

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5074378号
(P5074378)

(45) 発行日 平成24年11月14日(2012.11.14)

(24) 登録日 平成24年8月31日(2012.8.31)

(51) Int.Cl. F I
G 0 6 F 17/30 (2006.01)
 G 0 6 F 17/30 3 8 0 D
 G 0 6 F 17/30 1 8 0 Z

請求項の数 18 (全 22 頁)

(21) 出願番号	特願2008-504204 (P2008-504204)	(73) 特許権者	502208397
(86) (22) 出願日	平成18年3月24日 (2006. 3. 24)		グーグル インコーポレイテッド
(65) 公表番号	特表2008-535095 (P2008-535095A)		アメリカ合衆国 カリフォルニア州 9 4
(43) 公表日	平成20年8月28日 (2008. 8. 28)		0 4 3 マウンテン ビュー アンフィシ
(86) 国際出願番号	PCT/US2006/010965		アター パークウェイ 1 6 0 0
(87) 国際公開番号	W02006/104951	(74) 代理人	100079108
(87) 国際公開日	平成18年10月5日 (2006. 10. 5)		弁理士 稲葉 良幸
審査請求日	平成21年3月13日 (2009. 3. 13)	(74) 代理人	100093861
(31) 優先権主張番号	11/097, 689		弁理士 大賀 眞司
(32) 優先日	平成17年3月31日 (2005. 3. 31)	(74) 代理人	100109346
(33) 優先権主張国	米国 (US)		弁理士 大貫 敏史
		(72) 発明者	ホーグ, アンドリュウ, ウィリアム
			アメリカ合衆国, ニュージャージー州 O
			7 4 4 2, ポンプトン レイクス, サミッ
			ト リッジ 1 3 2

最終頁に続く

(54) 【発明の名称】 クエリ用語および返答用語を含む情報ソースからのスニペットを有するファクトクエリエンジンのユーザインタフェース

(57) 【特許請求の範囲】

【請求項 1】

1 つ以上のプロセッサと、1 つ以上のプログラムを記憶するメモリとを有するサーバシステムにおいて、前記 1 つ以上のプロセッサによる前記 1 つ以上のプログラムの実行によって実行される、ファクトのソースを表示する方法であって、

1 つ以上のクエリ用語を含む、ユーザにより作成されたファクトクエリを受け取ること

、
 ファクトレポジトリから、前記ファクトクエリへの返答であって、1 つ以上のファクト返答用語と、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むソースドキュメントを特定する情報と、を含む、前記ファクトクエリへの返答を特定すること、

前記返答を特定した後、前記ファクトレポジトリとは異なるドキュメントデータベース内の、前記ファクトレポジトリから特定された前記ソースドキュメントの少なくとも 1 つにアクセスすること、

前記ソースドキュメントの少なくとも 1 つに対して、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むスニペットを生成すること、および

前記スニペットを含む応答を生成すること、

を含み、

前記スニペット自体が、前記ファクトクエリに答えるファクト情報を提供する、方法。

10

20

【請求項 2】

前記応答を生成することは、前記ファクトクエリの前記 1 つ以上の用語および前記返答の前記 1 つ以上のファクト返答用語を前記スニペット内で強調表示することをさらに含む、請求項 1 に記載の方法。

【請求項 3】

前記ファクトクエリへの前記返答の表示された表現内のリンクのユーザ選択に応答することをさらに含む、請求項 1 に記載の方法。

【請求項 4】

前記応答は、前記ソースドキュメントのリストを要求するためのユーザ選択可能リンクを含み、

10

前記方法は、前記ソースドキュメントのリストを送ることにより前記リンクのユーザ選択に応答することをさらに含む、請求項 1 に記載の方法。

【請求項 5】

前記応答を生成することは、前記ソースドキュメントの少なくとも 1 つにおける前記 1 つ以上のクエリ用語および前記 1 つ以上のファクト返答用語の近似を判断することを含む、請求項 1 に記載の方法。

【請求項 6】

ファクトのソースを表示するシステムであって、

1 つ以上のプロセッサと、

メモリと、

20

前記メモリ内に記憶された 1 つ以上のプログラムと、

を備え、

前記 1 つ以上のプログラムは、

1 つ以上のクエリ用語を含む、ユーザにより作成されたファクトクエリを受け取るクエリレシーバと、

ファクトレポジトリから、前記ファクトクエリへの返答であって、1 つ以上のファクト返答用語と、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むソースドキュメントを特定する情報と、を含む、前記ファクトクエリへの返答を特定する返答特定器と、

前記返答特定器が前記返答を特定した後の実行のために、前記ファクトレポジトリとは異なるドキュメントデータベース内の、前記ファクトレポジトリから特定された前記ソースドキュメントの少なくとも 1 つにアクセスするソースドキュメント特定器と、

30

前記ソースドキュメントの少なくとも 1 つに対して、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むスニペットを生成するスニペットジェネレータと、

前記スニペットを含む応答を生成する応答ジェネレータと、

を実行させる命令を含み、

前記スニペット自体が、前記ファクトクエリに答えるファクト情報を提供する、システム。

40

【請求項 7】

前記応答ジェネレータは、前記ファクトクエリの前記 1 つ以上の用語および前記返答の前記 1 つ以上のファクト返答用語を前記生成されたスニペット内で強調表示する命令をさらに含む、請求項 6 に記載のシステム。

【請求項 8】

前記応答は、前記ソースドキュメントのリストを要求するためのユーザ選択可能リンクを含み、

前記クエリレシーバは、前記ソースドキュメントのリストを送ることにより前記リンクのユーザ選択に応答する命令をさらに含む、請求項 6 に記載のシステム。

【請求項 9】

前記スニペットジェネレータは、前記ソースドキュメントの少なくとも 1 つにおける前

50

記 1 つ以上のクエリ用語および前記 1 つ以上のファクト返答用語の近似を検出する近似検出器をさらに含む、請求項 6 に記載のシステム。

【請求項 10】

ファクトのソースを表示するシステムであって、

1 つ以上のプロセッサと、

メモリと、

前記メモリ内に記憶された 1 つ以上のプログラムと、

を備え、

前記 1 つ以上のプログラムは、

1 つ以上のクエリ用語を含む、ユーザにより作成されたファクトクエリを受け取る手段と、

ファクトレポジトリから、前記ファクトクエリへの返答であって、1 つ以上のファクト返答用語と、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むソースドキュメントを特定する情報と、を含む、前記ファクトクエリへの返答を特定する手段と、

前記返答を特定した後、前記ファクトレポジトリとは異なるドキュメントデータベース内の、前記ファクトレポジトリから特定された前記ソースドキュメントの少なくとも 1 つにアクセスする手段と、

前記ソースドキュメントの少なくとも 1 つに対して、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むスニペットを生成する手段と、

前記スニペットを含む応答を生成する手段と、

を実行させる命令を含み、

前記スニペット自体が、前記ファクトクエリに答えるファクト情報を提供する、システム。

【請求項 11】

コンピュータによる実行のために構成された 1 つ以上のプログラムを記憶するコンピュータ可読記憶媒体であって、

前記 1 つ以上のプログラムは、

1 つ以上のクエリ用語を含む、ユーザにより作成されたファクトクエリを受け取り、

ファクトレポジトリから、前記ファクトクエリへの返答であって、1 つ以上のファクト返答用語と、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むソースドキュメントを特定する情報と、を含む、前記ファクトクエリへの返答を特定し、

前記返答を特定した後、前記ファクトレポジトリとは異なるドキュメントデータベース内の、前記ファクトレポジトリから特定された前記ソースドキュメントの少なくとも 1 つにアクセスし、

前記ソースドキュメントの少なくとも 1 つに対して、前記ファクトクエリの 1 つ以上のクエリ用語および前記返答の 1 つ以上のファクト返答用語を含むスニペットを生成し、および

前記スニペットを含む応答を生成するための命令を含み、

前記スニペット自体が、前記ファクトクエリに答えるファクト情報を提供する、コンピュータプログラム製品。

【請求項 12】

前記 1 つ以上のプログラムは、前記ファクトクエリの前記 1 つ以上のクエリ用語および前記返答の前記 1 つ以上のファクト返答用語を前記スニペット内で強調表示するための命令をさらに含む、請求項 11 に記載のコンピュータ可読記憶媒体。

【請求項 13】

前記応答は、前記ソースドキュメントのリストを要求するためのユーザ選択可能リンクを含み、

10

20

30

40

50

前記 1 つ以上のプログラムは、前記ソースドキュメントのリストを送ることにより前記リンクのユーザ選択に応答するための命令をさらに含む、請求項 11 に記載のコンピュータ可読記憶媒体。

【請求項 14】

前記スニペットを生成するための前記命令は、前記ソースドキュメントの少なくともにおける前記 1 つ以上のクエリ用語および前記 1 つ以上のファクト返答用語の近似を判断するための命令を含む、請求項 11 に記載のコンピュータ可読記憶媒体。

【請求項 15】

前記返答を特定することは、1 つ以上の言葉または数を含む各ファクト返答用語を特定することを含み、

10

前記スニペットを生成することは、前記特定されたファクト返答用語を含む各ソースドキュメントのテキスト部分を特定することと、前記特定された各ファクト返答用語を含む前記特定されたテキスト部分を含めるよう前記スニペットを生成することとを含む、請求項 1 に記載の方法。

