

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2007-505403

(P2007-505403A)

(43) 公表日 平成19年3月8日(2007.3.8)

(51) Int. Cl.

G06F 17/30 (2006.01)

F I

G06F 17/30 340A

テーマコード (参考)

5B075

審査請求 未請求 予備審査請求 未請求 (全 31 頁)

(21) 出願番号 特願2006-526320 (P2006-526320)
 (86) (22) 出願日 平成16年9月10日 (2004. 9. 10)
 (85) 翻訳文提出日 平成18年4月19日 (2006. 4. 19)
 (86) 国際出願番号 PCT/US2004/029618
 (87) 国際公開番号 W02005/029367
 (87) 国際公開日 平成17年3月31日 (2005. 3. 31)
 (31) 優先権主張番号 10/661, 741
 (32) 優先日 平成15年9月12日 (2003. 9. 12)
 (33) 優先権主張国 米国 (US)

(71) 出願人 305029922
 グーグル・インク
 アメリカ合衆国、カリフォルニア州 94
 043、マウンテン・ビュー、アンフィシ
 アトル・パークウェイ 1600
 (74) 代理人 100083806
 弁理士 三好 秀和
 (74) 代理人 100095500
 弁理士 伊藤 正和
 (72) 発明者 トング、サイモン
 アメリカ合衆国 カリフォルニア州 94
 040 マウンテン ビュー デル メデ
 イオ アベニュー 541 ナンバー31
 9

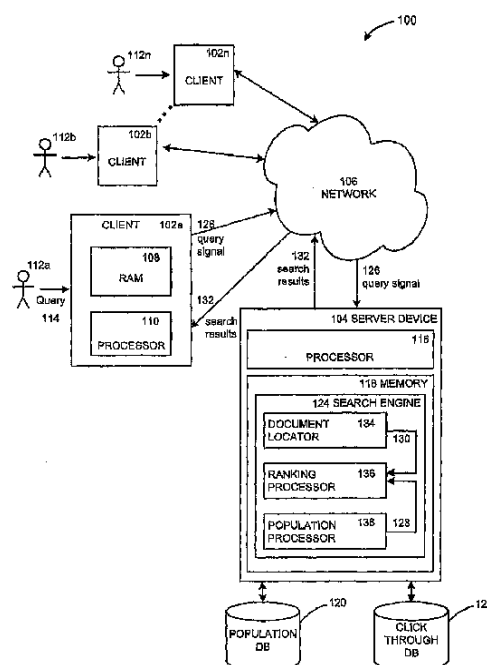
最終頁に続く

(54) 【発明の名称】 母集団情報を用いて探索ランク付けを改善する方法とシステム

(57) 【要約】

探索問い合わせの探索ランク付けを、その探索問い合わせに関する問い合わせと関連するデータを用いることによって改善するシステムと方法を説明する。1つの態様では、探索問い合わせが受信され、この探索問い合わせと関連する母集団が決定され、その探索問い合わせと関連する(ウェブページなどの)項目が決定されて、その母集団と関連するデータに少なくとも部分的に基づいたその項目のランク付け得点が決定される。このようなシステムと方法を実行する際に有用な母集団と関連するデータのアルゴリズムとタイプを説明する。

【選択図】図1



【特許請求の範囲】**【請求項 1】**

(a) 探索問い合わせを受信するステップと、
(b) 前記探索問い合わせと関連する第 1 の母集団を決定するステップと、
(c) 前記探索問い合わせと関連する第 1 の項目を決定するステップと、
(d) 前記第 1 の項目のランク付け得点を前記第 1 の母集団と関連するデータに少なくとも部分的に基づいて決定するステップと、
を含む、探索ランク付け得点を決定する方法。

【請求項 2】

前記探索問い合わせと関連する前記第 1 の母集団を決定するステップが、前記探索問い合わせの送出者と関連する人口統計学的データを決定するステップを含む請求項 1 に記載の方法。 10

【請求項 3】

前記送出者と関連する前記人口統計学的データを決定するステップが、前記送出者のありそうな地理的ロケーションを決定するステップを含む請求項 2 に記載の方法。

【請求項 4】

前記送出者のありそうな地理的ロケーションを決定するステップが、探索問い合わせの送信元であるインターネットプロトコルアドレス、探索エンジンにアクセスするために前記送出者が入力したアドレスおよび前記送出者が入力した人口統計学的データの内の少なくとも 1 つを決定するステップを含む請求項 3 に記載の方法。 20

【請求項 5】

前記送出者の前記人口統計学的データを決定するステップが、年齢、年齢層、性、人種、第 1 言語、第 2 言語、ロケーション、収入、収入層、大陸、領域、国、州、郡、市、性別、民族的グループ、グループ、特徴を共有する人物、興味を共有する人物、所定の選択によって分類された人物およびインターネットサービスプロバイダデータの内の少なくとも 1 つを決定するステップを含む請求項 2 に記載の方法。

【請求項 6】

前記探索問い合わせと関連する前記第 1 の母集団を決定するステップが、前記探索問い合わせと関連する人口統計学的データを決定するステップを含む請求項 1 に記載の方法。

【請求項 7】

前記探索問い合わせと関連する前記人口統計学的データを決定するステップが、前記探索問い合わせの言語を決定するステップと、前記探索問い合わせの前の送出者と関連するデータを決定するステップとの内の少なくとも一方を含む請求項 6 に記載の方法。 30

【請求項 8】

前記探索問い合わせと関連する前記第 1 の母集団を決定するステップが、前記探索問い合わせを送信するユーザと関連する自己識別データを決定するステップを含む請求項 1 に記載の方法。

【請求項 9】

前記自己識別データが、ユーザ登録データ、ユーザ嗜好データおよびユーザ選択データの内の少なくとも 1 つから選択される請求項 8 に記載の方法。 40

【請求項 10】

前記探索問い合わせと関連する前記第 1 の母集団を決定するステップが、前記探索問い合わせを送信するユーザと関連する自動識別データを決定するステップを含む請求項 1 に記載の方法。

【請求項 11】

前記自動識別データが、IP アドレス、ドメインおよび、ユーザと関連するアプリケーションから得られるデフォルトデータの内の少なくとも 1 つを含む請求項 10 に記載の方法。

【請求項 12】

前記第 1 の母集団と関連するデータが、前記第 1 の母集団の文脈における第 1 の項目の 50

選択得点を含む請求項 1 に記載の方法。

【請求項 1 3】

前記第 1 の母集団の文脈における前記第 1 の項目の選択得点が、前記第 1 の母集団のメンバーによる前記第 1 の項目に対するクリックの数を含む請求項 1 2 に記載の方法。

【請求項 1 4】

前記第 1 の母集団と関連するデータが、前記第 1 の母集団のメンバーの数を含む請求項 1 に記載の方法。

【請求項 1 5】

前記第 1 の母集団のメンバーの数が、前記第 1 の母集団の、前記探索問い合わせに対して返送された結果を選択したメンバーの数を含む請求項 1 4 に記載の方法。

10

【請求項 1 6】

前記第 1 の母集団のメンバーの数が、前記第 1 の母集団の、前記探索問い合わせを入力したメンバーの数を含む請求項 1 4 に記載の方法。

【請求項 1 7】

前記第 1 の母集団のメンバーの数が、前記第 1 の母集団の、前記探索問い合わせに対する探索結果が表示されたメンバーの数を含む請求項 1 4 に記載の方法。

【請求項 1 8】

前記第 1 の母集団と関連するデータが選択総得点を含む請求項 1 に記載の方法。

【請求項 1 9】

前記選択総得点が、前記第 1 の母集団の、前記第 1 の項目を選択したメンバーの総数を含む請求項 1 8 に記載の方法。

20

【請求項 2 0】

前記探索問い合わせと関連する第 2 の母集団を決定するステップをさらに含み、前記第 1 の項目の第 1 のランク付け得点を決定するステップがさらに、前記第 2 の母集団と関連するデータに少なくとも部分的に基づく請求項 1 に記載の方法。

【請求項 2 1】

(e) 前記探索問い合わせと関連する第 2 の項目を決定するステップと、

(f) 前記第 2 の項目の第 2 のランク付け得点を、前記第 1 の母集団と関連するデータに少なくとも部分的に基づいて決定するステップと、
をさらに含む請求項 1 に記載の方法。

30

【請求項 2 2】

前記第 1 の項目と前記第 2 の項目を、前記第 1 のランク付け得点と前記第 2 のランク付け得点に少なくとも部分的に基づいてランク付けするステップをさらに含む請求項 2 1 に記載の方法。

【請求項 2 3】

(a) 探索問い合わせを受信するプログラムコードと、

(b) 前記探索問い合わせと関連する第 1 の母集団を決定するプログラムコードと、

(c) 前記探索問い合わせと関連する第 1 の項目を決定するプログラムコードと、

(d) 前記第 1 の項目のランク付け得点を前記第 1 の母集団と関連するデータに少なくとも部分的に基づいて決定するプログラムコードと、
を含むプログラムコードを包含するコンピュータ読み取り可能媒体。

40

【請求項 2 4】

前記探索問い合わせと関連する第 1 の母集団を決定する前記プログラムコードが、前記探索問い合わせの送出者と関連する人口統計学的データを決定するように適用される請求項 2 3 に記載のコンピュータ読み取り可能媒体。

【請求項 2 5】

送出者と関連する人口統計学的データを決定するステップが、前記送出者のありそうな地理的ロケーションを決定するステップを含む請求項 2 4 に記載のコンピュータ読み取り可能媒体。

【請求項 2 6】

50

送出者のありそうな地理的ロケーションを決定するステップが、前記探索問い合わせの送信元であるインターネットプロトコルアドレス、探索エンジンにアクセスするために前記送出者が入力したアドレスおよび前記送出者が入力した人口統計学的データの内の少なくとも1つを決定するステップを含む請求項25に記載のコンピュータ読み取り可能媒体。

【請求項27】

送出者の前記人口統計学的データを決定するステップが、年齢、年齢層、性、人種、第1言語、第2言語、ロケーション、収入、収入層、大陸、領域、国、州、郡、市、性別、民族的グループ、グループ、特徴を共有する人物、興味を共有する人物、所定の選択によって分類された人物およびインターネットサービスプロバイダデータの内の少なくとも1つを決定するステップを含む請求項24に記載のコンピュータ読み取り可能媒体。

【請求項28】

前記探索問い合わせと関連する前記第1の母集団を決定するステップが、前記探索問い合わせと関連する人口統計学的データを決定するステップを含む請求項23に記載のコンピュータ読み取り可能媒体。

【請求項29】

前記探索問い合わせと関連する前記人口統計学的データを決定するステップが、前記探索問い合わせの言語を決定するプログラムコードと、前記探索問い合わせの前の送出者と関連するデータを決定するステップとの内の少なくとも一方を含む請求項28に記載のコンピュータ読み取り可能媒体。

【請求項30】

前記探索問い合わせと関連する前記第1の母集団を決定するステップが、前記探索問い合わせを送信するユーザと関連する自己識別データを決定するステップを含む請求項23に記載のコンピュータ読み取り可能媒体。

【請求項31】

前記自己識別データが、ユーザ登録データ、ユーザ嗜好データおよびユーザ選択データの内の少なくとも1つから選択される請求項31に記載のコンピュータ読み取り可能媒体。

【請求項32】

前記探索問い合わせと関連する前記第1の母集団を決定するステップが、前記探索問い合わせを送信するユーザと関連する自動識別データを決定するステップを含む請求項23に記載のコンピュータ読み取り可能媒体。

【請求項33】

前記自動識別データが、IPアドレス、ドメインおよび、ユーザと関連するアプリケーションから得られるデフォルトデータの内の少なくとも1つを含む請求項32に記載のコンピュータ読み取り可能媒体。

【請求項34】

前記第1の母集団と関連するデータが、前記第1の母集団の文脈における前記第1の項目の選択得点を含む請求項23に記載のコンピュータ読み取り可能媒体。

【請求項35】

前記第1の母集団の文脈における前記第1の項目の選択得点が、前記第1の母集団のメンバーによる前記第1の項目に対するクリックの数を含む請求項34に記載のコンピュータ読み取り可能媒体。

