

(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2009-252145

(P2009-252145A)

(43) 公開日 平成21年10月29日(2009.10.29)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/30 (2006.01)	G06F 17/30 180A	5B075
G06Q 30/00 (2006.01)	G06F 17/60 326	
G06F 19/00 (2006.01)	G06F 19/00 130	
	G06F 17/30 220Z	
	G06F 17/30 360Z	

審査請求 未請求 請求項の数 10 O L (全 13 頁)

(21) 出願番号 特願2008-102444 (P2008-102444)
 (22) 出願日 平成20年4月10日 (2008.4.10)

(71) 出願人 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (74) 代理人 100089118
 弁理士 酒井 宏明
 (72) 発明者 川村 隆浩
 東京都港区芝浦一丁目1番1号 株式会社東芝内
 (72) 発明者 山崎 智弘
 東京都港区芝浦一丁目1番1号 株式会社東芝内
 (72) 発明者 長野 伸一
 東京都港区芝浦一丁目1番1号 株式会社東芝内

最終頁に続く

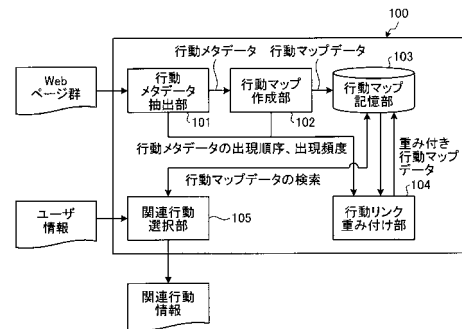
(54) 【発明の名称】 データ作成装置及び方法

(57) 【要約】

【課題】ユーザの現在の行動に関連する関連行動を提示する場合、セキュリティを確保してプライバシーを保護しつつ、ユーザにとって意外性のある行動を提示可能なデータ作成技術を提供する。

【解決手段】行動メタデータ抽出部101は、トピック毎に収集されたWebページから行動メタデータを抽出する。行動マップ作成部102は、行動オントロジー及び商品オントロジーを参照して、行動メタデータから、行動マップデータをトピック毎に作成する。行動リンク重み付け部104は、行動マップデータにおいてリンクに対して重み係数を設定したり新たなリンクを作成したりする。関連行動選択部105は、ユーザ情報に基づいて、行動マップデータを検索し、検索の結果得られた行動マップデータにおいて、ユーザの現在行動に関連する関連行動を選択し、これを示す関連行動情報を出力する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ネットワークを介してアクセス可能な文書から、少なくとも1つのトピックに関し、行動についての言語表現又は当該言語表現を含み係り受け関係のある2つ以上の言語表現を含む行動メタデータを抽出する抽出手段と、

行動についての言語表現をインスタンスとし、該インスタンスの概念をクラスとしてグラフにより表現する行動オントロジーに基づいて、前記行動メタデータに含まれる前記言語表現をクラスに各々変換する変換手段と、

変換された各前記クラスを各々表すと共に、前記クラス間の係り受け関係をリンクとして表す行動マップデータを作成するデータ作成手段とを備える

ことを特徴とするデータ作成装置。

10

【請求項 2】

前記クラス間の係り受け関係の出現頻度及び前記文書から抽出された複数の前記行動メタデータの出現順序のうち少なくとも一方に基づいて、前記リンクに対して重み係数を設定する設定手段を更に備える

