面30の具体例を示す図であり、図4は教材の主文書の関連文書の記載内容を上段に表示し取出された重要キーワードを下段に表示するディスプレイ画面40の具体例を示す図である。図3および図4の上段に示す文書は、法律入門の教材に関するものであり、学習者による学習者PC19の操作によりサーバ18から受信され情報検索装置1のハードディスク14内の教材用テキストファイル群141に格納される。主文書は知的財産権に関し、関連文書は独占禁止法に関する。各文書はそれぞれ別の記載内容を有する。以下、各文書中に………で示す部分は、記載内容を省略していることを示す。  
20

**【 0 0 3 1 】**

図3および図4において、上段に示すディスプレイ画面30および40は、下段左側に各文書を形態素解析して取出した自立語のうち名詞を抜き出した結果のキーワード(名詞)および各キーワードの重み(重要度)の算出結果を示し、下段右側に、算出したキーワードの重要度の中から選択した重要度の高い重要キーワードを示す。  
30

**【 0 0 3 2 】**

図3および図4は、それぞれ法律入門の教材であり、主文書および関連文書の記載内容と、形態素解析結果、すなわちキーワード(名詞、重み)および重要キーワードの取出し結果をそれぞれ示す。重みは、図3に示す文書31と図4に示す文書41とに出現するキーワード(ここでは名詞類)に対して、tf/idf値など統計処理で一般に用いられる重みの値を算出して用いる。重みの値が高いキーワードを重要キーワードとしてメモリ13に記憶する。

**【 0 0 3 3 】**

なお、図3および図4に示す重みは、主文書と関連文書とに出現するキーワード(ここでは名詞)に対して、tf/idf値など統計処理で一般に用いられる重みを算出して求めたものである。  
40

**【 0 0 3 4 】**

図5はキーワードテーブルを示す図であり、(A)は図3に示す教材の主文書におけるキーワードテーブルを示す図であり、(B)は図4に示す教材の主文書の関連文書におけるキーワードテーブルを示す図である。図5の(A)および(B)に示す例では、キーワードの重みが6.0以上のキーワードが重要キーワードに設定されている。この重要キーワードは、図中値1で示され、重要キーワードとみなされないものは値0で示される。それゆえ、このキーワードテーブルは重要語定義テーブルと称する。

**【 0 0 3 5 】**

この重要キーワードは、通信ユニット16から学習者PC19に出力された現在参照中の文書（第1学習領域）におけるキーワードのうち、質問文に関する情報の検索のために第1キーワードに加えるべき重要キーワードである。第2キーワード抽出手段は、重要語定義テーブルを参照して重要キーワードを検索のための第2キーワードとして抽出する。

#### 【0036】

重要語定義テーブル生成手段は、教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して重要度の値の高いキーワードを重要キーワードとして抽出可能にする重要語定義テーブルを生成する。

#### 【0037】

図6は学習対象の文書に関する質問の回答を得るのに適切な重要キーワードを加味した文書を提示する手順を示すフローチャートである。ここで、適切な文書とは所定の学習対象の文書に関する質問に対する回答を得るためのWEBページを意味する。

10

#### 【0038】

ステップS11で、学習者PC19から通信ユニット16を介して入力された図3に示す学習対象の文書に関する質問文を読取る。

ステップS12で、ステップS11で入力した質問文を形態素解析し質問文内のキーワードを取出す。

#### 【0039】

ステップS13で、ステップS12で取出したキーワードと学習対象の文書の重要キーワードであって、学習対象の文書に関連する学習者が現在参照中の図4に示す文書とを組合わせて、所望の情報を得るためにサーチエンジンに入力するキーワードで構成される検索式を生成する。現在参照中の文書は、学習者PC19と通信して得るか、質問文の読み取り時に読取る。

20

ステップS14で、ステップS13で生成した検索式で検索を実行し、質問に対する回答を得るための文書（WEBページ）を学習者に出力する。

#### 【0040】

図7は、学習者が学習対象の文書に関する文書を参照中に質問文を入力し、質問文の形態素解析を行っている所の画面70を示す図である。図7において、図6のステップS11で学習者が図4に示す文書を参照中に入力した質問文を図7の上段に示し、図6のステップS12で質問文の形態素解析した結果を図7の下段左側に示し、キーワードを取出した結果を図7の下段右側に示す。