【請求項36】

前記第1の母集団と関連するデータが、前記第1の母集団のメンバーの数を含ま請求項23に記載のコンピュータ読み取り可能媒体。

【請求項37】

前記第1の母集団のメンバーの数が、前記第1の母集団の、前記探索問い合わせに対して返送された結果を選択したメンバーの数を含ま請求項36に記載のコンピュータ読み取り可能媒体。

10

20

30

40

50

【請求項 38】

前記第 1 の母集団のメンバーの数が、前記第 1 の母集団の、前記探索問い合わせを入力したメンバーの数を含む請求項 36 に記載のコンピュータ読み取り可能媒体。

【請求項 39】

前記第 1 の母集団のメンバーの数が、前記第 1 の母集団の、前記探索問い合わせに対する探索結果が表示されたメンバーの数を含む請求項 36 に記載のコンピュータ読み取り可能媒体。

【請求項 40】

前記第 1 の母集団と関連するデータが選択総得点を含む請求項 23 に記載のコンピュータ読み取り可能媒体。

10

【請求項 41】

前記選択総得点が、前記第 1 の母集団の、前記第 1 の項目を選択したメンバーの総数を含む請求項 40 に記載のコンピュータ読み取り可能媒体。

【請求項 42】

前記探索問い合わせと関連する第 2 の母集団を決定するステップをさらに含み、前記第 1 の項目の第 1 のランク付け得点を決定するステップがさらに、前記第 2 の母集団と関連するデータに少なくとも部分的に基づく請求項 23 に記載のコンピュータ読み取り可能媒体。

【請求項 43】

(e) 前記探索問い合わせと関連する第 2 の項目を決定するプログラムコードと、

20

(f) 前記第 2 の項目の第 2 のランク付け得点を、前記第 1 の母集団と関連するデータに少なくとも部分的に基づいて決定するプログラムコードと、
をさらに含む請求項 23 に記載のコンピュータ読み取り可能媒体。

【請求項 44】

前記第 1 の項目と前記第 2 の項目を、前記第 1 のランク付け得点と前記第 2 のランク付け得点に少なくとも部分的に基づいてランク付けするステップをさらに含む請求項 43 に記載のコンピュータ読み取り可能媒体。

【発明の詳細な説明】

【技術分野】

【0001】

30

本発明は一般的には探索エンジンに関する。より詳しくは、本発明は、母集団情報を用いて探索ランク付けを改善する方法とシステムに関する。

【0002】

背景技術

WWWなどのネットワークされたコンピュータ環境でまたは個々のコンピュータ中で動作する従来の探索エンジンは、ユーザが探索問い合わせを入力するとそれに応答して探索結果を提供することが可能である。多くの例では、探索結果は、探索エンジンの得点付けやランク付けのシステムや方法にしたがってランク付けされる。たとえば、従来の探索エンジンは、特定の問い合わせに対する探索結果の文書を、キーワードまたは特定の語もしくはは句が探索結果中の各々の文書に顕れる回数によって得点またはランク付けする。文書には、たとえば、HTML、XML、XHTMLなどのさまざまなフォーマットを持つWebページと、携帯式文書フォーマット(PDF)ファイルと、ワープロやアプリケーションプログラムの文書ファイルと、がある。他の探索エンジンでは、文書の内容以上のものに基づいて得点付けやランク付けの結果を提供する。たとえば、Sergey BrinとLawrence Pageによる「大規模ハイパーテキスト探索エンジンの解剖学的構造(The Anatomy of a Large-Scale Hyper textual Search Engine)」という題名の項目に記載された1つの周知の方法では、Webページなどの文書に対して、そのWebページのリンク構造に基づいて重要度を割り当てる。他の従来の方法では、特定の問い合わせに対する探索結果におけるより高い得点やランクを、ユーザや顧客を自分のWebサイトにひきつけることを

40

50

希望している第三者に対して売却する。

【0003】

一部の例では、特定のロケーションにいるユーザが探索問い合わせを探索エンジンに入力してそのユーザに関連した探索結果を得る。たとえば、日本のユーザは、探索問い合わせを入力して、日本語のWebサイトを含む探索結果を得る。このような問い合わせに回答して、従来の探索エンジンでは信頼性のない探索結果を返送しかねないが、それは、その探索問い合わせに対してユーザにとって関係のあるまたは役に立つ、ユーザのロケーションにしたがって探索結果をランク付けしたり得点したりするためのデータが比較的少ないからである。

【0004】

従来の探索エンジンでは、この探索エンジンにアクセスするためにどのようなタイプのWebブラウザアプリケーションが用いられるかによってユーザに関連する情報を決定することが可能である。たとえば、ユーザがWebブラウザアプリケーションをインターネットからダウンロードする際に、ユーザには、自身の好みの言語に基づいて特定のバージョン、たとえば日本語バージョンやフランス語バージョンのアプリケーションをダウンロードするオプションがある。ユーザがインターネットを介して探索エンジンにアクセスするためにフランス語バージョンのWebブラウザアプリケーションを用いると、この探索エンジンは、フランス語バージョンのWebブラウザアプリケーションを用いていることを検出すると、それだけでこのユーザはフランスにいるのであろうとしばしば判定する。

【0005】

他の従来の探索エンジンは、特定のユーザが探索問い合わせで用いた国ドメイン接尾辞によってロケーション情報を得る。たとえば、日本のユーザは、日本語バージョンの探索エンジンを要求する際に、ドメイン名接尾辞 ".com" の代わりに国ドメイン接尾辞 ".co.jp" を持つ探索エンジンのWebアドレスを入力する。このような入力に基づいて、探索エンジンは、ユーザは日本にいるらしいと判定する。

【0006】

探索エンジンが探索問い合わせに対して2つ以上の探索結果を返送すると、このような探索結果は、これら探索結果と関連する文書に対するリンクのリストとして表示される。ユーザは、これら探索結果の内の1つ以上と関連するWebサイトをブラウズして訪問して、このWebサイトがユーザの探索問い合わせと関連しているかどうか評価する。たとえば、ユーザはマウスや別の入力デバイス进行操作して、特定の探索結果のところで「クリック」して、その探索結果と関連するWebサイトを閲覧する。多くの例では、ユーザは、探索結果中に提供されているいくつかのWebサイトをブラウズして訪問して、これらいくつかのWebサイトの各々と関連するリンクのところでクリックしてから、ユーザの探索問い合わせに対応する有用なまたは関連する情報を突き止める。

【0007】

1つの探索結果集合と関連する複数のWebサイトに対する複数のリンクをクリックすると、時間がかかりかねない。探索エンジンで用いられるランク付けアルゴリズムを改善して、したがって、ユーザに対してより良好な探索結果を提供することが望ましい。

【0008】

概要

本発明の実施形態は、探索問い合わせに対する探索ランク付けを、この探索問い合わせと関連する母集団情報を用いることによって改善するシステムと方法について記載する。本発明の1つの態様は、探索問い合わせを受信するステップと、この探索問い合わせと関連する母集団を判定するステップとを含む。このような母集団はさまざまな方法で定義され、判定される。本発明の別の態様は、探索問い合わせと関連する項目(Webページなど)を判定するステップと、母集団と関連するデータに少なくとも部分的に基づいて項目のランク付け得点を判定するステップとを含む。母集団情報を用いるさまざまなアルゴリズムがこのようなシステムと方法で適用される。

【0009】

10

20

30

40

50

本発明に関する上記のおよび他の特徴、態様および長所は、次の添付図面を参照して以下の詳細な説明を読めばより良好に理解されるであろう。

【0010】

詳細な説明

本発明は、母集団情報を用いることによって探索ランク付けを改善する方法とシステムを含む。以下に、本文と添付図面とに解説されているように本発明の例示の実施形態を詳細に参照する。図面全体と以下に説明にわたって、同じ又は同様の部品を参照するのに同じ参照番号を用いる。

【0011】

本発明によるさまざまなシステムが構築される。図1は、本発明の例示の実施形態が内部で動作する例示のシステムを示す図である。本発明は他のシステムでも動作し、また、実施される。

【0012】

図1に示すシステム100は、複数のクライアントデバイス102a~102n、サーバデバイス104およびネットワーク106を含む。図示するネットワーク106はインターネットを含む。イントラネットなどの他のネットワークが用いられる実施形態もある。そのうえ、本発明による方法は1つのコンピュータで動作する。図示するクライアントデバイス102a~102nはその各々が、プロセッサ110にカップリングされている実施形態中に、ランダムアクセスメモリ(RAM)108などのコンピュータ読み取り可能媒体を含む。プロセッサ110は、メモリ108に記憶されているコンピュータ実行可能プログラム命令の集合を実行する。このようなプロセッサには、マイクロプロセッサ、ASICおよび状態機械が含まれる。このようなプロセッサは、プロセッサによって実行されると、本書に記載するステップをプロセッサに実行させる命令を記憶する媒体、たとえばコンピュータ読み取り可能媒体を含むまたはこれらと通信する。コンピュータ読み取り可能媒体の実施形態には、これに限られないが、タッチセンシティブ入力デバイスと通信するプロセッサなどのプロセッサに対してコンピュータ読み取り可能命令を提供することが可能な電子式、光式、磁気式または他の方式の記憶デバイスや送信デバイスが含まれる。適当な媒体の他の例には、これに限られないが、フロッピディスク、CD-ROM、磁気ディスク、メモリーチップ、ROM、RAM、ASIC、設定済みプロセッサ、あらゆる光媒体、あらゆる磁気テープもしくは他の磁気媒体または、コンピュータが命令を読み取ることが可能な他の何らかの媒体が含まれる。また、有線と無線の双方の方式のルータ、プライベートネットワークもしくはパブリックネットワークまたは他の送信デバイスもしくは送信チャネルを含む他のさまざまな形態のコンピュータ読み取り可能媒体がコンピュータに対して命令を送信または搬送する。これらの命令は、たとえば、C、C++、C#、Visual Basic、JavaおよびJavaScriptを含むなんらかのコンピュータプログラム言語によるコードを含む。

【0013】

クライアントデバイス102a~102nはまた、マウス、CD-ROM、キーボード、ディスプレイまたは他の入力デバイスもしくは出力デバイスなどの多くの外部デバイスや内部デバイスが含まれる。クライアントデバイス102a~102nの例としては、パソコン、デジタルアシスタント、携帯情報端末、携帯電話、モバイル電話、スマート電話、ポケベル、デジタルタブレット、ラップトップコンピュータ、プロセッサベースのデバイスならびに類似のタイプのシステムおよびデバイスがある。一般的に、クライアントデバイス102a~102nは、ネットワーク106に接続され、1つ以上のアプリケーションプログラムと対話するなんらかのタイプのプロセッサベースのプラットフォームである。図示するクライアントデバイス102a~102nには、マイクロソフト社のInternet Explorer(商標)のバージョン6.0、ネットスケープ通信社のNetscape Navigator(商標)のバージョン7.1およびアップルコンピュータ社のSafari(商標)のバージョン1.0などのブラウザアプリケーションプログラムを実行するパソコンが含まれる。

【0014】

クライアントデバイス102a~102nを介して、ユーザ112a~112nは、ネットワーク106上で互いに通信したり、ネットワーク106に接続されている他のシステムやデバイスと通信したりすることが可能である。ユーザ112a~112nは互いに異なったロケーション、国または領域に置くことが可能である。図1に示すように、サーバデバイス104もまたネットワーク106に接続されている。図示する実施形態では、ユーザ112a~112nは、クライアントデバイス102a~102nのところから探索問い合わせを生成して、ネットワーク106を介してサーバデバイス104に送信することが可能である。たとえば、1つの国にいるユーザ112aは、テキストによる探索問い合わせ114をクライアントデバイス102aに表示されているWebページの問い合わせフィールド中にタイプする。すると、クライアントデバイス102aは、この探索問い合わせ114を表す関連の探索問い合わせ信号126をネットワーク106を介してサーバデバイス104に送信する。