ことを特徴とする請求項 1 に記載のデータ作成装置。

【請求項 3】

前記行動マップデータにおいて表される前記クラスと、当該クラスに近似するクラスとの間に新たなリンクを作成するリンク作成手段を更に備える

ことを特徴とするデータ請求項 1 又は 2 に記載のデータ作成装置。

20

【請求項 4】

前記新たなリンクに対して重み係数を設定する新設定手段を更に備える

ことを特徴とするデータ請求項 3 に記載のデータ作成装置。

【請求項 5】

前記行動マップデータを記憶する記憶手段を更に備える

ことを特徴とするデータ請求項 1 乃至 4 のいずれか一項に記載のデータ作成装置。

【請求項 6】

対象のトピックを示すトピック情報と、ユーザの現在の行動を示す行動情報とを取得する取得手段と、

前記トピック情報によって示されるトピックに対応する前記行動マップデータを前記記憶手段において検索して、該当の行動マップデータを得る検索手段と、

前記検索の結果得られた前記行動マップデータにおいて、前記行動情報によって示される前記行動の概念に相当する第1クラスに対する前記リンクに基づいて、前記行動に関連する関連行動を選択する選択手段と、

前記関連行動を示す関連行動情報を出力する出力手段とを更に備える

ことを特徴とするデータ請求項 5 に記載のデータ作成装置。

30

【請求項 7】

前記選択手段は、前記行動マップデータにおいて、前記第1クラスとの間に前記リンクがある他方のクラスを、前記関連行動として選択する

ことを特徴とするデータ請求項 6 に記載のデータ作成装置。

40

【請求項 8】

前記選択手段は、前記第1クラスに対する前記リンクのうち、第1所定値以上の重み係数が設定されたリンクに対する他方のクラスを、前記関連行動として選択する

ことを特徴とするデータ請求項 6 に記載のデータ作成装置。

【請求項 9】

前記選択手段は、前記第1クラスに対する前記リンクのうち、第2所定値以下の重み係数が設定されたリンクに対する他方のクラスを、前記関連行動として選択する

ことを特徴とするデータ請求項 6 に記載のデータ作成装置。

【請求項 10】

抽出手段と、変換手段と、データ作成手段とを備えるデータ作成装置で実現されるデー

50

タ作成方法であって、

前記抽出手段が、ネットワークを介してアクセス可能な文書から、少なくとも1つのトピックに関し、行動についての言語表現又は当該言語表現を含み係り受け関係のある2つ以上の言語表現を含む行動メタデータを抽出する抽出ステップと、

前記変換手段が、行動についての言語表現をインスタンスとし、該インスタンスの概念をクラスとしてグラフにより表現する行動オントロジーに基づいて、前記行動メタデータに含まれる前記言語表現をクラスに各々変換する変換ステップと、

前記データ作成手段が、変換された各前記クラスを各々表すと共に、前記クラス間の係り受け関係をリンクとして表す行動マップデータを作成するデータ作成ステップとを含むことを特徴とするデータ作成方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワークを介してアクセス可能な文書からユーザの行動の記録を抽出し、当該記録に基づいて、関連行動を提示するための行動マップデータを作成するデータ作成装置及び方法に関する。

【背景技術】

【0002】

従来より、例えば、「情報大航海プロジェクト」の1つとして、「マイ・ライフ・アシストサービス」などのライフログ系の研究がなされている(例えば、特許文献1参照)。この研究では、携帯電話やRFIDや各種センサーを活用して、ユーザの行動の記録から何らかの規則性を発見する試みが行われている。そして、発見された規則性に基づいて、ユーザの現在の行動(コンテキスト)やユーザが閲覧しているWebページや動画などのコンテンツに関連する行動(関連行動)を、推薦する行動として提示している。

20

【0003】

【特許文献1】特開2005-71026号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかし、特許文献1の技術のように、ユーザ個人の行動の記録のみに基づいて関連行動を提示する場合、その行動はそのユーザにとって意外性がない恐れがある。また、関連行動を提示するために他人の行動の記録を用いる場合、セキュリティやプライバシーなどの観点から問題があり得る。

30

【0005】

本発明は、ユーザの現在の行動に関連する関連行動を提示する場合、セキュリティを確保してプライバシーを保護しつつ、ユーザにとって意外性のある行動を提示可能なデータ作成装置及び方法を提供することを目的とする。

