

(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2016-134124

(P2016-134124A)

(43) 公開日 平成28年7月25日(2016.7.25)

(51) Int.Cl.

G06F 17/30 (2006.01)

F I

G06F 17/30 210A

G06F 17/30 170A

テーマコード (参考)

審査請求 未請求 請求項の数 18 O L (全 22 頁)

(21) 出願番号 特願2015-10177 (P2015-10177)
 (22) 出願日 平成27年1月22日 (2015.1.22)

(71) 出願人 314012076
 パナソニックIPマネジメント株式会社
 大阪府大阪市中央区城見2丁目1番61号
 (74) 代理人 100067828
 弁理士 小谷 悦司
 (74) 代理人 100115381
 弁理士 小谷 昌崇
 (74) 代理人 100118049
 弁理士 西谷 浩治
 (72) 発明者 藤原 菜々美
 大阪府門真市大字門真1006番地 パナ
 ソニック株式会社内
 (72) 発明者 堀井 則彰
 大阪府門真市大字門真1006番地 パナ
 ソニック株式会社内

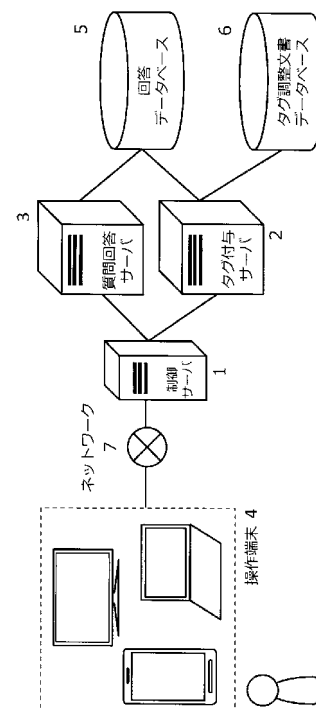
(54) 【発明の名称】 タグ付与方法、タグ付与装置、プログラム及び質問回答検索方法

(57) 【要約】

【課題】 ユーザの質問に関する回答文書を検索する際に、検索を行うユーザまたは検索の対象に関する所定の基準に適した回答文書の検索を行うためのタグを付与することを可能とするタグ付与方法を提供すること。

【解決手段】 本発明に係るタグ付与方法は、ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するタグ付与装置が行うタグ付与方法であって、前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択し、前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するタグ付与装置が行うタグ付与方法であって、

前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択し、

前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、

前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与する、
タグ付与方法。

10

【請求項 2】

前記所定の基準は、検索の対象に対するユーザの習熟度または知識レベルが高いかあるいは低いかである、

請求項 1 記載のタグ付与方法。

【請求項 3】

前記検索の対象は製品であり、前記回答文書は当該製品の取り扱いを説明する文書の少なくとも一部である、

請求項 2 記載のタグ付与方法。

【請求項 4】

前記検索の対象に対する習熟度または知識レベルが低いユーザによる検索に用いるタグを付与するための前記タグ調整文書として、一般知識に関する文書を選択する、

請求項 3 記載のタグ付与方法。

20

【請求項 5】

前記一般知識に関する文書とは、辞典、教科書、ニュース、ブログ、または短文投稿サービスのうち少なくとも 1 つである、

請求項 4 記載のタグ付与方法。

【請求項 6】

前記検索の対象に対する習熟度または知識レベルが高いユーザによる検索に用いるタグを付与するための前記タグ調整文書として、専門知識に関する文書を選択する、

請求項 3 記載のタグ付与方法。

30

【請求項 7】

前記専門知識に関する文書とは、工事取り付け専門書、取扱説明書、学術専門書、社内文書、または営業情報のうち少なくとも 1 つである、

請求項 6 記載のタグ付与方法。

【請求項 8】

前記検索の対象は製品であり、前記所定の基準は、当該製品の新機能に関する情報であるか旧機能に関する情報であるかのいずれか一方である、

請求項 1 記載のタグ付与方法。

【請求項 9】

前記回答文書は、前記製品の機能を説明する文書の少なくとも一部であり、前記タグ調整文書は、当該製品より新機種または旧機種の機能を説明する文書の少なくとも一部である、

請求項 8 記載のタグ付与方法。

40

【請求項 10】

前記検索の対象は製品であり、前記所定の基準は、当該製品の上位機種に関する情報であるか下位機種に関する情報であるかのいずれか一方である、

請求項 1 記載のタグ付与方法。

【請求項 11】

前記回答文書は、前記製品の機能を説明する文書の少なくとも一部であり、前記タグ調整文書は、当該製品より上位機種または下位機種の機能を説明する文書の少なくとも一部

50

である、

請求項 10 記載のタグ付与方法。

【請求項 12】

前記タグとして付与するキーワードの抽出は、文書を構成するトピックを自動で抽出するためのトピックモデルを用い、各文書の特徴を抽出することにより行う、

請求項 1 記載のタグ付与方法。

【請求項 13】

前記トピックモデルは、LDA (Latent Dirichlet Allocation)、LSI (Latent Semantic Indexing)、またはPLSA (Probabilistic Latent Semantic Analysis) の少なくとも 1 つである、

請求項 12 記載のタグ付与方法。

【請求項 14】

前記回答文書への前記タグの付与は、前記トピックモデルから得られた複数のタグ群の中で、最も多く回答文書を含むタグ群を選択することにより行う、

請求項 12 記載のタグ付与方法。

【請求項 15】

前記タグ調整文書の選択は、前記ユーザの行動履歴に基づいて行う、

請求項 1 記載のタグ付与方法。

【請求項 16】

検索対象の文書群からユーザの質問に適合する回答文書を検索する質問回答システムが行う質問検索方法であって、

前記検索を行うユーザまたは前記検索の対象に関する所定の基準を判定し、

前記所定の基準に基づき、前記回答文書に対してタグとして付与するトピックを調整するためのタグ調整文書を選択し、

前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、

前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与し、

前記ユーザの質問からキーワードを抽出し、当該キーワードと合致するタグが付与された前記回答文書を検索する、

質問回答検索方法。

【請求項 17】

ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するタグ付与装置であって、

前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択する文書選択手段と、

前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出するトピック抽出手段と、

前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与するタグ付与手段と、

を備えるタグ付与装置。

【請求項 18】

コンピュータを、ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するタグを付与するためのタグ付与装置として機能させるためのプログラムであって、

前記コンピュータに、

前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択し、

前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択さ

10

20

30

40

50

れたタグ調整文書を表わすトピックを抽出し、

前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与する、
ことを実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するためのタグ付与方法に関する。

【背景技術】

【0002】

近年、FAQ (Frequently Asked Questions) などに基づく質問回答システムが開発されている。質問回答システムとは、特定の分野や種類に対する質問をユーザが自然文で入力し、それに対してシステムが適当な回答を返すコンピュータシステムのことをいう。一般に、質問回答システムにおいて、回答対象となる文書群をデータベースに記憶しておく。

【0003】

このような質問回答システムにおいては、膨大な回答文書の中から、ユーザの意図に適した回答を適切に検索することが必要となる。これを実現するための一手段として、質問キーワードに関する多数のコメント文章を潜在的なトピックグループに分類した上で、トピックグループに含まれる文章群と類似する回答文を対応付ける技術が開示されている（例えば、特許文献1参照）。

【0004】

また、膨大な回答文書の中から、ユーザの意図に適した回答を適切に検索するための他の手段として、回答文書に予め検索タグを付与しておくことがある。検索タグとは回答文書の内容を適切に表す語のことであり、ユーザの質問文の中に検索タグが含まれていれば、該当する回答をユーザに提示する。これにより、膨大な文書群に対する検索精度を向上させる。

