

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5021640号
(P5021640)

(45) 発行日 平成24年9月12日 (2012.9.12)

(24) 登録日 平成24年6月22日 (2012.6.22)

(51) Int.Cl.	F I
G06F 17/30 (2006.01)	G06F 17/30 220B
G06F 17/21 (2006.01)	G06F 17/21 570R
G06F 13/00 (2006.01)	G06F 17/30 170Z
	G06F 17/21 590E
	G06F 17/30 220C
請求項の数 17 (全 23 頁) 最終頁に続く	

(21) 出願番号	特願2008-520267 (P2008-520267)	(73) 特許権者	500046438
(86) (22) 出願日	平成18年6月27日 (2006.6.27)		マイクロソフト コーポレーション
(65) 公表番号	特表2009-500747 (P2009-500747A)		アメリカ合衆国 ワシントン州 9805
(43) 公表日	平成21年1月8日 (2009.1.8)		2-6399 レッドモンド ワン マイ
(86) 国際出願番号	PCT/US2006/024847		クロソフト ウェイ
(87) 国際公開番号	W02007/005382	(74) 代理人	100077481
(87) 国際公開日	平成19年1月11日 (2007.1.11)		弁理士 谷 義一
審査請求日	平成21年6月17日 (2009.6.17)	(74) 代理人	100088915
(31) 優先権主張番号	11/172, 121		弁理士 阿部 和夫
(32) 優先日	平成17年6月29日 (2005.6.29)	(72) 発明者	スーザン ティー. デュマス
(33) 優先権主張国	米国 (US)		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 インターナショナル パテンツ内
最終頁に続く			

(54) 【発明の名称】 ユーザのアクティビティ、アテンション、および関心事のデータ活用手段の検知、格納、索引作成、および検索

(57) 【特許請求の範囲】

【請求項1】

コンピュータが実施するデータ操作システムであって、
前記システムは、
プロセッサと、
情報を提示するためのディスプレイ装置と、
前記プロセッサに動作可能に結合され、コンピュータ実行可能命令を格納するコンピュータ読み取り可能な記録媒体であって、該コンピュータ実行可能命令は該プロセッサによって実行される場合に複数のコンポーネントを実行する、コンピュータ読み取り可能な記録媒体と

を備え、

前記複数のコンポーネントは、

1つまたは複数のデータ項目を1つまたは複数のタグに自動的に関連付けるタグ付けコンポーネントであって、該関連付けられたタグは、重み付けられて、少なくとも1つのユーザによる前記データ項目の相互作用の量または性質、あるいは該ユーザによる前記データ項目の関心を指し示し、かつ少なくとも1つのユーザの、前記データ項目または関連付けられたデータ項目との後続の相互作用に回答して前記少なくとも1つのユーザに対して前記ディスプレイ装置を介してデータの提示を促進する、タグ付けコンポーネントと、

前記データ項目との少なくとも1つのユーザの相互作用に少なくとも部分的に基づいて、前記プロセッサを介して前記データ項目を処理してデータ項目のサブセットを決定する

操作ツールであって、前記データ項目に対する関心または相互作用を指し示すアノテーションの状態に依存して、検索に対して前記データ項目のサブセットを異なるように重み付けする索引作成手続きを行い、該索引は少なくとも監視されるユーザアクティビティに基づいて作成される、操作ツールと

を含み、

前記コンピュータ読み取り可能な記録媒体は、前記データ、前記関連付けられたタグ、または該タグの重み付けのうちの少なくとも一つを保持することを特徴とするシステム。

【請求項 2】

前記データ項目は、文書、ファイル、電子メールメッセージ、カレンダー予約、ウェブページ、前記データ項目内のサブセクション、または項目間の抽象概念を含むことを特徴とする請求項 1 に記載のシステム。

10

【請求項 3】

前記 1 つまたは複数のタグのうち第 1 のタグは、前記少なくとも 1 つのユーザが最後に前記関連付けられたデータ項目にアクセスした位置を表すか、あるいは前記少なくとも 1 つのユーザが前記関連付けられたデータ項目に対してアクセスまたは相互作用した回数の位置情報履歴を表すことを特徴とする請求項 1 に記載のシステム。

【請求項 4】

前記 1 つまたは複数のタグのうち第 1 のタグは、前記少なくとも 1 つのユーザが最後に前記関連付けられたデータ項目にアクセスした時間、前記関連付けられたデータ項目がアクセスされた総数を表すか、過去にさかのぼる期間内で前記関連付けられたデータ項目がアクセスされた頻度を表すか、または 1 つまたは複数の任意に指定した期間内で前記関連付けられたデータ項目がアクセスされた頻度を表すことを特徴とする請求項 1 に記載のシステム。

20

【請求項 5】

経時的なアクセス頻度の高次統計量を符号化するコンポーネントをさらに含むことを特徴とする請求項 1 に記載のシステム。

【請求項 6】

少なくとも 1 つのユーザに対して 1 つまたは複数のタグの機能に基づいて前記 1 つまたは複数のデータ項目のうちの少なくとも 1 つのデータ項目を検索可能とするビューワ、少なくとも 1 つのユーザに対して 1 つまたは複数のタグの機能に基づいて検索したデータ項目をソートまたはフィルタ可能とするビューワ、または 1 つまたは複数のタグの機能に基づいて前記 1 つまたは複数のデータ項目のうち検索したデータ項目を提示するビューワをさらに含むことを特徴とする請求項 1 に記載のシステム。

30

【請求項 7】

データ項目に対するユーザアクティビティの分析のために、コンピュータが実施する方法であって、

前記方法は、

プロセッサを使用して、コンピュータ読み取り可能な記録媒体に格納された、複数の動作を実行するためのコンピュータ実行可能命令を実行するステップを含み、

前記複数の動作は、

40

前記データ項目とのユーザの相互作用をモニタするステップと、

前記データ項目に対するユーザの相互作用の量または性質を反映するデータ項目を自動的にタグ付けするステップであって、該タグは相互作用の度合いを反映する関連付けられた重み付けを有する、ステップと、

前記タグを使用して、当該データまたはそのサブセットに関連する将来のユーザアクティビティに従って前記データ項目をデータ項目のサブセットにさらに処理するステップと、

前記データ項目に対する関心または相互作用を指し示すアノテーションの状態に依存して、検索に対してデータ項目のサブコンポーネントを異なるように重み付けする索引作成手続きを行うステップであって、該索引は少なくとも監視されるユーザアクティビティ

50

に基づいて作成される、ステップと、

将来のユーザアクティビティ、タグ、またはタグの重み付けのうちの少なくとも一つに基づいて当該データまたはそのサブセットをユーザに対して提示するステップとを含むことを特徴とする方法。

【請求項 8】

前記データ項目に関連付けられたアテンショナルなアノテーション内部のデータを、独立したデータベースに格納するステップ、または前記データ項目に埋め込まれるデータ構造内部に格納するステップをさらに含むことを特徴とする請求項 7 に記載の方法。

【請求項 9】

前記データ項目は、コンピュータ読み取り可能な記録媒体を構成することを特徴とする請求項 8 に記載の方法。

10

【請求項 10】

前記コンピュータ読み取り可能な記録媒体は、テキスト、グラフィック、および関連するデータコンポーネントを含むデータ項目であることを特徴とする請求項 9 に記載の方法。

【請求項 11】

前記索引作成手続きは、前記データ項目に対する関心または相互作用を指し示すアノテーションの状態に依存して、データ項目内の情報を見過ごすかまたは削除することを特徴とする請求項 7 に記載の方法。

【請求項 12】

20

コンピュータユーザがアテンションを向けないまたは相互作用しないコンポーネントを除去することで圧縮されるインデックスをさらに含むことを特徴とする請求項 7 に記載の方法。

【請求項 13】

データ項目のコンポーネントに対するアテンションまたは相互作用がより少ないコンポーネントを除去することで圧縮されるインデックスをさらに含むことを特徴とする請求項 12 に記載の方法。

【請求項 14】

検索に使用される順位付け点数を提供し、ユーザがアテンションを向けるデータ項目または相互作用するデータ項目のセクションに現れる単語またはオブジェクトにより多くの重み付けを生成するステップをさらに含むことを特徴とする請求項 12 に記載の方法。

30

【請求項 15】

アテンショナルなアノテーションを使用して、過去および/または現在にアテンションを向ける領域または相互作用する領域に基づいてクエリを自動的または半自動的に作成するステップ、またはアテンショナルなアノテーションを使用して、過去および/または現在にアテンションを向ける項目または相互作用する項目に対する差別的なアクセスまたは差別的な表示を提供するステップをさらに含むことを特徴とする請求項 12 に記載の方法。

【請求項 16】

データ項目に対するユーザアクティビティを分析するための、コンピュータ実施システムであって、

40

前記システムは、

プロセッサと、

ユーザインタフェースを提示するディスプレイと、

該プロセッサによって実行される場合に複数のコンポーネントを実行するコンピュータ実行可能命令を格納するコンピュータ読み取り可能な記録媒体と

を備え、

前記複数のコンポーネントは、

1 つまたは複数のデータ項目に関してユーザアクティビティを前記プロセッサによって決定する手段と、

50

前記ユーザアクティビティに少なくとも部分的に基づいて前記データ項目をタグ付けする手段であって、該タグは、ユーザによる前記データ項目との相互作用の度合いを反映する関連付けられた重み付けを有する、手段と、

前記データ項目に関連付けられた、前記タグに部分的に、および該タグの重み付けのレベルに基づいて前記データ項目をコンピュータ読み取り可能な記録媒体に格納する手段と

、
前記プロセッサによって前記データ項目およびその関連付けられた重み付けタグを処理する手段であって、該処理は、前記データ項目に対する関心または相互作用を指し示すアノテーションの状態に依存して、前記データ項目および該データ項目のサブセットを検索のために異なるように重み付けする索引作成手続きを行うことを含み、該索引は少なくとも監視されるユーザアクティビティに基づいて作成される、手段と、