【請求項 16】

前記返答特定器は、1 つ以上の言葉または数を含む各ファクト返答用語を特定するよう構成され、

前記スニペットジェネレータは、前記特定されたファクト返答用語を含む各ソースドキュメントのテキスト部分を特定し、前記特定された各ファクト返答用語を含む前記特定されたテキスト部分を含めるよう前記スニペットを生成するよう構成される、請求項 6 に記載のシステム。

20

【請求項 17】

前記スニペットを生成するための前記命令は、前記ファクトクエリの少なくとも 1 つのテキスト用語と前記返答の少なくとも 1 つのテキスト用語とを前記スニペットに含めるよう各ソースドキュメントからテキストを選択するための命令を含む、請求項 11 に記載のコンピュータ可読記憶媒体。

【請求項 18】

前記返答を特定するための前記命令は、1 つ以上の言葉または数を含む各ファクト返答用語を特定するための命令を含み、

前記スニペットを生成するための前記命令は、前記特定されたファクト返答用語を含む各ソースドキュメントのテキスト部分を特定し、前記特定された各ファクト返答用語を含む前記特定されたテキスト部分を含めるよう前記スニペットを生成するための命令を含む、請求項 11 に記載のコンピュータ可読記憶媒体。

30

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本出願は、参照により本明細書にそれぞれ援用される以下の出願に関連する。

米国特許出願第 11 / 097, 688 号明細書、名称「Corroborating Facts Extracted from Multiple Sources」、2005 年 3 月 31 日出願、

40

米国特許出願第 11 / 097, 676 号明細書、名称「Bloom Filters for Query Simulation」、2005 年 3 月 31 日出願、

米国特許出願第 11 / 097, 690 号明細書、名称「Selecting the Best Answer to a Fact Query from Among a Set of Potential Answers」、2005 年 3 月 31 日出願、および

米国特許出願第 11 / 024, 784 号明細書、名称「Supplementing Search Results with Information of Interest」、2004 年 12 月 30 日出願。

50

【 0 0 0 2 】

技術分野

開示する実施形態は、包括的にはファクトクエリに関し、詳細には、ファクトクエリエンジン並びにクエリ用語及び返答用語を有するソースのスニペットのユーザインタフェースに関する。

【 背景技術 】

【 0 0 0 3 】

ワールドワイドウェブ（「ウェブ」としても知られている）およびウェブ内のウェブページは、ファクト情報の膨大なソースである。ユーザはウェブページを見て、「ポーランドの首都はどこか」または「ジョージワシントンの生年月日はいつか」等のファクトクエリに対する答えを得ることができる。しかし、ウェブサーチエンジンは、一般に、上述したようなファクトクエリに対して単純で簡明な返答を提供しないため、これに関連してはユーザの助けにならないことがある。ウェブサーチエンジンは、単純で簡明な答えを提供せずに、クエリにマッチしたと判断されたウェブページリストをユーザに提供し、ユーザがマッチしたウェブページを調べて答えを見つける必要がある。

10

【 発明の開示 】

【 発明が解決しようとする課題 】

【 0 0 0 4 】

ファクトクエリに対して素早い返答を提供することができるサーチエンジンを構築しようというこれまでの試みには、それぞれ固有の欠点がある。例えば、サーチエンジンによっては、特定の百科事典等の単一ソースからファクトを引き出すものがある。これは、こういったエンジンが返答できるクエリの種類を制限する。例えば、百科事典をベースとしたサーチエンジンは、映画、歌等についてのクエリ等の人気のあるカルチャーに関する多くのクエリに答えられる可能性が低く、製品、サービス、小売業、および卸売業等についての多くのクエリに対して答えられる可能性も低い。仮に、このようなサーチエンジンが使用するソースセットが拡張されたとしたならば、このような拡張により、ファクトクエリに対して、いくつかが矛盾または曖昧であり得る複数の返答候補がもたらされる可能性がある。さらに、ソースの母集団が拡張するにつれて、信頼性の低いソースまたは信頼性が未知のソースから情報が引き出される可能性がある。

20

【 課題を解決するための手段 】

30

【 0 0 0 5 】

本発明の態様によれば、ファクトのソースを表示する方法は、1つ以上の用語を含むファクトクエリを受け取ること、ファクトクエリに対する、1つ以上の用語を含む返答を特定すること、クエリの1つ以上の用語および返答の1つ以上の用語を含む1つ以上のソースドキュメントを特定すること、少なくとも1つの情報ソースのスニペットを生成することであって、スニペットは、クエリの1つ以上の用語および返答の1つ以上の用語を含むこと、およびスニペットを含む応答を生成することを含む。

【 0 0 0 6 】

同様の参照番号は、全図面を通して対応する部分を指す。

【 発明を実施するための最良の形態 】

40

【 0 0 0 7 】

クエリエンジンは、多くの別々のソースから集められたファクト情報を記憶し、ユーザからのファクト情報（または「ファクトクエリ」）に対するクエリに回答して返答を返すことができる。多くのソースから情報を集めることにより、クエリエンジンで利用できるファクト情報の範囲が拡張するが、複数の返答候補が生じる可能性ももたらされる。クエリエンジンは、返答候補を特定し、返答候補から最良の返答を選択してユーザに提示するか、または返答候補のいずれもユーザに提示しないと判断することができる。クエリエンジンは、各ソースからのテキストの部分を含む返答のソースリストを提供することもできる。テキストの1つ以上の部分はスニペットと呼ばれ、ファクトクエリ用語および返答用語を含むことができる。スニペットは、サーチエンジンにより特定または選択された返答

50

を示すが、ソースリストは、返答のための基礎をユーザに提供し、ユーザが返答の正確度を評価するのを助けることができる。

【0008】

図1は、本発明のいくつかの実施形態によるネットワーク100を示す。ネットワーク100は、1つ以上のクライアント102およびクエリエンジン106を含む。クライアント102はクライアントアプリケーション（図示せず）を含むことができる。ネットワーク100は、これら構成要素を結合する1つ以上の通信ネットワーク104も含む。

【0009】

クライアントアプリケーションは、クエリエンジン106へのインタフェースをクライアント102のユーザ（図示せず）に提供する。ユーザは、クライアント102で実行されるクライアントアプリケーションを使用して、ドキュメントサーチ（たとえば、ウェブサーチ）およびファクトクエリをクエリエンジン106に提出し、クエリエンジン106からの応答を見ることができる。クライアントアプリケーションはウェブブラウザを含むことができる。ウェブブラウザの例として、ファイアフォックス（FIREFOX（商標））、インターネットエクスプローラ（INTERNET EXPLORER（商標））、およびオペラ（OPERA（商標））が挙げられる。

【0010】

クエリエンジン106は、ファクト情報を記憶し、ファクトクエリに応答するとともに、他の種類のサーチを処理するプラットフォームを提供する。クエリエンジン106は、ウェブサーチ等のドキュメントサーチおよびファクト情報のクエリを処理することができる。クエリエンジン106はクエリサーバ108を含む。クエリサーバ108は、フロントエンドをクエリエンジン106に提供する。クエリサーバ108は、クライアント102からクエリを受け取り、ファクトクエリ及び他のサーチを処理するクエリエンジン106の構成要素にクエリを送り、応答を生成し、応答をクライアント102に送る。クエリサーバ108は複数のコンピュータに分散してよい。他の実施形態では、クエリエンジンはより多数または少数の機能を処理してよい。たとえば、他の実施形態では、応答生成をクエリエンジン106内のどの場所で処理してもよい。

【0011】

クエリエンジン106は、ドキュメントサーチを処理するための第1のサーチコントローラ110、第1のキャッシュ112、ドキュメントインデックス114、およびドキュメントデータベース116を含む。いくつかの実施形態では、これら構成要素を複数のコンピュータに配備させて、大量のドキュメントの高速アクセスを提供することができる。たとえば、ドキュメントデータベース116をN台のサーバに配備させることができ、「モジュロN」関数等のマッピング関数を使用して、N台の各サーバにどのドキュメントが記憶されているかを判断する。Nは2以上の整数であってよく、たとえば、2～8196の整数である。同様に、ドキュメントインデックス114も複数のサーバに分散させることができ、第1のキャッシュ112も複数のサーバに分散させることができる。さらに、第1のサーチコントローラ110も複数のコンピュータに分散させることができる。

【0012】

第1のサーチコントローラ110はクエリサーバ108に結合される。第1のサーチコントローラ110は、第1のキャッシュ112、ドキュメントインデックス114、およびドキュメントデータベース116にも結合される。第1のサーチコントローラ110は、ドキュメントサーチクエリをクエリサーバ108から受け取り、クエリを第1のキャッシュ112、ドキュメントインデックス114、およびドキュメントデータベース116に送るように構成することができる。第1のキャッシュ112を使用して、前に見つけられたサーチ結果を一時的に記憶することによってサーチ効率を増大させることができる。

【0013】

第1のサーチコントローラ110は、ドキュメントサーチ結果を第1のキャッシュ112および/またはドキュメントインデックス114から受け取り、順序付きサーチ結果リストを構築する。次に、第1のサーチコントローラ110は、見つけられたドキュメント

10

20

30

40

50

のリストをクエリサーバ108に返し、クライアント102に転送する。第1のサーチコントローラ110が第1のキャッシュ112および/またはドキュメントインデックス114から受け取るドキュメントサーチ結果は、サーチ結果内の見つけられたドキュメントのスニペットを伴うことができる。

