

(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2009-512954

(P2009-512954A)

(43) 公表日 平成21年3月26日(2009.3.26)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30 340Z	5B075
	G06F 17/30 140	
	G06F 17/30 170A	

審査請求 有 予備審査請求 未請求 (全 35 頁)

(21) 出願番号	特願2008-537679 (P2008-537679)	(71) 出願人	507103802
(86) (22) 出願日	平成17年12月13日 (2005.12.13)		グーグル・インコーポレーテッド
(85) 翻訳文提出日	平成20年6月20日 (2008.6.20)		アメリカ合衆国・カリフォルニア・940
(86) 国際出願番号	PCT/US2005/045447		43・マウンテン・ビュー・アンフィシア
(87) 国際公開番号	W02007/046830		ター・パークウェイ・1600
(87) 国際公開日	平成19年4月26日 (2007.4.26)	(74) 代理人	100064908
(31) 優先権主張番号	11/257, 282		弁理士 志賀 正武
(32) 優先日	平成17年10月23日 (2005.10.23)	(74) 代理人	100089037
(33) 優先権主張国	米国 (US)		弁理士 渡邊 隆
		(74) 代理人	100108453
			弁理士 村山 靖彦
		(74) 代理人	100110364
			弁理士 実広 信哉

最終頁に続く

(54) 【発明の名称】 構造化データ上での検索

(57) 【要約】

本発明は、ユーザが構造化データ上でウェブベースの検索を実行することを可能にする。本発明の方法は、ユーザが検索エンジン内で照会用語を入力するステップと、照会用語による照会結果を決定するステップと、ユーザが照会結果内の少なくとも1つのデータ項目に関連した1つまたは複数の属性を指定するステップと、指定された属性による第2の照会結果を再決定するステップとを有する。ユーザは、ラベルまたは属性値が照会の結果をさらにフィルタリングするために使用されることを指定することによって構造化データ上での検索を詳細化することができる。

【特許請求の範囲】**【請求項 1】**

ユーザに構造化データ上でウェブベースの検索を実行することを可能にする方法であって、

前記ユーザに検索エンジン内で照会用語を入力することを可能にするステップと、

前記照会用語による照会結果を決定するステップと、

前記ユーザに前記照会結果内の少なくとも1つのデータ項目に関連した1つまたは複数の属性を指定することを可能にするステップと、

前記指定された属性による第2の照会結果を再決定するステップと

を有し、

前記属性は、属性名/属性値の対であり、

前記属性値は、属性タイプを含むことを特徴とする方法。

10

【請求項 2】

前記ユーザに照会用語を入力することを可能にするステップは、前記照会結果を決定するステップの前に、ユーザに属性および属性値を照会用語として指定することを可能にするステップを有することを特徴とする請求項1に記載の方法。

【請求項 3】

前記ユーザに前記1つまたは複数の属性を指定することを可能にするステップの前に、前記照会結果を表示するステップをさらに有することを特徴とする請求項1に記載の方法。

20

【請求項 4】

前記ユーザに前記1つまたは複数の属性を指定することを可能にするステップの前に、複数の属性名を表示するステップをさらに有することを特徴とする請求項1に記載の方法。

【請求項 5】

所定期間に最も人気のある属性を複数のコア属性へ追加するステップをさらに有することを特徴とする請求項1に記載の方法。

【請求項 6】

前記最も人気のある属性は、前記所定期間に検索を実行するユーザに最も人気があるものであることを特徴とする請求項5に記載の方法。

30

【請求項 7】

前記最も人気のある属性は、前記所定期間に前記照会結果内の最多数のデータ項目を有する属性であることを特徴とする請求項5に記載の方法。

【請求項 8】

前記最も人気のある属性は、所定期間にわたってユーザによって最も頻繁に選択された属性であることを特徴とする請求項5に記載の方法。

【請求項 9】

前記ユーザに前記照会結果内の少なくとも1つのデータ項目に関連した1つまたは複数の属性を指定することを可能にするステップは、

前記照会結果内のデータ項目に関連した1組の最も人気のあるコア属性名を決定するステップと、

前記照会結果内のデータ項目に関連したコア属性値の前記名前だけを表示するステップと

40

を有することを特徴とする請求項1に記載の方法。

【請求項 10】

1組の最も人気のあるコア属性を決定するステップは、

前記照会結果内のq個の最も関連のある結果を決定するステップと、

前記q個の最も関連のある結果について、それらの結果に関連したm個の最も人気のある属性を決定するステップと、

n個の最も人気のある属性について、上位m個の属性/値の対を決定するステップと

50

を有することを特徴とする請求項 9 に記載の方法。

【請求項 1 1】

前記検索は、ネットワークを介して前記照会を受信して、ウェブベースの検索エンジン内で実行されることを特徴とする請求項 1 に記載の方法。

【請求項 1 2】

ユーザに構造化データ上でウェブベースの検索を実行することを可能にする方法であって、

前記ユーザに検索エンジン内で照会用語を入力することを可能にするステップと、

前記照会用語による照会結果を決定するステップと、

前記ユーザに前記照会結果内の少なくとも 1 つのデータ項目に関連した複数のラベルのうち 1 つまたは複数を選択することを可能にするステップと、

前記選択された 1 つまたは複数のラベルによる第 2 の照会結果を再決定するステップとを有し、

前記ラベルは、値なしのタグであることを特徴とする方法。

【請求項 1 3】

前記ユーザに 1 つまたは複数のラベルを指定することを可能にするステップの前に、前記照会結果および前記複数のラベルを表示するステップをさらに有することを特徴とする請求項 1 2 に記載の方法。

【請求項 1 4】

前記ユーザに 1 つまたは複数のラベルを指定することを可能にするステップの前に、前記複数のラベルを表示するステップをさらに有することを特徴とする請求項 1 2 に記載の方法。

【請求項 1 5】

前記ユーザに 1 つまたは複数のラベルを指定することを可能にするステップは、前記ユーザにラベルを論理的に結合することを可能にするステップを有することを特徴とする請求項 1 2 に記載の方法。

【請求項 1 6】

所定期間に最も人気のあるラベルを複数のコアラベルへ追加するステップをさらに有することを特徴とする請求項 1 2 に記載の方法。

【請求項 1 7】

前記検索は、ネットワークを介して前記照会を受信して、ウェブベースの検索エンジン内で実行されることを特徴とする請求項 1 2 に記載の方法。

【請求項 1 8】

ユーザに構造化データ上でウェブベースの検索を実行することを可能にする装置であって、

前記ユーザに検索エンジン内で照会用語を入力することを可能にする手段と、

前記照会用語による照会結果を決定する手段と、

前記ユーザに前記照会結果内の少なくとも 1 つのデータ項目に関連した 1 つまたは複数の属性を指定することを可能にする手段と、

前記指定された属性による第 2 の照会結果を再決定する手段と

を具備し、

前記属性は、属性名 / 属性値の対であり、

前記属性値は、属性タイプを含むことを特徴とする装置。

【請求項 1 9】

ユーザに構造化データ上でウェブベースの検索を実行することを可能にする装置であって、

前記ユーザに検索エンジン内で照会用語を入力することを可能にする検索エンジンと、

前記ユーザに前記照会用語による第 1 照会結果内の少なくとも 1 つのデータ項目に関連した 1 つまたは複数の属性を指定することを可能にするユーザインターフェースと、

前記照会用語による前記第 1 照会結果を決定し、続いて前記指定された属性による第 2

10

20

30

40

50

照会結果を再決定する照会エンジンと

を具備し、

前記属性は、属性名/属性値の対であり、

前記属性値は、属性タイプを含むことを特徴とする装置。

【請求項 20】

コンピュータ可読媒体上に命令を有するコンピュータプログラム製品であって、

前記命令は、データ処理システムに、

前記ユーザに検索エンジン内で照会用語を入力することを可能にする過程と、

前記照会用語による照会結果を決定する過程と、

前記ユーザに前記照会結果内の少なくとも1つのデータ項目に関連した1つまたは複数の属性を指定することを可能にする過程と、

前記指定された属性による第2の照会結果を再決定する過程と

を実行させることができ、

前記属性は、属性名/属性値の対であり、

前記属性値は、属性タイプを含むことを特徴とするコンピュータプログラム製品。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、米国特許法第119条(e)に基づき、2005年10月23日に出願したReddyらによる「Search Over Structured Data」と題する米国特許出願第11/257,282号の優先権を主張するものである。本出願は、2005年10月23日に出願したReddyらによる「Adding Attributes and Labels to Structured Data」と題する米国特許出願第11/256,883号に関し、かつ引用により本明細書に組み込む。

【背景技術】

【0002】

従来の検索エンジンは、ワールドワイドウェブまたは非常に大きいデータベースなどの極めて大きい情報の集合体を検索することができる。検索されるべきデータ集合体のサイズが大きくなると、ユーザによって入力される照会用語に一致する照会結果を正しく返すことはできない。代わりに、ユーザが検索から返される大量のデータを分類する手伝いをする機構を提供することが望ましい。

【0003】

現在、いくつかの従来の検索エンジンは、様々な方法を使用して、照会結果に返されるデータを編成する。この種の編成方法の目的は、どの照会結果がユーザに最も興味を持たせるかを決定することである。概して、従来の検索エンジンは検索結果に優先順位を付けるために様々な技術を使用するが、これらの技術は、ユーザが検索している情報のタイプについて仮定しなければならないので理想的ではない。例えば、ユーザが「jobs」を入力する場合、ユーザは、求人募集、Steve Jobsの情報、特定の国に関する雇用統計、またはいくつかの他の項目を検索している可能性がある。したがって、従来の検索エンジンを使用するときユーザは、照会用語として単に「jobs」を入力しないだろう。おそらくユーザは、検索を絞り込んだ追加の照会用語も入力する。不幸にもユーザは、絞り込む用語を含まない関連するリストを取りこぼすこともありうる。

【0004】

現在、ワールドワイドウェブ上に格納されているかどうかわからない様々なデータにわたって検索することは困難である。通常、従来の検索エンジンは、少数のソースだけからのデータ上で動作する。例えば、通常、ウェブベースの検索エンジンによってユーザはワールドワイドウェブ上のページを検索することができる。ウェブ検索エンジンはしばしば、情報の集合体を検索可能にするために情報の集合体にインデックス付けする「バックエンド(back-end)」を有する。例えば、ウェブベースの検索エンジンは、ワールドワイドウェブを定期的にクロールし、クロールされたページおよびサイトのインデックスを生成する。他の検索エンジンによってユーザは既存のデータベースを検索することができる。

この種の検索エンジンは、所定のデータベースの編成に頼る。例えば、データベースが既知のフィールドおよび属性を有する場合、ユーザはそれらの属性内で検索することができる。例えば、XMLデータベースは適格なXML入力だけを受け入れる。検索されるべきデータがそれほど編成されていない場合、概してデータベースは、データを受け入れることができないか、または検索のためにデータを編成することができない。

【0005】

他の検索エンジンによって、ユーザは、データベースを検索するか、または均一な編成を有するテキスト文書を検索することができる。この種の検索エンジンは、データベースの編成およびその中の文書の編成について知らなければならない。データが格納されている位置および形式の多様性は、ユーザが必要な情報を見いだすために複数のデータベース内の複数の位置内で頻繁に検索しなければならないことを意味する。

10

【0006】

様々な種類の文書およびデータ形式を同時に含むとき、文書の集合体がウェブベースの検索エンジンを介して検索可能であり、したがってほとんどの人に容易にアクセス可能であることが望ましい。さらにそれは、検索可能な文書の集合体が、ユーザが自身の検索を微調整する助けをすることができる方法で編成されることが望ましい。

【特許文献1】米国特許出願公開第11/257282号明細書

【特許文献2】米国特許出願公開第11/256883号明細書

【発明の開示】

【課題を解決するための手段】

20

【0007】

記載された本発明の実施形態は、ラベルおよび属性値を検索されるべきデータ項目に関連付ける。プロバイダが、属性およびラベルをプロバイダのデータに関連付けるか、または属性およびラベルが、既存データに追加されることができる。好ましい実施形態によって、コンテンツプロバイダは、コンテンツプロバイダ独自のカスタムラベルおよびカスタム属性を項目に付加するか、または予め定義されたラベルおよび属性を使用することができる。プロバイダは、ユーザインターフェースまたはバルクアップロード機構を使用してデータをアップロードすることができる。ユーザは、ラベルまたは属性値が照会の結果をさらにフィルタリングするために使用されることを指定することによって検索を詳細化することができる。

30

【0008】

本発明の教示は、添付図面と共に下記の詳細な記載を考慮することによって容易に理解できる。同様の参照番号は、添付図面内で同様の要素のために使用される。

【0009】

図は、例示のみを目的として本発明の実施形態を示す。当業者は、本明細書で示される構造および方法の変形実施形態が本明細書に記載された本発明の原理から逸脱することなく使用されてよいことを以下の記載から容易に理解する。