30

#### 【0041】

学習者は図4に示す独占禁止法の文書を参照しており、独占禁止法の違反摘発事例を知りたいのであるが、質問文から取出されたキーワードは、「違反」、「摘発」および「事例」の3語だけである。これら3語を検索キーワードとして用いても、独占禁止法の違反または摘発の事例を検索することは困難である。

#### 【0042】

図8は質問文から取出されたキーワードを用いて検索した検索結果を示す図である。図8に示す検索結果の画面80は、質問文から取出されたキーワード、「違反」、「摘発」および「事例」の3語による検索結果は独占禁止法の違反または摘発の事例を含む文書情報（WEBページ）として不十分であることが示される。

40

#### 【0043】

図9は、本発明により学習者が学習対象の文書に関する質問内容からその学習に最適な文書を取得している所の画面90を示す図である。図9において、学習者が入力した質問文から取出されたキーワードを左上に示し、質問文の入力時に開いていた教材の文書、すなわち図4に示す文書に対応して予めメモリ13に格納された重要キーワードを右上に示し、質問文から取出されたキーワードと重要キーワードとを組合せた例を下段に示す。

#### 【0044】

図9に示されるように、図6のステップS13で、「違反」、「摘発」、「事例」、「メーカー」、「値段」、「競争」および「独占」をAND条件で検索式を生成することによ

50

り、独占禁止法の違反または摘発の事例を含む文書を検索結果の上位に得ることができる  
ことが判る。

【0045】

図10は学習者が入力した質問文から取出されたキーワードと質問文の入力時に開いていた教材の文書に対応する重要キーワードとを組合わせたキーワードを用いて検索した検索結果を示す図である。図10に示す検索結果の画面100は、独占禁止法の違反または摘発の事例を含む文書情報(WEBページ)として十分であることが示される。

【0046】

図11は図1に示す情報検索装置を用いて、複数の文書における各重要キーワードを取得する処理を示すフローチャートである。本フローチャートにおける処理は、学習者が質問文を入力しその回答を得る前に行う事前の準備段階における処理である。この処理を事前にを行うことで、質問文の入力後における処理の時間を短縮できる。  
10

【0047】

ステップS101で、教材を構成する複数の文書を登録する。すなわち、学習者PC19からの入力操作により学習対象となる主文書と主文書の関連文書を含む複数の文書を通信ユニット16を介して受信し、ハードディスク14内のテキストファイル群141に格納する。主文書は質問文から抽出されたキーワードを元に検索された文書であることが好みしい。また、主文書の関連文書は質問文から抽出されたキーワードを元に検索された文書であることが好みしい。

【0048】

ステップS102で、ステップS101で格納した各文書を形態素解析しキーワードを取出す。  
20

ステップS310で、キーワードテーブルを参照する。図5を用いて先に説明したキーワードテーブル(重要語定義テーブル)は、キーワード(本実施形態では名詞のキーワードと各キーワードの重み)および各キーワードが重要キーワードであるか否かの判定値を一覧表示するものである。

【0049】

キーワードテーブルの作成に際し、各文書に出現するキーワードに対しそれぞれtf/idf値など統計処理し、キーワードの重み(重要度)を算出する。より詳しくは、主文書と主文書の関連文書とを比較し、出現する名詞に対し重みを示す尺度である情報利得(Information Gain)やtf-idf値を計算する。情報利得やtf-idf値は、他の文書には出現せずにその文書にだけ出現する語(キーワード)ほど大きい。tf-idf値とは、キーワード抽出対象の文書から文書を代表するキーワードを抽出したときのキーワードの出現頻度を示す指標である。本実施形態では、一般的に用いられるtf-idf値を算出して重要度を示す指標とする。  
30

【0050】

ステップS104で、文書毎に重要度の高いキーワードを選択し、重要キーワードとしてメモリ13に保存する。

【0051】

以上説明した本発明の実施形態は、図2に示すフローチャートのステップS1～S5の処理および図6に示すフローチャートのステップS11～S14の処理を実行する手順を例に挙げて説明したが、他の実施形態として、例えば図11に示すフローチャートのステップS101～S104の処理を予め実行しておくことにより、ステップS101～S104の処理を予め実行するので事前の準備は必要であるが、学習対象の文書の形態素解析およびキーワードの抽出が事前に実行されるので、図6に示すフローチャートのステップS13の処理時間を短縮できる。  
40