【0005】

検索タグの付与において、膨大な文書群のタグ付けを人手で行うには多くのコストがかかるため、文書の特徴付ける語を自動的に抽出するような、自動検索タグ付与も検討されている（例えば、非特許文献1参照）。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2013-143066号公報

【非特許文献】

【0007】

【非特許文献1】加藤 亮、吉川大弘、古橋武「潜在的ディリクレ配分法を利用した文書への自動タグ付与に関する検討」、人工知能学会全国大会論文集 28、1-4、人工知能学会 2014年5月1日

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかし、自動検索タグ付与においては、質問の対象（すなわち、回答の検索対象）に対するユーザの知識レベルや、回答の検索対象に応じて適切なタグ付与ができるよう、さらなる改良が求められている。

【課題を解決するための手段】

【0009】

上記課題を解決するために、本発明に係るタグ付与方法は、ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するタグ付与装置が

10

20

30

40

50

行うタグ付与方法であって、前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択し、前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与する。

【発明の効果】

【0010】

本構成によれば、回答文書に対して、質問するユーザや、回答の検索対象（例えば、家電製品の機能に関する質問における家電製品）に関する所定の基準（例えば、家電製品に対するユーザの知識レベルや、家電製品の機種に搭載された機能であるか旧機種に搭載された機能であるか）に応じた検索タグを付与することが可能となる。その結果、ユーザからの質問に対して、質問するユーザの知識レベルや、回答の検索対象に応じて適切な回答文書を検索することが可能となる。

10

【図面の簡単な説明】

【0011】

【図1】本発明の実施の形態に係る質問回答システムの全体構成の一例を示す図である。

【図2】本発明の実施の形態に係る制御サーバ1の構成の一例を示すブロック図である。

【図3】本発明の実施の形態に係るタグ付与サーバ2の構成の一例を示すブロック図である。

【図4】本発明の実施の形態に係る質問回答サーバ3の構成の一例を示すブロック図である。

20

【図5】本発明の実施の形態に係る操作端末4の構成の一例を示すブロック図である。

【図6】本発明の実施の形態に係る質問回答システムの一連の処理の一例を示すシーケンス図である。

【図7】本発明の実施の形態に係るサーバ選択処理の手順の一例を示すフローチャートである。

【図8】本発明の実施の形態に係る文書選択処理の手順の一例を示すフローチャートである。

【図9】本発明の実施の形態に係るタグ抽出／紐付け処理の手順の一例を示すフローチャートである。

30

【図10】本発明の実施の形態に係る本実施形態によるタグ調整文書決定テーブルの一例を示す図である。

【図11A】本発明の実施の形態に係る表示画面の一例を示す図である。

【図11B】本発明の実施の形態に係る表示画面の他の例を示す図である。

【図12】本発明の実施の形態に係る回答データベース5のデータ構造の一例を示す図である。

【図13】本発明の実施の形態に係るタグ調整文書データベース6のデータ構造の一例を示す図である。

【図14】本発明の実施の形態に係るトピックモデルによるタグ付与の実験結果の一例を示す図である。

40

【発明を実施するための形態】

【0012】

（本発明の基礎となった知見）

本発明者は、タグ付けされた回答文書を検索することによりユーザの質問に回答する質問回答システムについて詳細に検討した。その結果、前述した従来技術では、ある質問に対する回答を検索する際に、質問の対象（すなわち、回答の検索対象）に対するユーザの知識レベルや回答の検索対象に応じて適切な回答文が返答されない場合があるという課題を見出した。この点について、検索対象に対するユーザの知識レベル（習熟度）の違いを例として説明する。

【0013】

50

例えば、ある回答文書に「アクセス、LAN、プロキシサーバ」と検索タグがついていた場合、ネットワーク技術に関する知識レベルの高いユーザが質問しうる「プロキシサーバを設定したい」という質問に対しては、質問中に含まれる「プロキシサーバ」という言葉が検索タグの一つと一致するため、この回答文書を検索し、回答文書の内容に基づく返答をすることができるのに対し、ネットワーク技術に関する知識レベルの低いユーザが質問しうる「インターネットの接続方法が知りたい」という質問に対しては、質問中に含まれるいずれの言葉も検索タグと一致しないため、この回答文書を検索することができず、ひいてはこの回答文書の内容に基づく返答をすることができないという課題が生じる。

【0014】

このような課題を解決するために、本発明に係るタグ付与方法は、ユーザからの質問に対する回答を構成する回答文書に対して、回答を検索するためのタグを付与するタグ付与装置が行うタグ付与方法であって、前記回答文書に対してタグとして付与するトピックを、前記検索を行うユーザまたは前記検索の対象に関する所定の基準により調整するためのタグ調整文書を選択し、前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与する。

【0015】

これにより、回答文書に対して、質問の対象（すなわち、回答の検索対象）に対するユーザの知識レベルや回答の検索対象に関する所定の基準に応じて適切な検索タグを付与することが可能となる。

【0016】

また、前記所定の基準は、検索の対象に対するユーザの習熟度または知識レベルが高いかあるいは低いかであることとしてもよい。

【0017】

これにより、回答文書に対して、検索対象に対するユーザの習熟度または知識レベルを基準として適切な検索タグを付与することが可能となる。

【0018】

また、前記検索の対象は製品であり、前記回答文書は当該製品の取り扱いを説明する文書の少なくとも一部であることとしてもよい。

【0019】

これにより、製品の取り扱いに関して、ユーザの習熟度または知識レベルを基準として適切な検索タグを付与することが可能となる。

【0020】

また、前記検索の対象に対する習熟度または知識レベルが低いユーザによる検索に用いるタグを付与するための前記タグ調整文書として、一般知識に関する文書を選択することとしてもよい。

【0021】

また、前記一般知識に関する文書とは、辞典、教科書、ニュース、ブログ、または短文投稿サービスのうち少なくとも1つであることとしてもよい。

【0022】

これにより、回答文書に対して、電子機器の機能を説明する文書よりも平易な言葉が記述された文書を用いて、知識レベルが低いユーザに適したタグを付与することが可能となる。

【0023】

また、前記検索の対象に対する習熟度または知識レベルが高いユーザによる検索に用いるタグを付与するための前記タグ調整文書として、専門知識に関する文書を選択することとしてもよい。

【0024】

また、前記専門知識に関する文書とは、工事取り付け専門書、取扱説明書、学術専門書、社内文書、または営業情報のうち少なくとも1つであることとしてもよい。

10

20

30

40

50

【0025】

これにより、回答文書に対して、電子機器の機能を説明する文書よりも高度な言葉が記述された文書を用いて、知識レベルが低いユーザに適したタグを付与することが可能となる。

【0026】

また、前記検索の対象は製品であり、前記所定の基準は、当該製品の新機能に関する情報であるか旧機能に関する情報であるかのいずれか一方であることとしてもよい。

【0027】

これにより、製品の機能に関する回答文書に対して、製品の新機能に関する情報であるか旧機能に関する情報であるかのいずれかを基準として適切な検索タグを付与することが可能となる。

10

【0028】

また、前記回答文書は、前記製品の機能を説明する文書の少なくとも一部であり、前記タグ調整文書は、当該製品より新機種または旧機種の機能を説明する文書の少なくとも一部であることとしてもよい。

【0029】

これにより、新機種に搭載された機能に関する文書、あるいは旧機種に搭載されていた機能に関する文書をもとに、製品の新機能に関する情報であるか旧機能に関する情報であるかのいずれかを基準として適切な検索タグを付与することが可能となる。

【0030】

また、前記検索の対象は製品であり、前記所定の基準は、当該製品の上位機種に関する情報であるか下位機種に関する情報であるかのいずれか一方であることとしてもよい。

20

【0031】

これにより、製品の機能に関する回答文書に対して、製品の上位機種に関する情報であるか下位機種に関する情報であるかのいずれかを基準として適切な検索タグを付与することが可能となる。