【0015】

図示するサーバデバイス104は、Google(商標)探索エンジンなどの探索エンジンアプリケーションプログラムを実行するサーバを含んでいる。クライアントデバイス102a~102nと類似して、図示するサーバデバイス104は、コンピュータ読み取り可能メモリ118にカップリングされているプロセッサ116を含んでいる。1つのコンピュータシステムとして表されているサーバデバイス104は、コンピュータプロセッサのネットワークとして実現される。サーバデバイス104の例には、サーバ、メインフレームコンピュータ、ネットワーキングされたコンピュータ、プロセッサベースのデバイスならびに類似のタイプのシステムおよびデバイスが含まれる。クライアントプロセッサ110とサーバプロセッサ116は、カリフォルニア州サンタクララ市のインテル社のプロセッサや、イリノイ州のシャウンブルグ市のモトローラ社のプロセッサなどの多くの周知のコンピュータプロセッサの内のどれでもよい。

【0016】

メモリ118は、探索エンジン124としても知られる探索エンジンアプリケーションプログラムを包含している。探索エンジン124は、ユーザ112a~112nからの探索問い合わせに応答して関連する情報を突き止める。

【0017】

サーバデバイス104または関連のデバイスは、前もってネットワーク106を探索してネットワーク106に接続されている他のデバイスやシステムのところに記憶されているWebページなどの項目を突き止めて、メモリ118や別のデータ記憶デバイス中のこの項目にインデックスを付けている。項目には、文書、たとえば、HTML、XML、XHTML、携帯式文書フォーマット(PDF)ファイルなどのさまざまなフォーマットを持つWebページならびに、ワープロ、データベースおよびアプリケーションプログラム文書ファイル、オーディオ、ビデオもしくは、(インターネットなどの)ネットワーク、パソコン、または他の計算手段もしくは記憶手段で利用可能である他のどのような情報でも含まれる。本書に説明する実施形態は一般に文書に関連して説明するが、どのようなタイプの項目に対しても実施形態は動作する。

【0018】

探索エンジン124は、探索問い合わせ114を表す関連の探索問い合わせ信号126に応答して、関連の情報または探索結果132の集合を、探索問い合わせ114の発信元であるクライアントデバイス102a~102nに対して返送する。

【0019】

図示する探索エンジン124は、文書ロケータ134とランク付けプロセッサ136と、母集団プロセッサ138とを含んでいる。図示する実施形態では、各々がメモリ118中に常駐するコンピュータコードを含んでいる。文書ロケータ134は、ユーザ112aからの探索問い合わせ114に対する応答である文書の集合を識別する。図示する実施形態では、これは、潜在的な探索問い合わせまたは探索条件にしたがってインデックス付

10

20

30

40

50

けされている文書のインデックスにアクセスすることによって遂行される。ランク付けプロセッサ 136 は、探索問い合わせ 114 および / または別のいずれかの判定基準に対する関連性に基づいて、突き止められた Web ページまたは文書の集合を含んでいる探索結果 132 をランク付けまたは得点付けする。母集団プロセッサ 138 は、ユーザ 112a ~ 112n と関連する母集団を表すまたは別様にこれに対して対応する母集団信号 128 などの母集団信号を判定または別様に測定する。なお、文書ロケータ 134、ランク付けプロセッサ 136 および母集団プロセッサ 138 の他の機能と特徴は以下にさらに説明する。

【0020】

サーバデバイス 104 もまた、図示する例中の母集団データベース 120 である母集団データ記憶エレメントや図示する例中の選択データデータベース 122 である選択データ記憶エレメントなどの他の記憶エレメントにアクセスする。図示する具体的な選択データベースはクリックスルー式データベースであるが、どのような選択データ記憶エレメントを用いてもよい。データ記憶エレメントには、配列、ハッシュテーブル、リストおよびペアを無制限に含むデータを記憶する方法の内のいずれか 1 つまたは組み合わせが包含される。他の類似したデータ記憶デバイスもサーバデバイス 104 からアクセス可能である。母集団データベース 120 は、ユーザ 112a ~ 112n が探索問い合わせを入力するとそれに関連する母集団情報を記憶する。ユーザ 112a ~ 112n と関連する母集団情報の例には、ユーザ 112a ~ 112n のロケーションに関する情報と、ユーザ 112a ~ 112n が関連する母集団についての情報と、ユーザ 112a ~ 112n が関連するグループについての情報が含まれる。

【0021】

ユーザのロケーションの例には、これに限られないが、大陸、領域、国、州または市が含まれる。1 例として、ユーザのロケーションは、フランス、ドイツ、日本および米国などの国として識別される。

【0022】

ユーザが関連するロケーションの例には、これに限られないが、性別、人口統計、民族的背景、大陸、領域、国、州、郡または市を含むことが可能である。1 例として、ユーザが関連する母集団は、「18 歳未満」、「18 ~ 24 歳まで」、「25 ~ 34 歳まで」、「35 ~ 49 歳まで」、「50 ~ 62 歳まで」および「63 歳以上」などのユーザの年齢幅で識別することが可能である。

【0023】

ユーザが関連するグループの例には、これに限られないが、性別、人口統計学的グループ、民族的グループ、特徴を共有する人物、興味を共有する人物および所定の選択によって分類された人物を含むことが可能である。1 例として、ユーザが関連することが可能なグループは、「古代の鯨の歯を収集することに興味を持つすべての人物」と「古代の鯨の歯には興味を持たないすべての人物」として識別することが可能である。

【0024】

母集団情報はまた、自己識別タイプデータまたは自動識別タイプデータを含むことが可能である。自己識別タイプデータには、これに限られないが、ユーザ登録データ、ユーザ嗜好データおよびユーザが選択した他のデータが含まれる。1 例として、自己識別データとは、ユーザがブラウザアプリケーションプログラムに入力する言語嗜好選択である。自動識別タイプデータには、これに限られないが、ユーザのロケーションのインターネットプロトコルアドレス、ユーザのブラウザアプリケーションプログラムから得られるデフォルトデータ、クッキーおよび、ユーザのアプリケーションプログラムが探索エンジンと対話しているときにこのアプリケーションプログラムから収集された他のデータが含まれる。1 例として、自動識別データは、インターネット上のユーザのネットワークアドレスのドメインを含むかまたは、ユーザのブラウザアプリケーションプログラムによって得られるまたはアクセスされる「クッキー」に記憶されている情報である。

【0025】

探索エンジン 124 は、母集団情報を判定するまたは別様に命令の集合を実行してユーザ 112a ~ 112n と関連する母集団情報を判定して、母集団タイプの情報を母集団データベース 120 に記憶する。代替例では、母集団プロセッサ 138 は母集団情報を判定するまたは別様に命令の集合を実行してユーザ 112a ~ 112n と関連する母集団情報を判定して、母集団タイプの情報を母集団データベース 120 に記憶する。

【0026】

本発明は、図 1 に示すアーキテクチャとは異なったアーキテクチャを有するシステムを含むことに注意すべきである。たとえば、母集団データベース 120 とプロセッサ 138 は探索エンジン 124 の一部ではなく、母集団データの修正や他の動作をオフラインで実行する本発明によるシステムがある。また、母集団プロセッサ 138 が文書ロケータ 134 または他のシステムの出力に影響するような実施形態もある。図 1 に示すシステム 100 は単に例示であり、図 2 ~ 3 に示す例示の方法を説明するために用いている。

10

【0027】

図示する実施形態では、母集団データベース 120 は、図 2 ~ 3 に示すような本発明の例示の方法を実行する以前に収集され記憶されたデータを包含する。図 1 をさらに参照すると、図示する母集団プロセッサ 138 は、探索問い合わせを入力し、この探索問い合わせの探索結果を選択するユーザのロケーションを獲得して分析することによって、ユーザ 112a ~ 112n と関連する母集団情報を記録する。たとえば、日本にいるユーザ 112a が「舟遊び」などの最初の探索問い合わせをタイプインすると、母集団プロセッサ 138 はユーザ 112a が日本にあるインターネットプロトコルアドレスから探索問い合わせを送信していること、また、ブラウザアプリケーションプログラムの日本語版を用いているものと判定する。さらに、ユーザ 112a が探索問い合わせに回答して探索エンジンが返送した特定の探索結果を選択すると、母集団プロセッサ 138 は、ユーザ 112a が特定の文書を日本語で選択するものと判定する。このように、日本にいるユーザ 112a が問い合わせ「舟遊び」を入力すると、日本にいるユーザ 112a に関連する "boat i n g . c o . j p " などの探索結果が返送される。他のタイプの母集団情報も本発明によって判定して母集団データベース 120 に記憶することが可能である。

20

【0028】

別の例として、母集団データベース 120 は、ユーザが母集団の亜母集団と関連しているという情報を記憶することが可能である。たとえば、母集団プロセッサ 138 は、世界におけるある領域であるヨーロッパにいるユーザ 112a がヨーロッパの国であるルクセンブルグから通信していると判定する。母集団データベース 120 はまた、ヨーロッパにいる別のユーザがヨーロッパの別の国であるフランスから通信しているという情報を記憶することが可能である。このように、各々のユーザは、「ヨーロッパ」という母集団と関連していると同時に、それぞれの亜母集団、すなわち、「ルクセンブルグ」と「フランス」と関連している。母集団と亜母集団は、これに限られないが、大陸、領域、国、州、郡、市、性別、人口統計学的グループ、民族的グループ、話す言語、ユニバーサルリソースロケータ、インターネットプロトコルアドレス、ドメイン名、インターネットサービスプロバイダ、グループ、特徴を共有する人物、興味を共有する人物および所定の選択によって分類された人物を含むことが可能である。ある母集団に対してさまざまなレベルの亜母集団が存在しえる。たとえば、「パリっ子」は「フランス」の亜母集団であり、「フランス」は「ヨーロッパ」の亜母集団であり、「ヨーロッパ」は「世界」の亜母集団である。亜母集団の情報は、フランスなどの特定のロケーションすなわち母集団からユーザが不十分な回数クリックすれば有用なものとなりえる。しかしながら、フランスは「ヨーロッパ」の亜母集団であるので、「ヨーロッパ」にいるユーザに対応するクリック情報は、フランスにいるユーザからの問い合わせに対するクリック情報を強化する目的で用いることが可能である。一般に、亜母集団に対するクリック情報がわかしくないまたは存在しない場合、より高いレベルの母集団からの情報をクリック情報を強化する目的で用いることが可能である。

30

40

【0029】

50

母集団情報と亜母集団情報を判定する際に、母集団プロセッサ１３８はまた、各々のタイプの情報に対して重み付けすることが可能である。たとえば、母集団プロセッサ１３８は、ユーザが特定の領域から通信していることに対して、ユーザが特定の国から通信していることに対するより少なく重み付けし、これで、後続の探索問い合わせに対するより良い探索結果が領域情報と国情報から得ることが可能となるようにすることが可能である。このように、ユーザが特定の領域（ヨーロッパ）から通信しているという情報を、ユーザが特定の国（フランス）から通信していることに対するよりも少なく重み付けすることが可能である。母集団タイプのデータに対する他のタイプの重み付けまたは類似の重み付けを本発明によって定義して母集団データベース１２０に記憶することが可能である。

【００３０】

10

図示する母集団データベース１２０は、特定の問い合わせに対するユーザロケーションのリストが包含されている。たとえば、「船遊び」という探索問い合わせに対して、「船遊び」という探索を入力したユーザの判定済みのロケーションなどの母集団情報は記憶されて、「船遊び」という探索問い合わせと関連付けられる。ユーザのロケーションは「フランス」や、「日本」や、「米国」であったりする。このようなロケーションが例示の目的のために用いられる。ロケーションの数が多かったり少なかったり、他の国や、ロケーションや、母集団や、亜母集団が用いることが可能な実施形態もある。