少なくとも1つのユーザの、前記データ項目または関連付けられたデータ項目との後続の相互作用に基づいて、前記少なくとも1つのユーザに対して前記ユーザインタフェースを介して前記データ項目を提示する手段であって、該データ項目は前記重み付けされたタグにしたがって提示される、手段と

を備えることを特徴とするシステム。

【請求項17】

前記データ項目に対するアテンショナルなアノテーションを符号化する手段と

前記データ項目内部のサブコンポーネントに対するアテンショナルなアノテーションを符号化する手段と

をさらに備え、

該アノテーションは、前記サブコンポーネントの各々のポイントまたはインジケーションと、その結果得られる、前記サブコンポーネントに対して受けたユーザのアノテーションとを取得することを特徴とする請求項16に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数のデータ型上でのデータ処理を可能とする様々なコンポーネント及びプロセスを含むシステムに関する。より詳細には、様々なコンポーネントおよびプロセスを提供して、複数のデータ型でのデータ操作を可能とし、ユーザアクティビティ、アノテーション、関心事、位置情報、またはデータと他の相互作用の履歴の態様を決定して、情報の記憶とアクセスを強化するシステムに関する。

【背景技術】

【0002】

コンピュータプラットフォームは、大規模且つ多様なデータセットを格納し、かつ処理する多数のツールを提供する。これらにはワードプロセッシングツール、データプレゼンテーションツール、コンピュータ支援グラフィックツール、電子メール処理ツール、カレンダーおよびスケジューリングツール、並びに多数のデータベース操作ツールが含まれる。このプラットフォーム上のデータの様々な利用方法を前提として、ややコンテンツ中心型のアプリケーションが長期に渡って開発されてきた。換言すると、データをコンピュータのデータベースに格納すると、実際の格納データのコンテンツに基づいた何らかの形でデータは引き続き検索され、かつ/または操作される。1つの具体的な例では、キーワードに基づいて電子メールの受信箱をサーチして以前に受信した電子メールをサーチすることができる。このキーワードは、その単語に関連する各電子メールに検索ツールに関連付ける。この場合、その単語は実際に格納した電子メールの内容に関連付けられる。従って、ユーザがキーワード「John」についてサーチしようとする場合、このキーワードに関連付けられた任意の電子メールが取り出されてユーザに提示され、それによりユーザは単語「John」に関連付けられた所望の電子メールについての取り出されたりリストを取捨選択する。得られたメールの一覧からユーザがサーチしている特定の電子メールが取り出されるかもしれないが、所望の電子メールを見つけるために多数の電子メールをサーチし

10

20

30

40

50

なければならない可能性がある（例えば、30個の電子メールが単語Johnを含む）。当然ながら、上の例で説明した電子メール処理を、多種のデータ処理およびファイル操作アクティビティを含めるように拡張することができる。例えば、これらは格納データの索引作成、格納データのプレゼンテーション、様々な種類の格納データのサーチ、データの順位付け、データの索引作成等を含むことができる。

【0003】

【特許文献1】国際公開第98/00787号パンフレット

【発明の開示】

【発明が解決しようとする課題】

【0004】

一般にコンテンツ中心型アプリケーションに関連して、読み手が検索、閲覧、および使用する「完全な」文書の或る共通認識は一般に、知識集約型作業を適切に支援するのに十分ではない。従って、ユーザまたはユーザグループは自身の情報を知識源（ナレッジ・ソース）に追加できるべきである。一例において、歴史家は詳細な分析を本のある章に追加したいかもしれない。別のユーザは、その分析から収集した知見で本の節に注釈を施したいかもしれない。

【0005】

実質的に全ての文書がウェブ上またはウェブを通して利用可能である一方、そのハイパーテキストの機能は、既存情報（例えば、本、論文、ウェブページ等）を直接的に修正および注釈付けする程広範囲には現在使用されていない。寧ろ、コンテンツが「完成している」と見なされると、コンテンツは一種のアーカイブ（例えば、デジタルライブラリ）に格納され、そこからコンテンツが最終的に単体のエンティティとして検索され、さらなるコンテンツの作成に使用される。さらに、情報検索のタスクは一般にコンテンツ開発作業と統合されていない。従って、ユーザは自分自身が作業に対して必要と信じる文書を検索し、発見した情報にコンテンツ開発の基礎を置かなければならない。新規文書サーチを常に手動で開始することができるが、コンテンツ開発と検索を統合すべきであるということにはるかに説得力のある見方である。例えば、ユーザが入力した新規テキストを継続的に走査し、かつ分析するシステムは、追加の関連情報をサーチしてこれをユーザに提示できるべきである。ユーザは次いで新規データを検査、統合し、相互参照（cross reference）を追加し、または提案された情報源を拒否することができる。

【0006】

情報源からの知識は一般に文書の作成者およびその読み手両方のコンテキストの記述なしには適用できないという別の態様がある。その2つのコンテキストフレームを明確に表現することによってのみそれらの間の（半自動的な）変換が可能となる。上の例では、古い知識を現代の標準および語彙に適應させることができるが、現在デジタル形式で生成および格納されている全ての文書が「歴史上の知識」となると、同様な問題が中期的および長期的な将来において次第に現れる可能性がある。

【0007】

現在、ユーザは一種の索引作成および順位付けシステム、即ち、プレーンなウェブページに対するウェブサーチエンジン、またはデジタルライブラリに対する一種の情報検索システム（歴史的には、これらのシステムの起源は様々であるが、現代の実装にはこれらの技術が幾分重複している）、を通して文書を取得する。何れの場合も、文書がウェブページ、論文、または本全体であっても、システムは通常は完全な文書を返す。これが、処理すべき情報源を事実上無限に有する多数のユーザが共通して感じる「情報過多」の主な理由の1つである。

【課題を解決するための手段】

【0008】

以下で、本明細書で説明する幾つかの態様の基本的な理解を与えるため、簡潔な要約を示す。本要約は広範囲な概要ではなく、本明細書で説明する様々な態様の主要/重要な要素を特定またはそれら態様の範囲を線引きすることを意図していない。本要約の唯一の目

10

20

30

40

50

的は、後述の詳細な説明に対する前置きとして幾つかの概念を簡潔な形で提示することである。

【0009】

純粋なコンテンツ中心型のデータ処理アプリケーションとは対照的に、ファイルまたはアプリケーションに関連付けられたメタデータタグを使用して効率的な情報の記憶および/または情報へのアクセスを容易にすることができる。各ファイルまたはアプリケーションに関連付けられるようなデータに対するユーザのアクティビティまたは相互作用（interaction）は、特に興味深く効率的な種類のメタデータを表し、多数のアプリケーションの焦点である。ユーザが長期に渡ってデータを処理すると、そのデータに対するユーザのアクティビティを監視し、アクティビティの種類および強さに従って重み付けすることができる。例えば、頻繁にファイルへのテキスト追加およびファイルからのテキスト削除を行うことでユーザが特定のファイルと大量に相互作用する場合、点数または重みをファイルに対してメタデータまたは他の形式で割り当て、上記アクティビティを示すことができる。

10

【0010】

別の事例では、ファイルとの相互作用が稀（例えば、年内に1回だけオープンする）である場合、このファイルとのアクションが相対的にないことによって低い重みを割り当て、ファイルの重要度が低い可能性があることをユーザに示すことができる。当然のことながら、点数または重みを広範囲のファイル使用アクティビティに対して割り当てることができる。作成、オープン、閲覧、スクローリング、編集、印刷、注釈付け、保存、転送等の複数の異なるアクティビティに対して割り当てることができる。次いでアクティビティの重みまたはパターンをデータ項目に関連付けること（例えば、データベース内の列にタグ付けすること）、項目のサブセクション、または項目のグループに関連付けることができる。次いでアクティビティの重みを例えば検索ユーティリティなどのデータ操作ツールにより後で利用して、大規模なデータ項目のセットをより小規模またはより管理可能な項目のセットに絞り込むことができる。例えば、単にデータ項目のセットをサーチしてコンテンツ中心型のキーワードを求めるのではなく、アクティビティで強化した手掛りを通して情報のサーチを補強して、関心対象の所望のデータをより効率的に検索することができる（例えば、特定のユーザに転送された全てのファイルを見つけること、他のアプリケーションに対して最も頻繁に利用されたプレゼンテーションのサブセットを見つけること、最後に編集されたパラグラフを決定すること等）。

20

30

【0011】

前述の目的および関連する目的を達成するために、本明細書では一定の実例的な態様を以下の説明および付属図面と関連させて説明する。これらの態様は実践可能な様々な方法を示し、それら方法の全てを本明細書で包含することを意図する。他の利点および新規特徴は、図面と関連させて考慮すると、以下の詳細な説明から明らかになるであろう。

【発明を実施するための最良の形態】

【0012】

本発明では、複数のデータ型上でのデータ処理を可能とする様々なコンポーネントとプロセスとが提供される。複数のデータ型において、データに対するユーザのアクティビティまたは相互作用を決定し、利用して、そのアクティビティに従って当該データをさらに処理する。例えば、アクティビティまたは相互作用を監視し、続いてデータ項目にタグ付けして（例えば、ファイル相互作用のアクティビティに重みを割り当て、データベース内の列に適用する）、データベース内に存在する様々なデータ項目（またはデータ項目のサブセット）のサーチ、索引作成、カタログ化、順位付け、または閲覧に後で利用することができる。本発明のある特定の態様では、データ操作システムが提供される。本システムには1つまたは複数のデータ項目が含まれ、それらのデータ項目は1つまたは複数のタグに関連付けられ、データ項目との少なくとも1つのユーザの相互作用を指し示す。操作ツール（例えば、サーチツール）はデータ項目を処理して、そのデータ項目とのユーザの相互作用に少なくとも部分的に基づいてデータ項目のサブセットを決定する。