【0052】

(付記1)

学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索プログラム  
50

であって、

コンピュータ装置を、

複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力手段と、

学習者からの質問文の入力を受け付ける質問文受付手段と、

前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出手段と、

前記出力手段から出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出手段と、

前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索手段と、

して機能させることを特徴とする情報検索プログラム。

#### 【0053】

(付記2)

前記コンピュータ装置を、

前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成手段、

として機能させることを特徴とする付記1に記載の情報検索プログラム。

#### 【0054】

(付記3)

学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索方法であつて、

コンピュータ装置が、

複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力ステップと、

学習者からの質問文の入力を受け付ける質問文受付ステップと、

前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出ステップと、

前記出力ステップで出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出ステップと、

前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索ステップと、

を行うことを特徴とする情報検索方法。

#### 【0055】

(付記4)

前記コンピュータ装置が、

前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成ステップ、

を行うこと特徴とする付記3に記載の情報検索方法。

#### 【0056】

(付記5)

学習者側のコンピュータから入力された質問文に応じて該質問文に関する情報を検索するためのキーワードを生成し該キーワードにより前記情報を検索する情報検索装置であつ

10

20

30

40

50

て、

コンピュータ装置が、

複数に区分された学習領域から構成される教材から抽出した第1学習領域を学習者端末に出力する出力手段と、

学習者からの質問文の入力を受け付ける質問文受付手段と、

前記質問文受付手段で受け付けた質問文を形態素解析して検索のための第1キーワードを抽出する第1キーワード抽出手段と、

前記出力ステップで出力された前記第1学習領域におけるキーワードのうち、前記質問文に関する情報の検索のために前記第1キーワードに加えるべき重要キーワードを格納した重要語定義テーブルを参照して該重要キーワードを検索のための第2キーワードとして抽出する第2キーワード抽出手段と、

前記第1キーワードと前記第2キーワードの双方を用いて前記質問文に関する情報の検索を行う検索手段と、

を備えたことを特徴とする情報検索装置。

【0057】

(付記6)

前記コンピュータ装置が、

前記教材を構成する複数の学習領域の各々に対し形態素解析を行い、出現頻度から重要度を判断して抽出したキーワードを前記重要キーワードとして前記重要語定義テーブルを生成する重要語定義テーブル生成手段、

を備えたこと特徴とする付記5に記載の情報検索装置。

【0058】

(付記7)

種々の情報を格納するサーバに接続され、該サーバから取得した所定の第1情報に関する質問文から質問に対する回答を得るための情報を検索する情報検索装置であって、

学習者側のコンピュータから受信した前記質問文を読み取り、形態素解析して第1キーワードを抽出する第1抽出手段と、

前記第1情報および該第1情報に関する第2情報の各文書をそれぞれ読み取り、形態素解析して第2キーワードを抽出する第2抽出手段と、

前記各文書に出現する前記第2キーワードについて重み付け処理を行い、各第2キーワードの重要度を示す重み値を求め、該重み値から重要度の高いキーワードを選択し、選択したキーワードを前記各文書に対する重要キーワードにそれぞれ設定する設定手段と、

前記第1キーワードと前記第2情報の文書に対する重要キーワードの組合せから検索式を生成する検索式生成手段と、

を備えることを特徴とする情報検索装置。

【0059】

(付記8)

前記第2情報の文書は、前記第1キーワードを元に検索された文書である、  
付記7に記載の情報検索装置。

(付記9)

前記第1情報の文書は、前記第1キーワードを元に検索された文書である、  
付記7に記載の情報検索装置。

【0060】

(付記10)