【0032】

また、前記回答文書は、前記製品の機能を説明する文書の少なくとも一部であり、前記タグ調整文書は、当該製品より上位機種または下位機種の機能を説明する文書の少なくとも一部であることとしてもよい。

30

【0033】

これにより、上位機種に搭載された機能に関する文書、あるいは下位機種に搭載されていた機能に関する文書をもとに、製品の上位機種に関する情報であるか下位機種に関する情報であるかのいずれかを基準として適切な検索タグを付与することが可能となる。

【0034】

また、前記タグとして付与するキーワードの抽出は、文書を構成するトピックを自動で抽出するためのトピックモデルを用い、各文書の特徴を抽出することにより行うこととしてもよい。

【0035】

また、前記トピックモデルは、LDA (Latent Dirichlet Allocation)、LSI (Latent Semantic Indexing)、またはPLSA (Probabilistic Latent Semantic Analysis) の少なくとも1つであることとしてもよい。

40

【0036】

これにより、トピックモデルを用いて、回答文書およびタグ調整文書から、検索タグとして用いるキーワードを抽出することが可能となる。

【0037】

また、前記回答文書への前記タグの付与は、前記トピックモデルから得られた複数のタグ群の中で、最も多く回答文書を含むタグ群を選択することにより行うこととしてもよい。

50

【0038】

これにより、回答文書とタグ調整文書から抽出されたキーワードのうち、最も適したものを検索タグとして付与することが可能となる。

【0039】

また、前記タグ調整文書の選択は、前記ユーザの行動履歴に基づいて行うこととしてもよい。

【0040】

これにより、ユーザの行動履歴をもとにユーザの知識レベル等を判定し、知識レベルに応じたタグ調整文書を選択することが可能となる。

【0041】

また、本発明に係る質問検索方法は、検索対象の文書群からユーザの質問に適合する回答文書を検索する質問回答システムが行う質問検索方法であって、前記検索を行うユーザまたは前記検索の対象に関する所定の基準を判定し、前記所定の基準に基づき、前記回答文書に対してタグとして付与するトピックを調整するためのタグ調整文書を選択し、前記回答文書および前記選択されたタグ調整文書から、当該回答文書および当該選択されたタグ調整文書を表わすトピックを抽出し、前記抽出されたトピックを、所定の基準のタグとして前記回答文書に付与し、前記ユーザの質問からキーワードを抽出し、当該キーワードと合致するタグが付与された前記回答文書を検索する。

【0042】

これにより、回答文書に対して、質問の対象（すなわち、回答の検索対象）に対するユーザの知識レベルや回答の検索対象に関する所定の基準に応じて適切な検索タグを付与した上で、ユーザの質問に適した回答文書を検索することが可能となる。

【0043】

なお、以下で説明する実施の形態は、いずれも本発明の一具体例を示すものである。以下の実施の形態で示される数値、形状、構成要素、ステップ、ステップの順序などは、一例であり、本発明を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、最上位概念を示す独立請求項に記載されていない構成要素については、任意の構成要素として説明される。また全ての実施の形態において、各々の内容を組み合わせることも出来る。

【0044】

（実施の形態）

以下、本発明の実施形態について、図面を参照しながら説明する。なお、本発明の実施の形態は、下記の実施例になんら限定されること無く、本発明の技術的範囲に属する限り、種々の形態をとり得る。

【0045】

図1は、本実施の形態における質問回答システムの全体構成の一例を示す図である。質問回答システムは、制御サーバ1、タグ付与サーバ2、質問回答サーバ3、操作端末4、回答データベース5、タグ調整文書データベース6、ネットワーク7を備える。

【0046】

制御サーバ1は、タグ付与サーバ2と質問回答サーバ3と操作端末4とデータの送受信を行う機能を有する。例えば、操作端末4から得られた入力に対して、タグ付与サーバ2または質問回答サーバ3のうち、処理を行うサーバを選択する。制御サーバ1の詳細な構成は後述する。

【0047】

タグ付与サーバ2は、制御サーバ1と回答データベース5とタグ調整文書データベース6の各々とデータの送受信を行う機能を有する。例えば、操作端末4から入力されたユーザの検索対象に対する知識レベルの情報を受信し、タグ調整文書データベース6から、ユーザの知識レベルを加味したタグ調整文書を選択するなどの処理を行う。ここでの検索対象とは、例えば、製品、システムまたはアプリケーションプログラムの使い方、故障に関する情報などである。また、検索対象は、サービスや金融商品、文書等を含め、有形無形

10

20

30

40

50

の商材であってもよい。ここで、製品とは、家庭用電気機器、情報通信機器、自動車、家具をはじめとして、取り扱いに説明を要する物品を指す。また、知識レベルとは、その検索対象に対して、ユーザが持っている知識の量、習熟度などを表すものである。ここに挙げたもの以外でも、ユーザの質問に対しての回答が用意できるものであればこれに限らない。以降、「ユーザの検索対象に対する知識レベル」は、「知識レベル」と表記することとする。タグ付与サーバ2の詳細な構成は後述する。

【0048】

質問回答サーバ3は、制御サーバ1と回答データベース5とデータの送受信を行う機能を有する。例えば、質問回答サーバ3は、制御サーバ1から入力されたユーザの質問を受信し、回答データベース5の内容を検索するなどの処理を行う。質問回答サーバ3の詳細な構成は後述する。

10

【0049】

操作端末4は、スマートフォンやタブレット端末、専用表示機器端末、パーソナルコンピュータ（PC）、又はテレビなどである。ここに挙げたもの以外でも、ユーザとデータのやりとりができる端末であれば何でもよい。端末の操作は、タッチパネルによる入力、画像や動画による入力、音声による入力、キーボードによる入力などが想定される。

【0050】

回答データベース5は、タグ付与サーバ2および質問回答サーバ3とデータのやりとりを行う。ここには質問に対する回答を構成する文書である回答文書が保持される。質問対象としては、例えば、製品、システムまたはアプリケーションプログラムの使い方や故障相談などが挙げられ、QAを提示することでユーザの質問に回答する。なお、これは一例であり、質問とそれに対する回答が用意できるものであれば、これに限らない。また、データの保存形式の一例として、XML（Extensible Markup Language）形式などが挙げられるが、データを構造的に保持できるものであれば、これに限らない。図12はXML形式の一例であるが、回答となる文書とともに自動的に抽出されたタグ保存されている。

20

【0051】

なお、以後の説明において、回答文書のことをタグ付け対象文書と称することもあるが、同じものを意味する。

【0052】

タグ調整文書データベース6は、タグ付与サーバ2とのデータのやりとりを行う。ここには、タグ付け対象文書にタグを付ける際、タグの語を所定の基準により調整するための文書であるタグ調整文書が保持される。ここで、所定の基準とは、質問を行う（すなわち、質問に対する回答を検索する）ユーザや質問対象（すなわち、回答の検索対象）に関する基準をいう。所定の基準がユーザの知識レベルである場合、タグ調整文書として、例えば、一般知識（辞典、教科書、ニュース、ブログ（Weblogの略）、Twitter（登録商標）のような短文投稿サービスなど）を有する文書、また、高度な専門知識（工事取り付け専門書、取扱説明書、学術専門書、社内文書、営業情報など）を有する文書が保持される。なお、辞典とは、例えばWikipediaのようなインターネット上の辞典や、書籍としての辞典の両者を指すものである。これらは一例であり、文書であればこれに限らない。また、データの保存形式の一例として、XML（Extensible Markup Language）形式などが挙げられるが、データを構造的に保持できるものであれば、これに限らない。

30

40

【0053】

ネットワーク7は、制御サーバ1、操作端末4が接続される。接続方法の一例として、有線、無線によるLAN接続などが挙げられるが、各構成要素を通信可能に接続するものであれば、これに限らない。