【００３１】

本発明のさまざまな実施形態によって実現される母集団データベースに記憶される情報の例には次のものがある：

20

【表１】

問い合わせ	ロケーション
Q1	日本、フランス、米国
Q2	ヨーロッパ、アジア、北アフリカ

30

【００３２】

上に示した表で、第１のカラムには固有の問い合わせがリストアップされており、第２のカラムにはそれに対応するユーザロケーションがリストアップされている。各々の問い合わせは、ユーザが入力した探索問い合わせを表している。対応するユーザロケーションは、それぞれの問い合わせを入力したユーザの判定済みのロケーションを表している。したがって、上に示した問い合わせ“Ｑ１”に対して、問い合わせ“Ｑ１”を前もって力したユーザの判明済みロケーションとして、「日本」、「フランス」および「米国」が含まれる。この表を本発明によって実現するには、探索エンジン１２４は、“Ｑ１”のような特定の問い合わせに対して“ロケーション”のカラム中のユーザの判明済みロケーションを見ればよい。ユーザが新たな問い合わせを入力すると、この新たな問い合わせは「問い合わせ」のカラムに挿入される。同様に、この新たな問い合わせに対するユーザのロケーションが判明したら、この新たなロケーションは「ロケーション」という名前の２番目のカラムに挿入される。

40

【００３３】

図示するクリックスルーデータベース１２２は、探索問い合わせ１１４に回答して突き止められた文書のリストからユーザが選択した探索結果などの探索エンジン１２４からユーザが選択した探索結果１３２と関連するデータを記憶する。たとえば、ユーザ１１２ａは、クライアントデバイス１０２ａ～１０２ｎから、マウスや別の入力デバイスを操作して、探索結果１３２のＷｅｂページまたは文書と関連する１つ以上のＵＲＬをクリックして入力情報を入力する。ユーザの「クリック」は、探索結果１３２中で突き止められた文

50

書をユーザが選択することによって発生する。ある文書をこのように「クリック」するとそれは、特定の探索問い合わせに対して返送される探索結果中にこの文書が存在することと関連する選択としてクリックスルーデータベース122に記憶される。探索結果中からユーザが文書を選択するとそれに関連するこのような他の多くの選択タイプのデータもそこに記憶される。

【0034】

ユーザがクリックしたものは「クリックスルー」データと呼ばれる。本実施形態では、探索エンジン124は、クリックスルーデータと、探索結果132中で突き止められた文書の各々と関連する他のデータとを測定して記憶する。

【0035】

クリックスルーデータは一般には探索結果における品質のインジケータである。品質信号またはクリックスルーデータには、これに限られないが、特定のURLもしくは文書が特定のユーザによってクリックされたかどうか、URL、文書もしくはWebページが一人以上のユーザによってどれほど頻繁にクリックされるか、特定のユーザが特定の文書やWebページをどれほど頻繁にクリックするかを含むことが可能である。ユーザ入力や観察タイプのデータなどのクリックスルーデータに類似した他のタイプの品質信号を、クリックスルーデータベース122や類似のデータ記憶デバイスで記憶することが可能である。

【0036】

クリックスルーデータベース122や他のデータ記憶デバイスに記憶することが可能な探索結果132中で突き詰められた文書に関連する他のデータには、これに限られないが、特定のURL、文書またはWebページがどれほどしばしば探索問い合わせ114に回答して示されたか、特定のロケーションのユーザ112a~112nによって特定の探索問い合わせ114が何回依頼されたか、特定の母集団のユーザ112a~112nによって特定の探索問い合わせ114が何回依頼されたか、特定のロケーションのユーザ112a~112nによって特定の文書が何回選択されたか、特定の母集団のユーザ112a~112nによって特定の文書が何回選択されたか、特定の探索問い合わせ114に対して特定の文書がユーザ112a~112nによって何回選択されたか、特定の文書がネットワーク106上に通知された年代または時間およびネットワーク106における特定の文書のソースのアイデンティティを含むことが可能である。

【0037】

母集団データベース120からの母集団情報と選択データベース(クリックスルーデータベース122として示されている)からの選択データとは、母集団プロセッサ138によって処理して後で使用されるように記憶することが可能である。たとえば、母集団プロセッサ138は特定の探索問い合わせのためのクリックスルーデータを検索する。特定の探索問い合わせ用のこのクリックスルーデータはユーザのロケーションに基づいて割り当てられる。探索エンジン124は、母集団データベース120にコールして、特定の探索問い合わせを入力してこの問い合わせの探索結果の文書を選択するすべてのユーザのロケーション情報を求める。母集団プロセッサ138が、日本、フランスおよび米国という3つのロケーションから通信するユーザが特定の問い合わせ114のための選択データを提出したと判定すると、各々のロケーションからのユーザの各々の集合に対するそれぞれの記号表示を母集団プロセッサ138で定義することが可能である。したがって、日本から通信するユーザを“J”と記号表示し、フランスからのユーザを“F”と記号表示し、米国からのユーザを“US”と記号表示することが可能である。次に、母集団プロセッサ138は、特定の文書集合に対してクリックスルーデータベース122によって収集されたクリック動作の数をそれぞれの記号表示に対してユーザロケーションに基づいて割り当てる。

【0038】

本発明の実施形態によって実現される母集団データベースに記憶される情報の1例を以下に示す。

10

20

30

40

50

【表 2】

問い合わせ	文書	すべての クリック動作 の総数	日本	フランス	米国
Q	D ₁	101	1	20	80
Q	D ₂	207	2	5	200
...	...	...	...	...	...
Q	D _i	#(Q,D _i ,A)	#(Q,D _i ,J)	#(Q,D _i ,F)	#(Q,D _i ,U)
...	...	...	...	...	...
Q	D _N	#(Q,D _N ,A)	#(Q,D _N ,J)	#(Q,D _N ,F)	#(Q,D _N ,U)
Q	D _{total}	1500	100	300	1100

10

【0039】

20

図示する母集団データベース120に対して上記された例では、探索問い合わせ“Q”に対して、文書“D₁”に対するユーザのクリック総数は“101”であった。日本にいるユーザの文書“D₁”に対するユーザクリック総数は“1”であり、フランスにいるユーザの文書“D₁”に対するユーザクリック総数は“20”であり、米国にいるユーザの文書“D₁”に対するユーザクリック総数は“80”であった。

【0040】

図1に示す実施形態では、サーバ104は母集団データベース120およびクリックスルーデータベース122と通信している。サーバ104はあるプロセスを実行するが、このプロセスによって、2つのデータベース120と122中のデータが用いられて、ユーザ112aからの探索問い合わせ114に応答して提供される探索結果が改善される。

30

【0041】

本発明にしたがってさまざまな方法が実行される。本発明による1つの例示の方法は：探索問い合わせを受信するステップと、この探索問い合わせと関連する母集団を決定するステップと、この探索問い合わせと関連する（Webページなどの）項目を決定するステップと、この項目のランク付け得点をこの母集団と関連するデータに少なくとも部分的に基づいて決定するステップと、を含む。このように、この母集団情報を用いて、項目のランク付け得点に影響を与える。この母集団情報は、探索問い合わせおよび/または項目に対して決定された母集団（たとえばグループや亜母集団）の動作を示すのが望ましい。たとえば、この情報は、同じの、類似したまたは関連した問い合わせに対する同じの、類似したまたは関連した母集団によって選択された好ましい項目を示す。

40

【0042】

探索問い合わせと関連する第2の項目、第3の項目、第4の項目などのランク付け得点もまた、母集団と関連した情報に少なくとも部分的に基づいて決定される。次に、これらの項目はランク付け得点に基づいて互いに対してランク付けされて、探索問い合わせを提出する人物に対してランク付けされた順序で提示される。この結果、最も関連する項目をユーザに対して最初に提供するようなランク付けとなることが望ましい。

【0043】

問い合わせと関連する母集団はさまざまな母集団の内の1つ以上であってよい。例としては、これに限られないが、年齢、年齢層、性、人種、第1言語、第2言語、ロケーション、収入、収入層、大陸、領域、国、州、郡、市、性別、民族的グループ、グループ、特

50

徴を共有する人物、興味を共有する人物および所定の選択によって分類された人物などの人口統計学的データと、インターネットサービスプロバイダデータ（またはこれらの内のいずれかに対するありそうなもしくは考えられるデータ）が含まれる。言い換えれば、母集団は、選択されたなんらかの特徴によって決定されたなんらかのグループである。

【0044】

探索問い合わせと関連する母集団は、さまざまな方法の内の1つ以上の方法で決定される。たとえば、探索問い合わせの送出者と関連する人口統計学的データは、目的とする母集団を決定するために決定され、また、このデータは上記の母集団または他の母集団の内の1つ以上であってよい。たとえば、探索問い合わせの送出者のありそうな地理的ロケーションは、探索問い合わせの送出元であるインターネットプロトコルアドレス、送出者が探索エンジンにアクセスするために入力したアドレスまたは送出者が入力した人口統計学的データを特定することによって判定される。探索問い合わせと関連する母集団もまた、探索問い合わせと関連する人口統計学的データを判定するなどの他の方法で判定される。この動作は、たとえば、探索問い合わせの言語を判定したり、探索問い合わせの前の送出者と関連するデータを判定したりすることによって遂行される。

10

【0045】

これから分かるように、探索問い合わせと関連する母集団を判定するステップは、自己識別データ、自動識別データもしくは他のデータまたは、ユーザ登録データ、ユーザ嗜好データおよびユーザ選択データなどの探索問い合わせを送信するユーザに関連する他のデータや情報を判定するステップを含むことがある。たとえば、会員資格を登録したりWebサイトにアクセスしたりする際に、ユーザは登録情報を入力する。ユーザはこのような登録プロセス中に嗜好（たとえば好みの言語）を示したり他の方法で嗜好（Webページやドメインの使用の選択など）を示したりする。自動識別データには、たとえば、IPアドレス、ドメイン、得られたデフォルトデータ、ユーザと関連するアプリケーションまたは他の自動的に調達されたもしくは自動的に提供された情報が含まれる。もちろん、自己識別データと自動識別データ間にはいくぶんかの重複が存在する可能性がある。

20

【0046】

母集団と関連するデータ自身も1つ以上の方法で判定される。たとえば、母集団という文脈におけるある項目の選択得点が判定される。例証として、当面の探索問い合わせと関連する母集団のある数のメンバーが、以前に当の項目をクリックしていると判定される。この選択得点は、母集団のメンバーが当の項目に対して持っている相対的な興味を示している。

30

【0047】

また、母集団情報を含むことが可能なさまざまなデータが存在する。他の例としては、母集団のメンバーの数、母集団の内で探索問い合わせに対して以前に返送された結果を選択したメンバーの数、母集団の内で探索問い合わせを入力したメンバーの数、母集団の内で探索問い合わせに対する探索結果が示された相手のメンバーの数、選択総得点、および母集団の内で項目を選択したメンバーの総数が含まれる。他にも多くの例がある。

【0048】

一部の例では、探索問い合わせと関連する2つ以上の母集団を、探索結果を改善するために決定されて用いられる。1つの実施形態では、たとえば、探索問い合わせと関連する第2の母集団を決定するが、項目のランク付け得点の決定はさらに、第2の母集団と関連するデータに少なくとも部分的に基づいて実行される。

40

【0049】

本発明の実施形態の上記の態様および他の態様を以下にさらに説明する。本発明の実施形態のこれらの例示の態様は、探索結果を改善するために繰り返されたり反復されたりする。そのうえ、本発明による方法で実施される上記のステップおよび他のステップは、プロセッサと結合しているメモリーやディスクなどのコンピュータ読み取り可能媒体または他のコンピュータ読み取り可能媒体中にプログラムコードという形態で記憶される。

【0050】

50

図 2 ~ 3 に、本発明による例示の方法 200 を詳細に示す。この例示の方法は例として提供するものであり、本発明による方法を実行するさまざまな方法がある。図 2 に示す方法 200 は、さまざまなシステムの内のいずれかによって実行したり別様に遂行したりすることが可能である。方法 200 を、図 1 に示すシステム 100 によって実行する様子を例として以下に説明するが、システム 100 のさまざまな部品を、図 2 ~ 3 の例示の方法を説明するに際して参照する。図示する方法 200 は、母集団情報を用いる探索ランク付けを改善するものである。