【0014】

クエリエンジン106は、第2のサーチコントローラ118、第2のキャッシュ120、ファクトインデックス122、およびファクトリポジトリ124も含む。いくつかの実施形態では、これら構成要素を複数のコンピュータに配備させて、大量のファクトへの高速アクセスを提供することができる。例えば、ファクトリポジトリ124をN台のサーバに配備させることができ、「モジュロN」関数等のマッピング関数を使用して、N台の各サーバにどのファクトが記憶されているかを判断する。Nは2以上の整数であってよく、例えば、2~8196の整数である。同様に、ファクトインデックス122も複数のサーバに分散させることができ、第2のキャッシュ120も複数のサーバに分散させることができる。さらに、第2のサーチコントローラ118も複数のコンピュータに分散させることができる。

【0015】

第2のサーチコントローラ118はクエリサーバ108に結合される。第2のサーチコントローラ118は、第2のキャッシュ20、ファクトインデックス122、およびファクトリポジトリ124にも結合される。第2のサーチコントローラ118は、ファクトクエリに対する返答のクエリを受け取り、クエリを第2のキャッシュ120およびファクトリポジトリ124に(ファクトインデックス122を介して)送るように構成することができる。第2のキャッシュ120を使用して、前に見つけられたサーチ結果を一時的に記憶することによってファクト検索効率を増大させることができる。

【0016】

第2のサーチコントローラ118は、第2のキャッシュ120および/またはファクトリポジトリ124からファクトクエリに対する返答候補であるファクトを受け取る。第2のサーチコントローラ118は、返答候補からユーザに提示する最良の返答として返答を選択する。この返答はクエリサーバ108に送られ、クエリサーバ108において、その返答を含む応答が生成されてクライアント102に送られ、ユーザに提示される。ファクトクエリに対する返答に、またはその横に表示されるアイコンをユーザが選択することに応答して、クエリサーバ108は、返答に関連するソースリストを特定し、そのソースリストを第1のサーチコントローラ110に送ることができる。第1のサーチコントローラ110は、ソースドキュメントの少なくともサブセットのソースおよびスニペットに対応するドキュメントにアクセスする。いくつかの実施形態では、スニペットは、クエリからの用語および返答からの用語を含む。

【0017】

ファクトリポジトリ124は、複数のドキュメントから抽出されたファクト情報を記憶する。特定のファクトを抽出できる元の文書が、そのファクトのソースドキュメント(または「ソース」)である。換言すれば、ファクトのソースは、その内容内にそのファクトを含む。ソースドキュメントは、制限ではなしにウェブページを含むことができる。そのファクトリポジトリ124内に、ファクトリポジトリ124がファクト情報を記憶できるエンティティ、概念等がオブジェクトにより表される。オブジェクトは、関連する1つ以上のファクトを有し得る。各オブジェクトはファクトの集まりであり、関連するファクトのないオブジェクト(空オブジェクト)は、ファクトリポジトリ124内に存在しないオブジェクトとして見る事ができる。各オブジェクト内で、そのオブジェクトに関連する各ファクトは、属性-値の対として記憶される。各ファクトは、それぞれの内容内のファクトおよびそこから抽出されたファクトも含む。ファクトリポジトリ内のオブジェクトおよびファクトについての詳細について、図2に関連してさらに詳細に後述する。

【0018】

ファクトリポジトリ124内の情報を調べるために、第2のサーチコントローラ118

10

20

30

40

50

は、サーチクエリ内の用語についてファクトインデックス 1 2 2 をサーチする。この結果として、サーチクエリの各種用語にマッチする（すなわち、ファクトまたはオブジェクトに対応する）ファクトリポジトリロケーションのリストが得られる。サーチクエリの論理構造（ブール表現またはツリーとして考えることができる）を利用して、次に、第 2 のサーチコントローラ 1 1 8 は、これらロケーションリストの論理的な組み合わせを形成して、サーチクエリにマッチするファクト候補（もしあれば）を特定する。

【 0 0 1 9 】

ファクトインデックス 1 2 2 は、インデックスをファクトリポジトリ 1 2 4 に提供し、ファクトリポジトリ 1 2 4 内の情報の効率的な索引に役立つ。ファクトインデックス 1 2 2 は、1 つ以上のパラメータに基づいてファクトリポジトリ 1 2 4 に索引付けることができる。例えば、ファクトインデックス 1 2 2 は、一意の用語をファクトリポジトリ内のロケーションに索引付けるインデックス（メインインデックスまたは用語インデックスと呼ぶことができる）を有することができる。ファクトインデックス 1 2 2 についてのさらなる詳細を、図 3 に関連して後述する。

【 0 0 2 0 】

クエリエンジン 1 0 6 の任意の構成要素は複数のコンピュータに分散してよいが、説明の便宜のために、クエリエンジン 1 0 6 の構成要素を単一のコンピュータで実施されるものとして考察することを理解されたい。

【 0 0 2 1 】

図 2 は、本発明のいくつかの実施形態によるファクトリポジトリ 1 2 4 内のオブジェクトの例示的なデータ構造を示す。上述したように、ファクトリポジトリはオブジェクトを含み、各オブジェクトは 1 つ以上のファクトを含むことができる。各オブジェクト 2 0 0 は、オブジェクト ID 2 0 2 等の一意の識別子を含む。オブジェクト 2 0 0 は 1 つ以上のファクト 2 0 4 を含む。各ファクト 2 0 4 は、ファクト ID 2 1 0 等のそのファクトに一意の識別子を含む。各ファクト 2 0 4 は、属性 2 1 2 および値 2 1 4 を含む。例えば、ジョージワシントンを表すオブジェクトに含まれるファクトは、「生年月日」および「没年月日」の属性を有するファクトを含むことができ、これらファクトの値はそれぞれ実際の生年月日および没年月日である。ファクト 2 0 4 は、ファクトリポジトリ 1 2 4 内の別のオブジェクトのオブジェクト ID 2 0 2 等のオブジェクト識別子である別のオブジェクトへのリンク 2 1 6 を含むことができる。リンク 2 1 6 により、オブジェクトはその値が他のオブジェクトであるファクトを有することができる。例えば、「米国」というオブジェクトについて、その値が、ファクトリポジトリ 1 2 4 内の別のオブジェクトである「ジョージ W ブッシュ」である「大統領」属性を有するファクトが存在できる。いくつかの実施形態では、値フィールド 2 1 4 はリンクされたオブジェクトの名称を記憶し、リンク 2 1 6 はリンクされたオブジェクトのオブジェクト識別子を記憶する。いくつかの他の実施形態では、ファクト 2 0 4 は、ファクト 2 0 4 の値 2 1 4 が別のオブジェクトへのリンクを記憶し得るため、リンクフィールド 2 1 6 を含まない。

【 0 0 2 2 】

各ファクト 2 0 4 は、1 つ以上のメトリック 2 1 8 を含むこともできる。メトリックは、ファクトの数量表示を提供することができる。いくつかの実施形態では、メトリックは、信頼レベルおよび重要レベルを含む。信頼レベルは、そのファクトが正確である可能性を示す。重要レベルは、同じオブジェクトの他のファクトと比較したそのオブジェクトに対するファクトの関連性を示す。換言すれば、重要レベルは、そのファクトが、オブジェクトにより表されるエンティティまたは概念の理解に対してどの程度重要であるかを測る。

【 0 0 2 3 】

各ファクト 2 0 4 は、ファクトを含み、ファクトが抽出されたソース 2 2 0 のリストを含む。各ソースは、ユニフォームリソースロケータ（URL）またはウェブアドレスにより特定することができる。

【 0 0 2 4 】

いくつかの実施形態では、いくつかのファクトは、ファクトを抽出したモジュールを特定するエージェントフィールド 222 を含むことができる。例えば、エージェントは、ファクトを特定のソースから抽出する特殊モジュールまたはウェブ全体を通してドキュメント内のフリーテキストからファクトを抽出するモジュール等であってよい。

【0025】

いくつかの実施形態では、オブジェクト 200 は、名称ファクト 206 および特性ファクト 208 等の 1 つ以上の特殊ファクトを有することができる。名称ファクト 206 は、オブジェクト 200 で表されるエンティティまたは概念の名称を伝えるファクトである。例えば、スペインという国を表すオブジェクトには、「スペイン」としてそのオブジェクトの名称を伝えるファクトがある。名称ファクト 206 は、一般ファクト 204 の特殊例であり、他の任意のファクト 204 と同じパラメータを含み、属性、値、ファクト ID、メトリック、ソース等を有する。名称ファクト 206 の属性 224 は、そのファクトが名称ファクトであることを示し、その値は実際の名称である。ネームは文字列であることができる。オブジェクト 200 は、多くのエンティティまたは概念が 2 つ以上の名称を有し得るため、1 つ以上の名称ファクトを有することができる。例えば、スペインを表すオブジェクトは、その国の一般的な名称「スペイン」および正式な名称「スペイン王国」を伝える名称ファクトを有することができる。別の例として、米国特許商標庁を表すオブジェクトは、エージェントの頭字語である「PTO」および「USPTO」ならびに正式名称「米国特許商標庁」を伝える名称ファクトを有することができる。