【0054】

図2は、本実施の形態における制御サーバ1の構成の一例を示すブロック図である。

【0055】

50

図2に示されるように、制御サーバ1は、入力検出部101、処理選択部102、制御部103、通信部104を備える。

【0056】

入力検出部101は、操作端末4を通じてユーザからの入力を検出する。

【0057】

処理選択部102は、ユーザの入力に対し、タグ付与サーバ2または質問回答サーバ3のうち、処理を行うサーバを選択する部分である。例えば、「知識レベル：1」といった、基準を意味する情報を取得した場合、タグ付与サーバ2での処理を選択し、「エアコンの効きが悪い」といった、質問を意味する情報を取得した場合、質問回答サーバ3での処理を選択する。

【0058】

制御部103は、処理選択部102で選択されたサーバに処理を行うコマンドを送信する。

【0059】

通信部104は、タグ付与サーバ2、質問回答サーバ3、および操作端末4とそれぞれ通信する。

【0060】

図3は、本実施の形態におけるタグ付与サーバ2の構成の一例を示すブロック図である。

【0061】

図3に示されるように、タグ付与サーバ2は、基準取得部201、文書選択部202、形態素解析部203、ベクトル生成部204、タグ自動生成部205、タグ紐付け部206、通信部207、形態素解析器208、回答データベース5、タグ調整文書データベース6を備える。

【0062】

基準取得部201は、例えばユーザの知識レベルといったような、タグ調整文書を選択するための基準となる情報を取得する。基準となる情報は、質問を行うユーザの入力により取得してもよく、自動で取得してもよい。例えば、基準となる情報がユーザの知識レベルである場合、ユーザの質問履歴などからユーザの知識レベルを自動で判定して取得することができる。

【0063】

文書選択部202は、タグ調整文書データベース6から、基準となる情報に適したタグ調整文書を、回答データベース5からタグ付け対象文書を選択する。例えば、ユーザの知識レベルを基準とする場合、図10に示すような、ユーザの知識レベルとタグ調整文書の難易度を対応させるテーブルを参照し、タグ調整文書を選択する。なお、図10に示した数値やタグ調整文書は一例であり、これに限らない。また、これらのタグ調整文書の決定方法は一例であり、これに限らない。たとえば、なんらかの数式に基づいてタグ調整文書を選択しても構わないし、機械学習などのアルゴリズムを用いて統計的に選択するようにしてもよい。

【0064】

形態素解析部203は、文書選択部202から文書を受け取り、形態素解析器208を介して文書の形態素解析を行う部分である。形態素解析とは、文書を意味のある単語に区切り、辞書を利用して品詞や内容を判別する技術をいう。「形態素」とは、文書の要素のうち、意味を持つ最小の単位のことである。形態素解析部203において文書は、各語に対して品詞情報を持ったデータとなる。

【0065】

ベクトル生成部204は、形態素解析部203で形態素解析された文書を受け取り、文書をベクトルで表現する部分である。この際、形態素解析部203から品詞付きの情報が得られるため、特定の品詞のみを採用してもよい。さらに、TF-IDF (Term Frequency - Inverse Document Frequency) によ

10

20

30

40

50

り、特定の文書だけに出現頻度の高い語に重みを付けても構わない。重み付けに関しては一例であり、ここに挙げたもの以外の手法を用いても構わない。

【0066】

タグ自動生成部205は、回答文書とタグ調整文書から、これらの文書を表わすトピックを抽出することにより、タグを生成する。文書を表わすトピックの抽出には、トピックモデルを用いることができる。トピックモデルとは、データに隠れた潜在的なトピックを統計的に推定するモデルのことである。具体的には、LDA (Latent Dirichlet Allocation) や、LSI (Latent Semantic Indexing)、PLSA (Probabilistic Latent Semantic Analysis) などが挙げられるが、その文書を表すトピックを抽出できるのであれば、これに限らない。

10

【0067】

タグ紐付け部206は、トピックモデルより抽出された複数のトピック（タグ群）の中から、回答文書を最も特徴付けるトピックをタグとして紐付ける。

【0068】

形態素解析器208は、形態素解析部203とデータのやりとりを行う。例えば、オープンソースの形態素解析エンジン「MeCab」を用いることができる。さらに、このエンジンによれば、形態素の品詞も解析することができる。例えば、形態素ごとに、「名詞」「動詞」・・・などである。なお、必ずしも形態素解析をする必要はなく、形態素解析部203、および形態素解析器208を通さず、用意されたキーワード辞書などを参照し、文書中の語を抽出する機構を用いてもよい。例えば、「エアコンの温度設定の方法」という文書に対して語を抽出する場合、「エアコン」「温度」「設定」「方法」がキーワード辞書に登録されていれば、それらの語を抽出するというものである。これらは一例であり、語を抽出できる手法であればこれに限らない。

20

【0069】

通信部207は、制御サーバ1、回答データベース5および、タグ調整文書データベース6とそれぞれ通信する。

【0070】

図4は、本実施の形態における質問回答サーバ3の構成の一例を示すブロック図である。

30

【0071】

図4に示されるように、質問回答サーバ3は、質問入力部301、形態素解析部302、キーワード抽出部303、質問検索部304、通信部305、形態素解析器306、回答データベース5を備える。

【0072】

質問入力部301は、操作端末4から入力されたユーザの質問を受け取る。

【0073】

形態素解析部302は、質問入力部301から文書を受け取り、形態素解析器306を介して文書の形態素解析を行う。なお、図3における形態素解析部203と同等の機能を有するものであれば望ましい。

40

【0074】

キーワード抽出部303は、形態素解析部302から得られた結果から、キーワードを抽出する。この際、形態素解析部302から品詞付きの情報が得られるため、特定の品詞の語を選択することで、キーワードとする。例えば、「名詞」「動詞」のみをキーワードとして取得する、などである。

【0075】

質問検索部304は、キーワード抽出部303より得られたキーワードを用いて、回答データベース5に回答の検索を行う部分である。

【0076】

形態素解析器306は、形態素解析部302とデータのやりとりを行う。なお、図3に

50

おける形態素解析器 208 と同様の機能を有するものあればよい。

【0077】

通信部 305 は、制御サーバ 1、回答データベース 5 とそれぞれ通信する。

【0078】

図 5 は本実施の形態における操作端末 4 の構成の一例を表すブロック図である。

【0079】

図 5 に示されるように、操作端末 4 は、入力検出部 401、制御部 402、表示部 403、通信部 404 を有する。

【0080】

入力検出部 401 は、ユーザからのタッチパネル、音声、およびボタンなどを用いた入力の情報を取得する。

10

【0081】

制御部 402 は、入力検出部 401、表示部 403 との情報の送受信と、通信部 404 とのやりとりを行う。例えば、ユーザから「知識レベル：1」といった基準を意味する情報や「エアコンの効きが悪い」といった質問を意味する情報を受け付けると、制御部 402 を通じて通信部 404 に送信する。

【0082】

表示部 403 は、制御部 402 から得られた情報を表示する。例えば、ユーザの質問に対する回答などである。

【0083】

20

通信部 404 は、入力検出部 401 の情報を、制御部 402 を介してタグ付与サーバ 2 または質問回答サーバ 3 に送信し、タグ付与サーバ 2 または質問回答サーバ 3 から、制御部 402 を通じて、画面表示に関する情報を受信する。

【0084】

図 6 は、本実施の形態における質問回答システムの一連の処理の一例を示すシーケンス図である。図 6 の例では、タグ調整文書を選択する基準として、ユーザの知識レベルを用いている。また、図 6 の例では、回答データベース 5 の更新、およびユーザの質問に対して操作端末 4 に回答を出力するまでの一連の通信処理のシーケンスを示している。このシーケンスは、操作端末 4 になんらかの入力が検出されたときに開始される。