【発明を実施するための最良の形態】

【0010】

以下の段落は、本発明による構造化データをアップロードし検索するためのシステムの様々な実施形態を説明する。

40

【0011】

図1aは、本発明の好ましい実施形態によるデータ処理システムを示す構成図100である。図1aは、複数のクライアントデータ処理システム110a~110n、ネットワーク130、およびサーバデータ処理システム120を含む。図中では、例示のユーザデータ処理システム110aは、プロセッサ140、ブラウザ150、およびメモリ160を含む。ユーザデータ処理システム110またはその構成要素は、それに限定されないが、パーソナルコンピュータ、有線ネットワークコンピュータ、無線ネットワークコンピュータ、携帯電話または携帯電話を内蔵した機器、ハンドヘルド機器、シンクライアント機器、それらのいくつかの組合せなどを含む任意の適切なデータ処理システムであってよい。ネットワーク130は、ユーザデータ

50

処理システム110のうちの1つまたは複数とサーバデータ処理システム120との間の通信を可能にする任意のネットワークであってよい。例えば、ネットワーク130は、それに限定されないが、インターネット、LAN、WAN、有線ネットワーク、無線ネットワーク、携帯電話ネットワーク、テキストメッセージを送信するネットワーク、およびそれらのいくつかの組合せであってよい。

【0012】

本発明の好ましい実施形態では、ユーザデータ処理システム110aは、ユーザがサーバシステム120と通信することができるためにプロセッサ140によって実行されるメモリ160内のブラウザソフトウェア150を含む。以下に詳細に記載されるように、この種のブラウザ150によって、ユーザは、照会用語をサーバデータ処理システム120に送信し、かつシステム120から照会結果を受信するために、サーバデータ処理システム120と通信することができる。以下にさらに詳細に記載されるように、ブラウザ150によって、ユーザは、照会結果に関連したラベルおよび属性を受信し、照会結果をさらに定義するためにラベルおよび属性を使用することができる。本明細書に記載される実施形態はブラウザベースであるが、本発明はブラウザベースの検索に限定されず、ユーザ110とサーバ120との間の通信に関する任意の適切な機構が、本発明の精神および範囲から逸脱することなく使用されてよい。

10

【0013】

本明細書に記載されるすべてのソフトウェアおよびコンピュータ実行可能命令のうちの一部は、それに限定されないが、データ処理システムのメモリ、CD ROM、フラッシュメモリ、およびフロッピー（登録商標）ディスクを含むコンピュータ可読媒体上にコンピュータプログラム製品として格納されること、またはネットワークを介する信号もしくはシステム構成要素間の信号として送信されることことができる。

20

【0014】

サーバデータ処理システム120は、サーバシステム120が照会用語に関する構造化データの集合体190を検索することができるように検索および照会エンジンソフトウェア185を実行するプロセッサ170を含む（検索および照会エンジン185は「検索エンジン」とも呼ばれる）。構造化データの一例はフィールド化されたデータ（すなわち、データ項目）であり、それぞれは1つまたは複数のデータフィールド（名前、アドレス、状態など）を有する。

【0015】

メモリ180は、構造化データ190内のデータ項目の一部またはすべてに関する属性（およびラベル）を格納する属性リポジトリ195も含む。リポジトリは、図5に関連して以下に記載される。リポジトリ195が構造化データの集合体190の一部であるとして示されているが、リポジトリ195はデータ集合体190から分離していてもよい。

30

【0016】

大きい検索エンジンおよび大きいデータ集合体は、それに限定されないが、分散データ処理システム、協働データ処理システム、ネットワークデータ処理システムなどを含んで多くの方法で格納されてよいが、検索エンジン185、リポジトリ195、および構造化データの集合体190のすべてが単一のメモリ180内にあるとして図1aに示されている。検索エンジン185は、ソフトウェア、ハードウェア、ファームウェア、またはこれらの任意の組合せであってよい。

40

【0017】

好ましい実施形態では照会用語は、複数のユーザシステム110のうちの1つまたは複数を通じてユーザによって入力され、ネットワーク130を介してサーバデータ処理システム120に送信される。データの集合体を受信し、インデックス付けし、かつ検索するためにサーバ120によって使用される方法の詳細が本明細書に詳細に記載される。

【0018】

図1bは、本発明の好ましい実施形態による他のデータ処理システムを示す構成図111である。図1bでは、ユーザはユーザの装置110上に個人のデータ集合体190を格納する。個人の検索エンジンは、ネットワーク130を介してユーザによって、および場合によっては他

50

のユーザによって、検索可能にするためにこのデータにアクセスし、かつ編成することが意図される。この種のシステムによってデータベースおよび他のタイプのデータ集合体が、中央の検索エンジンによってアクセス可能である検索可能な文書のプールに追加されることもできる。

【0019】

図1bの実施形態では、データ集合体190は、ユーザのデータ処理システム110またはエンタプライズサーバ(図示せず)上に格納され、ユーザのみ、より小さいサブセットのユーザのみ、またはデータ集合体190にアクセスする方法を認識しているすべてのユーザなどの選択されたグループの人または個人に対して利用可能であってよい。そのような場合、本明細書に記載されるように属性およびラベルを用いて検索をフィルタリングする機能は、コンピュータ上またはコンピュータのローカルネットワーク上で局部的に実行する個人の検索エンジン185の一部であってよい。例えば、Google社(Google, Inc. of Mountain View, CA)から入手できるGoogleデスクトップ検索ツールは、ユーザのデスクトップ上で実行し、ユーザのパーソナルコンピュータ上でデータにインデックス付けする検索ツールである。本発明を組み込んだGoogleデスクトップ検索の実装形態は、ユーザのデスクトップ上に格納されるまたはユーザのデスクトップからアクセス可能であるデータベースおよび他の種類のデータ集合体を検索する機能をユーザに提供する。

10

【0020】

その実装形態は、有用な属性およびラベルを用いてユーザのデータを編成する機能もユーザに提供する。例えば、大学図書館は、そのオンラインコレクションのすべてを大学の学生、教員、および卒業生に利用可能にすることができる。そのような場合、情報は、公的に利用できるサーバ上ではなく、大学のサーバ内に格納され、大学のデータプロバイダによってアクセスを許可された人(およびプログラム)に対してのみアクセス可能かつ検索可能となる。例では大学は、どのプロバイダがデータ集合体に追加する能力を有するかを制御することもできる。

20

【0021】

図1cは、本発明の好ましい実施形態によるアーキテクチャ図131である。記載された実施形態ではプロバイダは、データおよび属性をシステムに入力する3つの方法のうちの1つまたは複数を使用することができる。プロバイダが面するフロントエンド132(例えば、図6bを参照)によって、プロバイダは、その目的のために提供されるユーザインターフェースを使用してデータ項目および属性を入力することができる。プロバイダは、データ項目のバルクアップロード133も実行することができる(例えば、図8a~8dを参照)。プロバイダは、(例えば、FTPを使用して)特定のURLから項目をアップロードもできる(134)。検索および照会エンジン185は、すべてのデータのインデックス137を生成するためにデータ項目に関する入力された属性およびそれらの値を好ましくは含んでデータ集合体190内の項目にインデックス付けする。検索エンジン185によって、ユーザは、照会を入力することもできる(例えば、図4aを参照)。システムは、ソフトウェアプログラムが検索エンジン185を介してデータを照会することを可能にするアプリケーションプログラムインターフェース(API)も含む。

30

【0022】

図2aは、本発明の好ましい実施形態による検索可能なデータ項目の集合体の生成の概要を示す流れ図200である。図6a~6eおよび図8a~8dに関連して以下に記載されるように、サーバ120はデータ項目の集合体を受信する(202)。このデータは、標準のウェブクロールの結果として受信するか、またはそれらのデータが検索可能になることを望む1つまたは複数のプロバイダによって提供されるかのいずれかであってよい。受信されたデータ項目の集合体は、以下に記載されるようにラベル、属性、および属性値を抽出するように処理され、それらのラベル、属性および属性値は様々な情報タイプに関連付けられる。ある種の環境ではユーザは、入力されたデータの一部またはすべてに関する属性名および/または属性値を提供することになる。一例として、ユーザは、ユーザが医学雑誌の集合体を保持するために生成したデータベースをアップロードすることができる。ユーザは、「Jour

40

50

nal」、「year of publication」、「Journal Name」などの属性名を反映する値を用いてこれらの雑誌に対して属性を指定してよい。ユーザは、「Medical」、「Dental」、「From Harvard」など、雑誌ごとにゼロ以上のラベルも入力することができる。ラベルは、ラベルに関連した値を有しない特別の種類属性(値なしタグとも呼ばれる)である。要素204の詳細は図3aに関連して記載される。

【0023】

図2bは、本発明の好ましい実施形態による文書の集合体を検索し、検索を詳細化することの概要を示す流れ図210である。記載された実施形態ではユーザは、1つまたは複数の照会用語(図4aのスクリーンショット400での「cancer receptor」402など)を入力する(212)。

10

【0024】

ある種の実施形態ではユーザは、領域402内にタイプされる照会の一部として属性名および属性値も入力することができる。例えば、ユーザは、領域402内に下記をタイプすることができる。

Cancer receptor attr(JournalType: medical)

照会結果内のいくつかの項目がJournalTypeと名づけられた属性を有するが、その属性が属性のコアセットの一部でないことをユーザが知る場合、ユーザは医学雑誌だけを返すことを望む。

【0025】

システムは、図3bに関連して下記に詳細に記載されるように照会結果を決定する(213)。いくつかの実施形態では、照会結果がこの時点で表示される(213)。他の実施形態では、照会結果はまだ表示されないが、代わりにユーザは照会用語に固有のラベルおよび/または属性を選択することによって自身の検索をさらに詳細化するように求められる。例えば、図3dに示されるようにユーザは、ラベルおよび属性を指定することによってユーザの検索を詳細化することができる(214)。

20

【0026】

図3aは、データ項目の集合体からラベルおよび属性を抽出する方法を示す流れ図300である。この方法は、データの集合体が検索できるようにデータの集合体を編成するために使用される設定プロセスの一部である。

【0027】

データ項目が受信されると、情報タイプを有するデータ項目ごとにシステムは、この情報タイプに関するラベルおよび属性を決定する(304)。属性は、「journal」などの名前を有する名前/値の対であり、次いで名前/値の対は雑誌の名前の1つまたは複数の可能性のある値を有する。

30

【0028】

好ましい実施形態では、属性およびラベルはデータのプロバイダによって指定される。したがって、属性を決定することは、ユーザ提供の属性およびラベルを識別することの問題だけである。

【0029】

ある種の場合にはデータのプロバイダは、プロバイダの項目に関する属性およびラベルを指定しない。例えば、項目がウェブクローラによって探し出されるウェブページである場合、ウェブページの所有者はそれらのページに関する属性およびラベルを指定する機会を有しない。したがって、他の好ましい実施形態では、ラベルおよび属性はデータ集合体に関してソフトウェアによって得られる。ラベルおよび属性を得ることは、所定のリストのラベルおよび属性に対する可能性のある値がソフトウェアによってデータ集合体内で見いだされる完全に自動化されたプロセスを含むことができる。例えば、販売用項目のリスト(例えば、GoogleのFroogleシステム)では、所定の基準を満たす価格総額はその項目に関する「Price」属性の値として割り当てられる。他の実施形態では、ソフトウェアは、ソフトウェアが属性/値の対を提案するプロバイダとの対話式処理を実行し、次いで属性/値の対はプロバイダによって受け入れられるまたは拒否される。他の好ましい実施形態

40

50

では、htmlタグが走査され、見いだされた情報がタグを有するページに関する属性値を得るために使用される。一例として、ページがhtmlコメントを含む場合は以下ようになる。

```
<! Current price is at http://www.todayspricesforbigco.com?id=32423490 !>
```

ソフトウェアは、示されたURLから現在の価格を得て、それをそのウェブページに関するPrice属性の値とする。

【 0 0 3 0 】

属性およびラベルがデータ項目と関連付けられている(306)と、データ項目はそれらが検索できるようにインデックス付けされる(309)。第1の好ましい実施形態では、属性およびラベルならびにそれらの値もインデックス付けされるが、他の好ましい実施形態では、それらは個別に検索され、個別にインデックス付けされる。

10

【 0 0 3 1 】

図5aは、リポジトリ195内にラベルおよび属性を格納するために使用される形式500の一例を示す。各項目は、そのタイプに適した特定の属性およびラベルに関連付けられる。例えば、求人募集は属性の職務権限(製品管理)、雇用者(ABC Corporation)、職種(専門家)を有することができる。好ましい実施形態での属性およびラベルは、下記のタイプの値を有することができる。

BOOLEAN

INT

FLOAT

URL

STRING

LOCATION

DATE

DATE RANGE

20

【 0 0 3 2 】

属性およびラベルは、メタタグによって格納中に下記のように示される。

```
<start name>
```

```
  name
```

```
</end name>
```

```
<start value>
```

```
  value
```

```
</end value>
```

30

【 0 0 3 3 】

したがって、好ましい実施形態では各属性は、「journal」という属性名および「Journal of Inflammation」という「journal」属性に関する値などの名前/値の対である(図5bを参照)。各ラベルは、特定の雑誌が医学雑誌であることを示すであろう「Medical」などの名前だけを有する(図5cを参照)。好ましい実施形態では、データ項目の情報タイプもそのラベルのうちの1つの名前である。したがって、「Events and Activities」という情報タイプを有するデータ項目は、同一の名前を有するラベルも有するであろう。そのようにユーザは、データ項目の情報タイプと同一の名前を有するラベルを指定することによって特定の情報タイプを有するデータを検索することができる。

