

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-120129

(P2006-120129A)

(43) 公開日 平成18年5月11日(2006.5.11)

(51) Int. Cl.

G06F 17/30 (2006.01)

F I

G06F 17/30 220Z

テーマコード (参考)

5B075

審査請求 未請求 請求項の数 18 O L (全 18 頁)

(21) 出願番号 特願2005-272629 (P2005-272629)
 (22) 出願日 平成17年9月20日 (2005.9.20)
 (31) 優先権主張番号 10/969,567
 (32) 優先日 平成16年10月20日 (2004.10.20)
 (33) 優先権主張国 米国 (US)

(71) 出願人 500046438
 マイクロソフト コーポレーション
 アメリカ合衆国 ワシントン州 9805
 2-6399 レッドモンド ワン マイ
 クロソフト ウェイ
 (74) 代理人 100077481
 弁理士 谷 義一
 (74) 代理人 100088915
 弁理士 阿部 和夫
 (72) 発明者 ブレイク イー. アンダーソン
 アメリカ合衆国 98052 ワシントン
 州 レッドモンド ワン マイクロソフト
 ウェイ マイクロソフト コーポレーシ
 ョン内

最終頁に続く

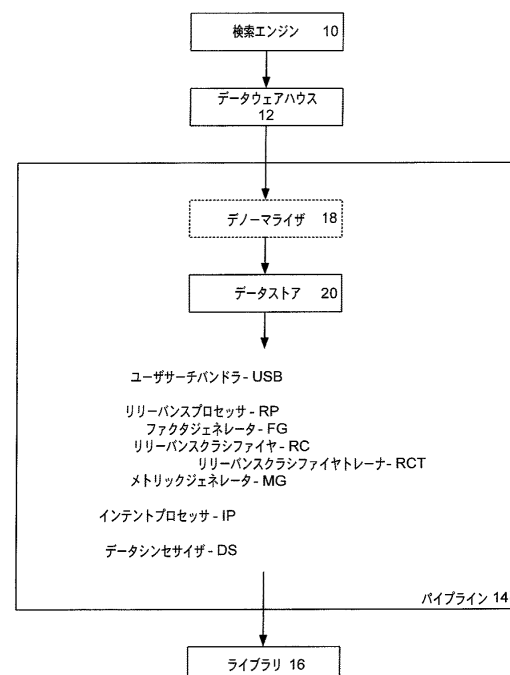
(54) 【発明の名称】 検索システムまたはその類似物からの動作データおよびその他のデータの分析

(57) 【要約】

【課題】 大量のデータ、特に検索システムおよびその類似物からの大量のデータを分析するシステムおよび方法を提供すること。

【解決手段】 システムは、検索エンジンからのデータを分析する。ユーザサーチバンドラは、ユーザサーチを分析し、類似するユーザサーチをユーザサーチバンドルにグループ化し、インテントプロセッサは、ユーザサーチバンドルに基づいてインテントを作る。ファクタジェネレータは、ユーザサーチおよび関連情報を検討して、ファクタを作り、各ファクタは、サーチリザルトのセットからの特定のリザルトに関する。リリーバンスクラシファイヤは、ファクタを受け取り、それに基づいて動作して、各リザルトのジャッジメントを作る。メトリックジェネレータは、ファクタおよびジャッジメントに基づいてメトリックを作り、データシンセサイザは、抽出されたデータをデータベースにフォーマットする。

【選択図】 図2



【特許請求の範囲】

【請求項 1】

検索エンジンからのデータを分析するシステムであって、そこでは前記検索エンジンは、要求するユーザから受け取ったクエリストリングに基づいてサーチリザルトのセットを生成し、前記クエリストリングおよび前記サーチリザルトは、集合的にユーザサーチを含み、前記サーチリザルトは、少なくとも 1 つのリザルトを含み、各リザルトは、前記クエリストリングに関係すると思われるコンテンツの特定のアイテムを参照し、これによって、一連の関連するユーザサーチは、セッションを含み、前記検索エンジンは、各ユーザサーチおよび関連情報をストアするシステムにおいて、

そのようなユーザサーチのうちで類似するユーザサーチを見つけ、そのような類似するユーザサーチをユーザサーチバンドルにグループ化するためにユーザサーチを分析するユーザサーチバンドラ (USB) と、 10

前記 USB からのユーザサーチバンドルに基づいてインテントを作るインテントプロセッサ (IP) であって、各インテントは、互いに関係すると思われる 1 つまたは複数のセッションのグループである、インテントプロセッサ (IP) と、

ファクタを作るためにユーザサーチおよび関連情報を検討するファクタジェネレータ (FG) であって、各ファクタは、サーチリザルトのセットからの特定のリザルトに関し、各ファクタは、1 つまたは複数のイベントに関係し、各イベントは、照会するユーザが実行した動作に関する 1 つの情報である、ファクタジェネレータ (FG) と、

前記リザルトのジャッジメントを作るために各リザルトについて前記 FG によって生成された前記ファクタを受け取り、それに基づいて動作するリリーバンスクラシファイヤ (RC) であって、前記ジャッジメントは、前記サーチリザルトからの前記リザルトへのアクセスを決定する際に前記リザルトを前記ユーザがどのように審判したかの判定を表す、リリーバンスクラシファイヤ (RC) と、 20

前記 FG によって生成された前記ファクタおよび前記 RC によって作られた前記ジャッジメントに基づいてメトリックを作るメトリックジェネレータ (MG) であって、各メトリックは、リザルト、ユーザサーチ、またはセッションに関する尺度である、メトリックジェネレータ (MG) と、

前記 USB、IP、FG、RC、および MG によって生成されたデータを抽出し、前記抽出されたデータを 1 つまたは複数のデータベースにフォーマットし、前記データベースをライブラリにストアし、これによって、フィードバックを提供するかレポートを生成するために前記データをレビューし、集計できるようにする、データシンセサイザ (DS) と 30

を含むことを特徴とするシステム。

【請求項 2】

前記検索エンジンは、各クエリストリングと前記対応するサーチリザルトと関連情報とをデータウェアハウス内に正規化された形でストアし、前記システムは、さらに、前記データウェアハウスから前記正規化されたデータを取り出し、これを正規化し、前記正規化されたデータをデータストアにストアするノーマライザを含むことを特徴とする請求項 1 に記載のシステム。 40

【請求項 3】

前記 USB は、クエリストリングの類似性およびサーチリザルトの類似性のうちの少なくとも 1 つについて前記ユーザサーチを分析することを特徴とする請求項 1 に記載のシステム。

【請求項 4】

各イベントは、ユーザが、特定のリザルトの選択およびクローズのうちの少なくとも 1 つを実行した時刻を含み、前記 FG は、ユーザがリザルトを見た時間の長さを表す「ドエルタイム」ファクタを計算し、前記ドエルタイムファクタは、対応するタイムスタンプ付きイベントによってそれぞれが表される、前記ユーザが前記リザルトを選択した時刻および前記ユーザが前記リザルトをクローズした時刻の間の差に基づくことを特徴とする請求 50