【0085】

30

ステップ S001 において、操作端末 4 はユーザの知識レベルに関する入力を取得する。

【0086】

ステップ S002 において、ステップ S001 で得られた情報から、回答データベース 5 に対してユーザの知識レベルに合ったタグを付与する処理か、ユーザの質問に回答する処理かを選択する。ここでは、知識レベルに関する情報が入力されたため、タグ付与の処理を行う。これらの処理は図 7 に示されているため、後述する。

【0087】

ステップ S003 において、S002 からユーザの知識レベルに関する情報を取得する。ステップ S004 において、回答データベース 5 からタグ付け対象の文書を取得する。また、図 10 に示すようなユーザの知識レベルとタグ調整文書との関係を示すテーブルを参照し、ステップ S005 において、適切なタグ調整文書をタグ調整文書データベース 6 から選択する。ここでの処理は、図 8 に示されているため、後述する。

40

【0088】

ステップ S006 において、ステップ S003、S004、S005 から選択された文書からタグを抽出し、タグ付け対象文書の内容を最も示しているタグを文書と紐付ける。ここでの処理は、図 9 に示されているため、後述する。

【0089】

ステップ S007 において、タグを回答データベース 5 に反映することでデータベースの更新を行う。

50

【0090】

ステップS008において、操作端末4はユーザの質問の入力を取得する。

【0091】

ステップS009において、ステップS008で得られた情報から、回答データベース5に対してユーザの知識レベルに合ったタグを付与する処理か、ユーザの質問に回答する処理かを選択する。ここでは、質問に関する情報が入力されたため、質問回答処理を行う。これらの処理は図7に示されているため、後述する。

【0092】

ステップS010において、ユーザの質問文からキーワードを抽出する。

【0093】

ステップS011において、ユーザの質問に対する回答検索のための処理を行う。

【0094】

ステップS012において、回答データベース5の検索を行う。

【0095】

ステップS013において、操作端末4に質問の回答を表示する。

【0096】

従来手法では、どのユーザに対しても同一のタグが付けられるものであった。ユーザの知識レベルが異なると質問に使われる語が異なるため、あらゆる人に対して正しい検索ができるとは限らなかった。本手法であれば、ユーザ自身が知識レベルを申告する、または、ユーザの質問回答システムの使用履歴により、ユーザの知識レベルに合った語をタグとして自動で付与することができる。このため、どのようなユーザでも、所望の回答を得ることができるようになる。また、質問回答システムの使用履歴から自動で更新することもできるので、知識レベルが向上した場合も自動的にタグを更新することが可能である。

【0097】

図7は、本実施の形態における、タグ付与サーバ2または質問回答サーバ3の選択処理の手順の一例を示すフローチャートである。この処理は、図6におけるサーバ選択処理（S002、S009）の詳細である。

【0098】

ステップS201において、操作端末4から得られたユーザの入力を取得する。

【0099】

ステップS202において、ステップS201から得られたデータ中に、知識レベルに関するデータ（すなわち、タグ調整文書を選択する基準に関する情報）が含まれているかを判断する。知識レベルに関する情報が含まれていると判断された場合には、ステップS203において、タグ付与サーバ2に知識レベルの情報を送信する。ステップS204において、ステップS201から得られたデータの中に、ユーザの質問が含まれているかどうかを判断する。質問に関する情報が含まれていると判断された場合には、ステップS205において、質問回答サーバ3に質問に関する情報を送信する。

【0100】

例えば、「インターネットの接続方法が知りたい」という情報が得られた場合、質問に関する情報であるため、質問回答サーバ3に情報が送信され、質問回答の処理が行われる。また、「知識レベル：1」のような情報が得られた場合、知識レベルに関する情報であるため、タグ付与サーバ2に情報が送信され、タグ付けの処理が行われる。図11は本実施形態による、ユーザの知識レベル入力時の操作端末4の表示画面の一例を示す図である。ただし、これらの表示方法は一例であり、これに限らない。

【0101】

図11Aおよび図11Bは、それぞれ、ユーザが質問を行う際に、操作端末4に表示される画面の一例を示す図である。例えば、図11Aのように、各製品や製品全体に対して自分の知識レベルを入力する。また、例えば、図11Bのように、ユーザの質問と同時に、質問対象の製品に自分の知識レベルを入力することで、ユーザにあったタグで検索が可能になる。

10

20

30

40

50

【0102】

なお、図11Aおよび図11Bの例では、ユーザが知識レベルを直接入力するようになっているが、ユーザの知識レベルに対応付けられる情報を入力するようにしてもよい。例えば、家電の使用方法に関する質問の場合、質問を行うユーザに対して、「若者」、「高齢者」のいずれであるかを選択する画面、あるいは、「一般ユーザ」、「営業マン」、「技術者」のいずれであるかを選択する画面を提示し、入力された情報に基づき、知識レベルを判定するようにしてもよい。また、製品が、特定の性別のユーザ、特定の年代のユーザ、または特定の地域に住むユーザによって主に使用されるものである場合、質問を行うユーザに対して、性別、年齢、または住所を入力する画面を提示し、入力された情報に基づき、知識レベルを判定するようにしてもよい。

10

【0103】

なお、図11Aのように、質問を受け付ける前に、回答提供者においてあらかじめ回答文書にタグ付けするためのオーサリングツールのような使い方もあれば、図11Bのように質問を問い合わせる際にタグを付与する方式でも構わない。質問問合せの際にタグを付与することで、オーサリングツールでは事前に対応できなかったものにも対応できることになる。さらに、そのユーザに合わせたタグだけを付与することができるため、余計なタグを検索することもない。例えば、オーサリングツールにおいて「インターネット、接続」というタグが付与されていても、質問時にユーザの知識レベルが向上していれば、質問時にタグ付けを行って、「インターネット、LAN、IPアドレス、ブラウザ」などと変更されていれば、より適切な回答を提示することができる。

20

【0104】

以上の処理により、各サーバで行う処理を切り分けることが可能である。これにより、ユーザの知識レベルを得て回答データベース5のタグを逐次的に更新しながら、ユーザの質問に対して適切な回答を行うことができる。

【0105】

図8は、本実施の形態における、タグ付け対象文書に対して、タグ調整文書を選択する処理の手順の一例を示すフローチャートである。この処理は、図6における文書選択処理(S003)の詳細である。

【0106】

ステップS301において、ユーザの知識レベルに関する情報（すなわち、タグ調整文書を選択する基準に関する情報）を取得する。ここで、図10のようなテーブルで読み取ることのできる情報に変換する。知識レベルを表す語として、例えば、「低・中・高」「易しい・普通・難しい」などが挙げられる。これらは一例であり、ここに挙げたもの以外を用いても構わない。これらを、テーブルが保持している情報に変換する機構を持っていることとする。図13は本発明の実施の態様に係るタグ調整文書データベース6のデータ構造の一例を示す図である。図13のように、タグ調整文書データベース6のデータの中に、複数の知識レベルの表現を含んでもよい。

30

【0107】

ステップS302において、ステップS301で取得した知識レベルと対応するタグ調整文書をタグ調整文書データベース6より選択する。この際、例えば図10に示すような、ユーザ習熟度とタグ調整文書のテーブルから対応関係を取得する。対応関係がわかればどのような形式でデータを保持してもよく、これに限らない。また、知識レベルに合わせた文章であれば、選択する文書の個数はいくつでもよい。

40

【0108】

ステップS303において、タグ付け対象文書と、ステップS302で選択されたタグ調整文書を組み合わせる。例えば、エアコンの取扱説明書のページにタグを付与したいとする。同時に、ユーザの知識レベルは「4（高い）」という情報が得られた場合、知識レベルとタグ調整文書の対応テーブルにより、タグ調整文書としてエアコン工事取り付け専門書が選ばれる。次に、タグ調整文書データベース6から、エアコン工事取り付け専門書に関する文書を取得し、タグ付け対象文書と組み合わせる。ここで、エアコン工事取り付