40

【 0 0 3 4 】

図5dは、情報タイプをそれらの属性にマッピングするための例示のデータ構造を示す。したがって、データの集合体190内の項目が「Product」という情報タイプを有する場合、項目の属性は、図5cのデータ構造にアクセスすることによって決定されてよく、そのデータ構造は情報タイプ「Product」に関する属性およびそれらの属性タイプを含む。

【 0 0 3 5 】

図5dに示されるように、各情報タイプは予め定義された属性を有する。属性の値は属性タイプである。図5eはいくつかの実際の値を示す。したがって、「Journal」という情報

50

タイプは、属性タイプの文字列の値を有する「Journal name」という属性およびヌル値を有する「Medical」というラベルを有する。例えば、この種の属性によって、ユーザが、特定の雑誌の題名を検索するまたはすべての医学雑誌を検索することができる。同様に、「Product」という情報タイプは、販売のために使用できるいくつかの特定の製品を示す「NumAvail」という属性を有し、整数の属性タイプを有する。すべての属性は任意選択である。プロバイダは、プロバイダに提案された属性のうちのいずれかを設定するために選択するまたはプロバイダ独自のものを生成することができる。

【0036】

図3bは、1つの受信した照会用語または複数の照会用語に応答して照会結果を表示する方法を示す流れ図310である。好ましい実施形態では、照会結果は検索エンジン185によって決定される。例えば、「cancer receptor」402の照会(図4aを参照)は、図4bに示されるものなどの属性404を有する項目の照会結果406を返すことができる(312)。上述のように、本発明のいくつかの実施形態は、この時点で照会結果406を決定するが表示しない。

10

【0037】

照会結果が照会に対して決定される(場合によっては表示される)と、照会結果に関する属性名およびラベルの少なくともいくつかは表示される(322)。データセット406内のデータ項目はある種の情報タイプを有する。最初に表示される属性404は、照会結果406内のデータ項目の情報タイプに関する属性の一部またはすべてである。照会結果は複数のデータ項目を有し、データ項目のそれぞれは様々な属性を有する。照会結果の上位に現れる属性は、照会結果内で最も一般的である属性であり、最も検索者によってクリックされているまたは詳細化されている属性である。例えば、照会「housing」は、属性としてbedrooms(寝室)およびbathrooms(浴室)を有する多数の項目を有し、検索者は、照会のhousing(家)に関する属性の「bathrooms」および「bedrooms」によって常に詳細化している。したがって、bedroomsおよびbathroomsは検索結果の上方の最上位の行に現れるべきである。

20

【0038】

図4bは、照会結果406ならびに複数の属性名およびラベル名404(「journal」、「pubmed」、「news source」、「authors」)を示す。各属性の後の数字は、それに関連した属性を有する照会結果406内の項目の数を示す。例えば、図4bでは照会結果406は、関連の「journal」属性/ラベルを有する2050項目を含む。したがって、特定の照会結果と共に示される属性の数および識別情報は、照会に依存しており、検索を絞り込むために後で選択される属性およびラベルにさらに依存している。

30

【0039】

図3cは、どの属性が所与の照会結果406のために表示されるべきかを決定する方法を示す流れ図340である。エンドユーザが検索を実行するとき、q個の最も関連のある結果が検索エンジン185によって決定され(341)、n個の最も人気のある属性がq個の最も関連のある結果に対して決定される(342)。上位のn個の属性名について、システムは、上位m個の属性/ラベル値を決定する(344)。次いでシステムは、関連のある結果の組内で一致する提供数を計算することによってヒストグラムまたは提供数を計算する(348)。値q、n、およびmはすべて設定可能である。限定を意図しない値の例は、qが1,000~100,000K(qは、特定の照会用語に一致するすべての結果に設定されてもよい)、nが数百の範囲であり、mが20~100の範囲である。

40

【0040】

好ましい実施形態では、属性はヒストグラムが決定される前に正規化される(346)。ある種の実装形態ではある量のデータ整理および正規化が、データがデータ集合体190内に最初に格納されるときに行われる。記載された実施形態ではデータの正規化は、検索されている照会用語に基づいて進行中に行われる(例えば、照会用語が「autos」であるとき、すべての「brand」属性を「make」に正規化することは意味をなすが、照会が「handbag」の場合、すべてのmake属性を「brand」に正規化することが意味をなす)。他の実施形態は、データがデータの集合体190に受信されるときにより多くの正規化を行うことができる。データの正規化は、下記によって好ましい実施形態内で実行される。

50

1. 語幹解釈 (例えば、restaurant=restaurants)
2. 省略 (例えば、sz=size)
3. 単位等価 (例えば、weight=ounces、lbsなど)
4. 入力されなかったスペルの補完

【 0 0 4 1 】

具体的には語幹解釈は、プロバイダが、語尾変化およびミススペルがデータ集合体190内に入り込むことを可能にしてプロバイダ独自の属性名を指定するシステムで有用である。例えば、語幹解釈によって、ユーザは、語幹解釈された属性の「Journals」の単一の選択を用いて「Journal」、「journasl」、「Journsl」などの属性名によってフィルタリングすることができる。

10

【 0 0 4 2 】

ある種の好ましい実施形態では、プロバイダによって追加される属性はタイプチェックがなされる。例えば、URL、DateTime、Number、String、Location、Boolean属性は、それらが有効値であるかどうかを知るためにチェックされる。いくつかの実施形態は、これが様々な実装形態にとって任意選択であるが、各URL値にピングしてそれがアクティブであるかどうかを知る。好ましい実施形態の場合、locationsが、それらが、例えばGoogleMapsなどのオンラインマッピングサービス上で参照できるようにジオコード化される。ある種の実施形態では、ジオコード化できない「location」という属性は無効であると見なされる。

人気のある属性およびラベルが決定され表示される(322)(図3b)と、ユーザは、照会結果に関する表示されたラベルおよび属性値のうちの1つまたは複数を選択することを許可される(324)(図3d参照)。

20

【 0 0 4 3 】

図4cは、ユーザが図4bから属性「journal」を選択し、ユーザが自身の検索を絞り込むことを望む雑誌名をフィールド408内に入力する用意をしている一例を示す。照会用語402はここで「cancer receptor filter:journal」であることに留意されたい。属性「journal」は記載されたコア属性404から姿を消した。

【 0 0 4 4 】

同様に、図4dではユーザは、第2の属性「year」410を選択し、ユーザが指定した雑誌の照会用語内で検索することを望む1年または範囲または複数年を入力する。属性yearは、属性タイプ「range of years」のことである。照会用語402はここで「cancer receptor filter:journal filter:year」であることに留意されたい。属性「year」は、記載された属性404から姿を消した。ユーザがGOボタン411を選択する場合、図4eで表示されるものなどの選択された属性をフィルタおよび表示として使用して、検索が再び実行される。したがって、ユーザは、表示された照会結果に関する1つまたは複数の人気のある属性を選択することができ、表示された属性(またはラベル)による最初の検索をフィルタリングすることができる。ユーザが属性値を空白にしたままである場合、すべての属性値が一致したものとされる。例えば、ユーザが属性Journalを選択するが雑誌名を入力しない場合、journalという属性(および同様に名づけられた属性)を有するデータ項目すべてが照会結果に関する可能性のある候補として選択される。Journalという属性を有しないデータ項目は照会結果に関して選択されない。

30

40

【 0 0 4 5 】

図4dは、ユーザが検索を絞り込むために2つ以上の属性またはラベルを選択した一例を示す。記載された実施形態では複数のラベルおよび属性が、属性およびラベル404の複数のラベルおよび属性上をクリックすることによって選択される。他の好ましい実施形態によって、ラベルおよび属性が検索ウィンドウ402内に入力されることができる。例えば、属性Priceが存在する場合、ユーザは照会用語として下記をタイプすることができる。

Attribute(Price:\$150)

この照会は、Priceという属性および\$150という属性値を有する現在の照会結果内のデータ項目を捜し出す。

50

【 0 0 4 6 】

他の例として、ユーザは下記をタイプすることができる。

Attribute(Price:\$150) AND Label(SmallerThanABreadBox)

この照会は、Priceという属性、および\$150という属性値、ならびにSmallerThanABreadBoxというラベルを有する現在の照会結果内のデータ項目を捜し出す。他の実施形態は、ユーザに属性とラベルを論理的に結合することを可能にするために他の適切なユーザインターフェース要素を使用する。

【 0 0 4 7 】

図4eは、図4dで指定されるように特定の年または年の範囲の特定の雑誌に絞り込まれた照会結果を示す。ユーザは、ユーザが雑誌内を検索412し続けたいかどうか、またはデータ項目の集合体全体を検索(例えば、「Search all of Googlebase」)413すべきかどうかを決定することを許可される。例ではユーザは、領域414内の複数のラベルの選択を提供される(照会結果406'内の30項目、15項目、および6項目にそれぞれ関連付けられる「biotechnology」、「medical」、および「photography」)。例ではユーザは、領域416内の属性(すなわち、Date、author、pubmed、citation)に対する値を指定するための選択をさらに提供される。ユーザは、relevance、date属性、またはユーザが定義した属性のいずれか(例えば、price、locationなど)によって照会結果406'を分類することのオプションも提供される(416)。

【 0 0 4 8 】

図4fではユーザは、図4eの領域414から属性「Date」を選択し、日付420を入力する機会を与えられる。ユーザがドロップダウンの演算子「between」を選択するとき、ユーザは(図示されたように)日付の範囲を選択するための機会を与えられる。属性「Date」は記載された属性418から姿を消した。この例では、「Author」という属性が属性414から姿を消した。属性は、それらが照会および照会結果にもはや関連しない場合に姿を消す。ここで、ユーザが雑誌によってフィルタリングしなかったということは、ユーザが限定された項目セットだけを見ていることであると仮定する。ユーザは、検索を実行するためにGoボタンを選択し、著者用語が再び現れる。

【 0 0 4 9 】

図4gは、著者名422を指定するユーザを示す。照会が詳細化されているとき、属性およびラベルが照会結果に基づいており、かつ照会結果が絶えず変化するので、新しい属性およびラベルが現れる。ユーザはGoボタン423を押すとき、他の検索が、ユーザによって指定される属性および属性値を反映するために照会結果をさらにフィルタリングして実行される。

【 0 0 5 0 】

下記の段落では、属性およびラベルを使用して検索中または検索の絞込み中の属性リポジトリ195のアクセスについて記載する。

【 0 0 5 1 】

照会およびその参照リポジトリ195のインデックス付けが下記の演算子をサポートすることが好ましい。

Number (Is、Between、Greater Than、Less Than、Number Rangeの提案)

String (Is、Has)

Date (Range、Before、After、Is)

Location (Within)

【 0 0 5 2 】

リポジトリ195は、少なくとも下記の方法で照会されてよい。

特定の属性の名とタイプとの対に一致するすべての項目をユーザに与える

属性値の値に基づいてこれらの項目を分類する

属性の下記のタイプに関する分類がサポートされる

・ DateTime

・ Number (Int、Float)

10

20

30

40

50

- ・String
- ・Location (位置を入力したユーザからの距離)

【0053】

この照会能力によって、ユーザは下記のタイプの属性照会を入力することができる。

特定の名前とタイプとの対を有するすべての項目をユーザに与える

属性値によってそれを分類するこれらの項目を与えられる(例えば、イベント日を有するすべての項目をユーザに与え、それを昇順で分類してユーザに与える)

特定の名前とタイプの属性に関する値1と値2との間にあるすべての項目をユーザに与える

【0054】

10

例

属性として料理タイプを有し、15と30との間の値(ここで単位は分)を有するすべての項目をユーザに与える

属性としてサイズを有し、単位の無い値1と15とを有するすべての項目をユーザに与える

イベント日を有し、本日未満の値を有するすべての項目をユーザに与える

発行日を有し、1925年での値を有するすべての項目をユーザに与える

【0055】

下記の演算子がサポートされる

数字の場合

20

int、float

Less than

Greater than

Between

日時の場合

Is

Before

After

Between

Scoring of Items(項目の点数付け)

30

【0056】

現在、項目が点数付けされる2つの重大な信号がある

照会依存ランク(主にIRスコア)

照会独立ランク(ページランクおよび項目ランクの混合)

【0057】

ページランクは、プロバイダのウェブサイトのページランクである。ページランクは、項目がデータ集合体190内でホストされるおよび/または項目が他の項目に関連付けられないもしくは接続されない場合には存在しない。

【0058】

項目ランクは、複数の要因によって決定されてよい。2つの主な信号は、

40

プロバイダ特定信号(例えば、評点)

提供特定信号(例えば、descの長さ、属性、ラベル、画像の数など)

項目ランクは下記の信号によって定義されてよい

Descの長さ

題名の長さ

ラベルの数

属性の数

画像

提供がスパムとして報告された回数

プロバイダの評価

50

提供の最新さ

項目は、(照会依存ランク*照会独立ランク)として点数付けされる。既定の分類についてランクは既定分類である。

【0059】

好ましい実施形態ではある種のパラメータがシステム内で設定されてよい。これらのパラメータは、プロバイダあたりの項目の最大数を含む。これは、特定のプロバイダによってページが密集するのを防止する。