項 1 に記載のシステム。

【請求項 5】

前記 R C は、「A c c e p t」ジャッジメント、「E x p l o r e」ジャッジメント、および「R e j e c t」ジャッジメントのうちの少なくとも 1 つと、前記ジャッジメントがどれほど正しい可能性があるかの信頼度を示す対応する値とを含むジャッジメントを作

【請求項 6】

前記 F G からエクスプリシットジャッジメントファクタを受け取り、それに基づいて前記 R C を生成するリリーバンスクラシファイアトレナをさらに含み、各エクスプリシットジャッジメントファクタは、対応するリザルトに関する前記ユーザからの明示的フィードバックを表し、前記 R C T は、どのファクタがどのジャッジメントを暗示するかを前記エクスプリシットジャッジメントファクタから学習し、これに基づいて前記 R C を生成することを特徴とする請求項 1 に記載のシステム。

10

【請求項 7】

前記 M G は、リザルトに関して、

前記ユーザは前記リザルトをどのようにランキングしたと審判したかに関するポジションメトリックと、

前記リザルトが前記サーチリザルト内でどのように位置付けられたかに関するリリーバンスポジションメトリックと、

前記ポジションメトリックおよび前記リリーバンスポジションメトリックに基づく、前記リザルトがあるべきであったところからどれほど「離れていた」かに関するミスランクトリザルトメトリックと

20

のうちの少なくとも 1 つを作ることとを特徴とする請求項 1 に記載のシステム。

【請求項 8】

前記 I P は、レビューされたユーザサーチバンドルに基づいて、セッションにまたがる共通のリザルトおよびセッションにまたがる共通のクエリタームを突き止めることによってセッション間の関係値を判定し、そのような共通のリザルトが見つかった場合にストレングスオブコモナリティを突き止め、そのようなストレングスオブコモナリティは、2 つのセッションが共通の目的を有することによって互いに関係する可能性がどれほどあるかを表し、前記 I P は、所定の閾値を超えるストレングスオブコモナリティを有するセッション対をインテントにバンドルすることを特徴とする請求項 1 に記載のシステム。

30

【請求項 9】

前記 D S は、前記抽出されたデータをリレーショナルデータベースにフォーマットすることを特徴とする請求項 1 に記載のシステム。

【請求項 10】

検索エンジンからのデータを分析する方法であって、そこでは前記検索エンジンは、要求するユーザから受け取ったクエリストリングに基づいてサーチリザルトのセットを生成し、前記クエリストリングおよび前記サーチリザルトは、集合的にユーザサーチを含み、前記サーチリザルトは、少なくとも 1 つのリザルトを含み、各リザルトは、前記クエリストリングに関係すると思われるコンテンツの特定のアイテムを参照し、これによって、一連の関連するユーザサーチは、セッションを含み、前記検索エンジンは、各ユーザサーチおよび関連情報をストアする方法において、

40

そのようなユーザサーチのうちで類似するユーザサーチを見つけ、そのような類似するユーザサーチをユーザサーチバンドルにグループ化するためにユーザサーチを分析することと、

U S B からのユーザサーチバンドルに基づいてインテントを作ることであって、各インテントは、互いに関係すると思われる 1 つまたは複数のセッションのグループである、作ることと、

ファクタを作るためにユーザサーチおよび関連情報を検討することであって、各ファクタは、サーチリザルトのセットからの特定のリザルトに関し、各ファクタは、1 つまたは

50

複数のイベントに関係し、各イベントは、照会するユーザが実行した動作に関する 1 つの情報である、検討することと、

前記リザルトのジャッジメントを作るために各リザルトについて生成された前記ファクタを受け取り、それに基づいて動作することであって、前記ジャッジメントは、前記サーチリザルトからの前記リザルトへのアクセスを決定する際に前記リザルトを前記ユーザがどのように審判したかの判定を表す、受け取ることと、

前記ファクタおよび前記ジャッジメントに基づいてメトリックを作ることであって、各メトリックは、リザルト、ユーザサーチ、またはセッションに関する尺度である、作ることと、

前記ユーザサーチバンドル、前記インテント、前記ファクタ、前記ジャッジメント、および前記メトリックを含むデータを抽出し、前記抽出されたデータを 1 つまたは複数のデータベースにフォーマットし、前記データベースをライブラリにストアし、これによって、フィードバックを提供するかレポートを生成するために前記データをレビューし、集計できるようにすることと

を含むことを特徴とする方法。

【請求項 1 1】

各クエリストリングと前記対応するサーチリザルトと関連情報とをデータウェアハウス内に正規化された形でストアすることを含み、さらに、前記データウェアハウスから前記正規化されたデータを取り出し、これを正規化し、前記正規化されたデータをデータストアにストアすることを含むことを特徴とする請求項 1 0 に記載の方法。

【請求項 1 2】

クエリストリングの類似性およびサーチリザルトの類似性のうちの少なくとも 1 つについて前記ユーザサーチを分析することを含むことを特徴とする請求項 1 0 に記載の方法。

【請求項 1 3】

各イベントは、ユーザが、特定のリザルトの選択およびクローズのうちの少なくとも 1 つを実行した時刻を含み、前記方法は、ユーザがリザルトを見た時間の長さを表す「ドエルタイム」ファクタを計算することを含み、前記ドエルタイムファクタは、対応するタイムスタンプ付きイベントによってそれぞれが表される、前記ユーザが前記リザルトを選択した時刻および前記ユーザが前記リザルトをクローズした時刻の間の差に基づくことを特徴とする請求項 1 0 に記載の方法。

【請求項 1 4】

「Accept」ジャッジメント、「Explore」ジャッジメント、および「Reject」ジャッジメントのうちの少なくとも 1 つと、前記ジャッジメントがどれほど正しい可能性があるかの信頼度を示す対応する値とを含むジャッジメントを作ることを含むことを特徴とする請求項 1 0 に記載の方法。

【請求項 1 5】

エキスプリシットジャッジメントファクタを受け取り、それに基づいてリリーバンスクラシファイヤ(RC)を生成することを含み、前記RCは、各リザルトについて生成された前記ファクタを受け取り、前記リザルトの前記ジャッジメントを作るためにこれに基づいて動作し、各エキスプリシットジャッジメントファクタは、どのファクタがどのジャッジメントを暗示するかを前記エキスプリシットジャッジメントファクタに基づいて学習できるように、対応するリザルトに関する前記ユーザからの明示的フィードバックを表すことを特徴とする請求項 1 0 に記載の方法。

【請求項 1 6】

リザルトに関して、

前記ユーザは前記リザルトをどのようにランキングしたと審判したかに関するポジションメトリックと、

前記リザルトが前記サーチリザルト内でどのように位置付けられたかに関するリリーバンスポジションメトリックと、

前記ポジションメトリックおよび前記リリーバンスポジションメトリックに基づく、前

10

20

30

40

50

記リザルトがあるべきであったところからどれほど「離れていた」かに関するミスランク
ドリザルトメトリックと

のうちの少なくとも1つを作ることを含むことを特徴とする請求項10に記載の方法。

【請求項17】

レビューされたユーザサーチバンドルに基づいて、セッションにまたがる共通のリザルト
およびセッションにまたがる共通のクエリタームを突き止めることによってセッション
間の関係値を判定し、そのような共通のリザルトが見つかった場合にストレングスオブコ
モナリティを突き止めることを含み、そのようなストレングスオブコモナリティは、2つ
のセッションが共通の目的を有することによって互いに関係する可能性がどれほどあるか
を表し、所定の閾値を超えるストレングスオブコモナリティを有するセッション対が、イ
ンテントにバンドルされることを特徴とする請求項10に記載の方法。