【0051】

図 2 ~ 3 に示す各々のブロックは、例示の方法 200 で実行される 1 つ以上のステップを表している。図 2 を参照すると、ブロック 202 において、例示の方法 200 が開始される。ブロック 202 の次にブロック 204 が実行されるが、ここで、母集団データベース 120 が提供される。これは、たとえば、このようなデータベースを構築したりこのようなデータベースとの通信を確立したりすることによって遂行される。図 1 を参照して説明したように、母集団データベース 120 は、探索問い合わせ 114 および他の探索問い合わせに対する探索結果中で選択された文書の母集団タイプの情報を記憶する。

10

【0052】

ブロック 204 の後に続いてブロック 206 が実行されるが、ここで、選択データベース、この場合ではクリックスルーデータベース 122 が提供される。これは、たとえば、このようなデータベースを構築したりこのようなデータベースとの通信を確立したりすることによって遂行される。図 1 を参照して説明したように、クリックスルーデータベース 122 は、ユーザ 112 a からの探索問い合わせ 114 に応答して提供された、文書のリストや Web ページなどの、探索エンジン 124 によって提供された探索結果 132 に対するユーザのクリックや入力と関連するデータを記憶する。

20

【0053】

ブロック 206 の後に続いてブロック 208 が実行されるが、ここで、探索信号という形態を持つ探索問い合わせがサーバによって受信される。図示する実施形態では、ユーザ 112 a はクライアントデバイス 102 a から探索問い合わせ 114 を発生する。クライアントデバイス 102 a、探索問い合わせ 114 を表す関連の探索問い合わせ信号 216 を、ネットワーク 106 を介してサーバデバイス 104 に送信する。探索エンジン 124 は、探索問い合わせ信号 126 を受信して、探索問い合わせ 114 を処理する。たとえば、ユーザ 112 a が探索問い合わせ「舟遊び」をブラウザウアプリケーションプログラムの探索ページの探索フィールドまたは問い合わせフィールドにタイプインすると、クライアント 102 a は、「舟遊び」というテキストまたは「舟遊び」を他に表示したり指示したりするものを含む探索問い合わせ信号 126 を送信する。探索エンジン 124 は、信号 126 を受信して、「舟遊び」が所望の探索問い合わせ 114 であると判定する。

30

【0054】

ブロック 208 の後に続いてブロック 210 が実行されるが、ここで、項目データ、図示する場合では、文書データが判定されて受信される。図示する実施形態では、このブロック 210 で、探索エンジン 124 は、以前にネットワーク 106 からインデックス付けされたことのある探索データベース（図示せず）またはメモリー 118 中の関連文書に対する探索を実行する。探索エンジン 124 は、ユーザ 112 a からの探索問い合わせ 114 を表す探索問い合わせ信号 126 に応答して文書データを探索データベースまたはメモリー 118 から受信する。文書データはまた、探索問い合わせ 114 に対する初期探索結果とも呼ばれる。文書データには、これに限られないが、文書や Web ページに対するリンクを提供するユニバーサルリソースロケータ（URL）または、文書もしくは Web ページを、ネットワーク 106 を介してユーザ 112 a によってそこから検索したり別様にアクセスしたりすることが可能なロケーションを含むことが可能である。文書データは、本明細書の本文全体を通じて、「文書」と呼ばれる場合があることに注意すべきである。代替例では、文書ロケータ 134 は、探索問い合わせ 114 を表す探索問い合わせ信号 126 に応答して文書データを獲得または別様に受信する。

40

50

【 0 0 5 5 】

たとえば、図示するブロック 2 1 0 では、図示する探索エンジン 1 2 4 は、「舟遊び」という探索問い合わせに応答して文書のリストを決定する。この文書リストは決定された文書データを含む。「舟遊び」に対する最初の探索結果リストは、1 5 の文書からなるリストを含むことがありえる。図示する実施形態では、文書データのこの初期決定は、従来のように探索エンジンに問い合わせして結果を返送することによって実施される。

【 0 0 5 6 】

ブロック 2 1 0 の後に続いてブロック 2 1 2 が実行されるが、ここで、目的とする項目毎に、母集団信号を決定する。図示する実施形態では、探索エンジン 1 2 4 は、ブロック 2 1 0 で決定された初期探索結果リストの各々の文書に対して母集団関数を用いて母集団信号 1 2 8 を生成する。たとえば、母集団信号は、ブロック 2 1 0 で決定された初期探索結果の中で特定された 1 5 の文書の各々に対して決定される。

10

【 0 0 5 7 】

母集団信号は目的とする項目のランク付けや得点を示しているが、このランク付けや得点は、探索ユーザと同じ母集団グループ中のユーザの相対的興味を表している。たとえば、ユーザ 1 1 2 a の入力と同じ問い合わせ “ Q ” を実行する際に問い合わせ側のユーザ 1 1 2 a と同じ母集団グループのユーザによって以前に選択された項目は、ユーザがメンバーとはなっていない母集団グループのユーザにしか以前に選択されていない項目よりも高い得点の母集団信号を受信する。しかしながら、単に 1 例であり、多くの変更例や置換例が用いられる。母集団信号に表されているこのランク付けや得点は、単独でまたは他の得点信号やランク付け信号と組み合わせられて用いられて、文書を得点したりランク付けしたりし、また、文書のグループをランク付けして比較して、たとえば、ユーザによって送られた問い合わせに対する探索結果を提供する。

20

【 0 0 5 8 】

図示する実施形態では、母集団信号は、母集団信号関数によって決定される。この母集団信号関数は、母集団信号を 1 つ以上の変数に基づいて計算するアルゴリズムを含んでいる。図示する実施形態における母集団信号は、母集団プロセッサ 1 3 8 によって処理される命令の集合を含んでいる。このアルゴリズムはメモリー 1 1 8 に記憶される。

【 0 0 5 9 】

さまざまな母集団信号関数の内のいずれかの 1 つ以上の関数が本発明のさまざまな実施形態で実現される。母集団信号関数に含まれる変数の例には、無制限に、次のものの内の 1 つ以上が含まれる：

30

- ・ 問い合わせ “ q ” に対する項目 “ d ” の選択総得点（たとえば、問い合わせ “ q ” の文脈で探索結果で返送された文書 “ d ” に対してあらゆるユーザがクリックした合計回数または文書 “ d ” を選択した母集団のメンバーの総数）、
- ・ 問い合わせ “ q ” に対する項目 “ d ” の選択得点間の関係の選択されたまたは計算された重み（たとえば、探索問い合わせ “ q ” の文脈における探索結果で返送された文書 “ d ” に対する母集団 “ p o p ” のすべてのユーザによるクリック動作の総数）、
- ・ 母集団 “ p o p ” のメンバーの文脈における問い合わせ “ q ” に対する文書 “ d ” の選択得点（たとえば、問い合わせ “ q ” に対する探索に対して文書が返送されたときの母集団のメンバーによる文書に対するクリック回数）、
- ・ クリック信号を信用するために必要とされるデータの量を表す平滑係数（たとえば、問い合わせ “ q ” に対する探索結果で返送された文書 “ d ” に対するあらゆるユーザのクリック回数の信頼性または信用を表す係数）、
- ・ 母集団に関係なくあらゆるユーザの文脈における問い合わせ “ q ” に対する文書 “ d ” の選択得点（たとえば、文書 “ d ” に対するクリック回数）、
- ・ 母集団に関係なくあらゆるユーザによる文書の集合の選択総得点（たとえば、問い合わせ “ q ” に対して返送されたあらゆる文書に対するクリック回数）、
- ・ 母集団 “ p o p ” のあらゆるユーザによる文書の集合の選択得点（たとえば、問い合わせ “ q ” に対して返送されたあらゆる文書に対するクリック回数）、

50

- ・ いずれかの問い合わせ “ q_i ” に対する母集団 “ pop ” のあらゆるユーザによる文書 “ d ” の選択総得点（たとえば、文書 “ d ” に対するクリック回数）、
- ・ いずれかの問い合わせ “ q_i ” に対する母集団に関係なくあらゆるユーザによる文書 “ d ” の選択得点（たとえば、文書 “ d ” に対するクリック回数）、
- ・ 母集団 “ pop ” のユーザによって問い合わせ “ q ” が入力された回数、
- ・ 母集団に関係なくあらゆるユーザによって問い合わせ “ q ” が入力された回数、
- ・ 母集団のメンバーの数（たとえば、特定の探索問い合わせを入力した、特定の探索問い合わせに対して返送された結果を選択した、または特定の探索問い合わせに対する探索結果を表示された母集団のメンバーの数）、
- ・ 考慮されている項目の、母集団に基づいた 1 つ以上の他のランク付け係数もしくは得点および / または他の係数。

10

【 0 0 6 0 】

他にもさまざまな変数が含まれるが、これらの変数は単なる例である。そのうえ、これらの変数および他の変数は、指定された時間期間、指定されたユーザ数、ある母集団中の自己識別されたまたは自動識別されたユーザの数によって、問い合わせ “ Q ” を入力したユーザによって、または他の制限や改良によって制限または定義される。母集団データと関連する変数、制限、定義または他のデータは一般に、母集団情報または母集団データと呼ばれる。

【 0 0 6 1 】

母集団信号関数の例を次に示す：

20

【 数 1 】

$$(1) \quad S(q, d_j) = \frac{\#(q, d_j, P) + \mu \left[\frac{\#(q, d_j)}{\sum_{i=1}^N \#(q, d_i) + \mu} \right]}{\sum_{i=1}^N \#(q, d_i, P) + \mu}$$

30

【 0 0 6 2 】

ここで：“ $S(q_i, d_j)$ ” は探索問い合わせ “ q ” に対する文書 “ j ” の、特定の母集団 “ P ” のユーザの母集団情報とクリックスルーデータとに基づいた計算された得点であり、

“ $\#(q, d_j, P)$ ” は探索問い合わせ “ q ” に対して母集団 “ pop ” のユーザによって文書 “ d ” がクリックされた回数であり、

“ μ ” は、問い合わせ “ q ” に対する母集団 “ P ” のユーザによるクリック数などのクリック信号を信用するに必要なデータの量を表す平滑係数であり、

“ $\#(q, d_j)$ ” は、問い合わせ “ q ” に対する母集団に関係なくあらゆるユーザによって文書 “ d ” がクリックされた回数であり、

40

“ $\#(q, d_i)$ ” は、問い合わせ “ q ” に対する文書 “ i ” に対するユーザがクリックした総数であり、

“ $\#(q, d_i, P)$ ” は、問い合わせ “ q ” に対する文書 “ i ” の母集団 “ P ” のあらゆるユーザのユーザクリック総数である。

【 0 0 6 3 】

母集団信号の別の例は次の通りである：

【数 2】

$$(2) \quad S(q, d_j) = \frac{\#(q, d_j, P) + \mu \left[\frac{\#(q, d_j)}{\sum_{i=1}^N SH(q, d_i) + \mu} \right]}{\sum_{i=1}^N SH(q, d_i, P) + \mu}$$

10

【0064】

ここで：“SH(q、d_i)”は、問い合わせ“q”に対して文書“i”が表示された回数であり、

“SH(q、d_i、P)”は、母集団“P”の場合に文書“i”が問い合わせ“q”に対して表示された回数であり、

他の変数は例(1)に関連して説明する。

【0065】

母集団信号の別の例は次の通りである：

【数 3】

$$(3) \quad S(q, d_j) = \frac{\#(q, d_j, P) + \mu_1 \left[\frac{\#(q, d_j)}{\sum_{i=1}^N \#(q, d_i) + \mu_2} \right]}{\sum_{i=1}^N \#(q, d_i, P) + \mu_1}$$