【0060】

ユーザが検索を絞り込むために属性および/またはラベルを選択するとき、システムは、ラベル、題名、説明および属性値を検索する。属性名はまた、完全な名前として検索可能であるべきである。句は、遠く離れて生じる語と比較して重く重み付けられる。ラベルは題名よりも重く重み付けられ、題名は説明よりも重く重み付けられる。属性値は、ラベルと同一に重み付けられる。各プロバイダによるマーチャント密集は、個別のプロバイダからの項目のページ数が、検索の結果として表示されるか否かを規制するためにユーザによってオンまたはオフされてよい。実行される検索に応じてマーチャント密集は望ましいかどうかはわからない。

【0061】

好ましい実施形態ではシステムは、同一のまたは類似のタイプの他の項目に関連した属性に基づいて新しい項目の特定のタイプの構造を定義する(例えば、情報タイプ「Jobs」のほとんどの項目がJob function、Job typeおよびEmployerという属性を有する場合、情報タイプ「job」のデータ項目に関する共通の属性構造は、職種、雇用者および職務権限であるように既定値を設定することになる)。検索者および他のプログラムは、「Give me all jobs whose employer is ABC Corporation and whose job-type is product management(雇用者がABC Corporationであり、職種が製品管理であるすべての職務を与える)」などの照会を用いてデータセットを照会することができる。

【0062】

本明細書に記載された例は人間のユーザを参照するが、当然のことながら、本発明の他の実施形態は、人口知能ソフトウェアプログラムなどの人ではないユーザと共に、または人もしくは人でないもののいずれかであってよいウェブを介して通信するエンティティと共に動作するように設計されてよい。人ではないユーザがソフトウェアプログラムである場合、本明細書に記載された結果および属性を表示する必要がなくてよい。代わりに、この種の実装形態は、照会結果を絞り込むために使用されうる可能性のある属性を伝達だけすることができる。この種の実施形態では、人ではない人口知能が選択されるべき多数の属性によって苦しめられないのでより多くの属性のオプションが表示されてよい。この種の実施形態では、ヒストグラムを決定することなどの方法の要素は必要でなくてよく、または方法の要素が、いくつかの利用できる属性選択に限定するためではなく、属性選択をランク付けするためだけに使用されてよい。

【0063】

当然のことながら、構造化データ190内の様々な情報タイプに関するコア属性が定期的に更新される必要がある。データが構造化データの集合体に追加されるとき、最初に人気なかったある種の属性が人気になりうる。例えば、配役写真がテレビ番組の何期からのものかを指定することができる整数の属性タイプを有する「Season」属性は、情報型「ワイドショー(TV shows)」に対する初期コア属性によって最初は意図していなかったが、ますます多くの配役写真がデータの集合体に追加されるにつれて人気になりうる。ある実施形態では、コア属性は、人気度および季節性に基づいて、スパムフィルタを通過後に自動更新もされる

【0064】

図3eは、任意の新しいプロバイダが提供する属性が情報タイプに関するコア属性に格上げされるべきかどうかを決定するために定期的に行われる方法350を示す。項目の情報タイプに関する属性のコアグループは、プロバイダが情報タイプの新しい項目を追加する

10

20

30

40

50

ときにはいつでも自動的に提供される属性である。好ましい実施形態ではコア属性だけが、プロバイダが表示される属性に強制的に入るために属性をスパムする可能性を低減するように提供される。情報タイプごとにその方法は、その情報タイプに関する最も人気のあるユーザが追加した属性を見て(352)、その情報タイプに関して最も人気のある属性をコア属性に格上げする。

【0065】

どの属性をコア属性に格上げすべきかを決定するために使用される「最も人気のあるもの」は、様々な実施形態に対して異なって定義される。例えば、最も人気のあるものは、コア属性内にはなく、例えば週または月などの所定の期間にわたってユーザによって最も頻繁に選択される属性であってよい(352)。他の例として、最も人気のあるものは、コア属性内にはなく、所定の期間にわたって照会結果内で最も頻繁に現れるデータ項目を有する属性であってよい。他の例として、最も人気のあるものは、コア属性内にはなく、所定の期間にわたって最多数のプロバイダのデータ内に現れる属性であってよい。最も人気のあるものは、コア属性に追加されることによって属性が検索を絞り込むのに有用である限り任意の適切な方法で決定されてよい。

10

【0066】

例えば、プロバイダは、記事がブログ内で言及されたことを示すために記事の項目の情報タイプに関して「blogged」の属性を追加し始めていてよい。この種の属性は、項目が言及されたブログのURLを示すURL属性タイプを有するであろう。固有のプロバイダまたはユーザの閾値354が情報タイプに関する特定の新しい属性を使用する場合、属性はその情報タイプに関する属性のコアグループに追加される(356)。好ましい実施形態では、閾値はシステムを使用するプロバイダの総数に基づいている。閾値は、2~3の低いものから始め、より大きい数字に増大されることになる。同様の方法が、人気のあるラベルをラベルのコアセットに追加するためにラベルに対して実行される。ある種の好ましい実施形態では、格上げされた属性は、人または適切なソフトウェアもしくはハードウェアが実装された方法によって健全にチェックされる。

20

【0067】

概して、前の段落には構造化データの集合体190内に入力されたデータを検索し更新する方法を記載した。以下の段落では、プロバイダが構造化データの集合体190にデータを入力するまたはデータを追加することができる方法を記載する。ある種の好ましい実施形態では、プロバイダはそれらのデータに関する新しい属性を指定することもできる。

30

【0068】

図6a~6eは、どのようにプロバイダがデータ集合体内で項目を編集することができるかを示す例示のスクリーンショットである。プロバイダは、データ集合体190に内容を追加するまたは追加することができる任意の人物である。記載された実施形態では、データの集合体190は、個人、非営利組織、または企業などの1つまたは複数のプロバイダによって所有されるデータである。実施形態によってこの種のプロバイダは、ウェブを介してプロバイダ独自の構造化データ(例えばデータベース)の集合体を設定し追加することができる。ウェブまたは同様のネットワークを介してそれらのデータの集合体を検索可能にすることができる。プロバイダは、データが他者によって検索されることができるよう、有料かまたはプロバイダの許可と引き換えかのいずれかで中央のリポジトリ内のデータを格納することを望んでいると考えられる。このような状態ではデータ集合体は、本明細書に記載された機能の一部またはすべてを含むバージョン内で、GoogleブラウザまたはGoogleデスクトップ検索エンジンなどのウェブまたはネットワークベースのブラウザを介して検索されてよい。

40

【0069】

図6a~6eは、プロバイダに、データを編集し、かつシステム内に入力することを可能にするユーザインターフェースを示す例示のスクリーンショットである。

【0070】

図6aは、プロバイダにデータの集合体190内のデータ項目を見て編集することを可能に

50

するユーザインターフェース600を示す。ユーザインターフェースは、データの集合体190に項目を追加するために使用されてもよい。領域602は、データの集合体190内の項目の一部のリストを含む。この例ではこのリストは、項目題名601、項目タイプ(情報タイプとも呼ばれる)603、状態605、有効期限、いくつかのインプレッション(項目が表示された回数)、いくつかの対象上のクリック、およびクリックスルー率(検索結果内で項目がクリックされた回数)を含む。例では、データ集合体内のすべての項目のサブセットが領域602に示されているが、プロバイダは、プロバイダの個人のデータ集合体620または全体のデータ集合体622のいずれかを検索することもできる。プロバイダは、非アクティブ項目616を見るまたはパルクファイル618をアップロードすることもできる。各データ項目は、関連の「edit」リンク619を有する。好ましい実施形態では、プロバイダはプロバイダ独自のデータ項目を編集だけすることができる。領域604によって、プロバイダは、既存の情報タイプ(Events and Activities、Housingなど)を示すドロップダウンメニューなどの選択デバイスを表示することができる。プロバイダが情報タイプを選択する場合、プロバイダは、プロバイダのデータに関する領域606内の情報タイプの説明を追加することができる。

10

20

30

40

50

【0071】

図6bは、プロバイダにデータの集合体190内のデータ項目を見て編集することができる(610)ユーザインターフェースを示す。項目は「News and articles」という情報タイプを有する。プロバイダが図6aの領域602内のデータ項目を選択した場合、その項目の情報は領域のフィールド611内に表示される。しかし、例ではプロバイダは項目を選択せず、したがってプロバイダは新しいデータ項目を自由に入力する。例では「News and articles」という情報タイプ610は以下のフィールドを含む。Title(題名)、Pictures(画像)、Description(説明)および照会結果内に表示されるべきリンク614(例えば、URL)。

【0072】

図6bのユーザインターフェースによって、プロバイダは、項目の属性およびラベルも編集することができる。各情報タイプは関連の属性を有するが、特定のタイプのすべてのデータ項目が、その情報タイプに関するすべての可能性のある属性に対する値を有するとは限らないことに留意されたい。例では、参照番号612によって示されるようにプロバイダは、数量「1」の項目が利用できるまたは存在することを示した。値が、この項目に関するAuthor属性またはNews Source属性に対して指定されていない。それらの属性のそれぞれは、「text」という属性タイプを有する。プロバイダは、個別のデータ項目の属性に対する値を自由に追加する。プロバイダは、領域613を使用して属性を追加することもできる。ここで、プロバイダは属性名および属性値を追加することができる。

【0073】

プロバイダは、領域618内の連絡先情報に関する属性値を提供することができる。プロバイダは、領域619内の位置情報に関する属性値を提供することができる。

【0074】

プロバイダは、領域619内の項目にラベルを追加することができる。ある種の実施形態では、情報タイプは既定の属性名である。ここで、情報タイプはNews and Articlesであり、これがラベルでもある。

【0075】

図6cは、プロバイダにデータの集合体190内のデータ項目を見て編集する(610)ことを可能にする図6bのユーザインターフェースを示す。例では、プロバイダは新しいプロバイダが定義した属性613に関する名前および値を追加することができる。既定の属性タイプは「text」であるが、プロバイダは、数字単位、数字、データ範囲、大きいテキスト、URL、ブル、および位置などの他の属性タイプを選択することができる。

【0076】

図6dは、プロバイダにデータ集合体190内のデータ項目を見て編集する(610)ことを可能にするユーザインターフェースを示す。項目は「Product」という情報タイプ630を有する。プロバイダが図6aの領域602内のデータ項目を選択した場合、その項目の情報は領域の

フィールド611内に表示される。しかし、例では、プロバイダは項目を選択せず、したがってプロバイダはユーザインターフェースを使用して新しい項目を自由に入力する(630)。例では、「Products」という情報タイプは以下のフィールドを含む。Title(題名)、Pictures(画像)、Descriptions(説明)および照会結果内に表示されるべきリンク634(例えば、URL)。

【 0 0 7 7 】

図6dのユーザインターフェースによって、プロバイダは、項目の属性およびラベルを編集することもできる。各情報タイプは関連の属性を有するが、特定の情報タイプのすべてのデータ項目がその情報タイプに関するすべての可能性のある属性に対する値を有するとは限らないことに留意されたい。例では、参照番号632によって示されるように、プロバイダは、項目あたり(例えば、ポンドあたりまたはダースあたりとは対照的に)\$150の価格を示している。数量「1」が指定される。Priceタイプは、プロバイダが設定している価格のタイプである(例えば、最高付け値、交渉可能、固定など)。この項目に関するPrice option(価格オプション)、Brand(ブランド)、Condition(条件)、およびProduct Type(製品タイプ)に対する値は指定されていない。それらの属性のそれぞれは、「text」という属性タイプを有する。この実施形態ではプロバイダは、プロバイダが指定したそれらの属性に関する属性タイプを変更することができる。プロバイダは、個別のデータ項目の属性に対する値を自由に追加することができる。プロバイダは、領域613を使用して属性を追加することもできる。ここで、プロバイダは属性名および属性値を追加することができる。

10

【 0 0 7 8 】

この実施形態では、プロバイダが追加する属性は、現在の情報タイプのプロバイダの項目のすべてに追加される。ここで例えば、タイプ「Products」のプロバイダの項目のすべてが、新しく追加された属性613が定義されると新しく追加された属性613を与えられる。項目ごとの値は一般に、個々に追加される。ある種の実施形態によって、プロバイダは、指定された情報タイプのプロバイダの項目のすべてに対して値を指定することもできる。上述のように、新しい属性が属性のコアセットに昇格することが可能である。他の実施形態では、新しい属性が情報のタイプのすべての項目に必ずしも追加されない。他の実施形態では、プロバイダは、定義されたグループのプロバイダのすべてが同一の属性を有することに合意することができ、その結果、1つのプロバイダが属性を追加するとき、グループ内の他のプロバイダも同一の属性を有することになる。

20

30

【 0 0 7 9 】

プロバイダは、領域618内の連絡先情報に関する属性値を提供することができる。プロバイダは、領域619内の位置情報に関する属性値を提供することができる。プロバイダは、領域638内のPayment(支払い)方法に関する属性値を提供することができる。