40

50

【 0 0 1 3 】

本出願で使用する場合、用語「コンポーネント」、「システム」、「タグ」、「モニタ」、「モデル」、「クエリ」等はコンピュータ関連のエンティティを指すよう意図され、ハードウェア、ハードウェアおよびソフトウェアの組合せ、ソフトウェア、または実行中のソフトウェアのいずれかである。例えば、コンポーネントはプロセッサ上で実行中のプロセス、プロセッサ、オブジェクト、実行可能ファイル、実行スレッド、プログラム、および/またはコンピュータであることができるが、これらに限らない。実例として、サーバ上で実行中のアプリケーション及びサーバの両方はコンポーネントであることができる。1つまたは複数のコンポーネントはプロセスおよび/または実行スレッド内に存在することができ、1つのコンポーネントを1つのコンピュータ上に局所化させること、および/または2つまたはそれ以上のコンピュータ間で分散させることが可能である。また、これらのコンポーネントを、様々なデータ構造を記憶した様々なコンピュータ読取可能な記憶媒体から実行することができる。コンポーネントは、例えば1つまたは複数のデータパケットを有する信号に従ってローカルおよび/またはリモートのプロセス経路で通信することができる(例えば、あるコンポーネントからのデータは別のコンポーネントと前記信号を通して、ローカルシステム内、分散システム内、および/または他システムを有するインターネットなどのネットワークに渡って相互作用する)。

10

【 0 0 1 4 】

本明細書で使用する場合、用語「推論」は一般に、事象および/またはデータを通して取得した観測結果のセットからシステム、環境、および/またはユーザの状態を論理的に考えるまたは推論するプロセスを指す。例えば、推論を使用して特定のコンテキストまたはアクションを識別するか、あるいは状態の確率分布を生成することができる。推論は確率的でありうる。即ち、関心となる対象の状態の確率分布の計算は、データおよび事象を考慮して行われる。推論は、事象および/またはデータのセットから高レベルの事象を構成するために利用する技術も指し示すことができる。上記の推論の結果、事象が時間的に近接して相関するか否かに関わらず、ならびに事象およびデータが1つまたは数個の事象およびデータの情報源から来るか否かに関わらず、観測した事象および/または格納した事象データのセットから新規の事象またはアクションが構築される。さらに、推論は論理モデルまたは規則に基づくことができ、それによりコンポーネントまたはデータ間の関係がデータ分析から決定され、結論が引き出される。例えば、あるユーザが他のユーザのサブセットとネットワーク上で相互作用することを観測することで、一度も相互作用しないかまたは相互作用が稀である複数の他ユーザとは対照的に、このユーザサブセットが前記ユーザの関心対象の所望の社会的ネットワークに属することを決定または推論することができる。

20

30

【 0 0 1 5 】

最初に図1を参照する。システム100は、ユーザのアクティビティデータまたは相互作用データを使用して様々なコンピュータ関連タスクを実施するデータ処理アーキテクチャを示す。モニタコンポーネント110は、1つまたは複数のデータ項目130を格納する1つまたは複数のデータベース120との経時的なデータ相互作用を監視する。モニタリングはバックグラウンドおよび/またはフォアグラウンドのコンポーネント(図示せず)を通して行うことができ、ユーザがデータ項目130と相互作用する時点の決定に利用することができる。例えば、モニタリングは、ファイルのオープンまたはクローズ、編集、ファイルに対する追加またはファイルからの削除、ファイルの読み出しや書き込み、ファイルの切り取り、貼り付け、ファイルの最終更新、転送、応答、送信、最終閲覧それぞれの時点、閲覧時間、時間範囲での相互作用時間、項目130または様々な項目のサブコンポーネントに着目した期間を示す注釈等の監視を含むことができる。

40

【 0 0 1 6 】

アプリケーションファイルを開き、特定のデータ項目130をアプリケーション内部から操作しているとき、様々な技術を利用してアプリケーション内部のアクティビティを決定することができる。これらは、特定のデータセットまたはデータサブセットに対

50

するユーザの検討時間、修正データまたは観測データ、データに影響を及ぼした頻度および期間等の監視を含むことができる。タグコンポーネント140は監視したデータアクティビティに従って重みまたは点数を割当てる。望むならばこれらを確認的に割当てる（または分類を重み付ける）ことができ、データまたはアプリケーションの所与部分とのユーザ相互作用の量またはパターンを反映することができる。例えば、ファイルの使用量が最小であると、広範囲にファイルを編集した場合よりも低い重みが生成される可能性がある。重みまたは点数が決定されると、アクティビティに対して決定した重みの情報を1つまたは複数のデータ項目に関連付けるかまたはタグ付けし、この重み情報をアクティビティデータ150として示す。この関連付けは例えばデータベース120の列（単数または複数）内の値を生成または修正する際にデータベース120内部で生じ、データベース120の行内で識別した特定項目の重み付けまたは重要度を、関心対象のメタデータタグに対して示すことができる。列は論理エンティティであり、明示的に格納するか、または使用時に動的に計算することができることは理解されるべきである。他種の関連付けは、1つまたは複数のデータ項目130に直接的または間接的に割り当てるメタデータ参照を含むことができる。

10

【0017】

タグおよびデータ項目130を1つまたは複数のデータ操作ツールに適応させることができる。そのデータ操作ツールは、タグから導出したユーザアクティビティ情報を使用して、情報の格納（例えば、効率的なインデックスの生成）、情報へのアクセス（例えば、データ項目のサーチ、フィルタリングまたは順位付け）および情報のプレゼンテーション（例えば、データ項目の編成、配置または提示）等の補強を行うことができる。符号170で、データ操作ツール160からの結果が自動的に生成される。データ操作ツール160からの結果は、データ項目130の大規模なセットより小規模なデータ項目のサブセットを含むことができる。本明細書で適用するとき、サブセットという用語はデータ項目130の全てまたは一部を含みうるわけではない。また、アクティビティ閾値をツール160内部にセットアップして、多少のデータ項目130を結果170内に含めることができる。

20

【0018】

他の態様では、コンテンツベースの点数を文書に割り当てることができる。例えば、点数をユーザのクエリの類似性に基づいて文書のコンテンツに割り当てることができる。従って例えば、順位付けの際、編集した文書または文書の部分またはユーザが長時間読んだ文書または文書の部分に現れる項目に対してより多くの重みを与える。システム100のさらに別の態様では、アクティビティデータ150を使用して関心対象の領域または単語を特定することができる。従って、文書の領域に差別的な重み付けを行うことで、以下のことを行うことができる。

30

- 1) インデックスを圧縮し、関心対象の領域内の用語を優先的に含めること、
- 2) 順位付けのため、関心対象の領域内の項目を差別的に重み付けすること、
- 3) 関連フィードバック (relevance feedback) のため、関心対象の領域内の項目を差別的に重み付けすること、
- 4) 現在のユーザが着目する領域に基づいて自動的にまたは半自動的にクエリを作成すること、および/または
- 5) 関心対象の項目または項目領域を（強調または他の技術により）差別的に提示すること。

40

【0019】

本明細書で説明するシステムおよび方法は複数のデータ処理アプリケーションを支援する。これには例えば、文書、ファイル、電子メールメッセージ、カレンダー予約、ウェブページ、データ項目の抽象概念または項目間の抽象概念内部のサブセクションといった、データ項目の処理が含まれる。タグをデータ項目に適用することで、ユーザが最後に項目にアクセスした位置を表現すること、またはユーザが項目に対してアクセスまたは相互作用した回数の位置情報履歴を表現することができる。タグはユーザが最後に項目にアクセス

50

した時点、項目がアクセスされた総数を表し、過去にまでさかのぼった期間内で項目がアクセスされた頻度を表し、あるいは1つまたは複数の任意の特定期間内で項目がアクセスされた頻度を表す。経時的なアクセス頻度の高次統計を符号化する他のコンポーネントを提供することができる。ある場合では、ビューによりユーザは1つまたは複数のタグの機能に基づいて項目を検索すること、1つまたは複数のタグの機能に基づいて検索した項目をソートまたはフィルタすることが可能となり、またはビューは1つまたは複数のタグの機能に基づいて検索した項目を表示する。別の場合では、タグはアクティビティまたは関心対象を確率的に示すことができる。

【0020】

様々なプロセスには、データ項目に対するユーザのアクティビティを分析することが含まれる。これには、データ項目がコンピュータユーザから受けた相互作用の量または性質を自動的にタグ付けすること、およびそのタグを使用して将来のデータアクティビティに従いデータ項目をさらに処理することが含まれる。プロセスに、独立したデータベース内のデータ項目に関連付けられた注釈内部にデータを格納することまたはデータ項目内に組み込んだデータ構造内部にデータを格納することを含めることができる。また、データ項目に対するアテンションまたは相互作用を示す注釈の状態に依存して、検索に対してデータ項目のサブコンポーネントを別々に重み付けする索引作成手続きを提供することができる。これには、データ項目に対するアテンションまたは相互作用を示す注釈の状態に依存して、データ項目内の情報を見逃すかまたは削除する索引作成手続きが含まれる。コンピュータユーザが関心を向けなかったまたはコンピュータユーザが相互作用しなかったコンポーネント、あるいは注目度が低かったまたはデータ項目のコンポーネントとの相互作用が少なかったコンポーネントを削除することで、インデックスを圧縮することができる。

【0021】

別の態様では、順位付け点数をデータ検索に対して利用して、ユーザが関心を向けたか、またはユーザが相互作用したデータ項目の部分に現れる単語またはオブジェクトに対してより多くの重みを与えることができる。これは、注釈を利用して、過去および/または現在において関心を向けたかまたは相互作用した領域に基づいて自動的にまたは半自動的にクエリを作成することを含むことができる。注釈を利用して、過去および/または現在において関心を向けたかまたは相互作用した項目に対して差別的なアクセスを提供するか、またはその項目の差別的な表示を提供することもできる。