20

【0066】

ここで：“μ₁”は、問い合わせ“q”に対する母集団“P”のユーザによるクリック数などのクリック信号を信用するために必要とされるデータの量を表す平滑係数であり、

“μ₂”は、問い合わせ“q”に対するあらゆるユーザのクリック数などのクリック信号を信用するために必要とされるデータの量を表す平滑係数であり、

他の変数は例(1)および/または例(2)に関連して説明する。

【0067】

解説目的で、例(1)のアルゴリズムは図2～3に示す本発明による例示の方法として実現されている。(1)、(2)および(3)に示す例のほかに他のアルゴリズムが本発明にしたがって用いられるが、アルゴリズム(1)は例を解説するために提供されている。このような他のアルゴリズムは例(1)、(2)および(3)に示す変数の内の一部を
40

【0068】

図3に、例(1)による図2に示す方法200を実行するサブルーチン212の例を示す。図示するサブルーチン212は、初期探索結果132中で受信された各々の文書に対して母集団信号128を提供する。他の実施形態では、このように分析される文書の数
は、受信したすべての文書の数より少ない値に制限される。サブルーチン212の例は次の通りである。

【0069】

図3を参照すると、例示のサブルーチン212はブロック300から開始される。ブロック300で、探索エンジン124と関連しているカウンタは“1”などの値に設定され
50

る。たとえば、母集団プロセッサ 1 3 8 は、関連しているメモリー 1 1 8 中の変数 “ i ” を “ 1 ” という初期値に設定することが可能である。カウンタすなわち変数 “ i ” は、サブルーチン 2 1 2 によって処理された文書の数のカウントするが、“ i ” の現在値は、文書データ中の文書のリスト上のどの文書が操作中であることを表している。

【 0 0 7 0 】

ブロック 3 0 0 の後に続いてブロック 3 0 2 が実行されるが、ここで、探索エンジン 1 2 4 と関連する別のカウンタが “ 1 ” などの値に設定される。たとえば、母集団プロセッサ 1 3 8 は関連するメモリー 1 1 8 中の変数 “ j ” を “ 1 ” という初期値に設定することが可能である。カウンタすなわち変数 “ j ” はサブルーチン 2 1 2 によって処理された文書のカウントするが、“ j ” の現在値は、文書データ中の文書のリスト上のどの文書が操作中であることを表している。

10

【 0 0 7 1 】

ブロック 3 0 2 の後に続いてブロック 3 0 4 が実行されるが、ここで、問い合わせ側のユーザ 1 1 2 a に対して母集団が決定される。探索エンジン 1 2 4 は母集団、この場合は “ P ” を決定する。たとえば、探索エンジン 1 2 4 は、探索問い合わせ “ Q ” を入力するユーザと関連する母集団を決定する。以前に述べたように、自己識別データまたは自動識別データまたは双方の組み合わせを探索エンジン 1 2 4 が利用して、ユーザと関連する母集団を決定することが可能である。本実施形態では、探索エンジン 1 2 4 は、ユーザのインターネットプロトコルアドレスから、ユーザは、母集団 “ F ” と記号表示されるフランスから通信しているらしいと判定し、したがって、探索エンジン 1 2 4 は、問い合わせ “ Q ” に対して以前に文書 “ D₁ ” を選択したフランスから通信しているユーザの母集団情報をクロスリファレンスする。探索エンジン 1 2 4 は、問い合わせ側ユーザの母集団 “ P ” を決定したら、母集団データベース 1 2 0 にコールして、対応する必要な限りのあらゆる母集団情報を得る。たとえば、探索エンジン 1 2 4 は、母集団データベース 1 2 0 からデータを検索して、フランス、すなわち母集団 “ F ”、にいる合計で “ 2 0 ” のユーザが以前に、探索問い合わせ “ Q ” に対して特定の文書 “ D₁ ” を選択したことがあるものと判定する。したがって、特定の母集団 “ P ” の母集団情報を探索エンジン 1 2 4 が適用して、本発明にしたがって文書 “ D₁ ” に対する母集団信号を計算することが可能である。

20

【 0 0 7 2 】

ブロック 3 0 4 の後に続いてブロック 3 0 6 が実行されるが、ここで、分析される文書数が決定される。ブロック 2 1 0 では、サーバ 1 0 4 は探索問い合わせ 1 1 4 と関連する文書データを検索した。決定されたデータの中には、探索問い合わせ 1 1 4 に応答して文書リストに載っている文書の合計数があった。

30

【 0 0 7 3 】

この文書数は変数 “ N ” によって特徴付けられる（また、“ N ” として設定される）。たとえば、すでに述べたように、探索問い合わせ「舟遊び」に対する探索結果は 1 5 の文書を有しており、サーバ 1 0 4 は “ N ” を “ 1 5 ” という値に設定する。

【 0 0 7 4 】

代替実施形態では、特定の問い合わせに対して母集団データベース 1 2 0 または別のデータ記憶デバイスで定義されているまたは別様に記憶されている探索問い合わせに対するいかなる文書合計数も、探索エンジン 1 2 4 または母集団プロセッサ 1 3 8 に対して送信するまたはこれらによって別様に判定することが可能であることに注意すべきである。さらに注意すべきは、特定の探索問い合わせに対する文書数は、母集団データベース 1 2 0 に前もって記憶されている母集団タイプの情報およびクリックスルーデータベース 1 2 2 に記憶されているクリックスルーデータまたは他のデータ記憶デバイスに記憶されている他の類似したタイプのデータによって異なることがあるということである。

40

【 0 0 7 5 】

ブロック 3 0 6 の後に続いてブロック 3 0 8 が実行されるが、ここで、変数 “ M ” が決定される。変数 “ M ” は、母集団信号が決定される文書数を表している。たいていの場合、“ M ” は変数 “ N ” と同じ値、すなわち、ブロック 3 0 6 で決定された文書数を有す

50

るが異なることもある。たとえば、すでに述べたように、探索問い合わせ「舟遊び」に対する探索結果は15の文書を有することがありえるが、この場合、サーバは変数 $N = 15$ と設定しており、サーバ104もまた“M”を“15”という値に設定する。

【0076】

ブロック308の後に続いてブロック310が実行されるが、ここでは、平滑係数が決定される。探索エンジン124は、特定の問い合わせに対する特定の母集団のユーザのユーザクリックなどのクリック信号を信用するために必要とされるデータの量を表す平滑係数を決定する。たとえば、母集団プロセッサ138は所定の方程式またはコンピュータ読み取り可能命令の集合を利用して、すべてのユーザまたは特定の問い合わせに対する特定の母集団のユーザのユーザクリックのクリック信号または品質を信用するために必要とされるデータの量を明らかにする平滑係数を決定する。この平滑係数は、特定の母集団のユーザのユーザクリックが非常に少ない場合や特定のソースのユーザクリックが信頼性がないまたは別様に信用するに足らないことが周知であるかまたは別様に感知されるような場合に特には有用である。この場合、平滑係数は、一定値または特定の母集団に関連するデータの重みまたは値に別様に影響を与えることが可能な値に設定することが可能である。ほとんどの場合、平滑係数は、母集団信号関数または母集団プロセッサ138によって処理されるコンピュータ読み取り可能命令の集合に適用される。

10

【0077】

たとえば、問い合わせ「舟遊び」に対して、母集団プロセッサ138は、フランスにいるユーザからのユーザクリックの数が不十分である場合に平滑係数を決定する。したがって、上記の表の例の場合には、“Q”が「舟遊び」でありフランスにいるユーザによる20のユーザクリックでは信頼するに足るクリック数としては不十分であれば、平滑係数が決定される。この平滑係数は、サブルーチン212中の上記の母集団信号関数中では“ μ ”で表される。この係数は、探索問い合わせ「舟遊び」に対するフランス在住のユーザによるクリックの数に対する信頼性または感知された信用を示す。

20

【0078】

別の例として、フランス在住のユーザが問い合わせ「クリケット」を発すると、フランス在住のユーザによるユーザクリックの数が不十分であれば、母集団プロセッサ138は平滑係数を決定する。しかしながら、フランスは「ヨーロッパ」の亜母集団であるので、母集団プロセッサ138は「ヨーロッパ」在住のユーザに対応するクリック情報を用いて、フランス在住のユーザからの問い合わせ「クリケット」のクリック情報を強化することが可能である。一般に、亜母集団のクリック情報がまばらであるまたは存在しない場合、より高い母集団レベルからの情報を用いて、クリック情報を強化することが可能である。一部の例では、このようにして用いることが可能な追加のレベルの母集団または亜母集団、すなわち、「パリっ子」は「フランス」の亜母集団であり、「フランス」は「ヨーロッパ」の亜母集団であり、「ヨーロッパ」は「世界」の亜母集団である、が存在する。

30

【0079】

場合によっては、一般的な平滑係数を決定することが可能であることに注意すべきである。探索エンジン124は、特定の問い合わせに対する特定の文書へのすべてのユーザによるユーザクリックに対する信頼性または信用を表す一般的平滑係数を決定する。図示する例では、ローカル母集団のユーザのクリックに対する信頼性または信用は一般的母集団のユーザのクリックに対する信頼性または信用と同じであり、上述したように決定された平滑係数を一般的平滑係数として用いることが可能であると仮定される。

40

【0080】

さらなる例として、一般的平滑係数は次のように決定することが可能である。母集団プロセッサ138は母集団データベース120または他のデータ記憶デバイスにアクセスして、ユーザクリックデータを検索する。所定の方程式またはコンピュータ読み取り可能命令の集合を用いて、母集団プロセッサ138は、特定の問い合わせに対する探索結果中のすべての文書に対するすべてのユーザによるユーザクリックの数を決定する。これで、一般的平滑係数を、特定の問い合わせに対する探索結果中のすべての文書に対するすべての

50

ユーザによるクリックまたは選択の数として設定することが可能である。

【0081】

特定の問い合わせに対する特定の文書の平滑係数、値または得点が母集団データベース120に前もって記憶されていれば、母集団プロセッサ138は、特定の問い合わせに対する特定の文書の平滑係数、値または得点を検索する。たとえば、母集団データベース120は、日本在住のたった一人のユーザしか問い合わせ“Q”に対して文書“D₁”を選択しなかった上記の表での場合のように、探索問い合わせ「舟遊び」に対する探索結果中の文書に対するすべてのユーザによるクリックの数が不十分であることを示したりする。この例では、探索問い合わせ「舟遊び」に対する探索結果中のすべての文書に対するすべてのユーザによるクリックの総数を用いるべきである。こうすれば、上述したように決定された一般的母集団平滑係数をアルゴリズム中で用いて、2つの結果の間を徐々に遷移することが可能である。探索問い合わせ「舟遊び」に対する探索結果中の文書に対するすべてのユーザによるクリックの数を用いるか、母集団情報と選択データとから改善された探索結果が得られるように探索問い合わせ「舟遊び」に対する探索結果中のすべての文書に対するすべてのユーザによるクリックの総数を用いるかを判断することが可能である。

10

【0082】

ブロック310の後に続いてブロック312が実行されるが、ここで、特定の問い合わせ“Q”に対する目的とする現行の文書（文書“i”）の場合の、特定の母集団“P”による選択数、この場合はクリック数が判定される。探索エンジン124は、文書“i”が問い合わせ“Q”に対する探索結果中で返送されると、文書“i”に対するクリック数を判定する。たとえば、母集団プロセッサ138は、母集団データベース120または他のデータ記憶デバイスに記憶されている母集団情報と、クリックスルーデータベース122に記憶されているもしくは母集団データベース120と前もって共有されているクリックスルーデータとにアクセスする。母集団プロセッサ138は所定の方程式またはコンピュータ実行可能命令の集合を母集団情報とクリックスルーデータとの一部またはすべてに適応して、“#(q、d_i、P)”とも呼ばれる探索問い合わせに対する文書への母集団によるクリックの数を決定する。たとえば、問い合わせ“Q”が「舟遊び」であれば、問い合わせ“Q”に対する文書“D₁”へのフランスでのユーザによるユーザクリック総数が20であったと判定することが可能である。