種々の情報を格納するサーバから取得した所定の第1情報に関する質問文から質問に対する回答を得るための情報を検索する情報検索方法であって、

学習者側のコンピュータから受信した前記質問文を読み取り、形態素解析して第1キーワードを抽出する第1抽出ステップと、

前記第1情報および該第1情報に関する第2情報の各文書をそれぞれ読み取り、形態素解析して第2キーワードを抽出する第2抽出ステップと、

10

20

30

40

50

前記各文書に出現する前記第2キーワードについて重み付け処理を行い、各第2キーワードの重要度を示す重み値を求め、該重み値から重要度の高いキーワードを選択し、選択したキーワードを前記各文書に対する重要キーワードにそれぞれ設定する設定ステップと、

前記第1キーワードと前記第2情報の文書に対する重要キーワードの組合せから検索式を生成する検索式生成ステップと、  
備えたことを特徴とする情報検索方法。

#### 【0061】

(付記11)

種々の情報を格納するサーバから取得した所定の第1情報に関する質問文から質問に対する回答を得るために情報を検索するための情報検索プログラムであって、

コンピュータに、

学習者側のコンピュータから受信した前記質問文を読み取り、形態素解析して第1キーワードを抽出する第1抽出ステップと、

前記第1情報および該第1情報に関する第2情報の各文書をそれぞれ読み取り、形態素解析して第2キーワードを抽出する第2抽出ステップと、

前記各文書に出現する前記第2キーワードについて重み付け処理を行い、各第2キーワードの重要度を示す重み値を求め、該重み値から重要度の高いキーワードを選択し、選択したキーワードを前記各文書に対する重要キーワードにそれぞれ設定する設定ステップと、

前記第1キーワードと前記第2情報の文書に対する重要キーワードの組合せから検索式を生成する検索式生成ステップと、

を実行させることを特徴とする情報検索プログラム。

#### 【符号の説明】

#### 【0062】

1 情報検索装置

10 バスライン

11 C P U

12 入力装置

13 メモリ

14 ハードディスク

15 ディスプレイ

16 通信ユニット

17 ネットワーク

18 サーバ

30、40、70、80、90、100 ディスプレイ画面(画面)