【0022】

注釈をデータ項目自体に対するアテンションとしてだけでなく、データ項目のサブコンポーネントに対するアテンションとしても符号化することができ、この場合、注釈は、サブコンポーネントの各々と受けたアテンションとに対するポイントまたは他の指標を取得する。例えば、211ページの文書などの大規模な文書を考える。文書は23回、オープンされ、相互作用され、関心を向けられた可能性があり、例えばそのことはその文書に対するある種の注釈として取得される。しかしながら、他の注釈は例えばユーザが4-6ページ、89-93ページ、123-124ページ、および198ページを繰り返し検証し、文書の他のページは素早く流し読みしたことを示す。従って、各サブコンポーネントを

【0023】

図2は、ユーザのアクティビティデータまたは相互作用データを使用するユーザインタフェースの実施例200を示す。本実施例において、操作ツール210(例えば、データベースに適用したユーザインタフェース)に出力またはディスプレイ220を関連付けることができる。ツール210は、1つまたは複数のデータベースからのデータを処理する多数の機能を含むことができる。例えば、ツール210は、データサーチの有効化、データの索引作成またはカタログ化、データの順位付け等に対する選択肢を含むことができる。上記データは例えば、XMLデータまたはASCIIデータなどのテキストデータを含むことができる。他のデータには例えば、イメージデータ、オーディオデータ、ビデオ

10

20

30

40

50

ータ、グラフィックデータ、および/または一連のスライド内に含まれるようなプレゼンテーションデータが含まれる。実質は、スプレッドシート、ユニバーサルリソースロケーション (URL) 情報、インターネットまたはウェブデータ等を含む任意のデータ型またはアプリケーションを利用することができる。図1に関連して上述したように、上記データを例えば列内またはファイルメタデータとしてタグ付けし、過去の使用または相互作用を示す点数または重みを示すことができる。操作ツールはその後、タグ付けしたデータをサーチ、取得または処理して、ユーザに対してより管理可能なデータサブセットを絞り込むことができ、または決定することができる。

【0024】

ツール210からの出力220はファイルまたは実際のユーザインタフェース表示であることができる。例えば、ツールをデータベース内部のサーチエンジンとして利用したならば、出力はリターンされた結果を表示するものになるであろう。リターンされた情報は符号230に示すように事実上、よりグローバルなものでありうる。これはグラフィックを強調して、あるいはグラフィックをファイルへ適用して、ユーザとのアクティビティの増加により1つのファイルまたはファイルのグループが選択されたことを示すことを含むことができる。例えば電子メールサーチツールの場合、ツール210を適用して、コンピュータなるキーワードを有し、且つ過去1ヶ月にファイルに関連付けられたグラフィカルイメージを少なくとも1つ有していたファイル全てをサーチすることができる。サーチを複数の方法で工夫することができ、サーチにコンテンツサーチ、アクティビティベースのサーチ、および/またはそれらの組合せのうちからいくつかを組み合わせたものを含めることができる。例えば、本例では、10個の電子メールのセットのうち3つの電子メールを或る色で強調し、別の色で区別した他の7つの電子メールより高いアクティビティ点数を有するものとするることができる。符号240での別の態様では、リターンされたファイルまたはデータセット内部の情報を強調または注釈付けして、使用アクティビティを示すことができる(例えば、ファイル内部のパラグラフを異なるフォントフォーマットで選択して、文書内部の使用領域を強調する)。

【0025】

図3は、ユーザのアクティビティデータまたは相互作用データを決定および適用するプロセス300を示す。説明を簡素化するため、手順を一連の動作または多数の動作で図示および説明するが、本プロセスに従って一部の動作を本明細書で図示および説明したものと異なる順序で行うことおよび/または他の動作と並列に行うことができるので、本プロセスは動作順序により限定されないことは理解されるべきである。例えば、当業者は、手順を代替的に状態図のような一連の相関する状態または事象として表現可能であることを理解するであろう。さらに、本プロセスに従う手順を実装するために、示した動作全てが必要であるわけではない。

【0026】

符号310に進む。1つまたは複数のローカルまたはリモートのデータベースとのデータ相互作用を監視する。当該監視はバックグラウンドおよび/またはフォアグラウンドのアプリケーションで行うことができ、ユーザがデータまたはファイルと相互作用する時点を決定するために利用される。例えば、これは、ファイルのオープンまたはクローズ、編集、ファイルへの追加またはファイルからの削除、ファイルの読み出しや書き込み等の時点の監視を含むことができる。アプリケーションがオープンして且つ特定のデータがアプリケーション内部で操作されているとき、様々な技術を利用してアプリケーション内部のアクティビティを決定することができる。これらは、データの特定セットまたはサブセットに対するユーザの検討時間、修正データまたは観測データ等を含むことができる。外部モニタリングもアプリケーションに関連付けて、データ、ファイルまたはアプリケーションとのユーザの相互作用を決定することができる。例えば、オーディオキュー(audio cue)、自動顔認識技術、またはデータセットが強くユーザに関連するという明示的なユーザ指示を利用することができる。

【0027】

符号 3 2 0 で、監視したアクティビティに対する重みを決定する。重みを確率的に割当てることができ、データまたはアプリケーションの所与部分に対するユーザの相互作用の量を反映することができる。例えば、文書を非常に軽く読む場合、文書を広範囲に編集する場合よりも低い重みが生成される可能性がある。符号 3 3 0 で、アクティビティに対して決定した重み情報を 1 つまたは複数のデータ項目に対して関連付けまたはタグ付けする。この関連付けは、例えばデータベースの列（単数または複数）にタグ付けする際にデータベースの境界内部で生じ、データベースの行内で識別した特定のデータ項目の重み付けまたは重要度を示すことができる。他の種類の関連付けは、1 つまたは複数のデータ項目に直接的または間接的に割り当てたメタデータ参照を含むことができる。符号 3 4 0 で、タグおよびデータ項目をデータ処理ツールに適応させることができる。これは、タグから導出したアクティビティ情報を利用して、データ項目のサーチ、データ項目の索引作成、データ項目の配置、データ項目の順位付け、データ項目の編成等を補強することを含むことができる。符号 3 5 0 で、データ処理アプリケーションからの結果を生成する。これは、大規模な結果セットをフィルタして小規模なサブセットにするような明示的動作、またはディスプレイに注釈を付与してディスプレイ上のファイルまたはデータを強調し、ユーザとの相互作用が多い項目を示すような捕らえにくい動作を含むことができる。

【 0 0 2 8 】

図 4 を参照する。例示的システム 4 0 0 は、ユーザアクティビティのデータ処理とともに使用可能な情報検索アーキテクチャを示す。システム 1 0 0 はサーチ結果をカスタマイズする（personalize）ための一般的な図を示すが、上述のように他の形態のデータ操作を実施することができる。カスタマイズ化コンポーネント 4 1 0 は、ユーザアクティビティに基づくユーザモデル 4 2 0 と、そのモデルを使用してクエリ 4 3 0 の修正および/またはサーチからリターンされた結果 4 4 0 の修正によりサーチ結果に影響を及ぼす処理コンポーネント（例えば、ユーザモデルに従って修正した検索アルゴリズム）とを含む。ユーザインタフェース 4 5 0 はクエリ 4 3 0 を生成し、カスタマイズ化コンポーネント 4 1 0 が提供するクエリ修正 4 7 0 および/または結果修正 4 6 0 に基づいて、修正またはカスタマイズした結果を受け取る。本明細書で利用する場合、用語「クエリ修正」は、クエリ 4 3 0 内の用語に関する変更と、クエリ 4 3 0 を文書に適合させてカスタマイズ化結果 4 4 0 を取得するアルゴリズムにおける変更との両方を指す。修正したクエリおよび/または結果 4 4 0 は、1 つまたは複数のローカルおよび/あるいはリモートのサーチエンジン 4 8 0 からリターンされる。ユーザ統計データのグローバルデータベース 4 9 0 を保守して、ユーザモデル 4 2 0 のアップデートを容易にすることができる。当然のことながら、ユーザモデル 4 2 0 および/またはグローバル統計データ 4 9 0 を、前述のようにユーザのアクティビティデータまたは相互作用データに関連付けてデータ操作またはデータ処理を容易にすることができる。

【 0 0 2 9 】

一般に、ユーザモデル 4 2 0 に基づいてサーチ結果を調整するアプローチが少なくとも 2 つある。一態様においては、クエリ修正が初期入力クエリを処理し、（ユーザモデルを通して）クエリを修正または再び作成して、カスタマイズ化した結果を生成する。関連フィードバックはこのプロセスを 2 回繰り返すものであり、（初期結果セットに関する明示的または暗黙的な判定を用いて）修正クエリに至る結果をクエリが生成する。その修正クエリは、クエリおよび結果セットに基づいて短期モデルにカスタマイズしたカスタマイズ化結果を生成する。長期ユーザモデルを関連フィードバックのコンテキストで使用することもできる。さらに、クエリ修正は、クエリの文書への適合に利用するアルゴリズム（単数または複数）において行う変更も指し示す。別の態様においては、結果修正はユーザの入力をそのまま取得して、結果を生成するクエリを作成し、次いでその結果を（ユーザモデルを通して）修正してカスタマイズ化結果を生成する。結果の修正には通常、決定したアクティビティデータからの検討または重み付けを含みうる代替手段の大規模なセットから、何らかの形で再順位付けおよび/または選択することが含まれることに留意されたい。結果の修正には、結果の全体またはサブセットを様々な形で集積および要約することも

10

20

30

40

50

含まれる。

【 0 0 3 0 】

結果の修正方法には、統計的類似性適合（ユーザの関心および満足度をベクトルで表し、項目に適合させる）、およびカテゴリマッチング（小規模な記述子セットを用いてユーザの関心および満足度を表し、項目に適合させる）が含まれる。上記のクエリ修正プロセスまたは結果修正プロセスを、独立に組み合わせること、または統合プロセス内で組み合わせることのいずれかが可能である。統合プロセスの場合、依存性を2つのプロセス間で導入して活用する。