【課題を解決するための手段】

【0006】

上述した課題を解決し、目的を達成するために、本発明は、データ作成装置であって、ネットワークを介してアクセス可能な文書から、少なくとも1つのトピックに関し、行動についての言語表現又は当該言語表現を含み係り受け関係のある2つ以上の言語表現を含む行動メタデータを抽出する抽出手段と、行動についての言語表現をインスタンスとし、該インスタンスの概念をクラスとしてグラフにより表現する行動オントロジーに基づいて、前記行動メタデータに含まれる前記言語表現をクラスに各々変換する変換手段と、変換された各前記クラスを各々表すと共に、前記クラス間の係り受け関係をリンクとして表す行動マップデータを作成するデータ作成手段とを備えることを特徴とする。

40

【0007】

また、本発明は、抽出手段と、変換手段と、データ作成手段とを備えるデータ作成装置で実現されるデータ作成方法であって、前記抽出手段が、ネットワークを介してアクセス

50

可能な文書から、少なくとも1つのトピックに関し、行動についての言語表現又は当該言語表現を含み係り受け関係のある2つ以上の言語表現を含む行動メタデータを抽出する抽出ステップと、前記変換手段が、行動についての言語表現をインスタンスとし、該インスタンスの概念をクラスとしてグラフにより表現する行動オントロジーに基づいて、前記行動メタデータに含まれる前記言語表現をクラスに各々変換する変換ステップと、前記データ作成手段が、変換された各前記クラスを各々表すと共に、前記クラス間の係り受け関係をリンクとして表す行動マップデータを作成するデータ作成ステップとを含むことを特徴とする。

【発明の効果】

【0008】

本発明によれば、ユーザの現在の行動に関連する関連行動を提示する場合、セキュリティを確保してプライバシーを保護しつつ、ユーザにとって意外性のある行動を提示可能である。

【発明を実施するための最良の形態】

【0009】

以下に添付図面を参照して、この発明にかかるデータ作成装置及び方法の最良な実施の形態を詳細に説明する。

【0010】

(1) 構成

本実施の形態のデータ作成装置100は、装置全体を制御するCPU (Central Processing Unit) 等の制御装置と、各種データや各種プログラムを記憶するROM (Read Only Memory) やRAM (Random Access Memory) 等の記憶装置と、各種データや各種プログラムを記憶するHDD (Hard Disk Drive) やCD (Compact Disk) ドライブ装置等の外部記憶装置と、これらを接続するバスとを備えており、通常のコンピュータを利用したハードウェア構成となっている。また、データ作成装置100には、情報を表示する表示装置と、ユーザの指示入力を受け付けるキーボードやマウス等の入力装置と、外部装置の通信を制御する通信I/F (interface) とが有線又は無線により各々接続される。

【0011】

次に、このようなハードウェア構成において、CPUが記憶装置や外部記憶装置に記憶された各種プログラムを実行することによりデータ作成装置100において実現される各種機能について説明する。尚、記憶装置や外部記憶装置に記憶される各種プログラムは、プログラムコードのみならず、プログラムコードの実行時に用いる各種データも含む。図1は、データ作成装置100の機能的構成を例示する図である。データ作成装置100は、行動メタデータ抽出部101と、行動マップ作成部102と、行動マップ記憶部103と、行動リンク重み付け部104と、関連行動選択部105とを有する。これらのうち、行動メタデータ抽出部101と、行動マップ作成部102と、行動リンク重み付け部104と、関連行動選択部105とは、CPUのプログラム実行時にRAMなどの記憶装置上に生成されるものである。行動マップ記憶部103は、例えば、外部記憶装置に記憶されるものである。また、図示しないが、外部記憶装置には、後述する行動オントロジー及び商品オントロジーが記憶されている。