【 0 0 8 0 】

プロバイダは、ラベルを領域616内の項目に追加することができる。ある種の実施形態では、情報タイプは既定の属性名である。ここで、情報タイプはProductsであり、これがラベルでもある。この実施形態では、プロバイダが追加するラベルは、現在のタイプのプロバイダの項目のすべてに追加されるわけではない(情報タイプであるラベルを除いて)。上述のように、新しいラベルがラベルのコアセットに昇格することが可能である。他の実施形態では、新しいラベルは、情報タイプのすべての項目に常に追加される。

40

【 0 0 8 1 】

図6eは、プロバイダにデータ集合体190内のデータ項目を見て編集する(630)ことが可能になる図6dのユーザインターフェースを示す。この例では、Contacts(連絡先)、Payments(支払い)、およびLocation(位置)がProduct(製品)情報タイプのすべての属性である。それらは、複雑なタイプ(単なる整数または簡単な文字列ではない)を有する属性である。例では、プロバイダは情報タイプ「Products」の項目に関する連絡先618に関する値を追加することができる。ここでプロバイダは、ニックネーム、電話番号、電子メールアドレスの一部またはすべてを指定する(プロバイダ情報のデータベースから取り込まれる可能性のある値、図示せず)。例ではプロバイダは、情報タイプ「Products」の項目に関するPay

50

ments(支払い)638に関する値を追加することができる。ここで、プロバイダはPayment(支払い)方法およびNotes(メモ)の一部またはすべてを指定する。例では、プロバイダは情報タイプ「Products」の項目に関する位置619に関する値を追加することができる。ここで、プロバイダはテキストメモの一部またはすべてを指定する(例えば、「Fremont,CA」)。この実施形態では、顧客がこの位置および配達区域から受け取ることができるかどうかを示すためのチェックボックスもある。

【0082】

例では、Contact(連絡先)の値、Payment(支払い)の値、およびLocation(位置)の値が項目ごとに個別に入力される。プロバイダが追加する値は、現在の情報タイプのプロバイダの項目のすべてに追加されるわけではない。ここで例えば、情報タイプ「Products」のプロバイダの項目のすべてが、図6eに示されるContact(連絡先)の値、Payment(支払い)の値、およびLocation(位置)の値を与えられるとは限らない。一般に、項目ごとの値は個々に追加される。ある種の実施形態によって、プロバイダは、指定された情報タイプのプロバイダの項目のすべてに対して値を指定することもできる。例えば、支払い情報は、プロバイダのすべての「Products」に関して同一であってよい。

10

【0083】

プロモータは、図6のUIを介してまたは図7および図8に示されるバルクアップロード方法を介してのいずれかで項目を入力することができる。

【0084】

図7は、バルクアップロードファイルを登録するためのユーザインターフェースを示す例示のスクリーンショット700である。バルクアップロードファイルは、データ集合体190を生成するまたはデータ集合体190に追加するために使用される。この例では、同一の情報タイプのすべてである項目のフラットファイルが追加されるべきである。例では、ファイル名712は「local inventory」である。プロバイダは、予め定義された情報タイプまたはカスタムの情報タイプであるデータタイプ714を選択する。プロバイダは、データ内のテキスト文字列に対する言語716を選択する。プロバイダがボタン「Register bulk upload file」(バルクアップロードファイルを登録する)718を選択するとき、ファイル名712を有するファイルが登録され、次いでプロバイダはファイルをアップロードすることが許可される。プロバイダは、ウェブベースのアップローディングインターフェースを使用してまたはFTP(ファイル転送プロトコル)もしくはRSSなどの他の機構を使用してファイルをアップロードすることができる。

20

30

【0085】

図8aは、バルクアップロードされるべきタブ区切りのファイルに関する形式801を示す。下記は、バルクアップロードファイルに対する形式の要求である。

タブ区切りされたプレーンテキスト

ファイルの第1の行はヘッダである (タブによって分離された属性名(以下に記載される)を含まなければならない)

行あたり1つの項目 (各属性はタブによって分離されるべきである)

最後の行に末尾タブを有しない

ファイルはLATIN1またはUTF-8符号化で保存されなければならない。ASCIIも、それがLATIN1のサブセットであるので許容可能である。

40

リンクおよび画像URLは完全に揃ってなければならない。つまり、それらはhttp://portion (例えば、http://www.example.com/image.gif)を含まなければならない。

タブ、キャリッジリターン、または改行文字 (これらのいずれかが属性内に現れる場合、その項目を表示することはできない)

HTMLタグ、コメント、およびエスケープシーケンス (htmlはバルクアップロードから除去されないが、最良の外観のためにHTMLは含まれるべきではない)

【0086】

好ましい実施形態では、データ項目は、属性も含むアップロードされるファイルの一部である。他の好ましい実施形態ではデータ項目および属性は構築される個別のファイル内

50

にアップロードされ、その結果、どの属性値がどのデータ項目に関係するかが明らかである。

【0087】

図8bは、バルクアップロードファイルを生成するためにプロバイダによって使用される例示の方法の流れ図800である。プロバイダは、人、またはハードウェアもしくはソフトウェアであってよい。

【0088】

要素802 (表計算プログラム内で新しいファイルを開く)

記載された方法は、バルクアップロードファイルを生成するためにMicrosoft Excelなどの表計算プログラムを使用する。Microsoft Excelのような表計算プログラムを使用すると、バルクアップロードを生成し、かつそれを適切な形式に変換することを容易にする。適切にフォーマットされたファイルをもたらす他の方法が使用されてよい。

【0089】

要素804 (ヘッダ行を生成する)

一例として、製品バルクアップロードに関するヘッダ行は図8c内の列832のように見える。プロバイダがサブミットしたい項目の情報タイプによりバルクアップロード内の列のそれぞれを指定する(図7の714を参照)。スプレッドシート832の第1の行上に、プロバイダがプロバイダの項目を説明するために含めたい属性のそれぞれの名前を入力する。これがヘッダ行である。ヘッダ行の内容は、サブミットされる情報の情報タイプに依存し、プロバイダが、定義された情報タイプを送信しているか、プロバイダがプロバイダ自体で生成したものを送信しているかに依存する。

【0090】

カスタム情報タイプ

バルクアップロードは、任意の情報タイプをサブミットするために使用されてよい。プロバイダがプロバイダ独自の情報タイプを送信している場合、プロバイダは予め定義された属性の任意の組合せを使用することができる。好ましい実施形態では、プロバイダが予め定義された属性を使用することが強く推奨される。プロバイダは無制限の数のカスタムの属性を含むこともできる。プロバイダは、プロバイダの項目を最良に説明する1組の属性を選ぶべきである。

【0091】

定義された情報タイプ

プロバイダは、定義された情報タイプのうちの1つに関するバルクアップロードを送信することができる。プロバイダがプロバイダのバルクアップロード内にそれらを含むことが強く推奨される。それらは、照会を検索するために項目のより正確な一致を可能にする。プロバイダが与える情報が多ければ多いほど、ユーザが項目を捜し出すことが容易になる。好ましい実施形態ではプロバイダは、プロバイダの項目が、行われた検索の重要な部分に現れることを可能にするためにこれらの推奨された属性を含まなければならない。

【0092】

要素806 (項目情報を入力する)

各行834上でプロバイダは、プロバイダのデータ集合体内の項目に関する情報を入力する。情報の各片は、それが存在する列のヘッダに反映すべきである。(例えば、製品の価格は「price」ヘッダの下に行くべきである)。各行は行あたり1つの項目だけを含む。図8c参照。

【0093】

要素808 (バルクアップロードをタブ区切りのプレーンテキストに変換する)

以前に登録されたファイル名を使用してスプレッドシートをタブ区切りのテキスト(.txt)ファイルに変換する(図7参照)。プロバイダがスプレッドシート内のすべての項目を入力した後、プロバイダはタブ区切りのテキスト(.txt)形式でスプレッドシートを保存する。登録されたファイル名は後続のアップロードのために再使用されてよい。アップロードされたファイルが登録されていない名前を有する場合、ファイル内の項目は、データ集合

10

20

30

40

50

体190に追加されない。好ましい実施形態では更新されたバルクアップロードは、項目がデータ集合体190内に残っていることを保証するために30日ごとに少なくとも1度送信されなければならない。

【0094】

要素810 (ファイルをアップロードする)

図8dはファイルをアップロードするためのユーザインターフェース840を示す。

【0095】

要素812 (誤りに関してバルクアップロードをチェックする)

プロバイダがバルクアップロードを送信した後、プロバイダは、中央のウェブサイトログインすることによってバルクアップロードの状態を知ることができる。結果が「Success」として記載される場合、バルクアップロードは変更される必要はない。そうでない場合、プロバイダは、誤りを訂正する方法についての情報を知るためにバルクアップロードのファイル名上をクリックすることができる。

10

【0096】

バルクアップロードがアップロードされた後、ファイルは、項目、属性、およびラベルをデータ集合体190および図5のデータ構造に追加するように処理されることになる。アップロードが承認されると同一のファイル名を有する任意の今後の更新が自動的に処理されることになる。

【0097】

本発明がいくつかの実施形態に関して上述されたが、様々な変更形態が本発明の範囲内で作成されてよい。例えば、ある種の好ましい実施形態は、無効なまたは「スパムの(spammy)」属性およびラベルを検出する方法およびシステムを含む。データ項目が検索の最上位になることを可能にする属性をプロバイダのデータに追加することはプロバイダにとって望ましくない。この種の属性を避けるために使用されるいくつかの方法は、ブラックリスト、特定のヒストグラム分布などを含む。

20

【0098】

他の好ましい実施形態では、表示された上位の属性およびラベルは、属性の主要なタイプの組およびラベルの人気度だけではなく値の分布(分布がより離散的であればあるほどよく、スキューが多ければ多いほどよい。例えば、属性に関する5つの人気のある値は一樣に分布された50個の値よりよい)に基づいて決定される。例えば、色が属性であり、赤、青および緑が上位の色として知られる場合、色は詳細化するのに良い属性であろう。一方、それぞれが3回生じる色に対して100個の値を有することはそれほど有用ではない。

30

【0099】

他の好ましい実施形態は、属性、各提供の項目ランク/提供ランクを使用するプロバイダの数に基づいて高度な信頼度点数付けを行う。

【0100】

他の好ましい実施形態は、どの属性をユーザに表示すべきかを決定するためにユーザからのクリック信号を使用する。属性およびラベルは、人気度ランクとして定義された何らかのものによって点数付けられる。

$PR = \text{照会結果の人気度} * \text{その特定の照会に関するCTR}$

40

【0101】

他の好ましい実施形態では、ユーザのALWAYS2属性が特定の照会を制限する場合(例えば、90%の割合でIpodは価格および位置について常に制限され、システムは、ユーザがipodをタイプするときに進む価格および位置によって制限する)、照会結果に既に適用されたそれらの制限を示す。

【0102】

したがって、本発明の開示は、添付の特許請求の範囲に記載された本発明の範囲の例示であって限定するものではない。

【図面の簡単な説明】

【0103】

50

【図 1 a】本発明の好ましい実施形態によるデータ処理システムを示す構成図である。

【図 1 b】本発明の好ましい実施形態による他のデータ処理システムを示す構成図である。

【図 1 c】本発明の好ましい実施形態によるアーキテクチャ図である。

【図 2 a】本発明の好ましい実施形態による検索可能なデータ項目の集合体の生成の概要を示す流れ図である。

【図 2 b】本発明の好ましい実施形態による文書の集合体を検索し、検索を詳細化することの概要を示す流れ図である。

【図 3 a】データ項目の集合体からラベルおよび属性を抽出する方法を示す流れ図である。

【図 3 b】照会用語を受信し、照会結果を表示する方法を示す流れ図である。

【図 3 c】どの属性が所与の照会結果に関して表示されるべきかを決定する方法を示す流れ図である。

【図 3 d】ユーザにラベル値および/または属性値を使用して表示される照会結果を詳細化することを可能にする方法を示す流れ図である。

【図 3 e】何らかの新しいプロバイダ提供の属性が情報タイプに関するコア属性に追加されるべきかどうかを決定するために定期的に行われる方法を示す図である。

【図 4 a】検索エンジンおよびユーザによって入力される照会用語の例示のスクリーンショットを示す図である。

【図 4 b】図4aの照会からの照会結果を示し、かつ照会用語に関する照会結果に関連するラベルおよび属性も示す例示のスクリーンショットを示す図である。

【図 4 c】追加の属性およびラベルならびにどのようにユーザが属性および/またはラベルを使用してユーザの検索を絞り込むことができるかを示す例示のスクリーンショットを示す図である。

【図 4 d】追加の属性およびラベルならびにどのようにユーザが属性および/またはラベルを使用してユーザの検索を絞り込むことができるかを示す例示のスクリーンショットを示す図である。

【図 4 e】追加の属性およびラベルならびにどのようにユーザが属性および/またはラベルを使用してユーザの検索を絞り込むことができるかを示す例示のスクリーンショットを示す図である。

【図 4 f】追加の属性およびラベルならびにどのようにユーザが属性および/またはラベルを使用してユーザの検索を絞り込むことができるかを示す例示のスクリーンショットを示す図である。

【図 4 g】追加の属性およびラベルならびにどのようにユーザが属性および/またはラベルを使用してユーザの検索を絞り込むことができるかを示す例示のスクリーンショットを示す図である。