20

【0083】

ブロック312の後に続いてブロック314が実行されるが、ここで、特定の問い合わせ（“Q”）に対する目的とする現行の文書（文書“i”）に対するすべてのユーザによる選択数、この場合ではクリック数、を判定する。探索エンジン124は、文書“i”が問い合わせ“Q”に対する探索結果中で返送されると、文書“i”に対するすべてのユーザによるクリック数を判定する。たとえば、母集団プロセッサ138は、母集団データベース120または他のデータ記憶デバイスに記憶されている母集団情報と、クリックスルーデータベース122に記憶されているもしくは母集団データベース120と前もって共有されているクリックスルーデータとにアクセスする。母集団プロセッサ138は所定の方程式またはコンピュータ実行可能命令の集合を母集団情報とクリックスルーデータとの一部またはすべてに適応して、“#(q、d_i)”とも呼ばれる探索問い合わせに対する文書への母集団に関係なくすべてのユーザによるクリックの数を決定する。たとえば、問い合わせ“Q”が「舟遊び」であれば、問い合わせ“Q”に対する文書“D₁”への母集団に関係なくすべてのユーザによるユーザクリック総数が101であったと判定することが可能である。

30

40

【0084】

ブロック314の後に続いて判断ブロック316が実行されるが、ここで、探索問い合わせに対するすべての文書が処理されたかどうか判断される。探索エンジン124は、カウンタすなわちブロック300で“1”という値に初期設定された変数“i”を探索問い合わせ“Q”に対して処理される文書の数にしたがってブロック306である値に設定された変数“N”と比較する。すべての文書が処理されたわけではない場合、カウンタすな

50

わち変数“i”は変数“N”とは等しくならず、「ノー」に分岐してブロック318に行く。代替実施形態では、分析される文書の最大数が設定される。たとえば、“N”は、ブロック306で決定された文書の数未満の最大数に設定される。

【0085】

ブロック318で、処理された文書数の後を追ってカウンタがインCREMENTされる。たとえば、“1”という値に初期設定されたカウンタすなわち変数“i”が、“2”などの次の値にインCREMENTされる。すると、サブルーチン212はブロック312に復帰して、次の文書の処理を継続する。次の文書がブロック312～314によって処理され、ブロック318でカウンタすなわち変数“i”は、すべての文書が処理されて、カウンタすなわち変数“i”が“N”に等しくなるまで、続いてインCREMENTされる。このようにして、「舟遊び」に対してすでに提供した例では、探索問い合わせ「舟遊び」に対する探索結果の15の文書のすべてがブロック312～314で処理される。

10

【0086】

文書すべてが処理されたら、判断ブロック318から「イエス」に分岐して、サブルーチン212はブロック320に続く。

【0087】

ブロック320で、特定の問い合わせ(“Q”)に対するすべての文書(文書“i”)への特定の母集団“P”による選択、この場合ではクリック、の数の合計と、特定の問い合わせ(“Q”)に対するすべての文書(文書“i”)へのすべてのユーザによる選択、この場合ではクリック、の数の合計とが決定される。図示する実施形態では、探索エンジン124は、探索“Q”と関連する文書に対する特定の母集団“P”によるクリックの総数を記述する特定の問い合わせ(“Q”)に対するすべての文書(文書“i”)への特定の母集団“P”による選択、この場合ではクリック、の数の合計を決定する。たとえば、母集団プロセッサ138は、探索問い合わせ“Q”の先行の探索に回答して返送された文書に対する特定の母集団“F”による一部のまたはすべてのユーザのクリックを表す合計値を決定する。これで、この合計値を探索エンジン124または母集団プロセッサ138が母集団信号関数またはコンピュータ実行可能命令集合に対して適用することが可能となる。ブロック312で決定された選択数が、すべての文書“i”に対してブロック320で合計されることに注意すべきである。

20

【0088】

さらにそのうえ、探索エンジン124は、探索問い合わせ“Q”と関連する文書に対する母集団に関係なくすべてのユーザによるクリックの総数を記述している特定の問い合わせ(“Q”)へのすべての文書(文書“i”)による選択、この場合はクリック、の数の合計を決定する。たとえば、母集団プロセッサ138は、探索問い合わせ“Q”の先行問い合わせに回答して返送された文書に対する母集団に関係なくすべてのユーザによる一部のまたはすべてのユーザクリックを表している合計値を決定する。これで、この合計値を探索エンジン124または母集団プロセッサ138が母集団信号関数またはコンピュータ実行可能命令集合に対して適用することが可能となる。ブロック314で決定された選択数もまた、すべての文書“i”に対してブロック320で合計されることに注意すべきである。

30

40

【0089】

ブロック320の後に続いてブロック322が実行されるが、ここで、平滑係数が決定される。探索エンジン124は、特定の問い合わせに対する特定の母集団のユーザによるユーザクリックなどのクリック信号を信用するために必要とされるデータの量を表す平滑係数“μ”を決定する。たとえば、母集団プロセッサ138は、特定の問い合わせに対するすべてのユーザによるユーザクリックや特定の母集団のユーザによるユーザクリックなどのクリック信号を信用するために必要とされるデータの量を表す平滑係数を決定するために、所定の方程式またはコンピュータ実行可能命令集合を利用する。この平滑係数は、特定の母集団のユーザのユーザクリックが非常に少ない場合や特定のソースのユーザクリックが信頼性がないまたは別様に信用するに足らないことが周知であるかまたは別様に感

50

知されるような場合に特には有用である。たとえば、上記の表における問い合わせ“Q”の例に適用すると、問い合わせ“Q”に対して文書“D₁”への日本在住のユーザによるクリックは1つだけであるので、平滑係数“ μ ”を“10”などの一定値に設定することが可能である。この場合、平滑係数は、特定の母集団と関連するデータの重みまたは値に作用するまたは別様に影響を与えるために用いることが可能である。ほとんどの場合、平滑係数は、母集団信号関数または母集団プロセッサ138によって処理されるコンピュータ読み取り可能命令の集合に適用される。

【0090】

場合によっては、一般的な平滑係数を決定することが可能であることに注意すべきである。探索エンジン124は、特定の問い合わせに対する特定の文書へのすべてのユーザによるユーザクリックなどのクリック信号を信号するために必要とされるデータの量を表す一般的な平滑係数を決定する。図示する例では、ローカル母集団のユーザのクリックを信用するために必要とされるデータの量は一般的な母集団のユーザのクリックを信用するために必要とされるデータ量と同じであり、上述したように決定された平滑係数を一般的な平滑係数として用いることが可能であると仮定される。

10

【0091】

さらなる例として、一般的な平滑係数は次のように決定することが可能である。母集団プロセッサ138は母集団データベース120または他のデータ記憶デバイスにアクセスして、ユーザクリックデータを検索する。所定の方程式またはコンピュータ読み取り可能命令の集合を用いて、母集団プロセッサ138は、特定の問い合わせに対する特定の文書へのユーザクリックの数を決定する。

20

【0092】

特定の問い合わせに対する特定の文書の平滑係数、値または得点が母集団データベース120に前もって記憶されていれば、母集団プロセッサ138は、特定の問い合わせに対する特定の文書の平滑係数、値または得点を検索する。たとえば、母集団データベース120は、探索問い合わせ「舟遊び」に対する探索結果中の文書に対するすべてのユーザによるクリックの数が不十分であること、したがって、探索問い合わせ「舟遊び」に対する探索結果中の文書に対するすべてのユーザによるクリックの総数を用いるべきであることを示したりする。一般的な母集団平滑係数は、2つの結果の間を徐々に遷移することが可能である。探索問い合わせ「舟遊び」に対する探索結果中の文書に対するすべてのユーザによるクリックの数をを用いるか、母集団情報と選択データとから改善された探索結果が得られるように探索問い合わせ「舟遊び」に対する探索結果中のすべての文書に対するすべてのユーザによるクリックの総数を用いるかを判断することが可能である。

30

【0093】

ブロック322の後に続いてブロック324が実行されるが、ここで、特定の問い合わせ(“Q”)に対する探索結果中で返送された各々の文書(文書“j”)の場合の、すべてのユーザによる選択、この場合はクリック、の数。探索エンジン124は、文書“j”が問い合わせ“Q”に対する探索結果中で返送されると、母集団に関係なく文書“j”に対するクリック数を判定する。たとえば、母集団プロセッサ138は、母集団データベース120または他のデータ記憶デバイスに記憶されている母集団情報と、クリックスルーデータベース122に記憶されているもしくは母集団データベース120と前もって共有されているクリックスルーデータとにアクセスする。母集団プロセッサ138は所定の方程式またはコンピュータ実行可能命令の集合を母集団情報とクリックスルーデータとの一部またはすべてに適応して、“#(q, d_j)”とも呼ばれる探索問い合わせに対する文書への母集団によるクリックの数を決定する。たとえば、問い合わせ“Q”が「舟遊び」であれば、問い合わせ“Q”に対する文書“D₁”へのすべてのユーザによるユーザクリック総数が101であったと判定することが可能である。

40

【0094】

ブロック324の後に続いてブロック326が実行されるが、ここで、特定の問い合わせ(“Q”)に対する探索結果中で返送された各々の文書(文書“i”)に対する母集団

50

“ P ”のすべてのユーザによる選択、この場合ではクリック、の数を判定する。探索エンジン 124 は、文書 “ j ” が問い合わせ “ Q ” に対する探索結果中で返送されると、文書 “ j ” に対する母集団 “ P ” のすべてのユーザによるクリック数を判定する。たとえば、母集団プロセッサ 138 は、母集団データベース 120 または他のデータ記憶デバイスに記憶されている母集団情報と、クリックスルーデータベース 122 に記憶されているもしくは母集団データベース 120 と前もって共有されているクリックスルーデータとにアクセスする。母集団プロセッサ 138 は所定の方程式またはコンピュータ実行可能命令の集合を母集団情報とクリックスルーデータとの一部またはすべてに適用して、“ $\#(q, d_j, P)$ ” と呼ばれる探索問い合わせに対する文書への母集団によるクリックの数を決定する。たとえば、問い合わせ “ Q ” が「舟遊び」であれば、問い合わせ “ Q ” に対する文書 “ D₁ ” へのフランスという母集団のすべてのユーザによるユーザクリック総数が 20 であったと判定することが可能である。

10

【 0095 】

ブロック 326 の後に続いて判断ブロック 328 が実行されるが、ここで、特定の探索問い合わせに対する文書の母集団信号が決定される。探索エンジン 124 は、探索結果 132 中の特定の文書に対する母集団信号 128 を決定する。たとえば、母集団プロセッサ 138 は、ある探索問い合わせが特定の母集団のユーザによって尋ねられた回数や、ある探索問い合わせがすべてのユーザによって尋ねられた回数や、特定の母集団のユーザによってある探索問い合わせが尋ねられて、特定の文書がクリックされた回数や、すべての母集団のユーザによってある探索問い合わせが尋ねられて、特定の文書がクリックされた回数や、いずれかの探索問い合わせに対して特定の母集団のユーザが文書をクリックした回数や、いずれかの探索問い合わせに対してすべてのユーザが文書をクリックした回数や、必要であれば、平滑係数や、必要であれば、探索結果中の特定の文書に対する母集団信号 128 を決定するための母集団クリックの重み、などの多くの係数を用いる。

20

【 0096 】

図示する実施形態では、この母集団信号 128 を、上に述べたブロック 300 ~ 324 で決定されたデータと例 (1) で示したアルゴリズムとを用いて計算する。問い合わせ「舟遊び」の先行する例に適用すると、母集団プロセッサ 138 は、探索結果 132 中の特定の文書の対する母集団信号 128 を決定する。サブルーチン 212 で上述したように母集団信号関数中で “ $S(Q, D_j)$ ” で表されるような特定の母集団 “ F ” のユーザによるクリックをカウントした後での文書 “ j ” へのユーザクリックの重み付け総数を表す重み付けされた値が母集団プロセッサ 138 によって決定される。これは、上述のアルゴリズムによって示されるような数学的関数を実行して文書 “ j ” の $S_j(Q, D_j)$ “ を計算することによって実行される。