【0012】

行動メタデータ抽出部101には、あるトピックに関して、ネットワークを介してアクセス可能な文書として収集された複数のWebページが入力される。ここで対象となるWebページは、例えば、ブログやSNSといったCGM (Consumer Generated Media) などである。ネットワークは、例えば、LAN (Local Area Network)、イントラネット、イーサネット (登録商標) 又はインターネットなどである。トピックとは、例えば、商品の名称や、人名や、話題などである。トピックに関するWebページの収集は、例えば、トピックをキーワードとして検索するキーワード型検索エンジンにより行われる。具体的には、例えば、「ハリー」というトピックに関し、これをキーワードとして含むWebページがキーワード型検索エンジンにより収集される。このようにして収集されたWebページが行動メタデー

10

20

30

40

50

タ抽出部 101 に入力されると、行動メタデータ抽出部 101 は、当該Webページから行動メタデータを抽出する。具体的には、行動メタデータ抽出部 101 は、各Webページに記載されている文章から、当該トピックに関し、行動についての言語表現又は当該言語表現を含み係り受け関係のある 2 つ以上の言語表現のリストを行動メタデータとして抽出する。言語表現とは、ターム (term) であり、主語、目的語、補語、及び述語のうち少なくとも 1 つになるものである。行動についてのタームとは、例えば、「見る」「借りる」「購入する」などであり、主に、述語に相当する。例えば、Webページに記載された「ハリリーのDVDが発売日だったってこと今日なんか気が付いて、仕事終わってから、×屋と行って借りてきた」という文では、「ハリー」、「DVD」、「×屋」、「借りてきた」という各タームについて、主語、目的語、補語及び述語という係り受け関係がある。このため、行動メタデータ抽出部 101 は、当該Webページから「ハリー、DVD、×屋、借りてきた」という行動メタデータを抽出する。尚、基本的には、行動メタデータは文単位で抽出されるが、行動についてのタームとして述語を少なくとも含んでいれば良い。

10

【0013】

行動マップ作成部 102 は、以下に説明する行動オントロジー及び商品オントロジーを参照して、行動メタデータ抽出部 101 が抽出した各行動メタデータに含まれる各タームをクラスに各々変換して、当該各クラスをノードとして各々表すと共に、クラス間の係り受け関係をノード間のリンクとして表す行動マップデータをトピック毎に作成する。

【0014】

図 2 は、行動オントロジーのデータ構成を例示する図である。行動オントロジーとは、行動表現をインスタンスとし、このインスタンスの概念をクラスとしてグラフにより表現したものである。同図においては、「do」というクラスに対して、「outer act (外的行動)」及び「inner act (内的行動)」の 2 つのクラスがあり、各クラスに対して更に各クラスが存在している。例えば、「outer act (外的行動)」には、「vision」、「買う」及び「行く」などの各クラスが存在し、「vision」には、更に、「見る」及び「録画する」などの各クラスが存在する。また、「inner act (内的行動)」には、「考える」、「好き」及び「計画する」などの各クラスが存在する。クラス間の関係は、上下関係であったり、前後関係であったり、並列関係であったり、一対一の関係であったり、一対多の関係であったりするが、ここではその関係は特に限定されない。また、「見る」、「買う」、「行く」、「考える」、「好き」及び「計画する」の各クラスには、その概念に収束される各行動表現であるインスタンスが存在する。例えば、「見る」というクラスには、「観る」、「みる」、「鑑賞する」などのインスタンスが存在する。即ち、「見る」という概念に収束される具体的な行動表現のバリエーションが、インスタンスとして存在する。

20

30

【0015】

図 3 は、商品オントロジーのデータ構成を例示する図である。商品オントロジーとは、商品名や商品名の細目や物の名称をインスタンスとし、このインスタンスの概念をクラスとしてグラフにより表現したものである。同図においては、例えば映画についての商品オントロジーを示しており、「映画」というクラスに対して、「洋画」及び「邦画」の 2 つのクラスが存在する。「洋画」というクラスに対して、例えば、映画名として「ハリー」、「ロード」及び「魔法使い」などの各クラスが存在する。そして、例えば、「ハリー」というクラスには、その映画の細目である「ハリー」、「謎」、「炎」及び「不死鳥」などのインスタンスが存在する。