【図 5 a】検索可能なデータの集合体に関する属性およびラベルを格納するために使用されるデータ形式を示す図である。

【図 5 b】図5aの形式を使用して格納される属性の一例を示す図である。

【図 5 c】図5aの形式を使用して格納されるラベルの一例を示す図である。

【図 5 d】情報タイプをそれらの属性にマッピングする例示のデータ構造を示す図である。

【図 5 e】その情報タイプに関するいくつかの例示の属性にマッピングされた情報タイプの一例を示す図である。

【図 6 a】プロバイダにデータを編集しシステムに入力することを可能にするユーザインターフェースを示す例示のスクリーンショットを示す図である。

【図 6 b】プロバイダにデータを編集しシステムに入力することを可能にするユーザインターフェースを示す例示のスクリーンショットを示す図である。

【図 6 c】プロバイダにデータを編集しシステムに入力することを可能にするユーザインターフェースを示す例示のスクリーンショットを示す図である。

10

20

30

40

50

【図 6 d】プロバイダにデータを編集しシステムに入力することを可能にするユーザインターフェースを示す例示のスクリーンショットを示す図である。

【図 6 e】プロバイダにデータを編集しシステムに入力することを可能にするユーザインターフェースを示す例示のスクリーンショットを示す図である。

【図 7】バルクアップロードファイルを登録するためのユーザインターフェースを示す例示のスクリーンショットを示す図である。

【図 8 a】どのようにプロバイダがデータおよび属性値のバルクアップロードを行うかを示す図である。

【図 8 b】どのようにプロバイダがデータおよび属性値のバルクアップロードを行うかを示す図である。

【図 8 c】どのようにプロバイダがデータおよび属性値のバルクアップロードを行うかを示す図である。

【図 8 d】どのようにプロバイダがデータおよび属性値のバルクアップロードを行うかを示す図である。

【符号の説明】

【0104】

- 100, 111 データ処理システム
- 110 a, 110 b, 110 n ユーザデータ処理システム
- 120 サーバデータ処理システム
- 130 ネットワーク
- 140 プロセッサ
- 150 ブラウザ
- 160 メモリ
- 170 プロセッサ
- 180 メモリ
- 185 検索および照会エンジン
- 190 構造化データの集合体
- 195 属性リポジトリ

10

20

【図 1 a】

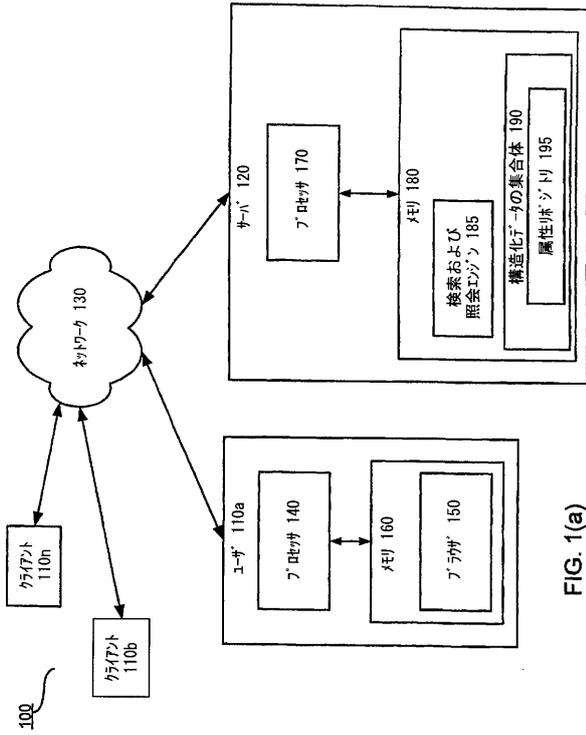


FIG. 1(a)

【図 1 b】

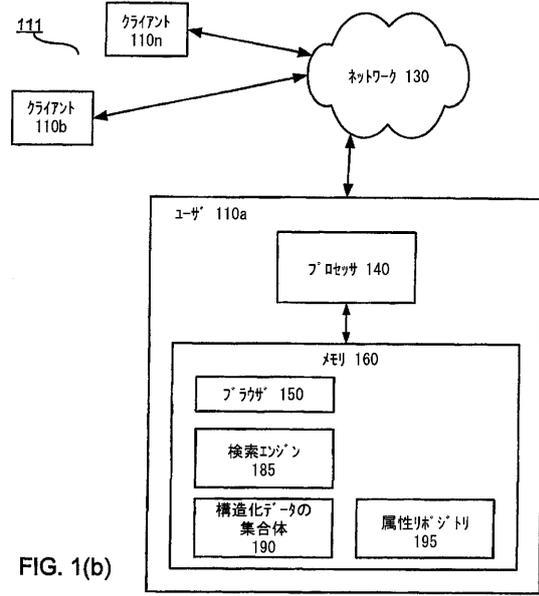


FIG. 1(b)

【図 1 c】

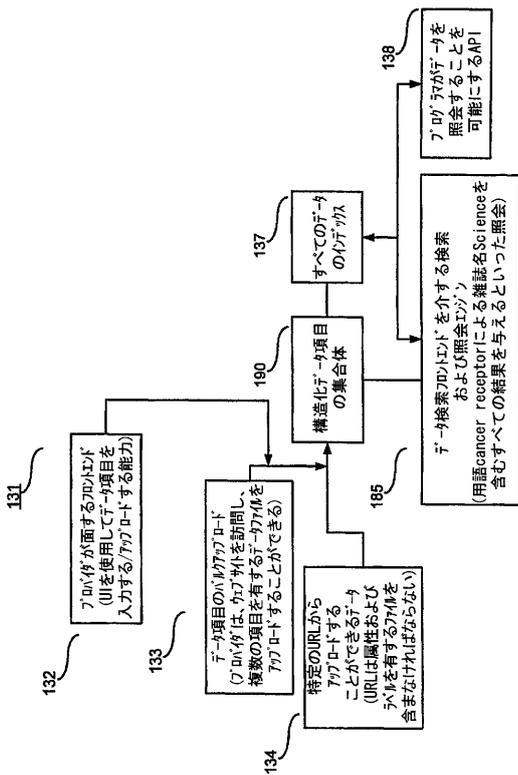


Fig. 1(c)

【図 2 a】

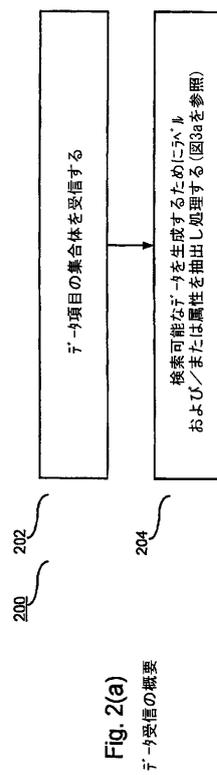


Fig. 2(a)

検索項目の概要

【 図 2 b 】

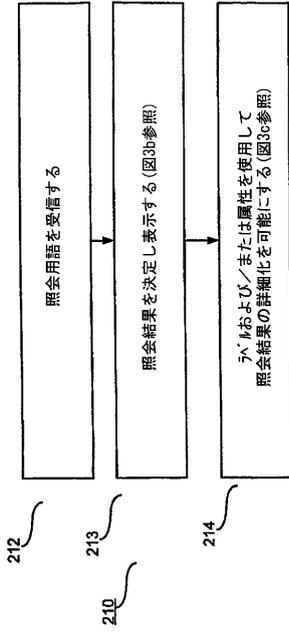


Fig. 2(b)
検索の概要

【 図 3 a 】

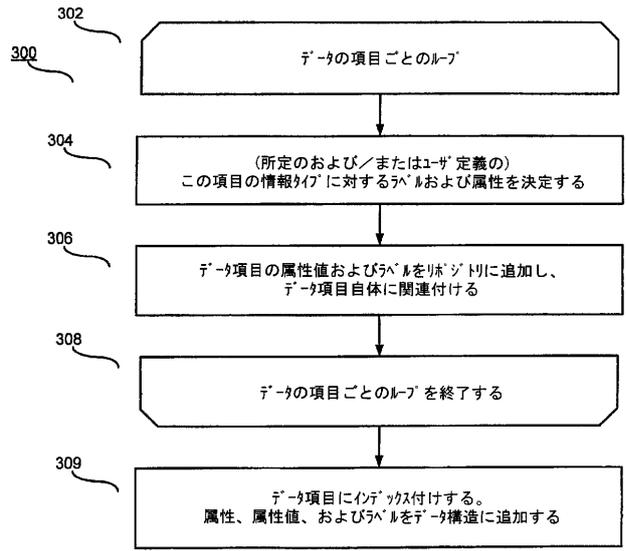


Fig. 3(a)

タブおよび/または属性を抽出し処理する(設定)

【 図 3 b 】

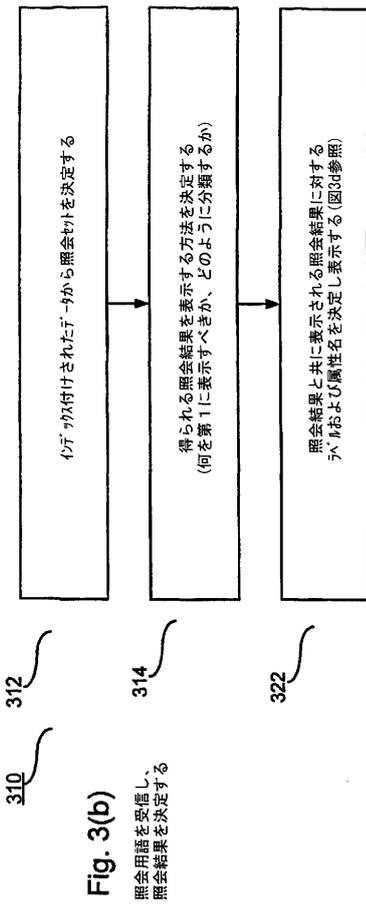


Fig. 3(b)
照会用語を受信し、照会結果を決定する

【 図 3 c 】

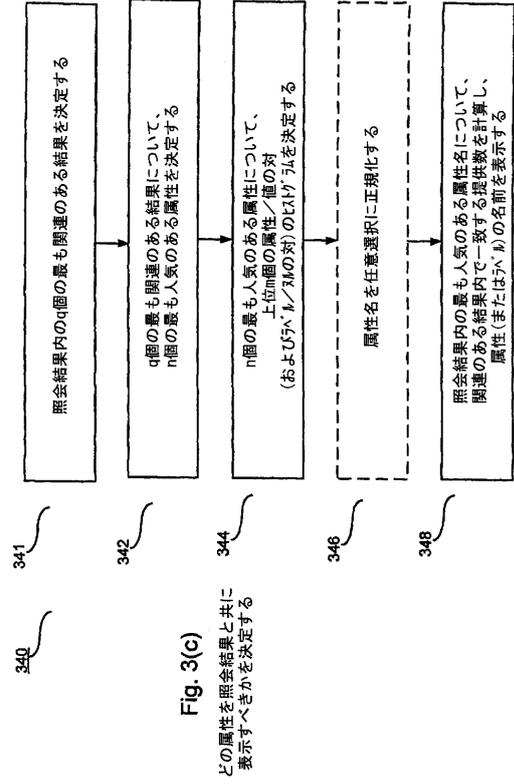
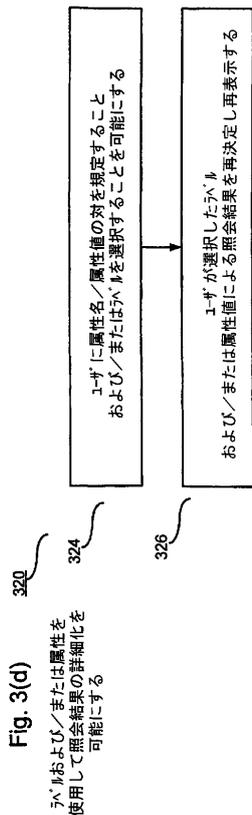


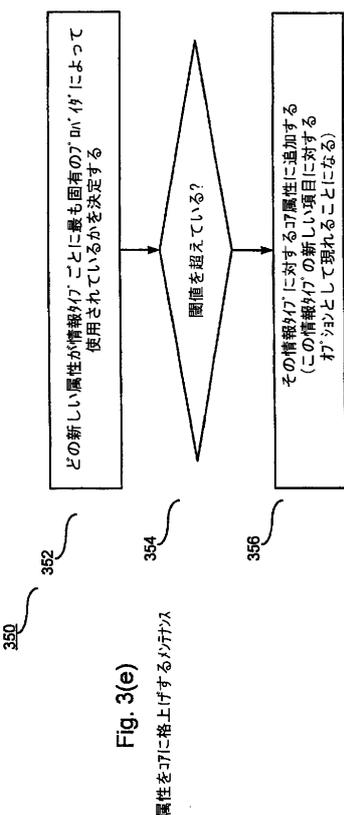
Fig. 3(c)