【 0 0 3 1 】

図5を参照する。図5は、ユーザのアクティビティデータまたは相互作用データとともに使用可能なユーザモデル500を示す。ユーザモデル500を使用して、カスタマイズしたサーチを汎用的なサーチと区別し、決定したアクティビティデータに従ってリッチな（rich）データ処理を容易にする。カスタマイズを成功させる1つの態様は、ユーザの関心を正確に反映し、且つ保守が容易で、長期的および短期的な関心に関する変化に適応が容易なユーザのモデルを構築することである。ユーザモデルを様々な情報源から取得することができ、その情報源は以下のことを含むがこれらに限らない。

1) ローカル、モバイル、またはリモートの情報源から得ることが可能な、符号510でのコンピューティングコンテキストのリッチな履歴（例えば、オープンしているアプリケーション、これらアプリケーションのコンテンツ、および位置を含む上記相互作用の詳細な履歴）

2) 符号520での、以前に遭遇したコンテンツのリッチなインデックス（例えば、文書、ウェブページ、電子メール、インスタントメッセージ、ノート、カレンダー予約等）

3) 最近のコンタクトまたは頻繁なコンタクト、キーワードから導出した関心対象のトピック、組織図内の関係、予約等を含む、符号530でのクライアント相互作用のモニタリング

4) 以前のサーチクエリの履歴を含む、符号540での訪問した以前のウェブページまたはローカルデータサイト/リモートデータサイトの履歴またはログ

5) 明示的に特定可能または暗黙的にバックグラウンドモニタリングを通して導出可能な、符号550でのユーザのインタレストプロファイル

6) 符号560での人口学的情報（例えば、位置、性別、年齢、経歴、職業区分等）

【 0 0 3 2 】

上記の例から、ユーザモデル500は様々な情報源に基づくことが可能であることが理解できる。例えば、モデル500は、全地球測位システム（GPS）などの装置で監視される、ユーザが長い間にわたり訪問した位置情報の履歴またはログを情報源とすることができる。GPSを用いて監視するとき、生の空間情報をテキスト形式の都市名、および郵便番号に変換することができる。例えば、ユーザが立ち止まった位置またはユーザが居住する位置またはユーザがGPS信号の損失を招いた位置に対して、生の空間情報をテキスト形式の都市名、および郵便番号に変換することができる。ユーザが立ち止まった位置またはユーザが居住する位置またはユーザがGPS信号の損失を招いた位置を、ビジネスおよび関心対象の地点のデータベースを通して特定し、テキスト形式のラベルに変換することができる。他の要素には、時刻または曜日を記録して関心対象の位置および地点を決定することが含まれる。

【 0 0 3 3 】

他の態様においては、タイプ、年齢、または他の組み合わせに基づいた、ユーザの情報コーパス（corpus）、予約、文書またはファイルの閲覧、アクティビティ、または位置の、サブセットへのグループ化方法、またはカスタマイズ化用の適合手続きにおける差別的な重み付け方法の制御パラメータを操作するコンポーネントを提供することができる。例えば、検索アルゴリズムを、クエリに関するユーザのコーパスの態様（例えば、クエリの用語を含むか、またはデータとの過去の相互作用を含む文書）に限定することができる。同様に、1ヶ月前のものから電子メールを分析することができ、ウェブアクセスは3日前

10

20

30

40

50

のものから分析することができ、去年に作成されたユーザのコンテンツを分析することができる。GPS位置情報を当日のみまたは他の期間から使用することが望ましい場合がある。(例えば、パラメータとユーザまたはシステムからのテスト応答とを変化させる最適化プロセスを通して)パラメータを自動的に操作してサブセットを作成可能であるか、またはユーザはユーザインタフェースを通してこれらのパラメータのうち1つまたは複数を変化させることが可能であり、この場合上記設定は、クエリの性質、時刻、曜日、あるいは他のコンテキストまたはアクティビティベースの観測値の関数でありうる。

【0034】

例えば個人または個人グループの間の類似性分析によってプロフィールを構築する協調型フィルタリング手法を通して、符号570で個人または個人グループに対してモデルを導出することができる。類似性の計算は、データ項目のコンテンツおよび/または使用量に基づくことができる。モデリングの基盤と関連する処理とは、クライアント、複数のクライアント、1つまたは複数のサーバ、あるいはサーバおよびクライアントの組合せに存在することができる。

10

【0035】

符号580で、機械学習技術を適用して、長い間にわたるユーザの特徴と関心事、およびユーザによるデータとの相互作用の方法と時期を学習する。学習モデルは、ユーザをモデル化し、嗜好と関心事とを決定するための統計的/数学的モデルおよびプロセスといった、実質的に任意の種類システムを含むことができる。この統計的/数学的モデルおよびプロセスには、例えばベイジアンネットワーク、ナイーブベイジアン分類器、および/またはサポートベクトルマシン(SVM)を含む他の統計的分類方法といった、ベイズ依存性モデルを生成可能なベイジアン学習の使用が含まれる。他の種類のモデルまたはシステムは、例えばニューラルネットワークおよび隠れマルコフモデルを含むことができる。精巧な推論モデルを利用できるが、他のアプローチも利用できることは理解されるべきである。例えば、より完全な確率的アプローチではなく、決定論的仮定を利用することもできる(例えば、特定のウェブサイトのサーチが最近X時間行われなことは、規則によって、ユーザがその各情報にもはや関心を示していないことを暗示することができる)。従って、不確実性の下での推論に加え、ユーザの状態、位置、内容、関心、焦点等に関して論理的な決定を行うこともできる。

20

【0036】

複数の異なるデータ源からデータを収集または集計するユーザ事象データ記憶(図示せず)から、学習モデルをトレーニングすることができる。当該データ源は、ユーザ事象データを記録する様々なデータ取得コンポーネント(例えば、携帯電話、マイクロフォンにより記録した音響アクティビティ、全地球測位システム(GPS)、電子カレンダー、視野監視装置、デスクトップアクティビティ、ウェブサイト相互作用等)を含むことができる。システムは、カスタマイズ化したクエリ処理および結果処理を支援する実質的に任意な方法で実装できることに留意されたい。例えば、システムをサーバ、サーバファームとして実装すること、クライアントアプリケーション(単数または複数)内部でシステムを実装すること、あるいはユーザインタフェースおよびサーチエンジンなどのサーチ機能と相互作用するウェブサービス(単数または複数)または他の自動化アプリケーション(単数または複数)を含むようにより一般化することが可能である。

30

40

【0037】

先に進む前に、符号570で適用したユーザモデル500の協調型フィルタ技術をより詳細に説明する。これらの技術は、協調型フィルタを使用してデータを分析し、ユーザに対するプロフィールを決定することを含むことができる。協調型フィルタリングシステムは一般に、ユーザの嗜好に関する集中データベースを使用して、ユーザが望む可能性のある追加のトピックを予測する。ユーザモデル500に協調型フィルタリングを適応させ、ユーザグループからの以前のユーザアクティビティを処理する。その以前のユーザアクティビティは所与のユーザに対する嗜好を示唆することができ、その嗜好がシステムの新規ユーザに対して考え得るまたは可能なプロフィールを予測する。相関係数、ベクトルベ

50

スの類似性計算、および統計ベイズ法に基づく技術を含む幾つかのアルゴリズムを使用することができる。

【0038】

図6を参照する。システム600は、ユーザのアクティビティデータまたは相互作用データに従うアクセススペースの情報検索を示す。システム600には、ユーザによりアクセスまたは考慮される1つまたは複数の情報源610が含まれる。これらの情報源610は同様であるか、または異種の情報コンテンツを有して本質的に別個であることができ、それにより例えば、情報源の幾つかは、ファイル、フォルダ、アプリケーション、イメージ、オーディオファイル、予約、電子メール等のようなローカルなデータ位置を表すことができ、他の情報源610はウェブ情報などのリモートな情報源を表すことができる。ユーザが経時的に異種の情報にアクセスすると、使用量分析器614はこの情報を、バックグラウンドタスクとしてクライアントマシン上でローカルに動作可能および/またはサーバに従ってリモートに動作可能なコンテンツ分析器620（またはモニタ）に渡し、アクセスしたデータを、データのコンテンツを解析するフィルタ624とアクセスしたデータ項目のコンテンツインデックス640（またはアクティビティタグ）を作成する自動インデクサ630とを通して処理する。

10

【0039】

一般に、分析器620はインデックス640内のアクセスしたデータの表現を生成する。例えば、ユーザがウェブページにアクセスすると、コンテンツ分析器620はウェブページのサムネイル表現を作成して、そのページおよびサムネイルに対してハイパーリンク参照をメタデータファイルの一部として関連付けることができる。前記表現をさらに調整して、ページに対するユーザの相互作用の詳細なパターンを反映することができる。別の場合は、ユーザがイメージを内部に含むテキスト文書にアクセスすると、分析器620はテキストまたはその一部を抽出し、ファイルパスなどのデータベースリンクをメタデータの一部として関連付けることができる。インデクサ630は次いで、コンテンツインデックス640内に2つの項目、即ちメタデータを含むサムネイル表現とテキスト文書表現とを有するインデックスを自動的に作成（または既存のインデックスに追加）する。一般に、フィルタは項目のコンテンツと項目に関連付けられたメタデータとを分析する。従って、例えばWord文書に対して、フィルタ624はファイル名、タイトル、著者、キーワード、作成日等のメタデータを文書内の単語とともに抽出する。これがインデックス640の構築に使用されるものである。望むならば、サムネイルの作成とイメージの分析をフィルタ624内にカプセル化することもできる。他の項目を含むアクティビティまたは相互作用のメタデータを使用することができる。当該他の項目には例えば、項目がコンテンツインデックス640内に格納されることを記述するユーザタグおよび/または暗黙的なタグがある。インデクサ630はフィルタ624の機能を実施する（例えば、インデクサがメタデータとフィルタしたコンテンツとを関連付ける）こともできることは理解されるべきである。