10

【請求項18】

前記抽出されたデータをリレーショナルデータベースにフォーマットすることを含むこ
とを特徴とする請求項10に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、検索の要求に応答するシステムなどのシステムの動作中にコンパイルされた
データを含むデータを分析するシステムおよび方法に関する。具体的には、本発明は、応
答の改善およびシステムパフォーマンスの維持を含む目的のために比較的大量のデータを
分析するシステムおよび方法に関する。

20

【背景技術】

【0002】

通常、検索エンジンなどのシステムに関して、これにアクセスするユーザは、多分ブ
ール演算子を伴う1つまたは複数の検索条件を含む検索ストリングまたはその類似物を入力
することによって検索を要求する。これに応答して、検索エンジンは、検索ストリングに
基づいて1つまたは複数のデータベースを検索し、それに基づいて検索結果のセットを生
成し、多分、ユーザがレビューできる情報または情報へのリンクのページの形で、要求す
るユーザにそのような検索結果を返す。特に後者の場合に、ユーザは、1つまたは複数の
リンクにアクセスして、特定の検索結果に関するコンテンツをレビューすることができ、
検索結果の1つまたは複数のリンクに関連するコンテンツがユーザに許容可能である場合
に、そのようなユーザは、通常、適当と思われる形で、そのような許容可能なコンテンツ
の使用に進行する。

30

【0003】

しかし、その代わりに、少なくともユーザの観点から、検索結果のコンテンツのどれも
が要求された検索を満足しないという点で、検索結果がユーザに許容可能でない場合があ
りえる。その場合に、ユーザは、新しい検索ストリングまたは前に入力した検索ストリン
グの変更を入力し、そのような新しいまたは変更された検索ストリングに基づく検索エン
ジンからの検索結果をレビューすると決定することができる。諒解されるように、そのよ
うなプロセスは、ユーザが許容可能な検索結果を突き止めるかあきらめるまで、検索セッ
ションの形で複数回繰り返すことができる。

40

【0004】

一般に、前述の検索エンジンなどのハイクオリティシステムでは、検索ストリングで示
されたユーザからの各クエリが、そのクエリに答えるコンテンツを表す検索結果に正確に
マッピングされなければならない。そのような目的は、よい検索経験を提供するために必
須であり、実際に、そのような目的を満たすことは、新しい検索セッションを持って検索
エンジンに戻る幸せな満足したユーザと、その代わりに別の検索エンジンを訪れる怒った
不満なユーザの間の差を表すことができる。

【発明の開示】

50

【発明が解決しようとする課題】

【0005】

そのような目標を達成するために、システム自体を調整または「チューニング」して、応答を改善し、システムパフォーマンスを維持しなければならないことが認識されている。諒解されるように、そのような調整は、排他的ではないが主に、システムの動作中にコンパイルされた動作データおよび他のデータに基づいて行うことができる。しかし、システムが、特に大きく、大量のトラフィックまたはその類似物を有する場合に、分析に使用可能なデータの量が、特に多くなる可能性があり、したがって、手におえないか、他の形で扱うことが困難である可能性がある。したがって、大量のデータ、特に検索システムおよびその類似物からの大量のデータを分析するシステムおよび方法の必要が存在する。

10

【課題を解決するための手段】

【0006】

前述の必要は、少なくとも部分的に、検索エンジンからのデータを分析するためにシステムが提供される本発明によって満足される。検索エンジンは、要求するユーザから受け取ったクエリストリングに基づいてサーチリザルト（検索結果）のセットを生成し、クエリストリングおよびサーチリザルトは、集合的に、ユーザサーチを含む。サーチリザルトに、少なくとも1つのリザルトが含まれ、各リザルトは、クエリストリングに関係すると思われるコンテンツの特定のアイテムを参照する。一連の関連するユーザサーチに、セッションが含まれ、検索エンジンは、各ユーザサーチおよび関連する情報をストアする。

【0007】

20

このシステムでは、ユーザサーチバンドラ（User Search Bundler、USB）が、ユーザサーチを分析して、そのようなユーザサーチのうちの類似するユーザサーチを見つけ、そのような類似するユーザサーチをユーザサーチバンドルにグループ化し、インテントプロセッサ（Intent Processor、IP）が、USBからのユーザサーチバンドルに基づいてインテントを作る。各インテントは、互いに関連すると思われる1つまたは複数のセッションのグループである。

【0008】

ファクタジェネレータ（Factor Generator、FG）が、ユーザサーチおよび関連情報を考慮して、ファクタを作り、各ファクタは、サーチリザルトのセットからの特定のリザルトに関する。各ファクタは、1つまたは複数のイベントに関し、各イベントは、照会するユーザが実行した動作に関する1つの情報である。リリーバンスクラシファイヤ（Relevance Classifier、RC）が、リザルトごとにFGによって生成されたファクタを受け取り、これに基づいて動作して、リザルトのジャッジメントを作り、ジャッジメントは、サーチリザルトからリザルトにアクセスすると判断する際にユーザがそのリザルトをどのように判断したかの判定を表す。メトリックジェネレータ（Metric Generator、MG）が、FGによって生成されたファクタおよびRCによって作られたジャッジメントに基づいてメトリックを作り、各メトリックは、リザルト、ユーザサーチ、またはセッションに関する尺度である。最後に、データシンセサイザ（data synthesizer、DS）が、USB、IP、FG、RC、およびMGによって生成されたデータを抽出し、抽出されたデータを1つまたは複数のデータベースにフォーマットし、データベースをライブラリにストアし、これによって、フィードバックを提供するかレポートを生成するためにデータをレビューし、集計することができる。

30

40

【0009】

前述の要約ならびに本発明の実施形態の次の詳細な説明は、添付図面と共に読まれる時によりよく理解される。本発明の例示のために、図面に、現在好ましい実施形態を示す。しかし、理解されるように、本発明は、図示の正確な配置および手段に限定されない。

【発明を実施するための最良の形態】

【0010】

コンピュータ環境

50

図 1 および次の議論は、本発明および/またはその一部を実施できる適切なコンピューティング環境の短い全般的な説明を提供することを意図された物である。必要ではないが、本発明を、クライアントワークステーションまたはサーバなどのコンピュータによって実行される、プログラムモジュールなどのコンピュータ実行可能命令の全般的な文脈で説明する。一般に、プログラムモジュールには、特定のタスクを実行するか特定の抽象データ型を実施する、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造体、およびその類似物が含まれる。さらに、本発明および/またはその一部を、ハンドヘルドデバイス、マルチプロセッサシステム、マイクロプロセッサベースまたはプログラマブルなコンシューマエレクトロニクス、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータ、およびその類似物を含む他のコンピュータシステム構成を用いて実践することができることを諒解されたい。本発明は、通信ネットワークを介してリンクされたりリモート処理デバイスによってタスクが実行される分散コンピューティング環境でも実践することができる。分散コンピューティング環境では、プログラムモジュールを、ローカルとリモートの両方のメモリストレージデバイスに置くことができる。