50

け専門書が2個選ばれた場合は、例えば、タグ付け対象文書1個と合わせて、3個分の文書が次の処理に送られることとなる。

【0109】

以上の処理により、タグ付け対象文書に組み合わせる文書を自動で選択する。これにより、ユーザが自分でタグ付け対象文書に合ったタグ調整文書を選ぶ必要がなく、ユーザの負担を軽くすることができる。

【0110】

図9は、本実施の形態における、タグ付け対象文書に対して、タグを抽出し紐付ける処理の手順の一例を示すフローチャートである。この処理は、図6におけるタグ抽出／紐付け処理(S006)の詳細である。

10

【0111】

ステップS601において、取得した文書の形態素解析を行う。形態素解析を行った文書は意味のある単語に区切られ、品詞情報を持ったデータとなる。

【0112】

ステップS602において、ステップS601から得られたデータを文書ごとにベクトル化する。この際、品詞付きの情報が得られているため、特定の品詞のみを採用してもよい。さらに、TF-IDF(Term Frequency - Inverse Document Frequency)により、特定の文書だけに出現頻度の高い語に重みを付けても構わない。重み付けに関しては一例であり、ここに挙げたもの以外の手法を用いても構わない。

20

【0113】

ステップS603において、トピックモデルを用いて文書のタグを抽出する。トピックモデルは、データに隠れた潜在的なトピックを統計的に推定するモデルのことである。ここでトピックとは話題や意味のまとまりのことを指す。用意された複数文書に対して、トピック数を指定しトピックモデルを適用すると、指定したトピックの数だけトピックが得られる。ここでは、トピックの名称が得られるわけではないが、文書間や単語間の意味の関係性を統計的に見つけ出し、それぞれの文書がどのようなトピックから構成されているかが分かる。さらに、ステップS604において、それぞれの文書に最も出現しているトピックをその文書を最も表すトピックとみなし、トピックに含まれている単語をタグとして付与する。トピックモデルとして、具体的には、LDA(Latent Dirichlet Allocation)や、LSI(Latent Semantic Indexing)、PLSA(Probabilistic Latent Semantic Analysis)などが挙げられるが、その文書を表すトピックを抽出できるものであれば、これに限らない。また、トピックの数は、文書の数もしくはそれ以上の数を指定するものとする。

30

【0114】

ステップS604において、抽出されたトピックをタグとして、タグ付け対象文書と紐付ける。この時、最もタグ付け対象文書を表しているトピックを選び、回答データベース5を更新する。例えば、最もタグ付け対象文書の語を含んでいるトピックを選ぶ、などが挙げられるが、これは一例であり、ここに挙げたもの以外の手法を用いても構わない。

40

【0115】

例えば、エアコンの取扱説明書1ページを1個と、タグ調整文書(知識レベル:1)2個の文書が得られたとする。形態素解析とベクトル化を行ったのち、タグの抽出を行ったとする。タグ抽出の結果、トピックとして、例えば、トピック1:「日当たり、センサー、調整」、トピック2:「テレビ欄、スマートフォン、テレビ」トピック3:「スポーツ、大会、選手」などのトピックが得られる。この時、タグ付け対象文書の1ページが日当たりセンサーの話題だった場合、トピック1の「日当たり、センサー、調節」がタグとして選択され、回答データベース5のタグを更新する。タグ調整文書の内容は、エアコンの内容と直接関係ない文章である可能性もあるため、トピック2の「テレビ欄、スマートフォン、テレビ」などの語は、タグ調整文書から抽出された語である可能性が高い。また、

50

タグ調整文書の知識レベルを高く設定した場合、「調節日射センサー、屋外温度、室外ユニット」といった、知識レベルの高い人が使うような語をタグとして割り当てることになる。

【0116】

以上の処理により、タグ調整文書の内容によって異なるトピックを抽出し、タグ付け対象文書に適切なトピックをタグとして付与する。これにより、ユーザの知識レベルから自動的にタグを抽出し文書に付与することで、ユーザに合わせた回答が可能になる。

【0117】

図12は、本実施の形態における、回答データベース5に保持されているXMLの一例を表したものである。例えば、このように、本文、タグなどが構造化された状態で保持されている。また、ユーザの知識レベルが更新されるごとに、このタグの内容も更新される。ここに挙げた項目は一例であり、これら以外の項目があっても構わない。また、データの保存形式の一例としてXML形式を挙げているが、データを構造的に保持できるものであれば、これに限らない。

【0118】

図13は、本実施の形態における、タグ調整文書データベース6に保持されているXMLの一例を表したものである。例えば、このように、本文、知識レベルなどが構造化された状態で保持されている。知識レベルの表現のゆらぎに関しては、図13のように知識レベルの項目に含んでおいてもよいし、別途、知識レベルの表現を変換する機構を持っていてもよい。ここに挙げた項目は一例であり、これら以外の項目があっても構わない。また、データの保存形式の一例としてXML形式を挙げているが、データを構造的に保持できるものであれば、これに限らない。

【0119】

図14は、本実施の形態における、LDAを用いたトピック抽出の結果の一例を表す図である。例えば、取扱説明書の、機器のインターネット接続に関する項目にタグを付与するため、本手法であるトピックモデルを用いてタグ付けを行うとする。タグ付け対象文書を機器のインターネット接続に関するページとし、タグ調整文書をインターネット接続と異なる内容の取扱説明書とした場合、図14の(a)に示すような語が抽出される。また、タグ付け対象文書を機器のインターネット接続に関するページとし、タグ調整文書をWikipediaから選んだ一般的な内容とした場合、図14の(b)に示すような語が抽出される。また、どちらの組み合わせでも得られる語は、共通のタグとして図14の(c)に示すような語が現れる。本来の手法では、タグとして、例えば図14の(a)に表すような語を一意に付与することになるが、本手法を用いると、図14の(b)のような語もタグとして付与することができるため、ユーザに合わせたタグ付けが可能になることを示している。上記のように、文書の組み合わせを変えると、トピックモデルにおいてモデルを構成する語が異なるため、得られるトピックやそれに含まれる語も異なるものになる。

【0120】

以上、本発明の一態様に係るタグ付与装置について、実施の形態に基づいて説明したが、本発明は、これらの実施の形態に限定されるものではない。本発明の趣旨を逸脱しない限り、当業者が思いつく各種変形を本実施の形態に施したもの、あるいは異なる実施の形態における構成要素を組み合わせて構築される形態も、本発明の範囲内に含まれる。

【0121】

例えば、ある家電製品に新しく搭載された機能のみを検索したいとする。このとき、ある家電製品の取扱説明書と後継機種（新機種）の取扱説明書を組み合わせてタグを抽出することで、例えば、後継機種（新機種）の特徴のみをタグとして付与することができる。同様に、ある家電製品の取扱説明書と旧機種の取扱説明書を組み合わせてタグを抽出することで、旧機種の特徴のみをタグとして付与することができるのは言うまでもない。また、上位機種・下位機種の組み合わせでタグを抽出することで、例えば、上位機種にしかない機能をタグとして付与することもできる。

【0122】

また、ある取扱説明書にタグを付与したい場合、タグ付け対象文書を取扱説明書の1ページとし、その他のページをタグ調整文書としてもよい。例えば、エアコンに関する取扱説明書の場合、タグ付け対象文書を風量調節に係るページとし、タグ調整文書をエアコンの風量調節以外のページとすることで、タグ付け対象文書にタグが付与される。さらに、これらを取扱説明書のすべてのページについて行うことで、取扱説明書の全ページにタグが付与される。なお、ページは一例であって、項目や目次単位でも構わない。