20

30

【0040】

コンテンツインデックス640内の情報項目に対するユーザクエリ654を受け取る、サーチコンポーネント650を提供する。サーチコンポーネント650をユーザインタフェースの一部として提供することができる、このユーザインタフェースはクエリ654に回答して、符号660のアクセス項目のリンクおよび/または表現をユーザにリターンする。例えば、ユーザは「去年のパフォーマンスレビューに関する項目」を問合せることができ、サーチコンポーネント650はクエリ654のコンテキストに関連する電子メール、同僚の評価、去年に発行された文書、ウェブページのイメージ、音声記録等の項目をインデックス640から抽出する。別の例では、暗黙的なクエリをクエリ654から導出することができる（例えば、この人から電話があるときはいつも、この人からの直近5個の電子メールを引き出す）。

40

【0041】

以下でより詳細に説明するように、アクセスした項目を、以前にアクセスした情報項目

50

の効率的且つタイムリーな検索を容易にするよう設計した複数の異なる形式で提示することができる。また、リンクおよび/または表現 660 は、例えば以前にアクセスした項目以外にユーザが見たい情報項目を提供することで、関心対象の他の項目を含めることができる（例えば、システムは手元のクエリに基づくかまたは手元のクエリから推論して、例えばパフォーマンスレビュー項目の表示に加えて、関心対象の他のコンテンツに対するリンクを提供し、ヒューマンリソースへのリンクを任意に提供する。このヒューマンリソースは、これらの項目をユーザが以前にアクセスしたか否かに関わらず、コンテンツの別のインデックスに基づいてレビューポリシーを記述する）。

【0042】

一態様において、事象コンポーネント（例えば、使用量分析器 614 に関連付けられたユーザアクティビティを監視するバックグラウンドタスク）を提供することができる（図示せず）。事象コンポーネントは例えば保存、読み出し、編集、コピー、情報に対するホバーリング、情報の選択、情報の操作および/またはファイルの削除などのユーザアクティビティを監視し、ユーザ動作に関連する決定を行う。これは、デスクトップアクティビティを監視してユーザの動作または目標を決定することとともに、マイクロフォン、カメラ、および他の装置などのセンサーを含むことができる。一例において、確率モデルおよび/または論理的決定を適用して、ユーザが情報を観察または熟考する時点などの事象を決定することができる。ユーザアクティビティのパターンに関連付けられた以下のエビデンスクラスの例を考慮した論理モデルおよび/または統計モデル（例えば、ベイズ依存性モデル、決定木、サポートベクトルマシン）を構築することができる。

・着眼点：

項目の選択および/または検討、文書のスクロール後に文書の一部または特定のサブテキストを検討すること

・イントロスペクション：

アクティビティ期間後の一時停止、または相互作用の速度の大幅な低下

・不要な情報：

一見した後に即座に文書を閉じること、情報へのアクセス動作後に以前の状態に戻ろうと試みること。これらの観察結果には、取り消し命令の発行、および項目の削除を含んだ直近動作の効果の取消が含まれる。

・ドメイン固有な構文内容および意味内容：

文書内容または文書構造における特別な差異と、これら特徴または項目とのユーザの相互作用方法とを考慮すること。これらには、タスクに関連付けられたドメイン固有な特徴が含まれる（例えば、ユーザが着目するメッセージの作成者から、電子メールメッセージの割合および頻度、および件名標目の年月日またはメッセージ数を考慮すること）。当然のことながら、事象コンポーネントを利用して、ユーザアクティビティに基づく多種の情報の索引作成をトリガすることができる。情報オブジェクトを伴うユーザのアクティビティを利用して、情報プレゼンテーションを改善することもできる。

【0043】

図7は様々な検索サービスアプリケーション700を示す。一態様において、明示的クエリ710および/または暗黙的クエリ714を支援することができる。ユーザは明示的クエリ710に指示して関心の対象となる情報を発見する（例えば、会議またはデートに関連する全てのデータ参照を示す）。暗黙的クエリ714は幾つの場合では明示的クエリ710から導出することができる。例えば、ユーザは自分自身のデスクトップ上の電話メッセージを自分自身の電子メールシステムまたは他のメッセージシステムに結び付けることができる。電話を選択した個人から着信する予定の場合、電子メールシステムはその個人に関連する電子メールを、暗黙的クエリ714を通して自動的に検索するであろう。別の例では、次の会議までの所定の合間に、ユーザのカレンダーシステムは過去の会議からのデータまたは次の会議に参加する個人に関する情報を呼び出すクエリをトリガすることができる。また、暗黙的クエリ714を、ユーザの現在のコンテキストまたはクエリ（例えば、最近読んだパラグラフ内の重要な単語から成るクエリ）に関連付けられた推論

プロセスに基づいて作成することができる。

【0044】

符号716に進む。他の種類のクエリがコンテキストセンシティブなクエリを支援する。これらの種類のクエリには、追加の選択オプションを提供してサーチを編集または絞り込むことが含まれる。例えば、クエリを特定の種類のアプリケーションまたは位置に向ける（例えば、本クエリをメールフォルダのみに適用する）ことができる。符号720で、クエリの実施時にアプリケーションのコンテキストを考慮することができる。例えば、写真アプリケーションを使用している場合、クエリをイメージに対するサーチのみに絞り込むことができる。符号724で、項目中心の統合を実施することができる。これには、マウスのクリック機能、項目のタグ付け、メタデータファイルの更新、項目の削除、項目またはコンテンツの編集等のインタフェース動作を支援するオペレーティングシステムの動作が含まれる。

10

【0045】

符号730で、ファイルを共有することができる。例えば、ユーザは、1人または複数人の他ユーザが彼らのクエリ/インデックスデータベースの全てまたはサブセットに対して検査またはアクセス可能であると指定することができる（例えば、自身のプロジェクトチームの全ユーザは、自身のプロジェクトメモにアクセスすることができる）。符号734で、インデックスの無効化を行うことができる。時間が経つにつれて、ユーザは1つまたは複数の項目を自身のインデックスから除去したいと思うかもしれない。このアクティビティによると、ユーザは除去すべき特定項目を指定するか、またはシステムが自動的に無効化することができる一般的なトピック領域を指定することができる（例えば、2年前の自分の誕生日に関するサムネイルを除去する）。論理的プロセスまたは推論プロセスに基づいて他の動作を行うことができる。例えば、項目へのアクセスが所定期間において一定回数未満である場合、望むならば項目を自動的に除去することができる。

20

【0046】

符号740で、効率的な時間計算が考慮される。例として、（ユーザへのデータのプレゼンテーション中に）ファイルについて関連性があるかまたは有益な日付はファイルの変更日であり、メールの提示日は通常はメールの配信日（従って、おおよそはユーザがメールを見た時点）であり、予約に都合のよい日は予約を行った日である。記録および索引作成した全ての時間情報と当該有用な日付情報とを、情報のプレゼンテーションに利用することに留意されたい。従って、予約に対して様々なタスクを行うことができる。そのタスクには例えば、メール送信時、（行うならば）メール更新時、ユーザが受理/拒否した時点、会議開催時を索引作成することがある。ただし、複数の時点を提供することができるが一般には1つの時点を選択して表示する。

30

【0047】

上述のように、ファイルタイプに関連付けられたファイル要素を分析することで、或る特定のデータを以前に観察したとして印付けすることができる。例えば、テキスト文書はファイルのオープン時または最終編集時を示す領域を含むことができる。しかしながらカレンダー予約に関しては、単にカレンダーの作成時点からインデックスを作成するだけでは効果が薄い可能性がある。なぜならば、会議が実際の会議日よりかなり前に開催される場合があるからである。このように、カレンダー予約について索引作成するとき、作成時点ではなく実際の会議データを追跡する場合がある。このように効率的に時間を考慮することで、ユーザはより記憶の呼び出しに適した方法で情報を検索することが可能となる。符号744で、データの不安定性を考慮し、処理する。この種の処理には、断続的な動作中にデータを永続的な形に索引作成することが含まれる。当然のことながら、様々な自動バックグラウンド動作が可能である。

40

【0048】

図8を参照する。本明細書で説明した様々な態様を実装する例示的な環境810にはコンピュータ812が含まれる。コンピュータ812には、処理ユニット814、システムメモリ816、およびシステムバス818が含まれる。システムバス818はシステムコ

50

ンポーネントを処理ユニット 8 1 4 に接続する。システムコンポーネントにはシステムメモリ 8 1 6 が含まれるがこれに限らない。処理ユニット 8 1 4 は様々な利用可能なプロセッサのうち任意のものであることができる。デュアルマイクロプロセッサおよび他のマルチプロセッサアーキテクチャを処理ユニット 8 1 4 として使用することもできる。

【 0 0 4 9 】

システムバス 8 1 8 は様々な種類のバスアーキテクチャ（単数または複数）のうち任意のものであり、それらにはメモリバスまたはメモリコントローラ、周辺バスまたは外部バス、および/あるいは任意の様々な利用可能なバスアーキテクチャを用いるローカルバスが含まれる。上記ローカルバスは、11ビットバス、業界標準アーキテクチャ（ISA）、マイクロチャンネルアーキテクチャ（MSA）、拡張ISA（EISA）、インテリジェントドライブエレクトロニクス（IDE）、VESALocalバス（VLB）、周辺機器相互接続（PCI）、ユニバーサルシリアルバス（USB）、拡張グラフィックスポート（AGP）、パーソナルコンピュータメモリカード国際協会バス（PCMCIA）、および小型コンピュータシステムインタフェース（SCSI）を含むがこれらに限らない。