【0026】

特性ファクト 208 は、オブジェクト 200 により表される、関心のあり得るエンティティまたは概念についてのステートメントを伝えるファクトである。例えば、スペインを表すオブジェクトでは、特性ファクトは、スペインがヨーロッパの国であることを伝えることができる。特性ファクト 208 は、一般ファクト 204 の特殊例であり、他のファクト 204 と同じパラメータ（属性、値、ファクト ID、メトリック、ソース等）も含む。特性ファクト 208 の属性フィールド 226 は、そのファクトが特性ファクトであることを示し、値フィールドは、関心のあるステートメントを伝える文字列である。例えば、スペインを表すオブジェクトでは、特性ファクトの値は、「ヨーロッパの国」という文字列であり得る。オブジェクト 200 は、ゼロ以上の特性ファクトを有することができる。

【0027】

図 2 に示し、上述したデータ構造が単なる例にすぎないことを理解されたい。ファクトトリポジトリ 124 のデータ構造は、他の形をとってもよい。他のフィールドをファクトに含めてよく、上述したフィールドのいくつかを省いてもよい。さらに、各オブジェクトは、名称ファクトおよび特性ファクトの他に、オブジェクトで表されるエンティティまたは概念を分類するタイプまたはカテゴリを伝えるファクト（例えば、人物、場所、映画、俳優等）等の特殊なファクトをさらに有してもよい。いくつかの実施形態では、オブジェクトの名称（複数可）および/または特性は、オブジェクトの属性 - 値の対に関連するファクトレコード 204 とは別のフォーマットを有する特別レコードで表される。

【0028】

図 3 は、本発明のいくつかの実施形態による例示的なファクトインデックスを示す。上述したように、ファクトインデックス 122 は、1 つ以上のパラメータに基づいてファクトトリポジトリに索引を付ける。いくつかの実施形態では、ファクトインデックス 300 はそのインデックスである。ファクトインデックス 300 は、ファクトまたはファクトトリポジトリ 124 内の情報ロケーションに一意の用語をマッピングする。本明細書において使用する用語とは、言葉（「スペイン」または「ジョージ」等）または数（「123」または「-9」等）である。いくつかの実施形態では、用語は、米国（United States）または生年月日（birth date）等の 2 つ以上の言葉を含む用語も含むことができる。ファクトインデックス 300 は、複数の用語セット 303 および関連する用語ロケーションレコードを含み、任意に、インデックス 300 についての情報（たとえば、インデックスサイズについての情報、セットを見つけるために使用されるマッピング

関数についての情報等)を有するインデックスヘッダ302を含むことができる。各セット303内には、用語304およびファクトリポジトリ124内の用語が出現する各ロケーションを特定する1つ以上のロケーションレコード306がある。各用語ロケーションレコードは、オブジェクト識別子308(用語が出現するオブジェクトを特定する)、ファクト識別子310(オブジェクト内のファクトを特定する)、ファクトフィールド識別子312(ファクト内のフィールドを識別する)、およびトークン識別子(フィールド内のトークンを特定する)を有する。これら4つのフィールドは、用語をファクトリポジトリ124内のロケーションにマッピングする。しかし、ファクトインデックス300は単なる例にすぎず、他の形のファクトインデックス300および他のファクトインデックスが可能なことを理解されたい。いくつかの実施形態では、用語ロケーションレコード306がオブジェクト全体を指す(例えば、用語がオブジェクトの名称である)場合、ファクト識別子310、フィールド識別子312、およびトークン識別子314は、予め定められた値またはヌル値を有し得る。

【0029】

図4は、本発明のいくつかの実施形態によるファクトクエリに対する返答候補の例示的なリストを示す。第2のサーチコントローラ118は、第2のキャッシュ120またはファクトリポジトリ124のサーチからファクトクエリに対する1つ以上の返答候補のリストを受け取り、返答候補リストから最良の返答を選択し、さらに処理するために最良の返答をクエリサーバ108に送る。これについての詳細を図5A~図5Cに関連してさらに後述する。図4は、返答候補400の例示的なリストを示す。リスト400は1つ以上の返答候補403を含む。各返答候補403は1つ以上のフィールドを有する。オブジェクトID404は、返答候補であるファクトを含むオブジェクトを特定する。オブジェクト名406は、オブジェクトID404により特定されるオブジェクトにより表されるエンティティまたは概念の名称を特定する。オブジェクト名406は、オブジェクトに含まれる名称ファクトの値であってよい(オブジェクトデータ構造の上記考察を参照)。ファクト属性408は、返答候補であるファクトの属性を特定する。ファクト値410は、返答候補であるファクトの値を特定する。返答フィールド412は、3つのフィールド、すなわちオブジェクト名406、ファクト属性408、またはファクト値410のうちのどれが、ファクトクエリに応答する実際の返答(すなわち、ユーザが探している返答の種類)を有するかを特定する。QAタイプ414は、ファクトクエリにより課された質問のタイプ(すなわち、ユーザが尋ねている質問の種類、問われている質問の種類に回答する返答)を特定する。スコア416は返答候補のスコアを示す。スコアは、正確で応答性のある返答としての返答候補の質を測ろうとするメトリックである。ファクトクエリ418は、そのような返答候補の特定に繋がる、第2のサーチコントローラ418により生成される内部クエリである。ファクトクエリ418は、ユーザクエリ(すなわち、クライアント102においてユーザにより入力されるクエリ)に基づいて生成される。QAタイプ414、スコア416、およびファクトクエリ418の詳細については、図5A~図5Cに関連してさらに後述する。いくつかの実施形態では、返答候補403はより多くまたはより少数の情報フィールドで表すことができる。いくつかの実施形態では、リスト400は、リスト400全体に当てはまる情報を含むリストヘッダ402を含む。例えば、ヘッダ402は、ユーザクエリのコピー、リスト400のトップエンティティへのポインタ、またはリスト400内のアイテムまたはレコードへのアクセスに役立つ他のデータ構造を含むことができる。

【0030】

図5A~図5Cは、ファクトクエリに対する最良の返答を選択し、その返答を提示する本発明のいくつかの実施形態による例示的なプロセスを示す。ファクトクエリへの返答は、ファクトクエリへの最良の応答として特定されるファクトリポジトリ124内のファクトである。クエリエンジン106は、ファクトクエリを受け取ると、クエリを処理し、返答候補を特定し、最良の返答を選択し、返答を含む応答を生成する。クエリエンジン106は、返答のソースリストを含む応答を生成することにもできる。

【 0 0 3 1 】

クエリエンジン 1 0 6 がクエリを受け取る (5 0 2)。クエリは、ユーザによりクライアント 1 0 2 において入力され、クライアント 1 0 2 によりクエリエンジン 1 0 6 に送られる。クエリは 1 つ以上の用語を含む。ユーザにより入力されたクエリはユーザクエリである。

【 0 0 3 2 】

ユーザクエリが処理される (5 0 4)。ユーザクエリは、第 1 のサーチコントローラ 1 1 0 および第 2 のサーチコントローラ 1 1 8 の両方に送られる。ユーザクエリは 1 つ以上の用語を含むため、ウェブサーチ等のドキュメントのサーチクエリとして扱われ、第 1 のサーチコントローラ 1 1 0 等のこのようなサーチを処理するシステム構成要素に送られる。ウェブサーチ等のドキュメントサーチは当分野において周知であり、これ以上説明する必要はない。

【 0 0 3 3 】

ユーザクエリは第 2 のサーチコントローラ 1 1 8 にも送られる。ユーザクエリは事前処理され、解析されて、そのユーザクエリが 1 つ以上の Q A タイプのいずれかに適合するかが判断される。事前処理は、「ストップワード」(定冠詞、不定冠詞、および前置詞) を落とし、それぞれの同義語または等価物を含むようにユーザクエリ内の単語および / または語句を拡張することを含むことができる。例えば、語句「生年月日 (b i r t h d a t e) 」を、(ストップワードである「 o f 」なしの) 同義語「生まれた日 (d a t e b i r t h) 」および「誕生日 (b i t h d a y) 」を含むように拡張することができる。この解析は、ユーザクエリの構文解析およびユーザクエリのテキストの解析を含むことができる。ユーザクエリがいずれかの Q A タイプに適合すると判断された場合、各 Q A タイプに対応するファクトクエリをユーザクエリに対して生成することができる。ファクトクエリは、クエリエンジン 1 0 6 内部のクエリであり、返答候補を求めて (ファクトインデックス 1 2 2 を介して) 第 2 のキャッシュ 1 2 0 およびファクトリポジトリ 1 2 4 にアクセスするために使用される。ユーザクエリがいずれの Q A タイプにも適合しないと判断される場合、第 2 のサーチコントローラ 1 1 8 の観点から、ユーザクエリがファクトクエリではないため、第 2 のサーチコントローラ 1 1 8 によるユーザクエリに対するそれ以上の処理を打ち切ることができる。ユーザクエリが 2 つ以上の Q A タイプに適合すると判断され、その結果、単一のユーザクエリに対して 2 つ以上のファクトクエリが生成され得ることを理解されたい。これら各ファクトクエリを使用して、返答候補を求めてファクトリポジトリ 1 2 5、第 2 のキャッシュ 1 2 0、およびファクトインデックス 1 2 2 にアクセスすることができる。

【 0 0 3 4 】

ユーザクエリは 2 つ以上の Q A タイプに適合し得る。Q A タイプは、どのファクトクエリがユーザクエリで尋ねられているかおよびファクトクエリに応答する返答の種類を示す質問 - 返答マッピングである。いくつかの実施形態では、3 つの一般的な Q A タイプ、すなわち、値への名称および属性 (「 N A - V 」)、名称への属性および値または特性 (「 A V - N 」)、ならびに特性、タイプ、または名称への名称 (「 N - P T N 」) がある。いくつかの実施形態では、特殊なタイプの質問を処理する特殊な Q A タイプがさらにあってよい。いくつかの実施形態では、こういった特殊な Q A タイプは、一般的な Q A タイプの特殊な例であってよい。