131 第1キーワード抽出手段

132 第2キーワード抽出手段

133 検索手段

134 重要語定義テーブル生成手段

141 教材用テキストファイル群

142 形態素解析辞書ファイル

10

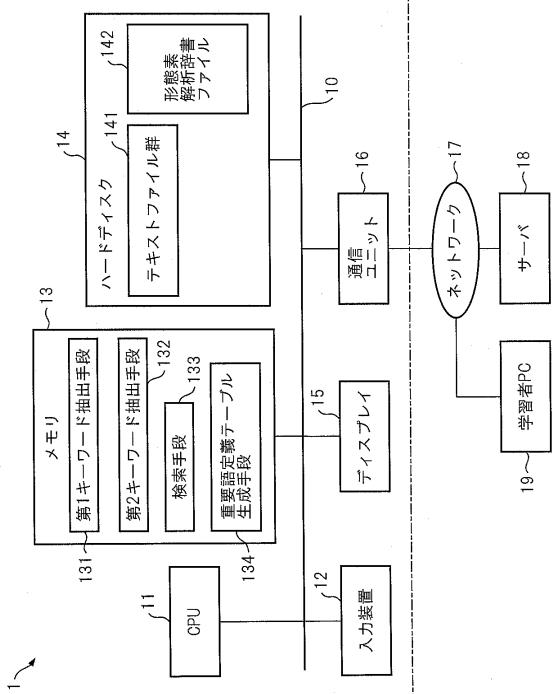
20

30

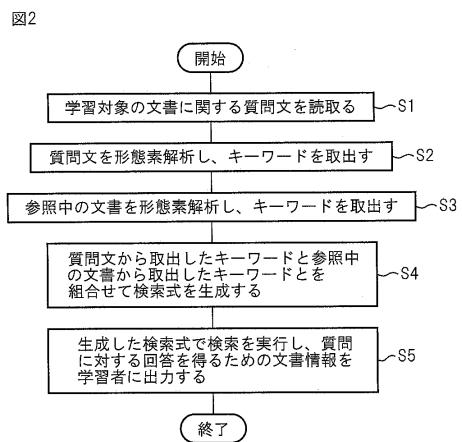
40

【図1】

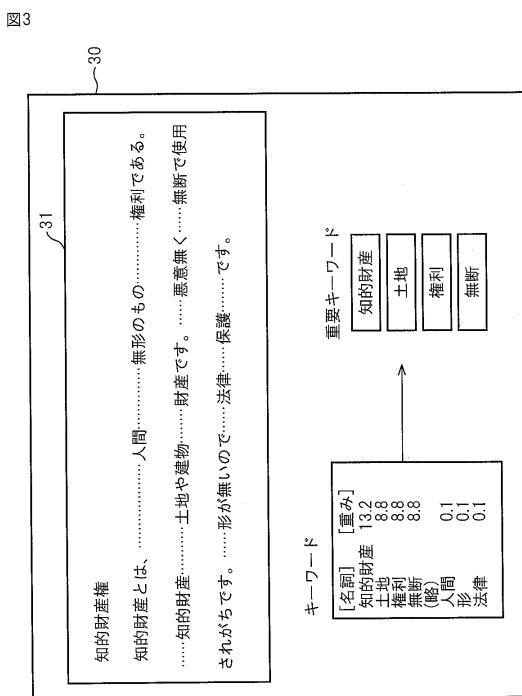
図1



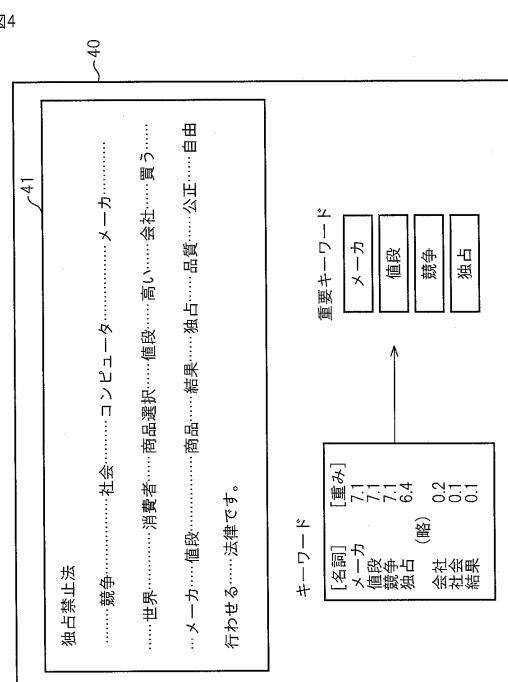
【図2】



【図3】



【図4】



【 四 5 】

【 四 6 】

5

(A)

キーワード		重要キーワード
名詞	重み	
知的財産	13.2	1
土地	8.8	1
権利	8.8	1
無断	8.8	1
⋮	⋮	⋮
人間	0.1	0
形	0.1	0
法律	0.1	0

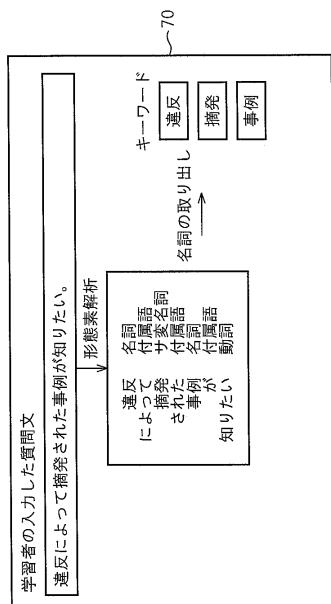
(B)

キーワード		重要キーワード
名詞	重み	
メーカー	7.1	1
値段	7.1	1
競争	7.1	1
独占	6.4	1
:	:	:
会社	0.2	0
社会	0.1	0
結果	0.1	0