10

20

30

40

50

【0011】

図 1 からわかるように、例示的な汎用コンピューティングシステムに、処理ユニット 121、システムメモリ 122、およびシステムメモリを含むさまざまなシステムコンポーネントを処理ユニット 121 に結合するシステムバス 123 を含む、普通のパーソナルコンピュータ 120 またはその類似物が含まれる。システムバス 123 は、メモリバスまたはメモリコントローラ、周辺バス、およびさまざまなバスアーキテクチャのいずれかを使用するローカルバスを含む複数のタイプのバス構造のいずれかとすることができる。システムメモリに、読取専用メモリ (ROM) 124 およびランダムアクセスメモリ (RAM) 125 が含まれる。スタートアップ中などにパーソナルコンピュータ 120 内の要素の間で情報を転送するのを助ける基本ルーチンを含む基本入出力システム 126 (BIOS) が、ROM 124 にストアされている。

【0012】

パーソナルコンピュータ 120 に、さらに、ハードディスク (図示せず) から読み取り、これに書き込むハードディスクドライブ 127、リムーバブル磁気ディスク 129 から読み取るか、これに書き込む磁気ディスクドライブ 128、CD-ROM または他の光学媒体などのリムーバブル光学ディスク 131 から読み取るか、これに書き込む光学ディスクドライブ 130 を含めることができる。ハードディスクドライブ 127、磁気ディスクドライブ 128、および光学ディスクドライブ 130 は、それぞれハードディスクドライブインターフェース 132、磁気ディスクドライブインターフェース 133、および光学ドライブインターフェース 134 によってシステムバス 123 に接続されている。ドライブおよびそれに関連するコンピュータ可読媒体は、コンピュータ可読命令、データ構造体、プログラムモジュール、および他のデータの不揮発性ストレージをパーソナルコンピュータ 120 に提供する。

【0013】

本明細書に記載の例示的な環境では、ハードディスク、リムーバブル磁気ディスク 129、およびリムーバブル光学ディスク 131 が使用されているが、コンピュータによってアクセス可能なデータをストアできる他のタイプのコンピュータ可読媒体も、例示的なオペレーティング環境で使用できることを諒解されたい。そのような他のタイプの媒体に、磁気カセット、フラッシュメモリカード、デジタルビデオディスク、ベルヌーイカートリッジ、ランダムアクセスメモリ (RAM)、読取専用メモリ (ROM)、およびその類似物が含まれる。

【0014】

オペレーティングシステム 135、1 つまたは複数のアプリケーションプログラム 136、他のプログラムモジュール 137、およびプログラムデータ 138 を含む複数のプログラムモジュールを、ハードディスク、磁気ディスク 129、光学ディスク 131、ROM 124、または RAM 125 にストアすることができる。ユーザは、キーボード 1

40 およびポインティングデバイス142などの入力デバイスを介してパーソナルコンピュータ120にコマンドおよび情報を入力することができる。他の入力デバイス(図示せず)に、マイクロホン、ジョイスティック、ゲームパッド、衛星パラボラアンテナ、スキャナ、またはその類似物を含めることができる。上記および他の入力デバイスは、しばしば、システムバスに結合されたシリアルポートインターフェース146を介して処理ユニット121に接続されるが、パラレルポート、ゲームポート、またはuniversal serial bus(USB)などの他のインターフェースによって接続することができる。モニター147または他のタイプのディスプレイデバイスも、ビデオアダプタ148などのインターフェースを介してシステムバス123に接続されている。モニター147のほかに、パーソナルコンピュータに、通常、スピーカおよびプリンタなどの他の周辺出力デバイス(図示せず)が含まれる。図1の例示的なシステムには、ホストアダプタ155、Small Computer System Interface(SCSI)バス156、およびSCSIバス156に接続された外部ストレージデバイス162も含まれる。

10

【0015】

パーソナルコンピュータ120は、リモートコンピュータ149などの1つまたは複数のリモートコンピュータへの論理接続を使用することによってネットワーク化された環境で動作することができる。リモートコンピュータ149は、もう1つのパーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ピアデバイス、または他の一般的なネットワークノードとすることができ、リモートコンピュータ149には、通常、パーソナルコンピュータ120に関して上で説明した要素の多数またはすべてが含まれるが、メモリストレージデバイス150だけが図1に示されている。図1に示された論理接続に、ローカルエリアネットワーク(LAN)151および広域ネットワーク(WAN)152が含まれる。そのようなネットワーク環境は、オフィス、会社全体のコンピュータネットワーク、イントラネット、およびインターネットでありふれたものである。

20

【0016】

LANネットワーキング環境で使用される場合に、パーソナルコンピュータ120は、ネットワークインターフェースまたはネットワークアダプタ153を介してLAN 151に接続される。WANネットワーキング環境で使用される場合に、パーソナルコンピュータ120に、通常、インターネットなどの広域ネットワーク152を介する通信を確立する、モデム154または他の手段が含まれる。モデム154は、内蔵または外付けとすることができるが、シリアルポートインターフェース146を介してシステムバス123に接続される。ネットワーク化された環境では、パーソナルコンピュータ120に関して図示したプログラムモジュールまたはその一部を、リモートメモリストレージデバイスにストアすることができる。図示のネットワーク接続が、例示的であり、コンピュータの間の通信リンクを確立する他の手段を使用できることを諒解されたい。

30

【0017】

データ分析パイプライン

前置きとして、また用語法の問題として、本発明に関して使用できる物などの検索エンジンに関して、これにアクセスする各ユーザが、多分ブール演算子を伴う1つまたは複数の検索条件を有する検索ストリングを含むクエリを入力することによって検索を要求することを諒解されたい。それに応答して、検索エンジンは、それに基づくサーチリザルトが実際に入手可能であるものとして、そのようなサーチリザルトのセットを生成し、そのようなサーチリザルトを要求元ユーザに返す。返されるサーチリザルトに、検索要求に関連すると思われるコンテンツの特定のアイテムまたはリザルトを含めることができるが、各特定のリザルトが、その代わりに、サーチリザルト内の対応するリンクによってアクセスされる可能性がより高い。

40

【0018】

特に返されたサーチリザルトが許容可能でない場合に、ユーザは、新しい検索ストリングまたは前に入力された検索ストリングの変更のいずれかを有する別のクエリストリング

50

を入力することができ、これによって、別のクエリに基づいて、検索エンジンからのサーチリザルトの別のセットが生成される。一連の関連するクエリに、全体的な検索セッションが含まれ、うまくいけば、そのようなセッションは、ユーザがサーチリザルト内で許可可能なりザルトを突き止めた場合に終了する。