40

【0016】

その他、地名オントロジーが外部記憶装置に記憶されておりこれを用いても良い。地名オントロジーとは、地域名や場所名をインスタンスとし、これらの存在する地名をクラスとしてグラフにより表現したものである。図 4 は、地名オントロジーのデータ構成を例示する図である。同図においては、「世界」というクラスに対して、「外国」及び「日本」の 2 つのクラスが存在する。「日本」というクラスに対しては、例えば、「神奈川」、「東京」、「千葉」などの各クラスが存在する。また、「神奈川」というクラスには、「川

50

崎」、「横浜」、「厨子」などの各クラスが存在する。そして、「川崎」というクラスには、「川崎」、「プラザ」、「××チーム」などのインスタンスが存在する。

【0017】

尚、各オントロジーにおいて、各クラスが概念が近似するか否か、即ち、各クラスが概念が意味的に近いかが予め設定されており、意味的に近いほど位置が近くなるよう各クラスがグラフにおいて配置される。

【0018】

行動マップ作成部102は、以上のような行動オントロジー及び商品オントロジーを少なくとも参照して、各行動メタデータに含まれる各タームについて、クラスに各々変換する。例えば、上述の「ハリー、DVD、××屋、借りてきた」という行動メタデータに含まれる各タームについては、行動マップ作成部102は、左から順に、「ハリー」、「DVD」、「レンタル店」、「見る」という各クラスに変換する。そして、行動マップ作成部102は、これらの各クラスをノードとして各々表すと共に、「ハリー」に対して「DVD」と「レンタル店」が繋がり、「DVD」と「レンタル店」が繋がり、「DVD」と「レンタル店」とに対して「借りる」が繋がるというクラス間の係り受け関係をノード間のリンクとして表す行動マップデータを作成する。

10

【0019】

図5は、行動マップデータのデータ構成を概念的に示す図である。同図に示される行動マップデータは、トピック「ハリー」に関するものである。同図においては、主語、目的語及び補語として、商品オントロジーに存在する各クラスがノードとして表され、述語として、行動オントロジーに存在する各クラスが、その言語表現によりノードとして表される。そして、クラス間の係り受け関係が、ノード間のリンクとして矢印により表されている。尚、ノード間のリンクは、上下関係であったり、前後関係であったり、並列関係であったり、一対一の関係であったり、一対多の関係であったりするが、ここではその関係は特に限定されない。

20

【0020】

行動リンク重み付け部104は、行動マップ作成部102が作成した行動マップデータについて、クラス間の係り受け関係の出現回数に基づいて、リンクに対して重み係数を設定する。例えば、以下の4つの行動メタデータがあるとする。

行動メタデータ1「ハリー、DVD、借りてきた」

行動メタデータ2「ハリー君、DVD、観た」

行動メタデータ3「ロード、DVD、鑑賞した」

行動メタデータ4「ロード、DVD、録画した」

これらに含まれる各タームは、例えば以下のように各クラスに変換される。

行動メタデータ1「ハリー、DVD、借りる」

行動メタデータ2「ハリー、DVD、見る」

行動メタデータ3「ロード、DVD、見る」

行動メタデータ4「ロード、DVD、録画する」

この場合、行動リンク重み付け部104は、各タームから変換された各クラスについてクラス間の係り受け関係の出現回数から、「ハリー」と「DVD」との間のリンク及び「DVD」と「見る」との間のリンクについて、重み係数を各々2倍にする。この重み係数の値は、リンクと対応付けられて行動マップデータにおいて記憶される。