どの属性を照会結果と共に表示すべきかを決定する

【 図 3 d 】



【 図 3 e 】



【 図 4 a 】

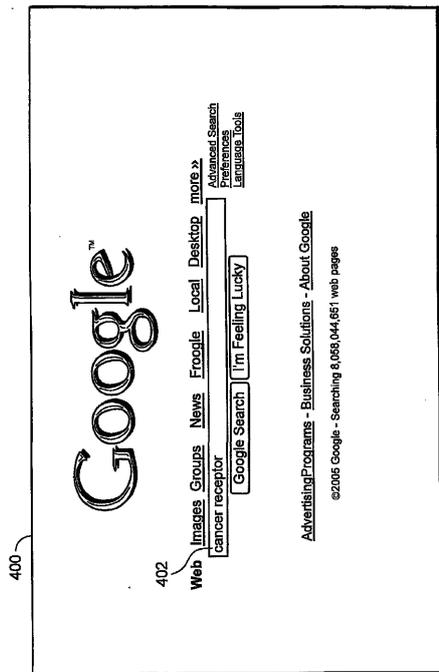


Fig. 4(a)

【 図 4 b 】

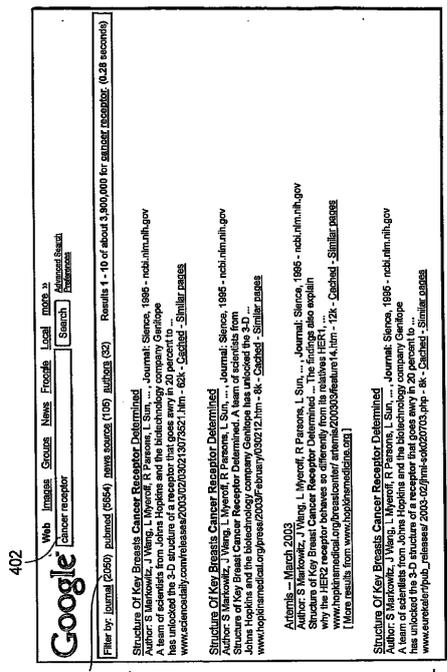


Fig. 4(b)

【 4 c 】

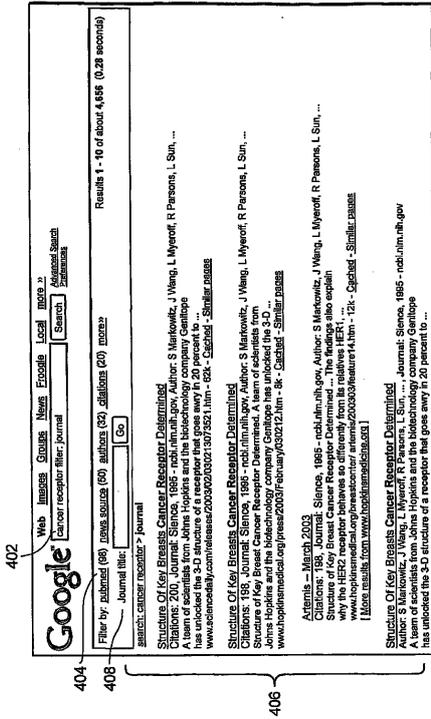


Fig. 4(c)

【 4 d 】

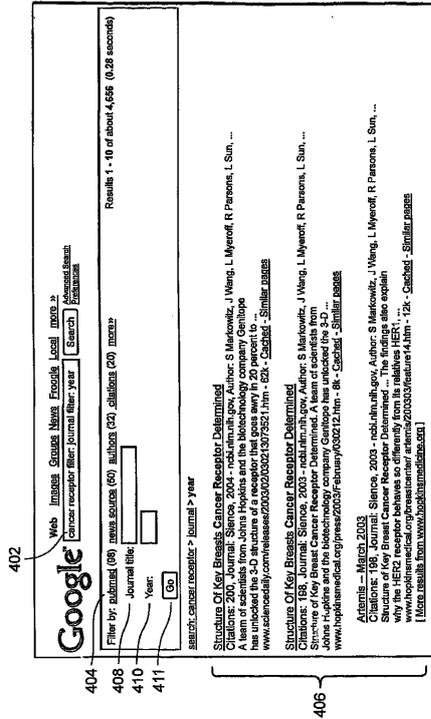


Fig. 4(d)

【 4 e 】

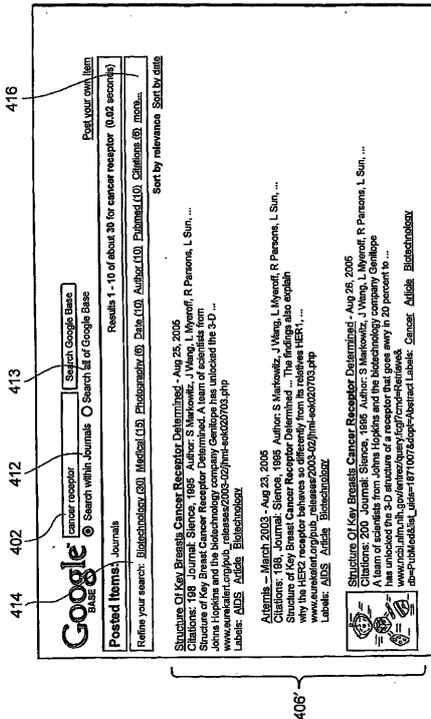


Fig. 4(e)

【 4 f 】

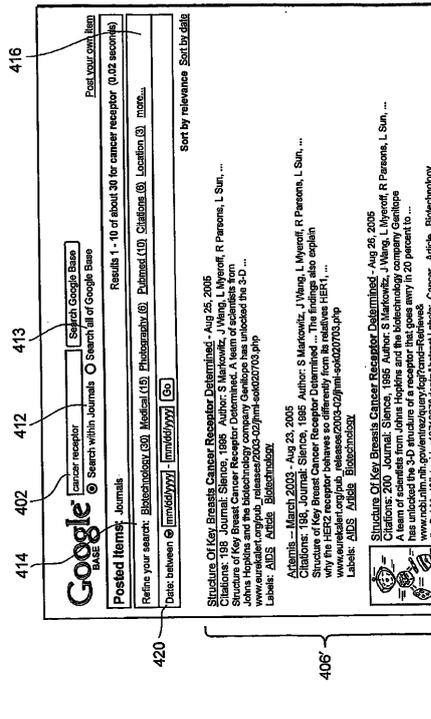


Fig. 4(f)

【 図 4 g 】

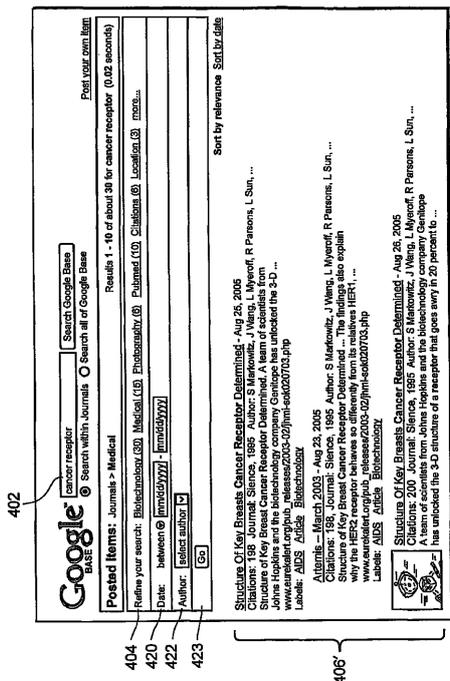


Fig. 4(g)

【 図 5 a 】

500 How we represent attributes in our data store.

- > Attributes and labels are one of the following types:
 - > - BOOLEAN
 - > - INT
 - > - FLOAT
 - > - URL
 - > - STRING
 - > - LOCATION
 - > - DATE
 - > - DATE RANGE
- > Attributes and labels are indicated by meta-tags as follows:
 - > <start name>
 - > name
 - > </end name>
 - > <start value>
 - > value
 - > </end value>

Fig. 5(a)

【 図 5 b 】

510

<start name> journalname </end name>
 <start value> Journal of Inflammation </end value>

Fig. 5(b)
属性データの例

【 図 5 c 】

520

<start name> Medical </end name> <start value>nullattribute </end value>

Fig. 5(c)
ラベルデータの例

【 図 5 d 】

530

データ情報タイプ	そのタイプのデータ項目のこのタイプに関する属性 / および属性タイプ
データ情報タイプ 1	Attr/type #1-1, Attr/type #1-2,...
...	
データ情報タイプ n	Attr/type #n-1, Attr/type #n-2,...

Fig. 5(d)
タイプと属性とのマッピング

【 図 5 e 】

540

データ情報タイプ	そのタイプのデータ項目のこのタイプに関する属性 / および属性タイプ
イベント および アクティビティ	Date/date, Holiday/holidayname
...	
製品	Numavail/integer

Fig. 5(e)
データ情報タイプ および それらの属性の例

【 6 a)

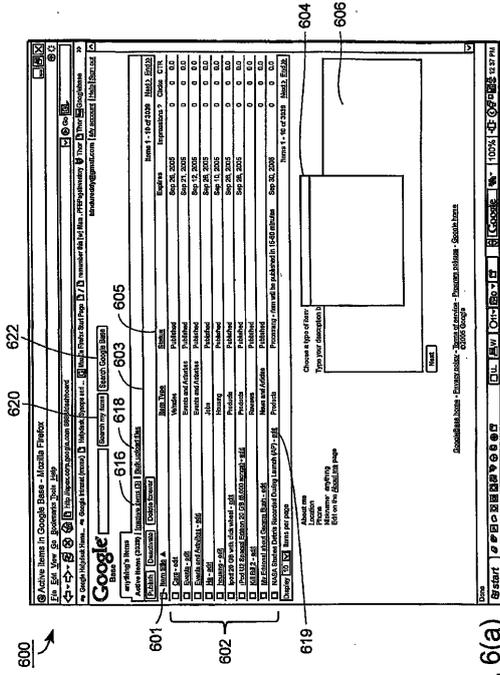


Fig. 6(a)

【 6 b)

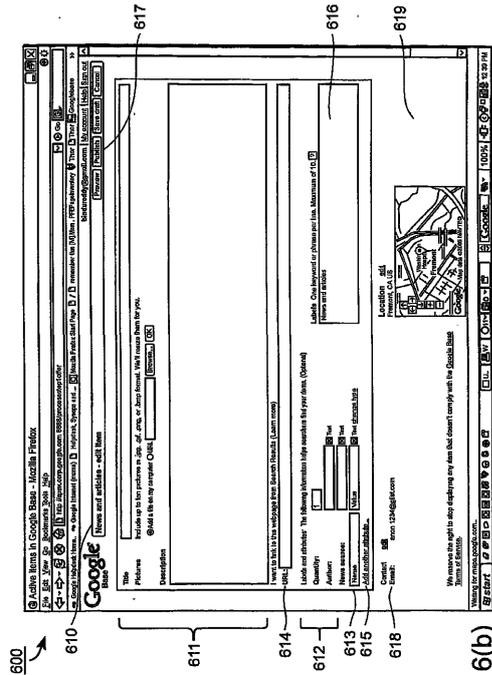


Fig. 6(b)

【 6 c)

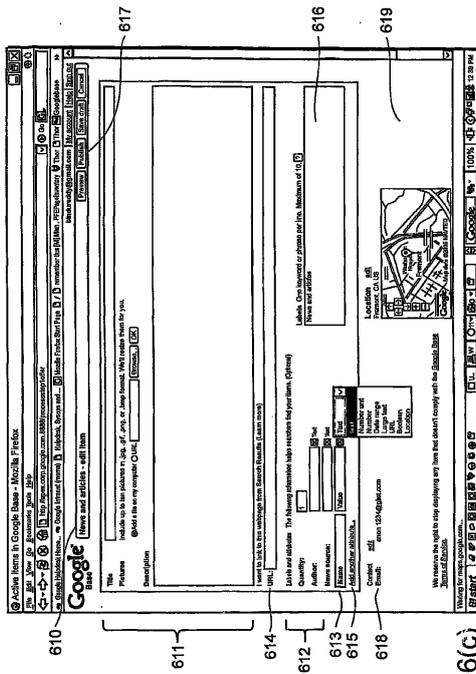


Fig. 6(c)

【 6 d)

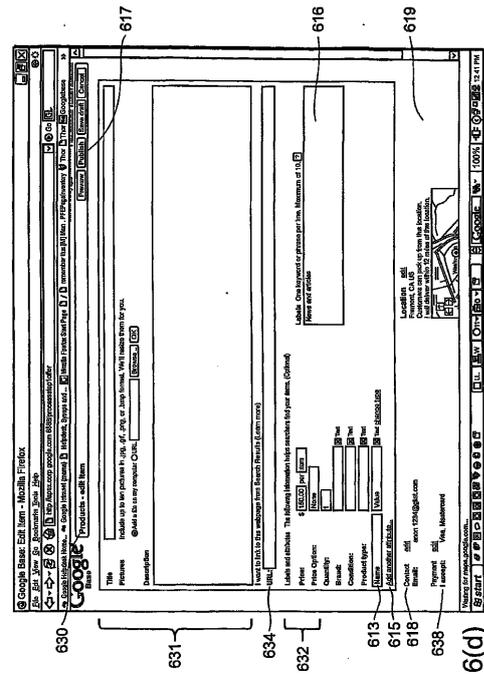


Fig. 6(d)