【 0 0 3 5 】

N A - V タイプでは、ユーザは、(ユーザクエリの入力により) オブジェクト名および属性を提供し、与えた名称を有するオブジェクトに対応する属性の値を知りたい。N A - V タイプの一例は、「ポーランドの首都はどこか」であり、「ポーランド」がオブジェクト名であり、「首都」が「ポーランド」の値が所望される属性である。このクエリに対する返答は、名称「ポーランド」を有し、かつ属性「首都」を有するオブジェクトに関連するファクトの値である。この場合、「首都」属性を有するファクトの値は文字列「ワルシャワ」である。この値は、名称「ワルシャワ」を有するオブジェクトのオブジェクト識別

10

20

30

40

50

子でもあることができ、この場合、名称「ワルシャワ」をオブジェクト識別子の代わりとし、返答候補として返すことができる。

【 0 0 3 6 】

A V - Nタイプでは、ユーザは属性および値（または、特性は上述したように単に特殊な属性 - 値の対であるため、特性）を提供し、与えた属性に対して与えた値を有する名称を欲する。ある意味、これは「逆引き」である。A V - Nタイプクエリの一例は、「首都がワルシャワである国はどこか」であり、「首都」が属性であり、「ワルシャワ」が値である。返答候補は、この属性 - 値の対を有するオブジェクトの名称、すなわち「ポーランド」である。

【 0 0 3 7 】

N - P T Nタイプでは、ユーザは名称を提供し、与えた名称に関連するオブジェクトの特性、タイプ、または代替名称を欲する。N - P T Nタイプクエリの一例は、「N R Aとは何か」である。「N R A」は、ユーザが特性、タイプ、または代替名称を欲するオブジェクトの名称である。「N R A」に対する1つの特性返答候補は、「憲法修正第2条権利支持団体」である。「N R A」での、オブジェクトにより表されるエンティティまたは概念のカテゴリを伝えるタイプ返答は「組織」であり、N R Aが、人物、書籍、映画等の他のタイプではなく組織であることを示す。「N R A」の代替名称は、（頭字語）名「N R A」を有するオブジェクトにより表されるエンティティの正式名称である「全米ライフル協会」である。

【 0 0 3 8 】

いくつかの実施形態では、ファクトクエリは追加の制約を含むことができる。例えば、ファクトクエリは、特定の用語が特定のフィールドでのみマッチし、他のフィールドでマッチしなくてもよいと指定することができる。別の制約は、返答候補がすべて特定のタイプ（人物、書籍等）にマッチしなければならないことであることができる。このような制約は、ユーザクエリの解析および処理中に第2のサーチコントローラ118により生成される。

【 0 0 3 9 】

ユーザクエリが処理され、1つ以上のファクトクエリが生成された後、返答候補を求めてファクトクエリを使用して（ファクトインデックス122を介して）ファクトリポジトリ124および第2のキャッシュ120にアクセスする（506）。返答候補とは、ファクトクエリ（複数可）にマッチするファクトである。返答候補にスコアが付けられる（508）。返答候補のスコアは、正確で応答性のある返答としての返答候補の質の指標を提供する。

【 0 0 4 0 】

いくつかの実施形態では、返答候補のスコアは複数の係数値の乗法積である。いくつかの実施形態では、係数値の1つ以上は0～1の（0および1を含む）正規化された値であってよい。スコアを決める際に実際に使用される係数は、返答候補にマッチするファクトクエリのQ Aタイプにより様々であり得る。いくつかの実施形態では、スコアは、0～1の（0および1を含む）の係数の積であり、0のままであるか、または0に向かって減少するが、決して増大し得ないため、任意の特定の返答候補のスコア計算を、その特定の返答のスコアが所定の閾値未満に低減した場合に打ち切ることができる。これは、返答候補が、それ以上のスコア計算が無駄であるような低い質であることを示すことができる。

【 0 0 4 1 】

いくつかの実施形態では、係数は、Q Aタイプ、ファクトクエリにマッチしたファクトのメトリック（信頼度メトリックおよび重要度メトリック等）、マッチするファクトを抽出したエージェント、ファクト内のフィールドがファクトクエリによりマッチした程度、ファクト内の特定のフィールドがファクトクエリに完全にマッチする程度等に基づくことができる。上記係数は単に例示にすぎず、他の係数を上記係数に加えて含めてもよく、上記係数のいくつかを省いてもよいことを理解されたい。

【 0 0 4 2 】

10

20

30

40

50

各返答候補のスコアが計算された後、返答候補を図 4 に関連して上述した返答候補リスト 400 等の返答候補リストに集めることができる。いくつかの実施形態では、所定数の最高スコアの返答のみが返答候補リスト 400 に集められる。例えば、返答候補リストは、スコアの上位 100 の返答候補のみを含むことができる。いくつかの実施形態では、返答候補リスト 400 のさらなる処理が、第 2 のサーチコントローラ 118 により行われる。

【0043】

図 5 B に続き、返答候補リスト 400 から、ある数の上位スコアの返答候補が特定される (510)。この数は、いくつかの上位スコア返答がさらに処理されるかを指定する所定数であることができる。処理する、特定された上位スコア返答候補がまだある限り (512 - no)、次の上位スコア返答が処理される。この処理は、各上位スコア返答をサポートする返答を特定すること (514) および各上位スコア返答のスコアおよびサポート返答のスコアに基づいて各上位スコアのサポートスコアを求めること (516) を含む。サポート返答の特定について以下においてより詳細に説明する。

10

【0044】

いくつかの実施形態では、サポートスコアは上位スコア返答およびそのサポート返答の各スコアをオッズ空間の値に変換することにより求められる。スコア s は、オッズ空間の値 x に変換される。

【数 1】

20

$$x = \left(\frac{s}{1-s} \right)$$

変換された値 (すなわち、スコアをオッズ空間に変換したもの) が合算されて、値 X が生成され、この和 X が再び確率空間の値に変換されて、上位スコア返答のサポートスコア S が得られる。

30

【数 2】

$$S = \left(\frac{X}{1+X} \right)$$

40

【0045】

特定された上位スコア返答のサポートスコアが求められた後 (512 - yes)、最高のサポートスコアを有する最高スコア返答 (以下、「最良サポート返答」) が特定される (518)。最良サポート返答について、最良サポート返答に矛盾する返答候補リスト内の上位スコア返答が特定される (520)。その矛盾する返答に対して、その矛盾する返答のサポートスコアである矛盾スコア C が求められる (522)。さらに、いくつかの実施形態では、最良サポート返答に関連しない返答候補リスト内の上位スコア返答が特定される (524)。その非関連返答に対して、その非関連返答のサポートスコアである非関連スコア U が求められる (526)。矛盾スコア C および非関連スコア U を求めるプロセスは S を求めるプロセス、すなわち、サポートする返答を特定し、そのスコアをオッズ空間

50

の値に変換し、オッズ空間の値を合算し、その和を再び確率空間の値に変換するプロセスと同じであることを理解されたい。矛盾返答および非関連返答の判断について後述する。

【 0 0 4 6 】

2つの返答候補がサポート返答であるか、矛盾返答であるか、それとも非関連返答であるかは、その2つの返答のフィールドの比較に基づく。各返答の対象となるフィールド、すなわち、名称、属性、および値が入力および出力にグループ化される。例えば、いくつかの実施形態では、NA - Vタイプクエリの場合、入力 は 名称 および 属性 であり、出力 は 値 である。AV - Nタイプクエリの場合、入力 は 属性 および 値 であり、出力 は 名称 である。2つの返答候補が、それぞれの入力フィールドおよび出力フィールドのペアワイズ比較を行うことにより比較される。この比較は、フィールド内のデータタイプ、すなわちフィールド内のデータが文字列であるか、日付であるか、それとも数等であることを考慮に入れる。返答のソースを考慮に入れることもできる。

10

【 0 0 4 7 】

いくつかの実施形態では、ペアワイズフィールド比較の結果は5つの分類のうちの1つである。これらは以下である。

- ・ 比較不可能：フィールドが異なるデータタイプを有する（例えば、文字列と日付）ため、比較できない。
- ・ 類似しない：フィールドは同じデータタイプであるが、全然同じではない。
- ・ いくらか類似する：フィールドはいくらかの類似性を有するが、同じことを意味しているか否かを結論付けるのは困難である。
- ・ かなり類似する：フィールドは略同じである。
- ・ 同一である：フィールドはまったく同じである。

20

【 0 0 4 8 】

フィールドが同じであるか否かの実際の判断は、データタイプにより異なり得る。例えば、数の場合、数が小さな整数である場合、同じものとして扱うためにはそれら数は厳密に同じでなければならない。数が非常に大きな整数または浮動小数点数である場合、互いに特定の割合以内にある場合にそれら数を同じものとして扱うことができる。