【 0 0 5 0 】

システムメモリ 8 1 6 には揮発性メモリ 8 2 0 および不揮発性メモリ 8 2 2 が含まれる。例えば起動時にコンピュータ 8 1 2 内部の要素間で情報を送信する基本ルーチンを含む基本入出力システム（BIOS）を、不揮発性メモリ 8 2 2 に格納する。限定ではなく例として、不揮発性メモリ 8 2 2 は読取専用メモリ（ROM）、プログラム可能ROM（PROM）、電気的プログラム可能ROM（EPROM）、電気的消去可能ROM（EEPROM）、またはフラッシュメモリを含むことができる。揮発性メモリ 8 2 0 にはランダムアクセスメモリ（RAM）が含まれ、RAMは外部キャッシュメモリとして動作する。限定ではなく例として、RAMは、同期RAM（SRAM）、ダイナミックRAM（DRAM）、同期DRAM（SDRAM）、ダブルデータレートSDRAM（DDR SDRAM）、拡張SDRAM（ESDRAM）、シンクリンクDRAM（SLDRAM）、およびダイレクトランバスRAM（DRRAM）などの多数の形で利用可能である。

【 0 0 5 1 】

コンピュータ 8 1 2 には取り外し可能/取り外し不能、揮発性/不揮発性コンピュータ記憶媒体も含まれる。図 8 は、例えばディスクストレージ 8 2 4 を示す。ディスクストレージ 8 2 4 は、磁気ディスクドライブ、フロッピー（登録商標）ディスクドライブ、テープドライブ、ジャズドライブ、ジップドライブ、LS-100ドライブ、フラッシュメモリカード、またはメモリスティックなどの装置を含むがこれらに限らない。さらに、ディスクストレージ 8 2 4 は他の記憶媒体とは独立した記憶媒体、または他の記憶媒体と組み合わせた記憶媒体を含むことができる。当該他の記憶媒体は、コンパクトディスクROMデバイス（CD-ROM）などの光ディスクドライブ、CD記録可能ドライブ（CD-Rドライブ）、CD書き換え可能ドライブ（CD-RWドライブ）またはデジタル多用途ディスクROMドライブ（DVD-ROM）を含むがこれらに限らない。ディスク記憶装置 8 2 4 のシステムバス 8 1 8 への接続を容易にするため、インタフェース 8 2 6 などの取り外し可能または取り外し不能インタフェースを一般に使用する。

【 0 0 5 2 】

図 8 は、ユーザと適切な動作環境 8 1 0 内に記載した基本的なコンピュートリソースとの間の媒介物として動作するソフトウェアを説明することは理解されるべきである。このようなソフトウェアにはオペレーティングシステム 8 2 8 が含まれる。オペレーティングシステム 8 2 8 をディスクストレージ 8 2 4 に格納することができ、オペレーティングシステム 8 2 8 はコンピュータシステム 8 1 2 のリソースを制御し、かつ割り当てるよう動作する。システムアプリケーション 8 3 0 は、システムメモリ 8 1 6 またはディスクストレージ 8 2 4 のいずれかに格納したプログラムモジュール 8 3 2 およびプログラムデータ 8 3 4 を通して、オペレーティングシステム 8 2 8 によるリソース管理を生かす。本明細書で説明した様々なコンポーネントを、様々なオペレーティングシステムまたは複数のオ

10

20

30

40

50

ペレーティングシステムの組合せにより実装可能であることは理解されるべきである。

【 0 0 5 3 】

ユーザは入力装置（単数または複数）836を通してコンピュータ812に命令または情報を入力する。入力装置836は、マウスなどのポインティングデバイス、トラックボール、スタイラスペン、タッチパッド、キーボード、マイクロフォン、ジョイスティック、ゲームパッド、パラポラアンテナ、スキャナ、TVチューナカード、デジタルカメラ、デジタルビデオカメラ、ウェブカメラ等を含むがこれらに限らない。これらおよび他の入力装置を、インタフェースポート（単数または複数）838経由でシステムバス818を通して処理ユニット814に接続する。インタフェースポート（単数または複数）838には例えば、シリアルポート、パラレルポート、ゲームポート、およびユニバーサルシリアルバス（USB）が含まれる。出力装置（単数または複数）840は、入力装置（単数または複数）836と同種のポートの一部を使用する。従って例えば、USBポートを使用してコンピュータ812に inputs を提供し、コンピュータ812からの出力情報を出力装置840に提供することができる。出力アダプタ842を提供して、他の出力装置840の中でも特別なアダプタを要するモニタ、スピーカ、およびプリンタのような幾つかの出力装置840があることを示す。限定ではなく例として、出力アダプタ842には、出力装置840とシステムバス818との間の接続手段を提供するビデオカードおよびサウンドカードが含まれる。他の装置および/または装置のシステムがリモートコンピュータ（単数または複数）844などの入力機能と出力機能の両方を提供することに留意されたい。

10

20

【 0 0 5 4 】

コンピュータ812は、リモートコンピュータ844などの1つまたは複数のリモートコンピュータに対する論理接続を用いてネットワーク化環境で動作することができる。リモートコンピュータ844はパーソナルコンピュータ、サーバ、ルータ、ネットワークPC、ワークステーション、マイクロプロセッサベースの機器、ピアデバイスまたは他の共通ネットワークノード等であることができ、一般にコンピュータ812に関して説明した構成要素の多くまたは全てを含む。説明を簡潔にするため、メモリ記憶装置846のみをリモートコンピュータ844とともに示してある。リモートコンピュータ844はネットワークインタフェース848を通してコンピュータ812に論理的に接続し、次いで通信接続850を通して物理的に接続する。ネットワークインタフェース848には、ローカルエリアネットワーク（LAN）および広域ネットワーク（WAN）などの通信ネットワークが含まれる。LAN技術には光ファイバ分散データインタフェース（FDDI）、銅分散データインタフェース（CDDI）、イーサネット（登録商標）/IEEE 802.3、トークンリング/IEEE 802.5等が含まれる。WAN技術はポイントツーポイント接続、統合サービスデジタルネットワーク（ISDN）およびその変形のネットワークなどの回線交換ネットワーク、パケット交換ネットワーク、およびデジタル加入者線（DSL）を含むがこれらに限らない。

30

【 0 0 5 5 】

通信接続（単数または複数）850は、ネットワークインタフェース848をバス818に接続するために使用するハードウェア/ソフトウェアを指し示す。説明を明確にするため、通信接続850をコンピュータ812内部に示しているが、通信接続850はコンピュータ812の外部にあってもよい。例示の目的としてのみ、ネットワークインタフェース848への接続に必要なハードウェア/ソフトウェアには、通常の電話用モデム、ケーブルモデムおよびDSLモデムを含むモデム、ISDNアダプタ、ならびにイーサネット（登録商標）カードといった、内部および外部の技術が含まれる。

40

【 0 0 5 6 】

図9は、使用可能なコンピューティング環境の例900の略ブロック図である。システム900には1つまたは複数のクライアント910が含まれる。1つまたは複数のクライアント910はハードウェアおよび/またはソフトウェア（例えば、スレッド、プロセス、コンピューティングデバイス）であることができる。システム900には1つまたは複

50

数のサーバ930も含まれる。1つまたは複数のサーバ930もハードウェアおよび/またはソフトウェア(例えば、スレッド、プロセス、コンピューティングデバイス)であることができる。サーバ930は例えば本明細書で説明したコンポーネントを使用することで、スレッドを内蔵してトランスフォーメーションを実施することができる。クライアント910およびサーバ930間の1つの可能な通信は、2つまたはそれ以上のコンピュータプロセス間で送信するよう調整したデータパケットの形であることができる。システム900には通信フレームワーク950が含まれ、通信フレームワーク950を使用してクライアント910およびサーバ930間の通信を容易にすることができる。クライアント910を1つまたは複数のクライアントデータストア960に動作可能に接続する。クライアントデータストア960を使用して、クライアント910に局所的な情報を格納することができる。同様に、サーバ930を1つまたは複数のサーバデータストア940に動作可能に接続する。サーバデータストア940を使用して、サーバ930に局所的な情報を格納することができる。

10

【0057】

上述したことには様々な例示的態様が含まれる。勿論、これら態様を説明するために全ての考えうる構成要素または手順の組合せを説明することは不可能であるが、当業界の技術者は多数の異なる組合せおよび置換が可能であることを認識できる。従って、本明細書で説明した態様は、添付した特許請求の範囲の要旨および範囲内にある全ての上記変更、修正および変形を包含することを意図するものである。さらに、発明の詳細な説明または特許請求の範囲のいずれかにおいて用語「含む」を使用する範囲では、このような用語は、特許請求の範囲において使用した用語「備える」のように、用語「備える」と同様な形で包括的であることを意図するものである。

20

【図面の簡単な説明】

【0058】

【図1】ユーザのアクティビティデータまたは相互作用データを使用するデータ処理システムを示す略ブロック図である。

【図2】ユーザのアクティビティデータまたは相互作用データを使用するユーザインタフェースの例を示す図である。

【図3】ユーザアクティビティの決定とプロセスを示すフロー図である。

【図4】ユーザアクティビティのデータ処理とともに使用可能な情報検索アーキテクチャのシステムの例を示す図である。

30

【図5】ユーザのアクティビティデータまたは相互作用データとともに使用可能なユーザモデルの例を示す図である。

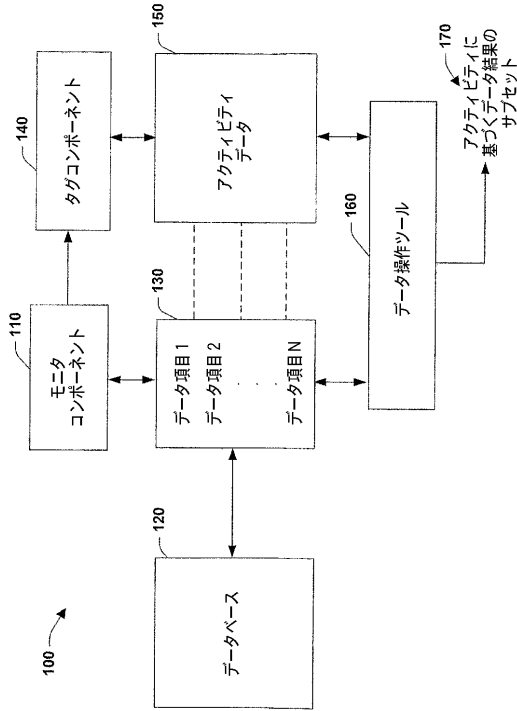
【図6】ユーザのアクティビティデータまたは相互作用データに従ってアクセスベースの情報検索を示すシステム図である。