【0019】

ユーザがクエリストリングを入力し、サーチリザルトをレビューする過程に、検索エンジンまたは関連するエンティティが、そのようなセッションに関するデータを識別し、ストアすることができ、しばしば実際にそうする。具体的に言うと、セッションを識別することに加えて、検索エンジンまたは関連するエンティティ（以下では「検索アナライザ」）は、とりわけ、セッションの各クエリストリングおよび各クエリストリングについて返されたサーチリザルトを識別し、ストアすることができる。さらに、検索アナライザは、サーチリザルトの各返されたリザルトのリンクごとに、とりわけ、ユーザがそれによってリザルトにアクセスしたかどうかと、ユーザがそのリザルトをレビューして過ごした時間の長さを識別し、ストアすることができる。したがって、諒解されなければならないように、検索アナライザまたはそのような情報を有する別のエンティティは、セッションから返されたサーチリザルトの各セットに関してユーザがどれほど満足または「幸せ」であるかの、定量的でないとしても定性的な尺度を展開することができる。

10

【0020】

前述の機能を実行する検索アナライザが、当業者に既知であるか明白であらねばならず、したがって、詳細に述べる必要がないことに留意されたい。したがって、任意の適当な検索アナライザを、本発明に関連して使用することができる。

20

【0021】

セッションを形成するクエリストリングのグループの例として、次の複数のクエリストリング「Cars」、「Fords」、「Ford Edsel」、「Cheap vacations」、および「London trip prices」を検討されたい。各クエリストリングが、検索の場合にユーザが入力した実際のテキストであることを諒解されたい。さらに、各クエリストリングは、リザルトへのリンクを有するサーチリザルトのセットを生成し、検索アナライザによって、各リンクが選択されたかどうか、リザルトを見て過ごされたドエルタイム、リザルトに関して行われたスクロールまたは他のアクション、および他の類似するユーザ挙動などの関連データを関連付けられることができる。

30

【0022】

前述の情報のすべてに基づいて、ユーザが2つの異なる意図を有したように見えるという理由から、最初の3つのクエリストリング（すなわち、「Cars」、「Fords」、および「Ford Edsel」）が、第1セッションの一部であり、最後の2つのクエリストリング（すなわち、「Cheap vacations」および「London trip prices」）が、第2セッションの一部であることを理解されたい。すなわち、セッションが、単一のユーザによって時間的に連続して行われた同一の意図または目的を有するクエリストリングおよびそれに対応するサーチリザルト（すなわち、ユーザサーチ）のグループであることを諒解されたい。

【0023】

図2に移ると、本発明が、検索リクエストから要求された検索を実行し、適当な検索結果によってこれに応答する検索エンジン10などのシステムを仮定することがわかる。そのような検索エンジン10は、もちろん、本発明の趣旨および範囲から逸脱せずに、すべての適当な検索エンジンとすることができる。そのような検索エンジン10の動作は、一般に、当業者に既知であるか、明白であらねばならず、したがって、他の形で示される場合を除いて、本明細書で詳細に説明する必要はない。より一般的に、このシステムを、やはり本発明の趣旨および範囲から逸脱せずに、検索エンジン10以外のシステムとすることができる。

40

【0024】

やはり図2からわかるように、データウェアハウス12が、システム/検索エンジン1

50

0と共に維持されることが仮定される。諒解されるように、データウェアハウス12は、多分検索アナライザによって作られた前述のデータを含むがこれに限定されない、検索エンジン10の動作に関するデータをストアする。データウェアハウス12によってストアできるそのようなデータは、もちろん、本発明の趣旨および範囲から逸脱せずに他の適当なデータとすることができ、データウェアハウス12およびその編成を、同様に、本発明の趣旨および範囲から逸脱せずに任意の適当なデータウェアハウスおよび編成とすることができる。

【0025】

諒解されるように、データウェアハウス12内のデータを、無限に増やすことができ、あるいは、周期的に満了させ、削除することができる。したがって、本発明では、パイプライン14を使用して、そのようなデータを分析し、下で詳細に示すように、すべてのデータまたはデータのスライディングウィンドウを含む、データウェアハウス12内のデータの任意の部分の分析することができる。理解されるように、パイプライン14によって実行されるそのようなデータ分析に、複数のコンポーネントを含めることができ、このコンポーネントのそれぞれは、ウェアハウスからデータの一部を受け取り、これを操作する。本発明の一実施形態では、パイプライン14の出力が、ライブラリ16に置かれ、出力データが、リレーショナルデータベース、多次元テーブル、またはその類似物などの形に編成される。したがって、適当なコントロールを用いて、ウェブブラウザまたはその類似物を使うユーザが、そのような編成されたデータを見ることができ、多分さらにこれ进行分析することができる。

10

20

【0026】

諒解されるように、パイプライン14によって実行される処理は、検索エンジン10がどのように使用されているかに関する統計を提供する。さらに、本発明の一実施形態では、そのように実行される処理が、検索エンジン10がそのユーザによって使用されているパターン（特にそのようなパターンが真に貴重な情報を提供できる限り）を識別するのに使用される。

【0027】

特に、ウェアハウス12内のデータの正規化によって新しいデータを効率的な形で追加できるようになることを諒解できる限り、ウェアハウス12内のデータを非常に正規化することに留意されたい。しかし、そのような正規化されたデータは、分析の前に非正規化されなければならない。したがって、データパイプライン14に、その入力としてデノーマライザ（*de-normalizer*）18を含めることができる。そのようなデノーマライザ18は、任意の適当なデノーマライザとすることができ、任意の適当な形で動作することができる。そのようなデノーマライザ18の動作は、当業者に既知であるか明白であるはずであり、したがって、本明細書で詳細に示す必要はない。

30

【0028】

本発明のパイプライン14で使用するデノーマライザ18は、ウェアハウス12内のデータを非正規化し、後の処理のために非正規化されたデータのバッチを作成することもできる。バッチ化は、任意の適当な判断基準に従って実行することができる。たとえば、バッチを、データの以前のバッチ以降に到着したすべてのデータとすることができ、ここで、バッチを、デノーマライザ18が毎日1回動作する場合に1日のデータとすることができ、同様に、動作するのは週1回であるが、その週の1時間ごとのバッチを生成するようにデノーマライザ18をセットすることもできる。

40

【0029】

理解されるように、デノーマライザ18によって出力されたバッチは、パイプライン14の他のコンポーネントによるさらなる処理のためにデータストア20にストアされる。そのようなデータストア20およびその中のバッチの編成は、本発明の趣旨および範囲から逸脱せずに、任意の適当なストアおよび編成とすることができ、図2からわかるように、データストア20内のデータのバッチを使用するコンポーネントに、ユーザサーチバンドラ（*USB*）、リリーバンスプロセッサ（*RP*）、インテントプロセッサ（*IP*）、