【 0 0 4 9 】

ペアワイズフィールド比較に基づき、2つの返答の関係が分類される。

- ・ 返答が同じソースからのものである場合、2つの返答は「補足的」に分類される。返答Aに対して補足的な返答は無視される。
- ・ 返答が同一またはかなり類似する入力を有するが、出力がいくらか類似するにすぎない場合、2つの返答は「サポートし得る」（すなわち、返答Aが返答Bを「サポートし得る」）に分類される。返答Bを「サポートし得る」返答Aも無視される。
- ・ 返答が同一またはかなり類似する入力および同一またはかなり類似する出力を有する場合、2つの返答は、同じソースからのものではない限り、「サポートしている」に分類される。「サポートしている」返答のスコアは、サポートスコアの決定の一部である。
- ・ 入力 が 同一 または かなり 類似 する が、出力 が 類似 しない、または 比較 不可能 である 場合、2つの返答は「矛盾」返答である。
- ・ 入力 が 類似 しない、または 比較 不可能 である 場合、2つの返答は「非関連」返答である。

30

40

【 0 0 5 0 】

最良サポート返答のサポートスコアSが所定の閾値Tと比較される（528）。閾値Tは、これ以降も最良サポートスコアとみなすためにサポートスコアSが達しなければならない最小スコアである。SがT以下の場合（528 - no）、図5Bに示す処理を実行しているプロセッサ（複数可）（たとえば、第2のサーチコントローラ118またはクエリエンジン106）は、クエリエンジン106が返答を提供できないことを示す応答を生成することができる（534）。例えば、第2のサーチコントローラ118は、返答を入手できないことを示す応答をクエリサーバ108に送り、クエリサーバ108は、そのことに対する応答を生成してクライアント102に送り、ユーザに提示することができる。

50

【 0 0 5 1 】

S が T よりも大きい場合 (5 2 8 - y e s)、最良サポート返答のサポートスコア S が、矛盾返答の最良サポートスコア C を少なくとも第 1 の所定のマージンで超えるか否かが調べられる。一実施形態では、この調査は、S を、矛盾スコア C に定数 を乗算した数と比較することで行われる (5 3 0)。定数 は、最良サポート返答をファクトクエリに対する最良返答として選択するために実現しなければならない S と C の最小比を表す。換言すれば、S は、少なくとも矛盾スコア C の 倍である必要がある。S が C 未満である場合 (5 3 0 - n o)、図 5 B に示す処理を実行しているプロセッサ (複数可) (たとえば、第 2 のサーチコントローラ 1 1 8 またはクエリエンジン 1 0 6) は、クエリエンジン 1 0 6 が返答を提供できないことを示す応答を生成することができる (5 3 4)。

10

【 0 0 5 2 】

S が C 以上である場合 (5 3 0 - y e s)、最良サポート返答のサポートスコア S が非関連返答の最良サポートスコア U を少なくとも第 2 の所定のマージンで超えるか否かがさらに調べられる。一実施形態では、この調査は、S を、非関連スコア U に定数 を乗算した数と比較することで行われる (5 3 2)。定数 は、最良サポート返答をファクトクエリに対する最良返答として選択する前に実現しなければならない S と U の最小比を表す。換言すれば、S は、少なくとも非関連スコア U の 倍である必要がある。S が U 未満である場合 (5 3 2 - n o)、図 5 B に示す処理を実行しているプロセッサ (複数可) (たとえば、第 2 のサーチコントローラ 1 1 8 またはクエリエンジン 1 0 6) は、クエリエンジン 1 0 6 が返答を提供できないことを示す応答を生成することができる (5 3 4)。S が U 以上の場合 (5 3 2 - y e s)、最良サポート返答がファクトクエリに対する返答として選択され、さらに処理される。これについて、図 5 C に関連してさらに詳細に後述する。

20

【 0 0 5 3 】

図 5 C に続き、最良サポート返答がファクトクエリに対する最良返答として選択された後、クエリサーバ 1 0 8 は応答を生成する (5 3 6)。応答は最良サポート返答を含むことができる。応答は、最良サポート返答のソースの識別子および / またはハイパーリンク (たとえば、URL へのハイパーリンク) を含むことができる。いくつかの実施形態では、応答は、ユーザがクライアント 1 0 2 でクリックしたときに最良サポート返答のソースリスト要求を生成することができるリンクを含むこともできる。いくつかの実施形態では、応答は、ユーザクエリに基づいたウェブサーチ等のドキュメントサーチの結果を含むこともできる。ドキュメントサーチ結果は、第 1 のサーチコントローラ 1 1 0 等の、このようなサーチを処理するクエリエンジン 1 0 6 の構成要素からクエリサーバ 1 0 8 に送ることができる。応答はクライアント 1 0 2 に送られて、ユーザに提示される (5 3 8)。ユーザクエリを使用したドキュメントサーチの返答および結果を含む例示的な応答について、図 6 に関連してさらに詳細に後述する。

30

【 0 0 5 4 】

ユーザは、クライアント 1 0 2 に提示された応答を見て、その返答のソースリストを要求することができる。いくつかの実施形態では、ユーザは、クリックされたときにソースリスト要求を生成する、応答に含まれる (上述したような) リンクをクリックすることでその要求を行うことができる。

40

【 0 0 5 5 】

クエリエンジン 1 0 6 は、返答のソースをリストする要求を受け取る (5 4 0)。返答のソースが特定される (5 4 2)。いくつかの実施形態では、返答のソースは、ファクトリポジトリ 1 2 4 内の返答ファクトのソース 2 2 0 (図 2) を調べることで特定することができる。スニペット生成要求がソースリスト、ユーザクエリ、返答にマッチしたファクトクエリ 4 1 8、および返答と共に第 1 のサーチコントローラに送られる。スニペット生成要求は、第 1 のキャッシュ 1 1 2、ドキュメントインデックス 1 1 4、および / またはドキュメントデータベース 1 1 6 に送られる。いくつかの実施形態では、ソースリストが所定の限度よりも長い場合、第 1 のサーチコントローラ 1 1 0 は、ソースリストのスニペ

50

ットを選択して第1のキャッシュ112、ドキュメントインデックス114、および/またはドキュメントデータベース116に送ることができる。第1のキャッシュ112、ドキュメントインデックス114、および/またはドキュメントデータベース116、またはスニペット生成要求が提出される1つ以上のプロセッサは、リストされた各ソースのスニペットを生成する(544)。各スニペットは、各ソースからのテキストの1つの連続部分または複数の非連続部分を含むことができる。特定のスニペットでは、スニペットに含めるものとして選択されたテキストの部分がソース内で連続していない場合、その部分を楕円で分けることができる。

【0056】

各スニペットは、可能な限り多くのユーザクエリ用語および/またはファクトクエリ用語および返答用語を含むように生成される。ソースを、クエリ用語および返答用語のばらつき(すなわち、クエリ用語および返答用語がソースドキュメント内でどのように点在するか)について解析して、スニペットの生成を助けることもできる。クエリ用語および返答用語のばらつきが最小のテキスト部分(複数可)が、スニペットに含めるものとして選択される。

【0057】

スニペットを含む応答が生成される(546)。応答は、ソースリストおよびユーザ/ファクトクエリ用語および返答用語を含む各ソースのスニペットを含む。応答は、返答、ユーザクエリ、および各ソースへのハイパーリンクを含むこともできる。応答はクライアント102に送られ、ユーザに提示される(548)。

【0058】

いくつかの実施形態では、各スニペット内のユーザ/ファクトクエリ用語および返答用語が強調表示されて、応答がユーザに提示されたときにより目立つようにされる。本明細書において使用するスニペット内の用語の強調表示は、用語を太字にする、用語に下線を付ける、用語を斜体にする、用語のフォントカラーを変える、および/または背景カラーを用語の局所エリアに加えることを含むがこれらに限定されない、ユーザに提示されたときに用語をより目立たせる任意の様式を指す。ソースリストおよびスニペットを含む例示的な応答について、図7に関連してさらに詳細に後述する。

【0059】

いくつかの実施形態では、クエリエンジン106に提出されるクエリは、ユーザクエリ用語および1つ以上の他のパラメータを含むURLで表すことができる。例えば、「britney spears parents」という用語のクエリを、URL「http://www.google.com/search?hl=en&q=britney+spears+parents」で表すことができる。いくつかの実施形態では、「&fsrc=1」等の追加のパラメータをクエリURLに加えることにより、返答のソースリストを表示する要求を行うことができる。したがって、上記URLでは、クエリ「britney spears parents」への返答のソースリストが所望される場合、クエリURLは、「http://www.google.com/search?hl=en&q=britney+spears+parents&fsrc=1」のようになり得る。いくつかの実施形態では、返答を含む応答内の、ユーザがクリックしたときに返答のソースリスト要求をトリガするリンクは、追加パラメータが加えられたユーザクエリのクエリURLである。

【0060】

いくつかの実施形態では、クエリエンジン106は、ファクトクエリと共に、クエリを使用して見つけられたドキュメントリストをドキュメントサーチの入力として最初に返さずに、クエリエンジン106にファクトクエリへの返答を見つけ、返答および返答のソースリストを返すように命令する所定の特別な演算子を受け入れることができる。例えば、ユーザは「Z: X of Y」を入力することができ、「Z:」は、クエリエンジン106にファクトクエリ「X of Y」への返答および返答のソースリストを見つけるように命令する特別な演算子である。ある意味、クエリと共に演算子を使用することで、クエ

10

20

30

40

50

りが、そのクエリに対して見つけられた返答が何であれ、そのソースリスト要求と併合される。いくつかの実施形態では、上述したように、返答のソースリスト要求を生成する応答内のリンクは、ユーザがそのリンクを選択した（たとえば、クリックした）ときに、特別な演算子をオリジナルのクエリに加え、特別な演算子を有するクエリをクエリエンジン 106 に提出する。