【図7】ユーザのアクティビティデータまたは相互作用データとともに使用可能な検索サービスアプリケーションを示す図である。

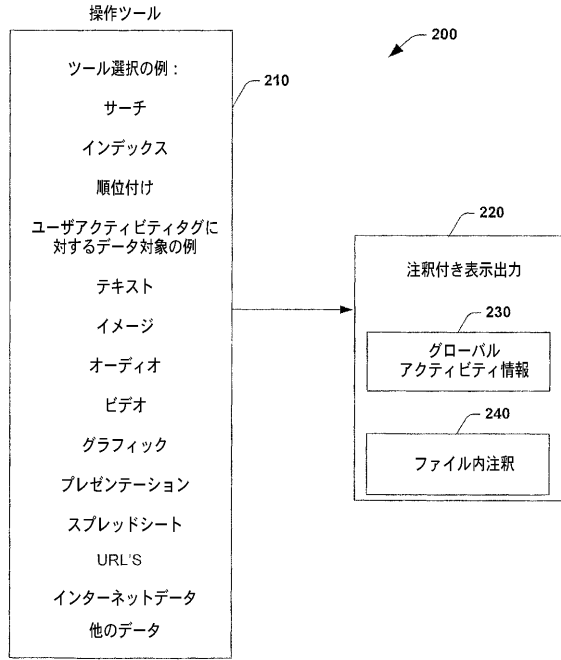
【図8】適切な動作環境を示す略ブロック図である。

【図9】コンピューティング環境の例の略ブロック図である。

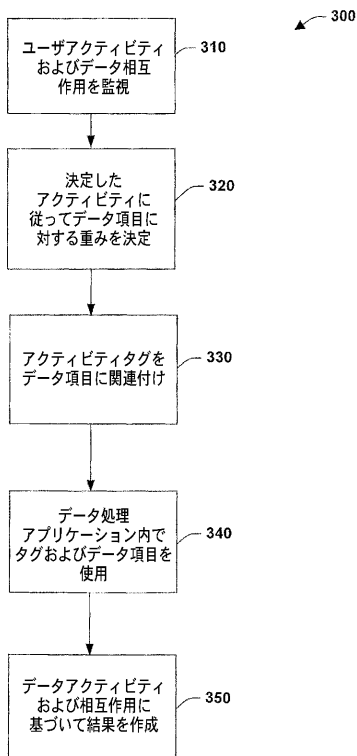
【図1】



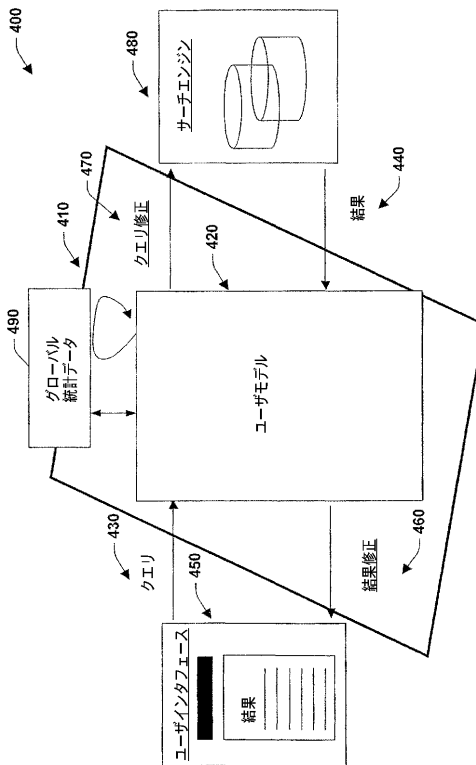
【図2】



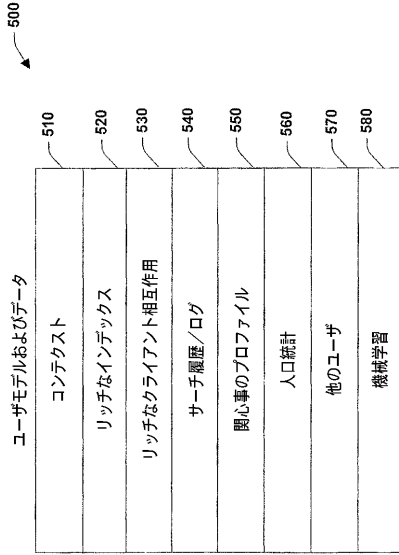
【図3】



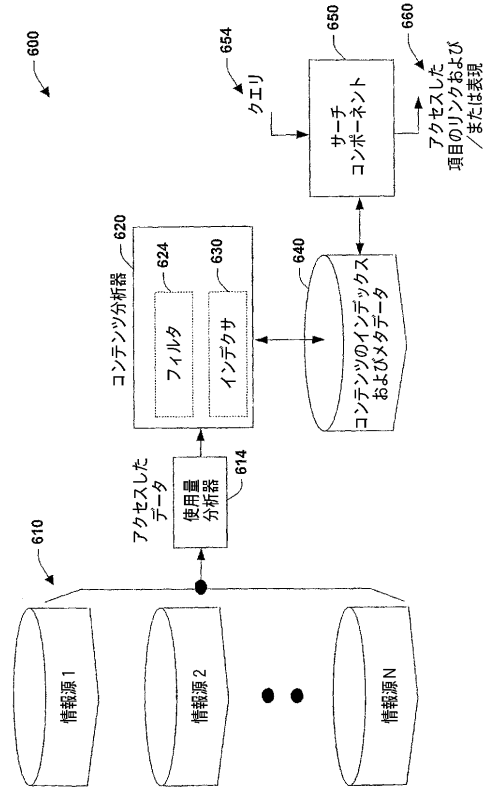
【図4】



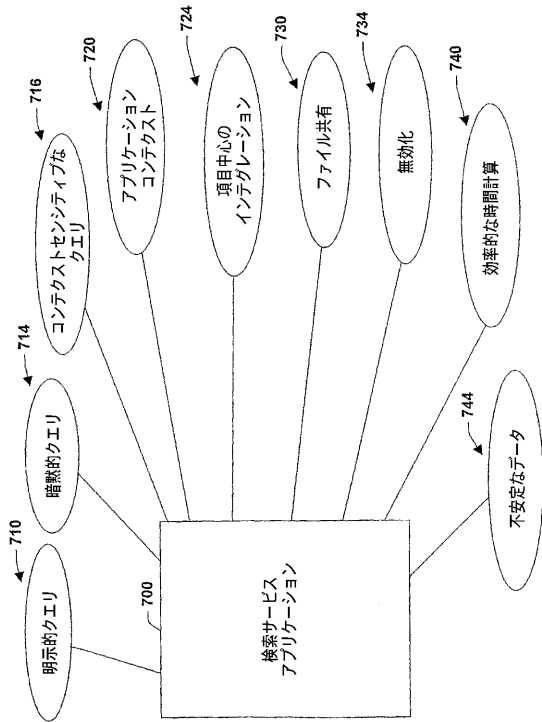
【図5】



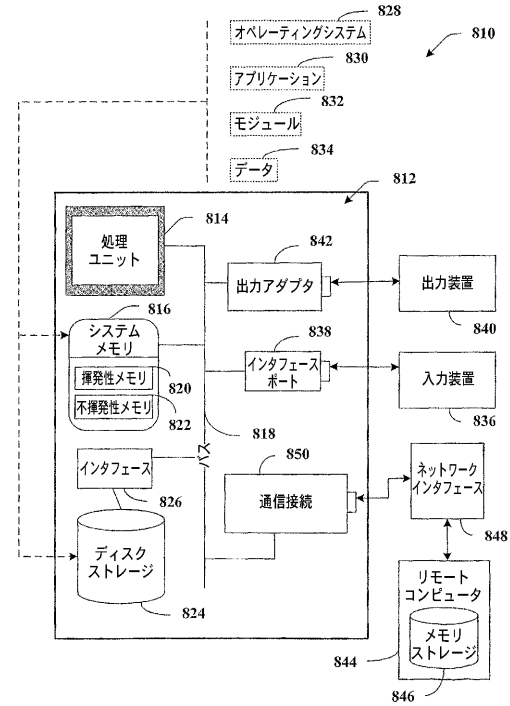
【図6】



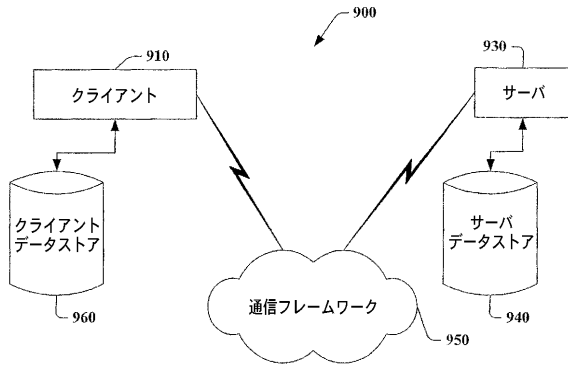
【図7】



【図8】



【図9】



フロントページの続き

(51)Int.Cl. F I
G 0 6 F 13/00 5 4 0 E

(72)発明者 エリック ジェイ・ホルビッツ
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 インターナショナル パテント内

審査官 野崎 大進

(56)参考文献 特開2000-090109(JP,A)
特開2005-031906(JP,A)
特開2003-044511(JP,A)
特開2005-018530(JP,A)

(58)調査した分野(Int.Cl., DB名)
G06F 17/30
G06F 13/00
G06F 17/21
JSTPlus(JDreamII)