50

データシンセサイザ（ＤＳ）、およびその類似物を含めることができる。

【００３０】

データパイプライン１４のユーザサーチバンドラ（ＵＳＢ）は、ユーザサーチを分析し、このユーザサーチのそれぞれは、検索エンジン１０に対してクエリを呼び出したユーザのインスタンスであり、とりわけ、クエリストリングおよび返されたサーチリザルトの両方に関する情報を含む。同一のクエリストリングを呼び出す２つのユーザは、２つのユーザサーチを作り、同一のクエリストリングを２回呼び出した単一のユーザも、２つのユーザサーチをもたらす。本発明の一実施形態では、ＵＳＢは、「そっくりに見え」したがって類似するユーザサーチを見つけ、そのような類似するユーザサーチをユーザサーチバンドルにグループ化することを試みる。ＵＳＢは、本発明の趣旨および範囲から逸脱せずに、任意の適当な形で動作することができる。ＵＳＢを動作させる方法は、当業者に既知であるか明白であるはずであり、したがって、本明細書で詳細に示す必要はない。

10

【００３１】

たとえば、ＵＳＢは、クエリストリングを分析する「ルックアライク（そっくりに見える）」アルゴリズムを使用することができる。たとえば、２つのユーザのそれぞれが、「Wilkes-Barre, PA」の検索を呼び出す場合に、これらのクエリストリングが、そっくりに見える。より興味深いことに、あるユーザが、「duck」を検索し、別のユーザが、「ducks」を検索する（一方は複数、一方は単数）場合に、この２つのクエリストリングは、アルゴリズムがステミング（stemming）、大文字小文字正規化、および他の類似するクエリ比較技法を使用する限り、「そっくりに見える」。

20

【００３２】

しかし、ＵＳＢは、類似性に関するクエリストリングの分析に制限されない。その代わりに、本発明の一実施形態では、ＵＳＢが、返されたサーチリザルト分析アルゴリズムを使用することもできる。たとえば、ＵＳＢは、返されたサーチリザルト分析アルゴリズムに基づいて、対応するクエリストリングが全く似ていない場合であっても、２つのユーザサーチが似ており、それぞれがサーチリザルトの類似するセットを作ると結論することができる。それに加えてまたはその代わりに、ＵＳＢは、コンテンツ満足分析アルゴリズムに基づいて、やはり対応するクエリストリングが全く似ていない場合であっても、２つのユーザサーチのそれぞれが、ユーザが満足と判定した特定のコンテンツまたはリザルトを作る場合に、その２つのユーザサーチが似ていると結論することができる。同様に、ＵＳＢは、コンテンツ不満足分析アルゴリズムに基づいて、２つのユーザサーチのそれぞれが、ユーザが不満と判定した特定のリザルトを作る場合に、２つのユーザサーチが似ていると結論することができる。

30

【００３３】

ＵＳＢの出力は、バンドルされたユーザサーチのデータベースとしてライブラリ１６にストアすることができる。その代わりに、そのような出力を、データストア２０にストアして、パイプライン１４の他のコンポーネントから使用可能にすることができる。

【００３４】

本発明のリリーバンスプロセッサ（ＲＰ）は、ファクタジェネレータ（ＦＧ）、リリーバンスクラシファイヤ（ＲＣ）、およびメトリックジェネレータ（ＭＧ）に副分割することができる。ただし、ＲＰのそのように副分割された要素が、その代わりに、ＲＰの下でグループ化されずに、本発明のパイプライン１４に個別に存在できることに留意されたい。

40

【００３５】

ファクタジェネレータ（ＦＧ）は、ユーザサーチに関する情報を考慮し、ファクタを作り、このファクタは、リリーバンスクラシファイヤ（ＲＣ）への入力として適用される。各ファクタは、サーチリザルトのセットからの特定のリザルトに関する。ファクタは、必ずではないがしばしば、複数のイベントを考慮することによって作られ、各イベントは、通常は、照会するユーザが実行した動作に関する１つの情報である。たとえば、特定のリザルトを選択することを、イベントとすることができ、リザルトの表示、クローズ、印刷

50

、特定のリストへの追加、再オープン、およびその類似物をイベントとすることができる。イベントに、提供される場合に明示的なユーザフィードバックも含めることができる。

【0036】

本発明の一実施形態では、F Gが、イベントに基づいて値を計算する。たとえば、F Gは、ユーザがリザルトを見た時間の長さを表すファクタとして「ドエルタイム」を計算することができる、対応するタイムスタンプ付きのイベントによってそれぞれが表される、ユーザがリザルトを見るのを開始した時刻とそれを停止した時刻の間の差に基づいてそれを行うことができる。ただし、他の適当なタイムスタンプ付きイベントによって表される、ユーザがリザルトを見ることに割り込んだかどうかを含む、ドエルタイムファクタを生成する場合に、他のイベントをF Gによって考慮することもできることに留意されたい。

10

【0037】

F Gによって生成されるもう1つのファクタを、ユーザがリザルトをお気に入りリストなどの特定のリストに追加したかどうかに関するファクタとすることができる。その場合に、これを、そのような「お気に入り」ファクタに真の値をセットすることとすることができる。諒解されるように、イベントに基づくファクタのタイプは、本発明の趣旨および範囲から逸脱せずに、任意の適当なファクタとすることができる。そのようなファクタは、当業者に既知であるか明白であるはずであり、したがって、本明細書で詳細に示す必要はない。

【0038】

ファクタが、イベント以外のことから来る場合があることに留意されたい。たとえば、ファクタを、リザルトをオーサリングしたユーザの識別子とすることができる。したがって、ファクタを、リザルトのプロパティにすることができる。他のファクタに、クエリストリングのロケール（たとえば、米国英語、カナダ英語、ブラジルポルトガル語）を含めることができ、ユーザサーチ内の情報から導出することができる。より一般的に、ファクタは、本発明の趣旨および範囲から逸脱せずに、使用可能な任意の情報からF Gまたは別のコンポーネントによって作ることができる。

20

【0039】

F Gおよび多分他所によって生成されたファクタは、ライブラリ16内の適当なデータベースにストアすることができる。その代わりに、そのようなファクタを、パイプライン14の他のコンポーネントから使用可能なデータストア20にストアすることができ、あるいは、そのようなコンポーネントに直接に配送することができる。

30

【0040】

リリーバンスプロセッサ(RP)のリリーバンスクラシファイヤ(RC)は、F Gおよび多分他所によって生成されたファクタを受け取る。RCは、それへの入力として受け取られたファクタに基づいて動作する、マシン生成の判断木である。RPによって開始される時に、RCは、データストア20または他所からリザルトのファクタを読み取り、リザルトのジャッジメントを作る。そのようなジャッジメントは、一般に、サーチリザルトのセットからのリザルトにアクセスすることを決定する際にユーザがそのリザルトをどのように審判したかの判定である。

【0041】

ジャッジメントは、本発明の趣旨および範囲から逸脱せずに、任意の適当な審判システムに従って表すことができる。たとえば、ジャッジメントを、数値のまたは文字のグレードスコアとすることができる、「Accept」（すなわち、ユーザがリザルトに満足した）、「Explore」（すなわち、ユーザがリザルトに満足でも不満足でもなかった）、および「Reject」（すなわち、ユーザがリザルトに満足しなかった）またはその類似物のうちの1つとすることができる。さらに、特定のリザルトのジャッジメントに、RCによって決定される、ジャッジメントが正しい可能性がどれほどかの信頼度を示す値を含めることができる。