【0061】

図6は、入力としてファクトクエリを使用して返答およびドキュメントサーチ結果を含む、ファクトクエリに対してクライアント102でユーザに提示される本発明のいくつかの実施形態による例示的な応答を示す。応答600は、オリジナルのユーザクエリを有するサーチボックス602を示すことができる。応答600は、クエリに対する返答604、返答のソースへのハイパーリンク606、および、ユーザがクリックしたときに、返答に対するソースのリストを要求するトリガとなるリンク608を含む。いくつかの実施形態では、リンク608は、上述したように、ソースリスト要求パラメータが追加されたユーザクエリのクエリURLであってよい。いくつかの実施形態では、返答ファクト604がファクトリポジトリ124に1つのソースしか有さない場合、ユーザがクリックしたときに返答のソースリスト要求をトリガするリンクを省いてよい。応答は、入力としてファクトクエリを使用したウェブサーチ等のドキュメントサーチの結果リスト610を含むこともできる。

【0062】

図7は、本発明のいくつかの実施形態による返答のソースリスト要求に対する例示的な応答を示す。応答700は、オリジナルのユーザクエリを有するサーチボックス702を含むことができる。いくつかの実施形態では、サーチボックス702は、上述したように、ファクトクエリと共に使用されてソースリストを要求できる特別な演算子を含むこともできる。例えば、サーチボックス702内で「factsources:」は特別な演算子であり、「britneyspearsparents」はオリジナルのユーザクエリである。いくつかの他の実施形態では、例えば、ソースリスト要求のトリガが、ソースリスト要求パラメータを有するクエリURLを含んだリンク608等のリンクのユーザクリックである場合、演算子をサーチボックス内のクエリの提示から省いてもよい。応答は、ファクトクエリへの返答704と、各ソースのURL、ハイパーリンク、およびスニペット708を有する、返答の1つ以上のソースリスト706とを含むこともできる。いくつかの実施形態では、各スニペット708内でクエリ用語および返答用語を強調表示することができる。スニペット708内で、クエリ用語および返答用語は太字により強調表示される。

【0063】

図8は、本発明のいくつかの実施形態によるファクトクエリ返答システム800を示すブロック図である。システム800は、通常、1つ以上の処理装置(CPU)802、1つ以上のネットワークまたは他の通信インタフェース810、メモリ812、およびこれら構成要素を相互接続する1つ以上の通信バス814を含む。システム800は、任意に、表示装置806およびキーボード/マウス808を備えるユーザインタフェース804を含むことができる。メモリ812は、DRAM、SRAM、DDR RAM、または他のランダムアクセス固体状態メモリ装置等の高速ランダムアクセスメモリを含み、1つ以上の磁気ディスク記憶装置、光ディスク記憶装置、フラッシュメモリ装置、または他の不揮発性固体状態記憶装置等の不揮発性メモリを含むこともできる。メモリ812は、任意に、CPU802から離れて配置された1つ以上の記憶装置を含むことができる。いくつかの実施形態では、メモリ812は以下のプログラム、モジュール、およびデータ構造またはこれらのサブセットを記憶する。

- ・各種基本システムサービスを処理し、ハードウェア依存タスクを実行する手続きを含むオペレーティングシステム816

- ・クエリを受け取り、クエリを構文解析してQAタイプを判断してファクトクエリを生成するなどのクエリを処理を実行するクエリ受信・処理モジュール818

- ・ファクトクエリに対する返答候補を特定する返答特定モジュール 8 2 0
- ・返答のスコアおよびサポートスコアを求める返答スコア計算モジュール 8 2 2
- ・返答を比較してサポート返答、矛盾返答等を判断する返答比較モジュール 8 2 4
- ・ユーザに提示する返答として返答候補を選択する返答選択モジュール 8 2 5
- ・返答ソースを特定するソース特定モジュール 8 2 6
- ・ドキュメントサーチ時にドキュメントインデックスとインタフェースするドキュメントインデックスインタフェース 8 2 8
- ・スニペットを要求し受け取るときにドキュメントストレージシステムとインタフェースするドキュメントストレージインタフェース 8 3 0
- ・ファクトサーチ時にファクトインデックスとインタフェースするファクトインデックスインタフェース 8 3 2
- ・ファクトストレージシステムとインタフェースするファクトストレージインタフェース 8 3 4
- ・クライアント 1 0 2 に送る応答を生成する応答生成モジュール 8 3 8

【 0 0 6 4 】

いくつかの実施形態では、システム 8 0 0 のメモリ 8 1 2 は、ファクトインデックスへのインタフェース 8 3 2 に代えてファクトインデックスを含む。システム 8 0 0 は、そのうちのいくつかが返答ファクトのソースであり得るドキュメントの内容を記憶するドキュメントストレージシステム 8 4 0 も含む。ドキュメントストレージシステムは、ドキュメントの内容にアクセスし、内容からスニペットを生成するスニペットジェネレータ 8 4 2 およびスニペット内のクエリ用語および返答用語を強調表示するスニペット用語強調表示モジュール 8 3 6 を含む。システム 8 0 0 は、ファクトを記憶するファクトストレージシステム 8 4 4 も含む。ファクトストレージシステム 8 4 4 に記憶される各ファクトは、各ファクトが抽出された対応するソースリストを含む。

【 0 0 6 5 】

上記特定した各要素は、上述したメモリ装置の 1 つ以上に記憶することができ、上述した機能を実行する命令セットに対応する。上記特定したモジュールまたはプログラム（すなわち、命令セット）は、別個のソフトウェアプログラム、手続き、またはモジュールとして実施される必要はなく、そのため、各種実施形態において、これらモジュールの各サブセットを組み合わせたか、または再配置することが可能である。いくつかの実施形態では、メモリ 8 1 2 は、上で特定したモジュールおよびデータ構造のサブセットを記憶することができる。さらに、メモリ 8 1 2 は、上述していない追加のモジュールおよびデータ構造を記憶することもできる。

【 0 0 6 6 】

図 8 にはファクトクエリ返答システムが示されているが、図 8 は、本発明において述べた実施形態の構造的な概略よりも、サーバセットに存在し得る各種特徴の機能的な説明を意図する。実際に、そして当業者に認識されるように、別個に図示されるアイテムを組み合わせてもよく、かついくつかのアイテムを分けてもよい。たとえば、図 8 に別個に示されるいくつかのアイテムを単一のサーバで実施してもよく、かつ単一のアイテムを 1 つ以上のサーバで実施してもよい。ファクトクエリ返答システムの実施に使用される実際のサーバの数および特徴がサーバにどのように割り振られるかは、実施態様によって異なり、システムがピーク使用期間ならびに平均使用期間中に処理しなければならないデータトラフィック量に部分的に依存し得る。

【 0 0 6 7 】

説明を目的とした上記説明は特定の実施形態を参照してなされた。しかし、上記の例示的な考察は、網羅的である、すなわち本発明を開示した厳密な形態に限定する意図はない。多くの変更および変形が上記教示に照らして可能である。実施形態は、本発明および当業者が意図する特定の用途に適した各種変更を有する各種実施形態を最良に利用できるように、本発明の原理およびその実際の応用を最良に説明するために選択され、説明されたものである。

10

20

30

40

50

【図面の簡単な説明】

【0068】

【図1】本発明のいくつかの実施形態によるネットワークを示す。

【図2】本発明のいくつかの実施形態によるファクトリポジトリ内のオブジェクトおよび関連ファクトのデータ構造を示す。

【図3】本発明のいくつかの実施形態によるファクトインデックスのデータ構造を示す。

【図4】本発明のいくつかの実施形態による返答候補リストのデータ構造を示す。

【図5A】ファクトクエリに対する返答を選択し、返答および返答のソースを表示する、本発明のいくつかの実施形態によるプロセスの流れ図である。

【図5B】ファクトクエリに対する返答を選択し、返答および返答のソースを表示する、本発明のいくつかの実施形態によるプロセスの流れ図である。

【図5C】ファクトクエリに対する返答を選択し、返答および返答のソースを表示する、本発明のいくつかの実施形態によるプロセスの流れ図である。

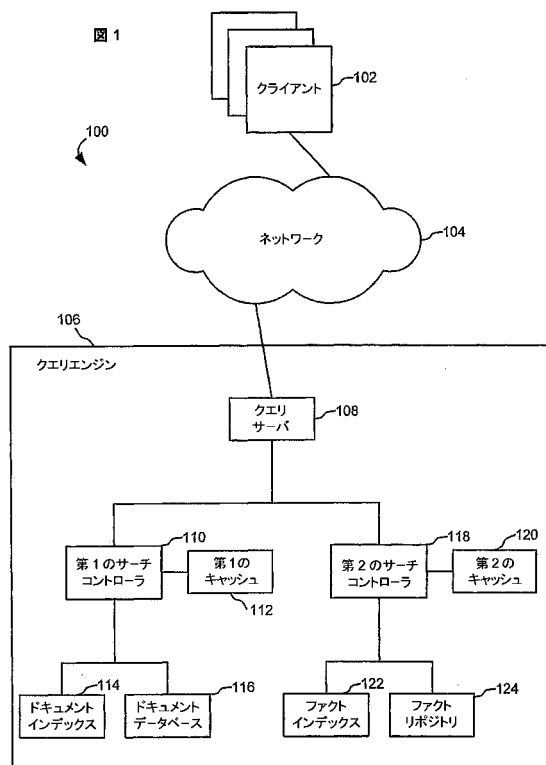
【図6】本発明のいくつかの実施形態によるファクトクエリに対する返答の提示を示す。