30

40

【0021】

また、行動リンク重み付け部104は、行動メタデータの出現順序に基づいて、リンクに対して重み係数を設定する。行動メタデータの出現順序とは、同一のWebページ内に記載された文章から抽出された複数の行動メタデータについてその元になる文の記載順のことである。例えば、上述の行動メタデータ1と行動メタデータ2とが同一のWebページ内にその順に記載された文から抽出されたものである場合、行動リンク重み付け部104は、その順序から、行動メタデータ1を基にしたノード間のリンクの重み係数及び行動メタデータ2を基にしたノード間のリンクの重み係数を各々1.5倍にする。

50

【0022】

更に、行動リンク重み付け部104は、各クラスの意味的な近さに基づいて、新たなリンクを作成する。例えば、行動メタデータ3に述語として含まれるタームが変換されたクラスである「見る」と、行動メタデータ4に述語として含まれるタームが変換されたクラスである「録画する」とは、行動オントロジーにおいて意味的に近いと予め設定されているものとする。この場合、行動リンク重み付け部104は、「見る」というクラスと、「録画する」というクラスとに繋がりがあるとしてリンクを新たに作成する。尚、リンクとしての前後関係は任意又は所定の方法により決定すれば良い。ここでは、行動リンク重み付け部104は、「録画する」というノードに対して「見る」というノードにリンクを新たに作成する。尚、行動リンク重み付け部104は、新たに作成したリンクに対して、更に重み係数を設定するようにしても良い。

10

【0023】

行動リンク重み付け部104は、このようにリンクに重み係数を設定したりリンクを新たに作成したりすることにより更新した行動マップデータを行動マップ記憶部103に記憶する。図6は、更新された行動マップデータのデータ構成を概念的に示す図である。同図においては、例えば、「映画」というノードと「見る」というノードとを繋ぐ矢印が他の矢印より太く示されており、当該ノード間のリンクに重み係数が設定されていることが示されている。

【0024】

関連行動選択部105には、ユーザ情報が入力される。尚、ユーザ情報とは、ユーザが現在閲覧しているWebページのトピックを示すトピック情報や、ユーザが現在行っている行動（現在行動という）を示す行動情報や、現在位置を示す位置情報や、それらの過去の履歴を示す履歴情報や、ユーザの年齢や性別などのプロフィール情報などである。位置情報は、例えば、GPS(Global Positioning System)により取得され得る。但し、位置情報を含め行動情報やプロフィール情報や履歴情報の取得方法は限定されず、これらが、任意の方法により取得され、例えば、通信I/Fを介して外部装置からデータ作成装置100に入力される。このようなユーザ情報が入力されると、関連行動選択部105は、当該ユーザ情報に基づいて、行動マップデータを検索し、検索の結果得られた行動マップデータにおいて、ユーザの現在行動に関連する行動（関連行動）を選択し、これを示す情報を関連行動情報として出力する。

20

30

【0025】

具体的には、関連行動選択部105は、まず、ユーザ情報に含まれるトピック情報によって示されるトピックを検索キーとして行動マップデータを検索する。そして、関連行動選択部105は、検索の結果得られた行動マップデータにおいて、ユーザ情報に含まれる行動情報や位置情報や履歴情報やプロフィール情報に基づいて、現在行動に関連する行動（関連行動）を選択する。具体的には例えば、関連行動選択部105は、行動マップデータにおいて、リンクを辿り、当該現在行動の概念に相当するノードや現在位置に対応するノードを特定し、特定したノードに対するリンクのうち、所定値以上の重み係数が設定されたリンクに対する他方のノードを関連行動として選択する。尚、行動マップデータにおけるリンクの辿り方については特に限定されない。そして、関連行動選択部105は、関連行動を示す関連行動情報を出力する。

40

【0026】

(2) 動作

次に、本実施の形態にかかるデータ作成装置100の行う行動マップデータ作成処理の手順について図7を用いて説明する。データ作成装置100にあるトピックについて収集されたWebページが入力されると（ステップS1）、データ作成装置100は、各Webページから行動メタデータを抽出する（ステップS2）。次いで、データ作成装置100は、行動オントロジー及び商品オントロジーを参照して、ステップS1で抽出された各行動メタデータに含まれる各タームをクラスに各々変換して、当該各クラスをノードとして各々表すと共に、クラス間の係り受け関係をノード間のリンクとして表す行動マップデータを