40

【0042】

RCおよび多分他所によって生成されたジャッジメントを、ライブラリ16内の適当な

50

データベースにストアすることができる。その代わりに、そのようなファクタを、パイプライン 14 の他のコンポーネントから使用可能なデータストア 20 にストアすることができる。あるいは、そのようなコンポーネントに直接に配送することができる。

【0043】

RC の判断木をマシン生成するために、RP が、エクスプリシットジャッジメントファクタを有する各リザルトを考慮する場合がある。エクスプリシットジャッジメントは、リザルトに関するユーザからの明示的なフィードバックを表すタイプのファクタである。エクスプリシットジャッジメントファクタを有する各リザルトをとり、そのリザルトに関する他の使用可能なファクタを分析することによって、RP のリリーバンスクラシフィカトリーナ (RCT) または他所が、どのファクタがどのジャッジメントを暗示するかを「学習」でき、それに基づいて RC を構築することができる。どのファクタが暗示するかの学習およびそれからの RC の構築は、本発明の趣旨および範囲から逸脱せずに、任意の適当な形で行うことができる。それを行うことは、当業者に既知であるか明白であるはずであり、したがって、本明細書で詳細に示す必要はない。

10

【0044】

RP のメトリックジェネレータ (MG) は、メトリックを作り、このメトリックは、リザルト、ユーザサーチ、セッション、またはその類似物に関する尺度である。一般に、MG は、FG および多分他所によって生成されたファクタ、RC によって作られたジャッジメント、および使用可能な他の関連情報に基づいてそのようなメトリックを作る。そのようなメトリックの生成は、本発明の趣旨および範囲から逸脱せずに、任意の適当な形で行うことができる。それを行うことは、当業者に既知であるか明白であるはずであり、したがって、本明細書で詳細に示す必要はない。

20

【0045】

MG は、リザルトに関する次のメトリックを作ることができる。

【0046】

- Accept、Explore、Reject それぞれが、リザルトのジャッジメントから導出され、信頼度を含めることができる。

【0047】

- ミスランクドリザルト リザルトがサーチリザルト内でどのように位置付けられたか、およびユーザがリザルトをどのようにランキングしたと審判したかに基づく、リザルトがあるべきであったところからどれほど「遠い」かの尺度。

30

【0048】

- ポジション ユーザがリザルトをどのようにランキングしたと審判したか。

【0049】

- リリーバンスポジション リザルトがサーチリザルト内でどのように位置付けられたか。

【0050】

MG は、ユーザサーチに関する次のメトリックを作ることができる。

【0051】

- リザルトセットランキングスコア リザルトごとに作られるミスランクドリザルトメトリックに似ているが、ユーザサーチのすべてのリザルトに関する。この値は、リザルトセットがそうあるべきであった物とどれほど離れているかを反映することを試みる。

40

【0052】

- Accept、Explore、Reject の要約 それぞれ、ユーザサーチ内のリザルトごとの、すべての Accept メトリック、Explore メトリック、および Reject メトリックの要約。

【0053】

MG は、セッションに関する次のメトリックを作ることができる。

【0054】

- コンテンツクオンティティ セッションが、ユーザが探している使用可能な情報の欠

50

如を示したかどうか。

【0055】

- インテントデターミネーション セッションが、ユーザが探している物を判定できないことを示したかどうか。

【0056】

M Gおよび多分他所によって生成されたメトリックを、ライブラリ16内の適当なデータベースにストアすることができる。その代わりに、そのようなメトリックを、パイプライン14の他のコンポーネントから使用可能なデータストア20にストアすることができる、あるいは、そのようなコンポーネントに直接に配送することができる。

【0057】

パイプライン14のインテントプロセッサ(I P)は、インテントを作り、各インテントは、ユーザが各インスタンスで同一のサーチリザルトを探しているので互いに関連すると思われる1つまたは複数のセッションのグループである。すなわち、インテントは、共通の目的を共有するセッションのグループである。

【0058】

I Pは、各セッション、各セッションの各ユーザサーチ、および各ユーザサーチの各リザルトを考慮することによってインテントを作る。本発明の一実施形態では、共通のリザルトを有するセッションが関係する可能性が高いことを念頭において、A c c e p tのジャッジメントを有するリザルトだけを調査する。しかし、他のジャッジメントを有するリザルトも、本発明の趣旨および範囲から逸脱せずに使用することができる。

【0059】

一般に、I Pは、レビューされたユーザサーチバンドルに基づいて、セッションにまたがる共通のリザルトおよびセッションにまたがる共通のクエリタームを突き止め、見つかった場合にストレングスオブコモナリティを確かめることによって、セッションの間の関係値を判定する。そのようなストレングスは、2つのセッションが、共通の目的を有することによって互いに関係する可能性がどれほどあるかを表す。ある決定された閾値を超えるストレングスを有するセッション対を、I Pによってインテントにバンドルすることができる。

【0060】

I Pおよび多分他所によって生成されたインテントを、ライブラリ16内の適当なデータベースにストアすることができる。その代わりに、そのようなインテントを、パイプライン14の他のコンポーネントから使用可能なデータストア20にストアすることができる、あるいは、そのようなコンポーネントに直接に配送することができる。

【0061】

最後に、これまでに作られたすべてのデータと共に、パイプライン14には、データストア20、ライブラリ16、または他所からそのようなデータを抽出し、そのようなデータを、ライブラリ16にストアされる1つまたは複数のデータベースにフォーマットするデータシンセサイザ(D S)が含まれる。そのようなフォーマットは、特に異なるコンポーネントがデータを異なるテーブル、データベース、またはその類似物に書き込んだ場合に必要である。たとえば、U S Bが、各ユーザサーチに関するデータをユーザサーチバンドルドテーブルに追加した場合があり、M Gが、データをメトリックテーブルに追加した場合がある。したがって、D Sは、そのようなテーブルおよび他者を、そこからそのようなデータの検索、集計、提示、およびその類似物を行える意味のある形に適当に接合する。

【0062】

D Sをプログラムして、データの一部だけを適当にフォーマットできることに留意されたい。たとえば、D Sは、ストアされたデータの一部が有用でないと思われる場合に、そのようなデータを省略することができる。その代わりに、D Sが、同一のデータを複数のテーブルにコピーするか、データを複数のテーブルに分割することができる。より一般的に、本発明の趣旨および範囲から逸脱せずに、D Sを、パイプライン14からのデータ