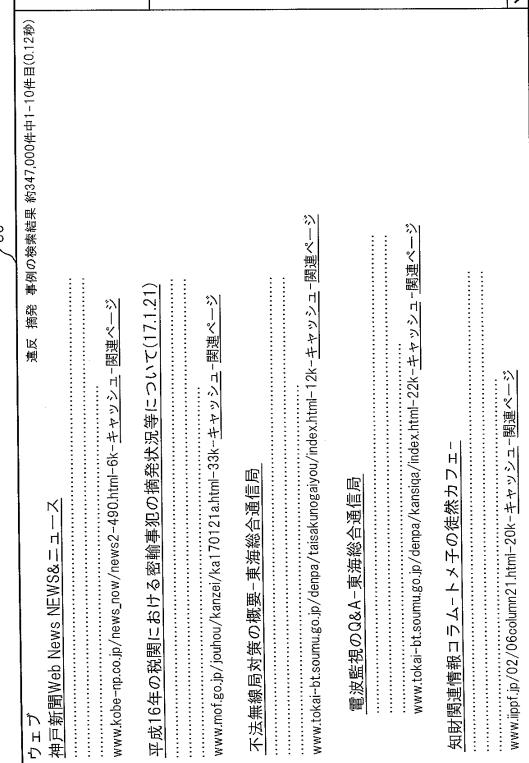
( 7 )

( 8 )

図7

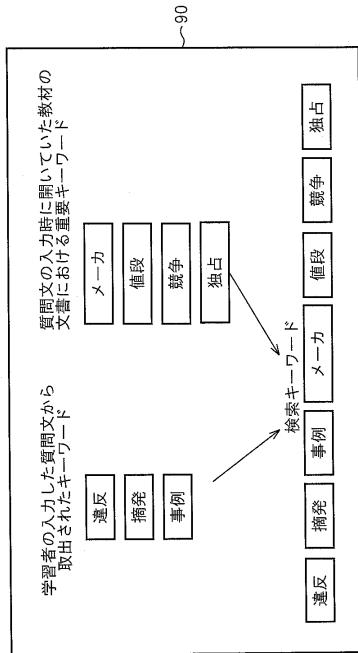


四〇

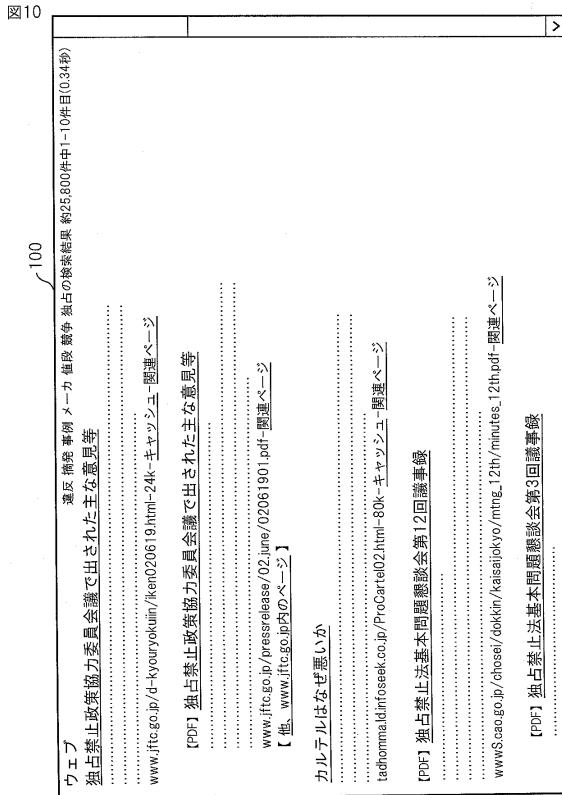


【図9】

図9

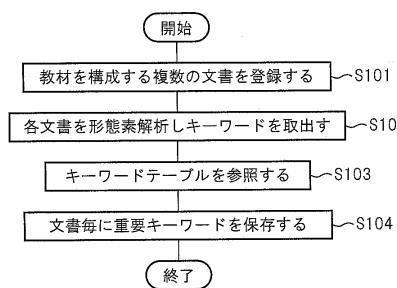


【図10】



【図11】

図11



---

フロントページの続き

審査官 小太刀 慶明

(56)参考文献 特開2005-301855(JP,A)

米国特許出願公開第2006/0026013(US,A1)

特開平10-074210(JP,A)

特開2004-348768(JP,A)

河重 貴洋 外3名，“質問修正と再ランキングを用いた文脈依存Web検索”，DEWS20

06論文集 [online]，日本，電子情報通信学会データ工学研究専門委員会，2006

年 6月30日，p. 1-7，Internet<URL:<http://www.ieice.org/iss/de/DEWS/DEWS2006/3C-ii4.pdf>>

(58)調査した分野(Int.Cl., DB名)

G06F 17/30