【0123】

また、タグ調整文書データベース6を常時ネットワークにつなぐことで、データベースの内容を追加・更新する機構を持ってもよい。タグ調整文書データベース6を更新する内容としては、例えば、Wikipediaのようなインターネット上の辞典、ニュース、ブログ、またはTwitter（登録商標）のような短文投稿サービスなどの内容を自動的に取り入れて、データベースの内容を更新することができる。また、新しい製品のプレスリリースなどの情報を取り入れてもよい。これらは一例であり、文書データであればこれに限らない。

10

【0124】

また、システム構成全体を表す図1において、制御サーバ1、タグ付与サーバ2、質問回答サーバ3は1つのサーバにより構成してもよいし、図1に示すものと異なる構成であっても構わない。また、例えば、オーサリングツールとして使う場合は、質問回答サーバ3がなくても構わない。回答データベース5、タグ調整文書データベース6は一つのデータベースとして構成されていても構わない。

20

【0125】

また、本実施形態の例では、操作端末4からユーザの知識レベルを入力すると回答データベース5の回答文書にタグが付与されるが、質問回答システムにおいて、ユーザの行動履歴から知識レベルを取得することができれば、回答データベース5の文書のタグを自動、または手動で更新してもよい。例えば、ユーザのウェブの閲覧履歴やメールのやりとりで使われている語などを取得すれば、例えば機械学習を使って、統計的に知識レベルを算出してもよい。また、過去の質問の履歴からどのような語を使って質問しているのか、質問に対する回答は適切だったのかなどの情報を元に、回答データベース5のタグを更新してもよい。これらは一例であって、ユーザの知識レベルを得ることができる情報であればこれに限らない。

30

【0126】

なお、本実施形態の例における知識レベルにおける難易度の部分は、どのような基準を用いても構わない。例えば、何らかの技術に対するスキルレベル、営業マンまたは技術者が使う語の違い、若者と高齢者が使う言葉の違いなどの観点で、タグ付けする語を変えても構わない。また、これらは一例であり、基準によって語が変わるようなものであればこれに限らない。

【0127】

また、本実施の形態においては、家電の使用方法に関する質問回答システムを例として説明したが、他の分野への応用も可能である。例えば、医療診断システムにおいて、まず、質問を行うユーザに対して、医療分野あるいは質問しようとする症状に関する知識レベルを問い合わせ、得られた知識レベルに応じて、症例や治療方法に関する回答文書に適切なタグを付与するといったことが可能である。ユーザからの知識レベルの取得にあたっては、図11Aまたは図11Bと同様の画面により、知識レベルに相当する数値等の入力を受け付けるようにしてもよいし、一般患者であるか、医療関係者であるかといったような、知識レベルに相当する情報を受け付けるようにしてもよい。また、タグ調整文書については、例えば、知識レベルが高い場合は医学書や医療分野の論文を、知識レベルが低い場合は一般知識に関する文書を選択するようにすればよい。

40

【0128】

また、実施の形態におけるタグ付与装置が備える構成要素の一部または全部は、1個の

50

システムLSI (Large Scale Integration: 大規模集積回路) から構成されているとしてもよい。

【0129】

システムLSIは、複数の構成部を1個のチップ上に集積して製造された超多機能LSIであり、具体的には、マイクロプロセッサ、ROM (Read Only Memory)、RAM (Random Access Memory) などを含んで構成されるコンピュータシステムである。前記ROMには、コンピュータプログラムが記憶されている。前記マイクロプロセッサが、前記コンピュータプログラムに従って動作することにより、システムLSIは、その機能を達成する。

【0130】

なお、ここでは、システムLSIとしたが、集積度の違いにより、IC、LSI、スーパーLSI、ウルトラLSIと呼称されることもある。また、集積回路化の手法はLSIに限るものではなく、専用回路又は汎用プロセッサで実現してもよい。LSI製造後に、プログラムすることが可能なFPGA (Field Programmable Gate Array)、あるいはLSI内部の回路セルの接続や設定を再構成可能なリコンフィギュラブル・プロセッサを利用してもよい。

【0131】

さらには、半導体技術の進歩又は派生する別技術によりLSIに置き換わる集積回路化の技術が登場すれば、当然、その技術を用いて機能ブロックの集積化を行ってもよい。バイオ技術の適用等が可能性としてありえる。

【0132】

また、本発明は、このような特徴的な処理部を備える、タグ付与装置として実現することができるだけでなく、タグ付与装置に含まれる特徴的な処理部をステップとするタグ付与方法などとして実現することもできる。また、そのような方法に含まれる特徴的な各ステップをコンピュータに実行させるコンピュータプログラムとして実現することもできる。そして、そのようなコンピュータプログラムを、CD-ROM等のコンピュータ読取可能な非一時的な記録媒体あるいはインターネット等の通信ネットワークを介して流通させることができるのは、言うまでもない。

【産業上の利用可能性】

【0133】

本発明に係るタグ付与方法は、文書に対して自動でタグを付与するタグ付与装置、文書に付されたタグにより質問に対する回答を検索する質問回答システム、文書に付されたタグにより文書を管理する文書管理システムなどに適用することができる。

【符号の説明】

【0134】

- 1 制御サーバ
- 2 タグ付与サーバ
- 3 質問回答サーバ
- 4 操作端末
- 5 回答データベース
- 6 タグ調整文書データベース
- 7 ネットワーク
- 101 入力検出部
- 102 処理選択部
- 103 制御部
- 104, 207, 305, 404 通信部
- 201 基準取得部
- 202 文書選択部
- 203, 302 形態素解析部
- 204 ベクトル生成部

10

20

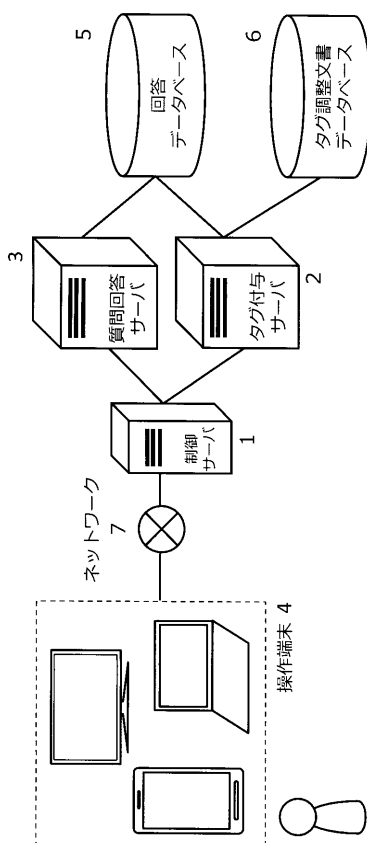
30

40

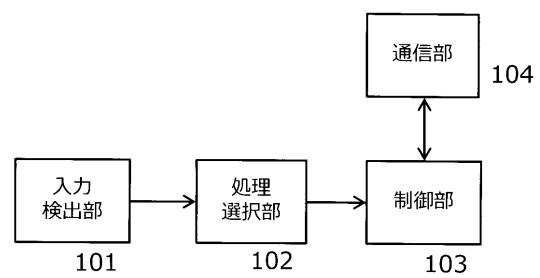
50

- 205 タグ自動生成部
- 206 タグ紐付け部
- 208, 306 形態素解析器
- 301 質問入力部
- 303 キーワード抽出部
- 304 質問検索部
- 401 入力検出部
- 402 制御部
- 403 表示部