【図7】本発明のいくつかの実施形態によるファクトクエリに対する返答のソースリストの提示を示す。

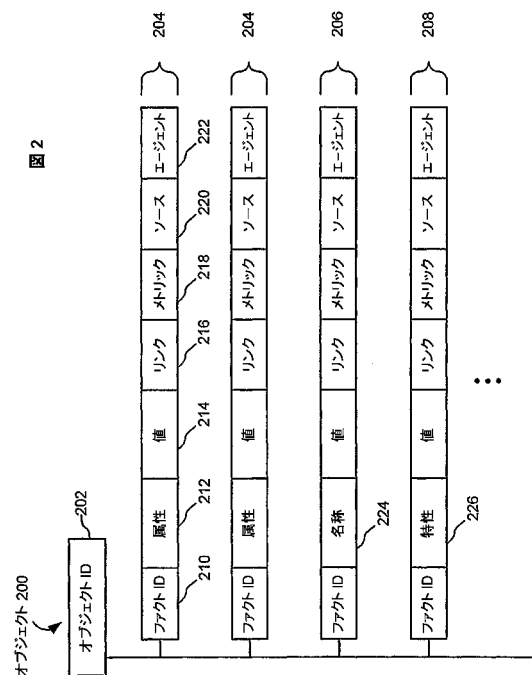
【図8】本発明のいくつかの実施形態によるファクトクエリに対する返答を選択し、返答および返答のソースリストを表示するシステムを示す。

10

【図1】



【図2】



【 図 3 】

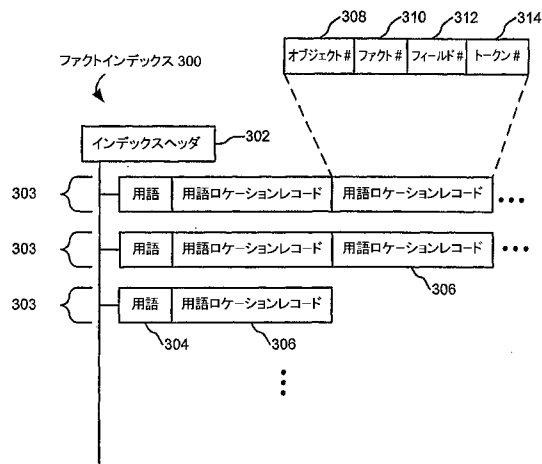
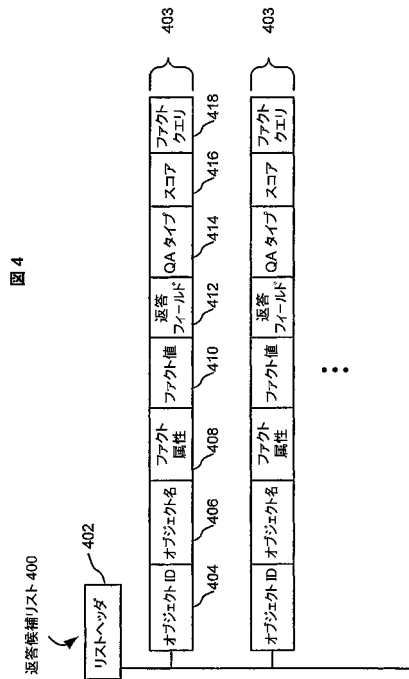


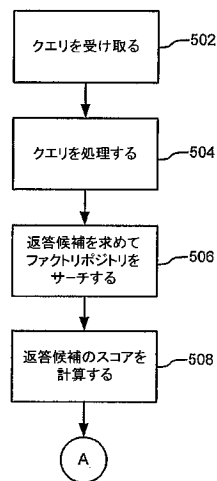
图 3

【 図 4 】

4. ☒

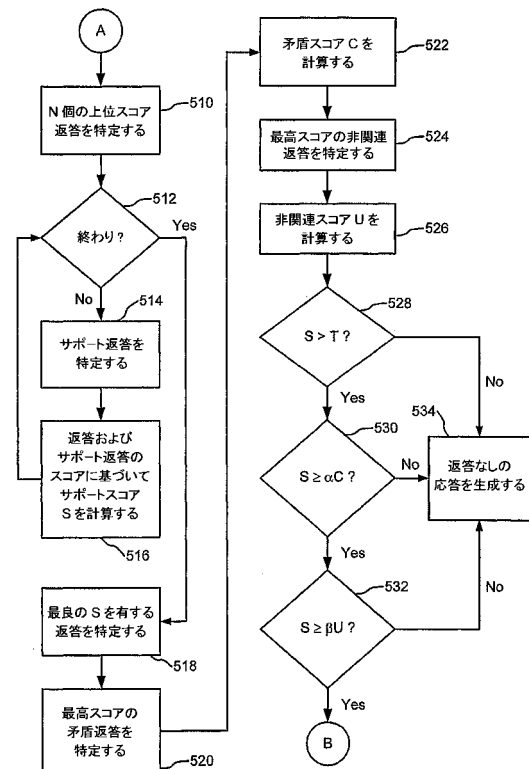
【 図 5 A 】

5A

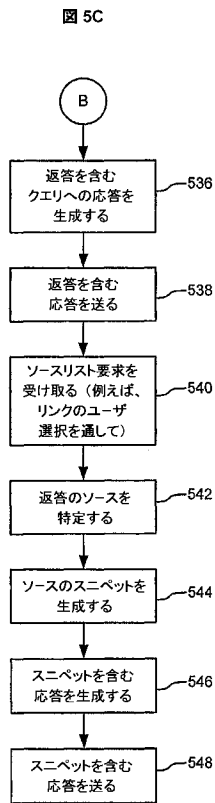


【 図 5 B 】

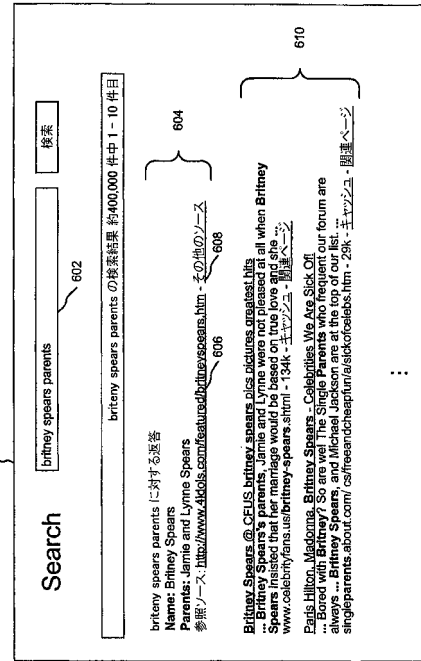
图 5B



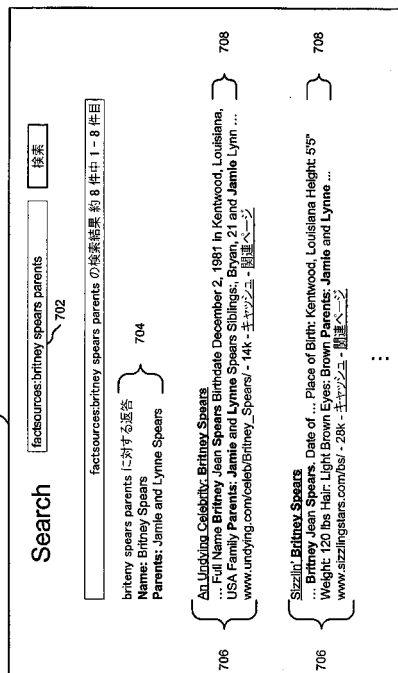
【図 5 C】



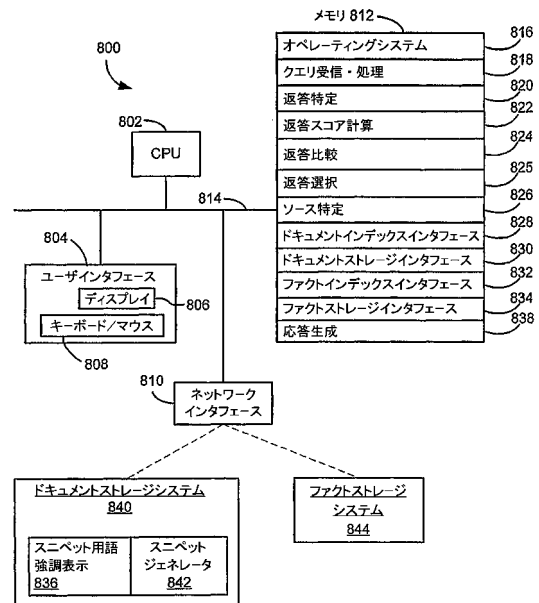
【図 6】



【図 7】



【図 8】



フロントページの続き

審査官 野崎 大進

- (56)参考文献 特開 2 0 0 2 - 1 5 7 2 7 6 (J P , A)
特開平 1 1 - 2 6 5 4 0 0 (J P , A)
特表 2 0 0 2 - 5 4 0 5 0 6 (J P , A)
特開 2 0 0 3 - 2 8 1 1 7 3 (J P , A)
米国特許出願公開第 2 0 0 3 / 0 0 6 9 8 8 0 (U S , A 1)
国際公開第 2 0 0 4 / 1 1 4 1 6 3 (W O , A 1)

- (58)調査した分野(Int.Cl. , D B 名)
G06F 17/30
JSTPlus(JDreamII)