10

20

30

40

50

を任意の適当なフォーマットにフォーマットするようにプログラムすることができる。

【0063】

本発明の一実施形態で、DSは、データを、SQLデータベースまたはその類似物などのリレーショナルデータベースにフォーマットする。したがって、データは、中央のファクトテーブルおよびそれを囲むディメンションテーブルを有する「スター」形で表すことができる。諒解されるように、それを行うことによって、ファクトテーブルおよびディメンションテーブルを、OLAPキューブまたはその類似物など、検索エンジン10の管理者または似たような者が異なる判断基準に基づいてデータを効率的に集計できるようにするデータキューブにすることができる。たとえば、そのような管理者または似たような者は、キューブに尋ねて、あるデータ範囲に存在するすべてのセッションのインテントデータミネーションメトリックの平均値を見つけ、これをインテントによって集計することができる。同様に、キューブに対してレポートを実行させることができ、それからの結果を、そのような管理者または似たような者に転送することができる。もちろん、そのような管理者または似たような者は、そのようなキューブに対してほとんどすべての他の適当なクエリを提案することもでき、あるいは、そのようなキューブに対して実行されたすべての他のレポートを受け取ることができる。

10

【0064】

結論

本発明を、検索エンジン10を含むまたは他の形のほとんどすべての適当なシステムからのデータの分析に関して実践することができる。諒解されるように、本明細書に記載の本発明を用いると、応答を改善し、システムパフォーマンスを維持し、システムが満足に動作することを他の形で保証するように、システムを調整または「チューニング」することを含む任意の目的のために、システムのデータ表現を分析することができる。

20

【0065】

本発明に関連して実行される処理をもたらすのに必要なプログラミングは、比較的単純であり、当業者に明白であるはずである。したがって、そのようなプログラミングは、本明細書に添付しない。任意の特定のプログラミングを使用して、本発明の趣旨および範囲から逸脱せずに、本発明を実現することができる。

【0066】

前述の説明で、本発明には、大量のデータ、特に検索エンジン10またはその類似物からの大量のデータを分析する新しい有用な装置が含まれることがわかる。そのような装置は、検索エンジン10が特に大きく、大量のトラフィックまたはその類似物を有し、分析に使用可能なデータの量が特に大きく、したがって手におえないか、他の形で扱うことが困難である場合に、特に有用である。

30

【0067】

上で説明した実施形態の発明的概念から逸脱せずに、上で説明した実施形態に変更を行うことができることを諒解されたい。一般に、したがって、本発明が、開示された特定の実施形態に限定されず、請求項によって定義される本発明の趣旨および範囲に含まれる変更を含むことが意図されていることを理解されたい。

【図面の簡単な説明】

40

【0068】

【図1】本発明の態様および/またはその一部を組み込むことができる汎用コンピュータシステムを表すブロック図である。

【図2】本発明の実施形態による、検索エンジンなどの大きいシステムからのデータを分析するデータ分析パイプラインを示すブロック図である。

【符号の説明】

【0069】

120 コンピュータ

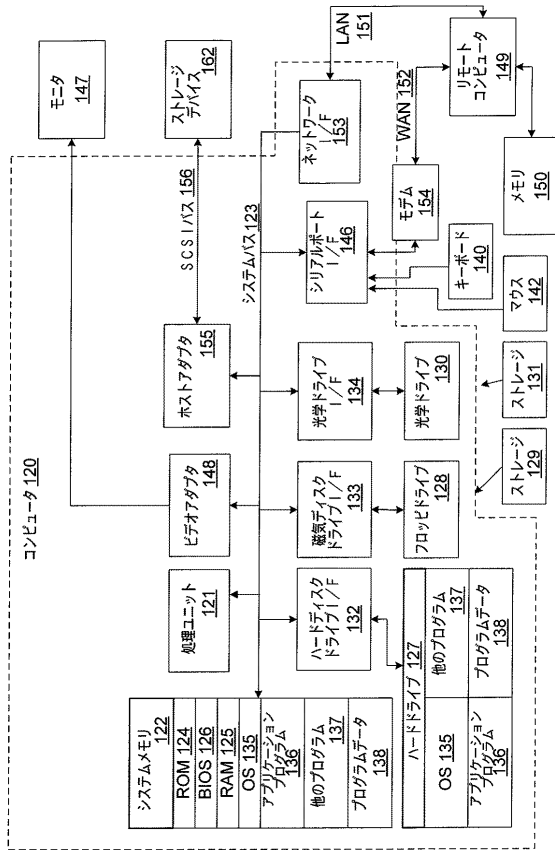
122 システムメモリ

124 ROM

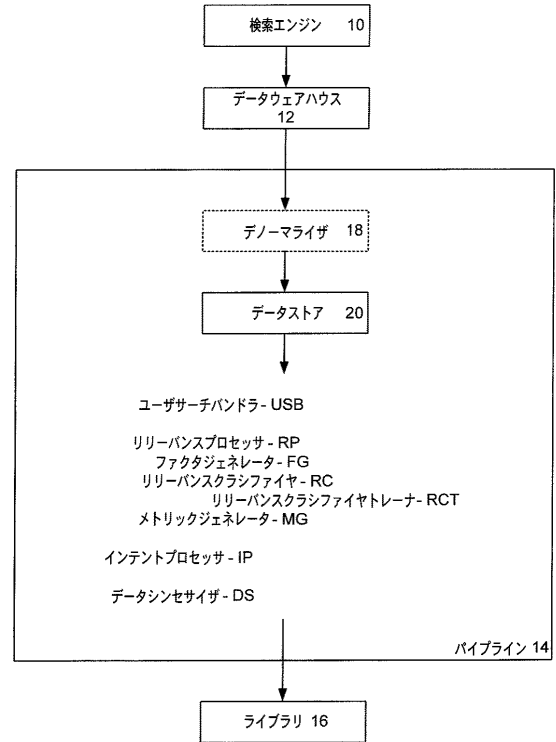
50

1 2 6	B I O S	
1 2 5	R A M	
1 3 5	O S	
1 3 6	アプリケーションプログラム	
1 3 7	他のプログラム	
1 3 8	プログラムデータ	
1 2 1	処理ユニット	
1 4 8	ビデオアダプタ	
1 5 5	ホストアダプタ	
1 4 7	モニタ	10
1 5 6	S C S I バス	
1 6 2	ストレージデバイス	
1 2 3	システムバス	
1 3 2	ハードディスクドライブ I / F	
1 3 3	磁気ディスクドライブ I / F	
1 3 4	光学ドライブ I / F	
1 4 6	シリアルポート I / F	
1 5 3	ネットワーク I / F	
1 5 1	L A N	
1 2 7	ハードドライブ	20
1 3 5	O S	
1 3 6	アプリケーションプログラム	
1 3 7	他のプログラム	
1 3 8	プログラムデータ	
1 2 8	フロッピドライブ	
1 3 0	光学ドライブ	
1 2 9	ストレージ	
1 3 1	ストレージ	
1 4 2	マウス	
1 4 0	キーボード	30
1 5 4	モデム	
1 5 0	メモリ	
1 5 2	W A N	
1 4 9	リモートコンピュータ	

【図 1】



【図 2】



フロントページの続き

- (72)発明者 フレデリック エイチ ベア ジュニア
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 ジェームス シー・フィンガー
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 ジェニファー ジェー・マースマン
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 クルディーブ カーナワット
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 マーク ビー・マイドランド
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 ポール エム マロレプシー
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 タケシ シミズ
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 トーマス ディー・ホワイト
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内
- (72)発明者 イン チャン
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ
イクロソフト コーポレーション内

F ターム(参考) 5B075 NR12 PP23 QM08