30

【 0097 】

ブロック 328 の後に続いて判断ブロック 330 が実行されるが、ここで、ある探索問い合わせに対してすべての文書が処理されたかどうか判断される。探索エンジン 124 は、カウンタすなわちブロック 302 で “ 1 ” という値に初期設定された変数 “ j ” を探索問い合わせに対して処理される文書の数にしたがってある値に設定された変数 “ M ” と比較する。すべての文書が処理されている場合、カウンタすなわち変数 “ j ” は変数 “ M ” と等しく、「イエス」に分岐してブロック 332 に行く。代替実施形態では、分析される文書の最大数が設定される。たとえば、“ M ” は、ブロック 302 で決定された文書の数未満の最大数に設定される。

40

【 0098 】

ブロック 332 で、サブルーチン 212 は終了する。

【 0099 】

しかしながら、判断ブロック 330 で、文書のすべてが処理されているわけではなく、カウンタすなわち変数 “ j ” が変数 “ M ” と等しくない場合、「ノー」に分岐してブロック 334 に行く。

【 0100 】

50

ブロック 3 3 4 で、処理された文書数の後を追ってカウンタがインCREMENTされる。たとえば、“ 1 ” という値に初期設定されたカウンタすなわち変数 “ j ” が、“ 2 ” などの次の値にインCREMENTされる。すると、サブルーチン 2 1 2 はブロック 3 2 4 に復帰して、次の文書の処理を継続する。次の文書がブロック 3 2 4 ~ 3 2 8 によって処理されて、ブロック 3 3 4 でカウンタすなわち変数 “ j ” は、すべての文書が処理されて、カウンタすなわち変数 “ j ” が “ M ” に等しくなるまで続いてインCREMENTされる。このようにして、「舟遊び」に対してすでに提供した例では、探索問い合わせ「舟遊び」に対する探索結果の 1 5 の文書のすべてがブロック 3 2 4 ~ 3 2 8 で処理される。

【 0 1 0 1 】

文書すべてが処理されたら、判断ブロック 3 3 2 から「イエス」に分岐して、サブルーチン 2 1 2 はブロック 3 3 2 で終了する。

10

【 0 1 0 2 】

再度図 2 を参照すると、サブルーチン 2 1 2 の後に続いてブロック 2 1 4 が実行されるが、ここで、文書毎の母集団信号がランク付けプロセッサに提供される。たとえば、図示する実施形態では、各々の文書 “ 1 ~ N ” に対する得点の計算値 “ $S(Q, D_j)$ ” が “ N ” 個の母集団信号に含まれる。文書毎の母集団信号 1 2 8 は、ランク付けプロセッサ 1 3 6 に送信されて、他の探索問い合わせに応答して探索結果の以降のランク付けや得点を決定するようにする。ランク付けプロセッサ 1 3 6 は、母集団プロセッサ 1 3 8 からの母集団信号 1 2 8 および / または他の出力を組み込んだランク付けもしくは得点付け機能またはコンピュータ実行可能命令集合を含んでいる。たとえば、サブルーチン 2 1 2 から発生した重み付けされた値はランク付けプロセッサ 1 3 6 に送信されると、このプロセッサは、重み付けされた値などの母集団信号 1 2 8 を利用して、以降の探索結果をランク付けしたり別様に得点付けしたりする。探索エンジン 1 2 4 または別のシステムもしくは方法によって文書毎に生成された他の信号 1 3 0 もまた、以降の探索結果をランク付けまたは得点付けするためにランク付けプロセッサ 1 3 6 に送信することが可能である。

20

【 0 1 0 3 】

ブロック 2 1 4 の後に続いてブロック 2 1 6 が実行されるが、ここで、探索結果が提供される。ランク付けプロセッサ 1 3 6 は、探索問い合わせ 1 1 4 に応答して探索結果 1 3 2 中にある各々の文書のランクや得点を生成する。重み付けされた値などのブロック 2 1 4 からの母集団信号 2 1 8 を用いて、ランク付けプロセッサ 1 3 6 は、探索結果 1 3 2 中にある 1 つ以上の文書のランクや得点に作用する。ランク付けプロセッサ 1 3 6 は、母集団信号 1 2 8 と一緒に 1 3 0 と図 1 に示すような信号などの他の信号を用いて、探索結果 1 3 2 の文書をランク付けしたり別様に得点付けしたりすることに注意すべきである。一部の例では、ランク付けプロセッサ 1 3 6 は、さらに、探索結果 1 3 2 のランクや得点を処理している間に特定の母集団信号 1 2 8 および / または他の信号 1 3 0 を利用すべきであるかどうかを判断することが可能である。

30

【 0 1 0 4 】

ブロック 2 1 6 の後に続いてブロック 2 1 8 が実行され、ここで方法 2 0 0 は終了する。

【 0 1 0 5 】

40

本発明の他の実施形態では、方法 2 0 0 を繰り返し利用して、母集団データベース 1 2 0 および / またはクリックスルーデータベース 1 2 2 または他のデータ記憶デバイス中に新たなデータやデータ変更が受信されたり別様に得られたりした場合にはいつでも、新たなまたは更新された母集団信号が決定されたりする。新たなまたは更新された母集団信号が決定されたら、次にこの信号をランク付けプロセッサ 1 3 6 に送信して、探索結果 1 3 2 のランクや得点を変更したり更新したりすることが可能である。

【 0 1 0 6 】

多くの細目を上述したが、これらの細目は本発明の範囲に対する制限と解釈すべきではなく、開示する実施形態の単なる例示と解釈すべきである。当業者は、本発明の範囲内にある他の多くの潜在的な変更例を想像しえるであろう。

50

【図面の簡単な説明】

【 0 1 0 7 】

【図 1】 本発明の一実施形態によるシステムのブロック図を示す図である。

【図 2】 本発明の一実施形態による方法のフローチャートを示す図である。

【図 3】 図 2 に示す方法のサブルーチンのフローチャートを示す図である。

【 図 1 】

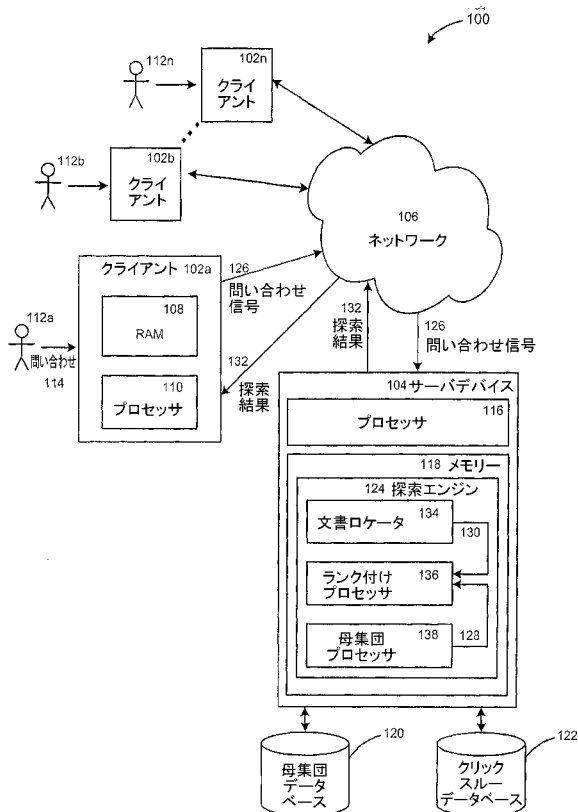


FIG. 1

【 図 2 】

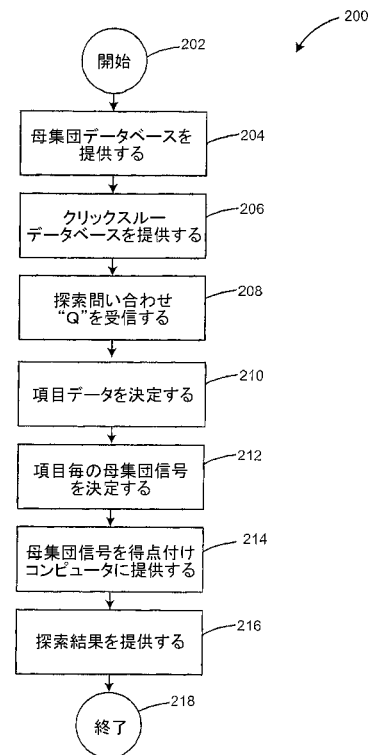
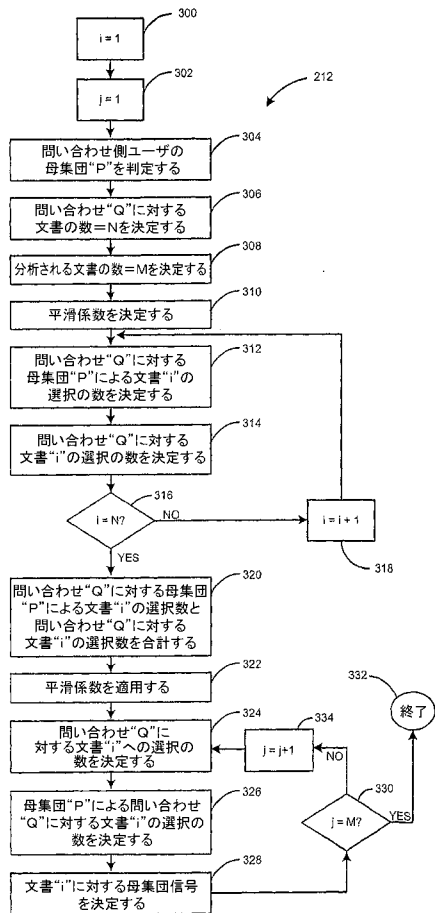


FIG. 2

【 図 3 】

FIG. 3



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No PCT/US2004/029618
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 G06F17/30		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 539 377 B1 (CULLISS GARY A) 25 March 2003 (2003-03-25) abstract column 1, line 24 - column 2, line 52 column 3, line 11 - column 3, line 65 column 5, line 21 - column 6, line 29 column 7, line 8 - column 8, line 39 column 11, line 35 - column 11, line 38	1-44
X	WO 01/67297 A (BEZEM SHLOMO ; JACOBSON RON (IL); KLEINBERGER PAUL (IL); TZUNAMI INC) 13 September 2001 (2001-09-13) abstract page 5, line 3 - page 7, line 17 page 8, line 20 - page 10, line 6 page 12, line 1 - page 13, line 9 page 13, line 25 - page 14, line 13 page 20, line 1 - page 20, line 13 ----- -/--	1-44
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
10 December 2004		21/12/2004
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Boyadzhiev, Y

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US2004/029618

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 353 849 B1 (LINSK STEVEN J) 5 March 2002 (2002-03-05) abstract column 1, line 12 - column 2, line 42 column 3, line 11 - column 4, line 45	1-44
A	WO 00/77689 A (TRIOGO INC) 21 December 2000 (2000-12-21) page 6, line 1 - page 7, line 7 page 9, line 2 - page 9, line 7 page 10, line 1 - page 11, line 1 page 12, line 8 - page 13, line 24 page 18, line 1 - page 18, line 22	1-44

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US2004/029618

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 6539377	B1	25-03-2003	US 6182068 B1	30-01-2001
			US 6078916 A	20-06-2000
			US 6014665 A	11-01-2000
			US 6006222 A	21-12-1999
			US 2003187837 A1	02-10-2003
			AU 2035199 A	22-02-1999
			WO 9906924 A1	11-02-1999
WO 0167297	A	13-09-2001	AU 4100001 A	17-09-2001
			WO 0167297 A1	13-09-2001
US 6353849	B1	05-03-2002	US 6138142 A	24-10-2000
WO 0077689	A	21-12-2000	AU 7823400 A	02-01-2001
			CA 2376789 A1	21-12-2000
			EP 1203323 A1	08-05-2002
			WO 0077689 A1	21-12-2000

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

1 . J A V A

(72)発明者 ピアソン、 マーク

アメリカ合衆国 カリフォルニア州 9 4 7 2 0 パークレー ピエドモント アベニュー 2 2
9 9 インターナショナル ハウス ルーム 4 1 1

Fターム(参考) 5B075 KK07 PQ75 PR08