【 8 c 】

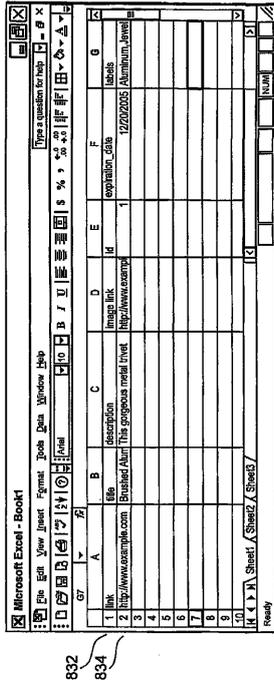


FIG. 8(c)

【 8 d 】

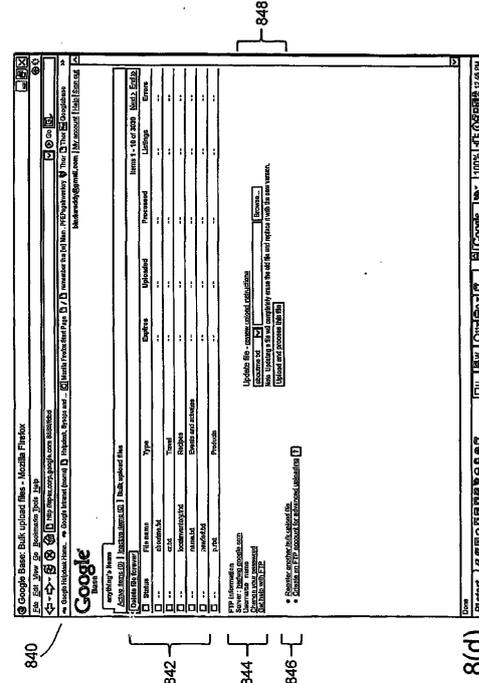


Fig. 8(d)

- 【 手続補正書 】
- 【 提出日 】平成20年12月15日 (2008.12.15)
- 【 手続補正 1 】
- 【 補正対象書類名 】特許請求の範囲
- 【 補正対象項目名 】全文
- 【 補正方法 】変更
- 【 補正の内容 】
- 【 特許請求の範囲 】
- 【 請求項 1 】

複数の被所有データ項目のインデックスを保持する段階と、

複数のプロバイダのそれぞれから、個々のプロバイダの所有するデータ項目に属性を関連付けるデータを受信する段階と、

クライアントから検索照会を受信する段階と、

前記検索照会を満たす複数の第1データ項目を選択する段階と、

2つ以上の前記第1データ項目のそれぞれへの個別の参照を含む第1検索エンジン照会結果を生成する段階と、

前記複数のプロバイダによって前記複数の第1データ項目に関連付けられている複数の第1属性のすべてよりも少ない数の第1属性を選択する段階と、

前記第1検索エンジン照会結果を含む検索エンジン結果ページを出力する段階とを有し、

各被所有データ項目は、それぞれのプロバイダを有し、

各プロバイダは、前記複数の被所有データ項目を集团的に提供する複数のプロバイダのうちの一つであり、

各被所有データ項目は、前記それぞれのプロバイダから入手可能なデータ項目であり、

前記複数の第1データ項目は、前記複数の第1属性を共通して有することを特徴とする

コンピュータで実行される方法。

【請求項 2】

各第 1 属性が、
属性名と、
前記属性名に関連する値を反映した属性値と
を有することを特徴とする請求項 1 に記載の方法。

【請求項 3】

前記属性値が、データ項目のソースを記述していることを特徴とする請求項 2 に記載の方法。

【請求項 4】

前記複数の第 1 属性のすべてよりも少ない数の第 1 属性を選択する段階の前に、前記第 1 検索エンジン照会結果が選択されることを特徴とする請求項 1 に記載の方法。

【請求項 5】

前記選択された複数の属性のすべてよりも少ない数の属性のユーザ選択を受信する段階と、
前記選択された複数の属性のすべてよりも少ない数の属性にそれぞれ関連する前記複数の第 1 データ項目のサブセットを選択する段階と、
前記複数の第 1 データ項目の 2 つ以上の前記選択されたサブセットのそれぞれへの個別の参照を含む第 2 検索エンジン照会結果を生成する段階と、
前記第 2 検索エンジン照会結果を含む第 2 検索エンジン結果ページを出力する段階と
をさらに有することを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記複数の第 1 属性のすべてよりも少ない数の第 1 属性を選択する段階が、
以前の検索実行による前記複数の第 1 属性の過去選択頻度を決定する段階と、
前記決定に基づいて、前記複数の第 1 属性のすべてよりも少ない数の第 1 属性を選択する段階と
をさらに有することを特徴とする請求項 1 に記載の方法。

【請求項 7】

前記複数の第 1 属性のそれぞれが、前記データ項目自体からは自動的に生成されないユーザ提供属性を有することを特徴とする請求項 1 に記載の方法。

【請求項 8】

前記複数の被所有データ項目のインデックスを保持する段階が、データ収集のためのウェブクローリングを実行する段階をさらに有し、
前記データ収集では、前記複数のデータ項目と、各データ項目の情報タイプを記述した前記複数の関連する属性との格納が行われることを特徴とする請求項 1 に記載の方法。

【請求項 9】

コンピュータプログラム製品でコード化され、実行されたときに、コンピュータに複数の演算を実行させるように動作する命令を有するコンピュータ可読媒体であって、
前記複数の演算は、
複数の被所有データ項目のインデックスを保持する段階と、
複数のプロバイダのそれぞれから、個々のプロバイダの所有するデータ項目に属性を関連付けるデータを受信する段階と、
クライアントから検索照会を受信する段階と、
前記検索照会を満たす複数の第 1 データ項目を選択する段階と、
2 つ以上の前記第 1 データ項目のそれぞれへの個別の参照を含む第 1 検索エンジン照会結果を生成する段階と、
前記複数のプロバイダによって前記複数の第 1 データ項目に関連付けられている複数の第 1 属性のすべてよりも少ない数の第 1 属性を選択する段階と、
前記第 1 検索エンジン照会結果を含む検索エンジン結果ページを出力する段階と
から成り、

各被所有データ項目は、それぞれのプロバイダを有し、
各プロバイダは、前記複数の被所有データ項目を集团的に提供する複数のプロバイダの
うちの1つであり、

各被所有データ項目は、前記それぞれのプロバイダから入手可能なデータ項目であり、
前記複数の第1データ項目は、前記複数の第1属性を共通して有することを特徴とする
コンピュータ可読媒体。

【請求項10】

各第1属性が、
属性名と、
前記属性名に関連する値を反映した属性値と
を有することを特徴とする請求項9に記載のコンピュータ可読媒体。

【請求項11】

前記複数の属性のすべてよりも少ない数の属性が選択される前に、前記第1検索エンジ
ン照会結果が選択されることを特徴とする請求項9に記載のコンピュータ可読媒体。

【請求項12】

前記複数の演算が、
前記選択された複数の属性のすべてよりも少ない数の属性のユーザ選択を受信する段階
と、
前記選択された複数の属性のすべてよりも少ない数の属性にそれぞれ関連する前記複数の
第1データ項目のサブセットを選択する段階と、
前記複数の第1データ項目の2つ以上の前記選択されたサブセットのそれぞれへの個別
の参照を含む第2検索エンジン照会結果を生成する段階と、
前記第2検索エンジン照会結果を含む第2検索エンジン結果ページを出力する段階と
をさらに有することを特徴とする請求項9に記載のコンピュータ可読媒体。

【請求項13】

前記複数の第1属性の段階が、
以前の検索実行による前記複数の第1属性の過去選択頻度を決定する段階と、
前記決定に基づいて、前記複数の第1属性のすべてよりも少ない数の第1属性を選択す
る段階と
をさらに有することを特徴とする請求項9に記載のコンピュータ可読媒体。

【請求項14】

1つ以上のコンピュータと、
前記1つ以上のコンピュータに接続されたコンピュータ可読媒体と
から成り、
前記コンピュータ可読媒体は、複数の命令を格納しており、
前記複数の命令は、前記1つ以上のコンピュータによって実行されたときに、前記1つ
以上のコンピュータに複数の演算を実行させ、
前記複数の演算は、
複数の被所有データ項目のインデックスを保持する段階と、
複数のプロバイダのそれぞれから、個々のプロバイダの所有するデータ項目に属性を関
連付けるデータを受信する段階と、
クライアントから検索照会を受信する段階と、
前記検索照会を満たす複数の第1データ項目を選択する段階と、
2つ以上の前記第1データ項目のそれぞれへの個別の参照を含む第1検索エンジン照会
結果を生成する段階と、
前記複数のプロバイダによって前記複数の第1データ項目に関連付けられている複数の
第1属性のすべてよりも少ない数の第1属性を選択する段階と、
前記第1検索エンジン照会結果を含む検索エンジン結果ページを出力する段階と
から成り、
各被所有データ項目は、それぞれのプロバイダを有し、

各プロバイダは、前記複数の被所有データ項目を集团的に提供する複数のプロバイダのうちの1つであり、

各被所有データ項目は、前記それぞれのプロバイダから入手可能なデータ項目であり、前記複数の第1データ項目は、前記複数の第1属性を共通して有することを特徴とするシステム。

【請求項15】

各第1属性が、属性名と、前記属性名に関連する値を反映した属性値とを有することを特徴とする請求項14に記載のシステム。

【請求項16】

前記属性値が、データ項目のソースを記述していることを特徴とする請求項15に記載のシステム。

【請求項17】

前記複数の第1属性のすべてよりも少ない数の第1属性が選択される前に、前記第1検索エンジン照会結果が選択されることを特徴とする請求項14に記載のシステム。

【請求項18】

前記複数の第1属性のすべてよりも少ない数の第1属性を選択する段階が、以前の検索実行による前記複数の第1属性の過去選択頻度を決定する段階と、前記決定に基づいて、前記複数の第1属性のすべてよりも少ない数の第1属性を選択する段階とをさらに有することを特徴とする請求項14に記載のシステム。

【請求項19】

前記属性が、前記データ項目自体からは自動的に生成されないユーザ提供属性を有することを特徴とする請求項14に記載のシステム。

【請求項20】

前記複数の被所有データ項目のインデックスを保持する段階が、データ収集のためのウェブクローリングを実行する段階をさらに有し、

前記データ収集では、前記複数のデータ項目と、各データ項目の情報タイプを記述した前記複数の関連する属性との格納が行われることを特徴とする請求項14に記載のシステム。

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US05/45447
A. CLASSIFICATION OF SUBJECT MATTER IPC: G06F 7/00(2006.01);G06F 17/30(2006.01) USPC: 707/3 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) U.S. : 707/3, 1, 5, 53 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 practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/0083039 A (Ferrari et al) 27 June 2002 (27.06.2002), paragraphs [0014], [0053], [0056], [0057], [0062], [0067]-[0070], [0078].	1-20
Y	US 2004/0093321 A (Roustant et al) 13 May 2004 (13.05.2004), paragraph [0043].	1-20
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
"A"	document defining the general state of the art which is not considered to be of particular relevance	"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
"E"	earlier application or patent published on or after the international filing date	"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
"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)	"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
"O"	document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search 07 May 2007 (07.05.2007)		Date of mailing of the international search report 21 JUL 2007
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201		Authorized officer Apu Moffiz Telephone No. 571-272-7220 DEBORAH A. THOMAS PARALEGAL SPECIALIST <i>DMT</i>

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, IS, IT, LT, LU, LV, 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, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(72) 発明者 ビンドゥ・レディ
アメリカ合衆国・カリフォルニア・94103・サン・フランシスコ・ハワード・ストリート・1097・#305

(72) 発明者 ジョナサン・ブランスマン
アメリカ合衆国・カリフォルニア・95134・サン・ホセ・デスカンソ・ドライブ・20・#1218

(72) 発明者 ニン・モスバーガー
アメリカ合衆国・カリフォルニア・94536・フレモント・ランゲル・レーン・35706

(72) 発明者 ゴーラヴ・ラヴィンドラ・バヤ
アメリカ合衆国・カリフォルニア・94086・サニーヴェイル・サウス・ウォルフ・ロード・925・#38

(72) 発明者 サラ・シラジュディン
アメリカ合衆国・カリフォルニア・94086・サニーヴェイル・サウス・ベルナルド・アヴェニュー・385

(72) 発明者 デヴィッド・ケイル
アメリカ合衆国・カリフォルニア・94306・パロ・アルト・サウス・コート・3488

(72) 発明者 ジェニファー・エル・コゼンスキ
アメリカ合衆国・カリフォルニア・94087・サニーヴェイル・スーザン・ウェイ・1202

(72) 発明者 アーヴィンド・サンダラジャン
アメリカ合衆国・カリフォルニア・94103・サン・フランシスコ・ハワード・ストリート・1097・#305

(72) 発明者 パニート・アガーワル
アメリカ合衆国・カリフォルニア・94040・マウンテン・ビュー・デイル・アヴェニュー・1200・#42

Fターム(参考) 5B075 KK02 ND03 ND34 UU40