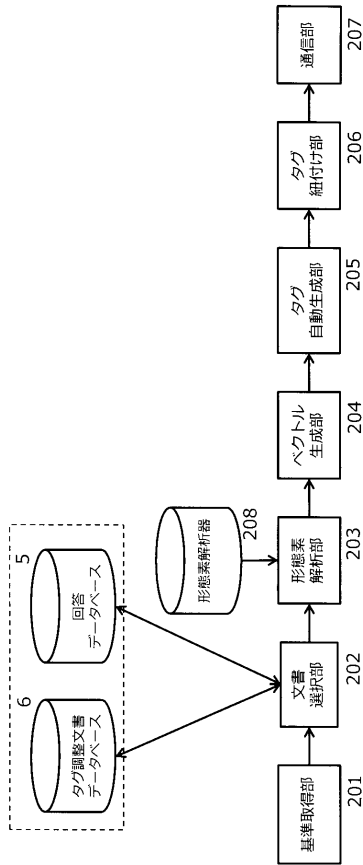
【図 1】



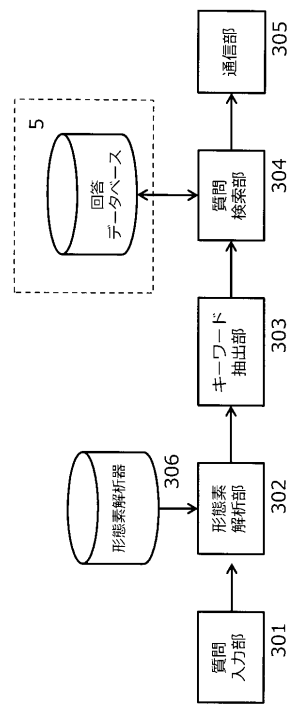
【図 2】



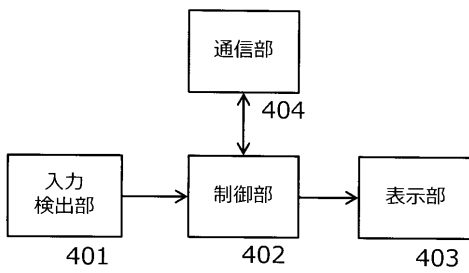
【図 3】



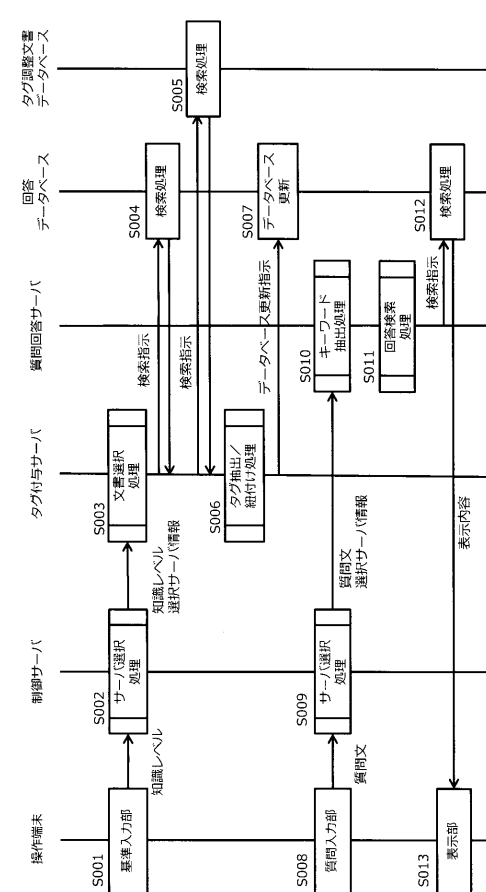
【図 4】



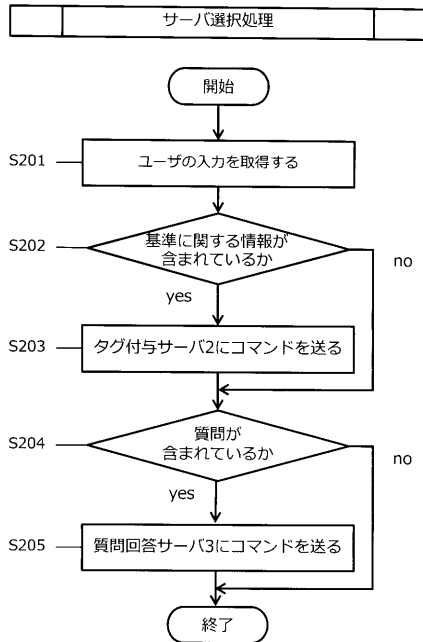
【図 5】



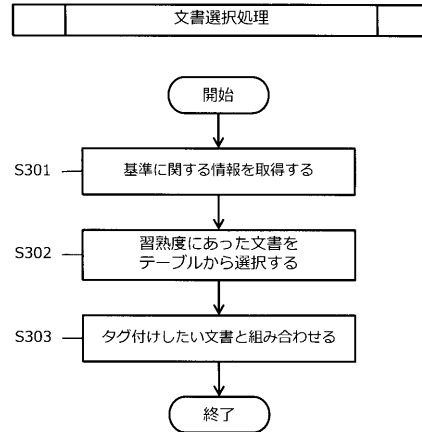
【図 6】



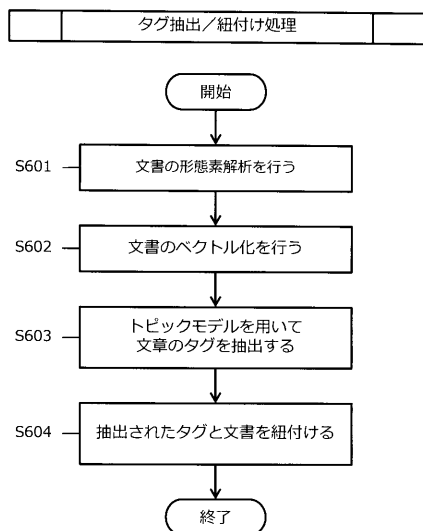
【図 7】



【図 8】



【図 9】



【図 1 1 A】

家電製品全般に関する知識レベル（習熟度）

☐ 1 ☐ 2 ☒ 3 ☐ 4 ☐ 5

エアコンに関する知識レベル（習熟度）

☒ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

レコーダーに関する知識レベル（習熟度）

☐ 1 ☒ 2 ☐ 3 ☐ 4 ☐ 5

【図 1 1 B】

お困りごとはなんですか？下に入力してください。

エアコンに関する知識レベル（習熟度）

☒ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

【図 1 0】

知識レベル (4段階)	タグ調整文書
1	レビューやブログ
2	Wikipedia
3	取扱説明書
4	工事取り付け専門書

【図 1 2】

<取扱説明書>
<本文>
ご使用前にふたロックと解除方法「すずぎ」「脱水」「乾燥の前半」「乾燥の前半」および「槽乾燥」時は、ふたロックがかかります。ふたロック中は、無理に開けないでください。
(ふたが破損することがあります)
※槽回転が止まるまで約15秒かかります。
を押し、「カチッ」と音がするまでお待ちください。
ふたロックを解除する場合(ふたロック中)点灯ふた・・・
</本文>
<タグ>ふたロック、音、すずぎ、脱水、乾燥の前半、槽乾燥</タグ>
</取扱説明書>

【図 1 3】

図 1 0 の知識レベル 2 (Wikipedia における文書例)

<調整文>
<本文>
ひまわり (向日葵, sunflower)
北アメリカのテキサスやカリフォルニアを原産地とする。8月～9月頃に黄色い大きな花を咲かせる。また、高さ2～3メートルに達してまっすぐに成長する。
名前は、花が太陽に向かって回るといふことからつけられた。
また、ひまわりの花は、小さな花がいくつも集まって大きな花を形成している。
ひまわりの呼び方は様々で、日本では別名として、日車、日輪草、天神葵などがあることが知られている。
</本文>
<知識レベル> 2、中、易しい</知識レベル>
</調整文>

【図 1 4】