50

トピック毎に作成する(ステップS3)。ここで作成される行動マップデータのデータ構成の概念的なものは図5に示したとおりである。次いで、データ作成装置100は、ステップS3で作成した行動マップデータについて、係り受け関係のある2つ以上のクラスの出現回数や行動メタデータの出現順序に基づいて、リンクに対して重み係数を設定したり、各クラスの意味的な近さに基づいて、新たなリンクを作成したりする(ステップS4)。そして、データ作成装置100は、行動マップデータを、ステップS4でリンク重み係数を設定したり新たにリンクを作成したりすることにより更新した行動マップデータを行動マップ記憶部103に記憶する(ステップS5)。ここで更新された行動マップデータのデータ構成の概念的なものは図6に示したとおりである。

【0027】

次いで、データ作成装置100が、行動マップ記憶部103に記憶された行動マップデータを用いて関連行動情報を出力する処理の手順について図8を用いて説明する。データ作成装置100にユーザ情報が入力されると(ステップS10)、データ作成装置100は、当該ユーザ情報に含まれるトピック情報によって示されるトピックを検索キーとして、行動マップ記憶部103に記憶された行動マップデータを検索する(ステップS11)。そして、データ作成装置100は、検索の結果得られた行動マップデータにおいて、ユーザ情報に含まれる、現在行動を示す行動情報に基づいて、当該現在行動に関連する行動(関連行動)を選択する(ステップS12)。そして、データ作成装置100は、当該関連行動を示す関連行動情報を出力する(ステップS13)。

【0028】

例えば、トピック情報によって示されるトピックが「ハリー」であり、行動情報によって示される現在行動が「観る」ことであり、位置情報によって示される現在位置が「×プラザ」であるとする。「×プラザ」は、図4に例示した地名オントロジーにおいて「川崎」というクラスに相当するものとする。「観る」という概念に相当するクラスは「見る」である。この場合、「ハリー」というノードに対して「川崎」というノードが繋がり、「川崎」というノードに対して「見る」というノードが繋がることになる。この場合、例えば、図9に示されるような行動マップデータにおいて、各ノードと、ノード間のリンクとが特定できる。そして、同図に示されるように、ここで特定された「見る」というノードに対するリンクのうち、「寝る」というノードに対するリンクがあるとする。この場合、データ作成装置100は、「寝る」というノードを関連行動として選択する。即ち、「ハリーを×プラザで見ている」と推測されるユーザに対して、それに関連する関連行動が選択されるのである。尚、データ作成装置100は、リンクに設定された重み係数が所定値以上である場合にのみ当該ノードを関連行動として選択するようにしても良い。そして、データ作成装置100は、「寝る」という関連行動を示す関連行動情報として、例えば、「そろそろ寝ますか？」などのメッセージを出力する。この関連行動情報をデータ作成装置100は、例えば、通信I/Fを介して携帯情報端末などの外部装置に出力するようにしても良いし、表示装置に出力するようにしても良い。

【0029】

以上のようにして、データ作成装置100が、CGMなどユーザ自身のみならず不特定多数のユーザが記載したブログなどの複数のWebページから行動メタデータを抽出することにより、ノイズの少ないデータを抽出することができる。そして、この行動メタデータを用いてデータ作成装置100が行動マップデータを作成することにより、行動のルールなどを手動で登録する必要がなくなる。このため、システムの構築や保守にかかるコストを低減することができる。

【0030】

また、商品オントロジーや行動オントロジーを参照して、行動メタデータを用いて、同一の概念に収束されるものをクラスとしてグループ化して行動マップデータを作成することにより、個人を特定する情報や秘匿されるべき情報を含むことなく行動マップデータを作成することができ、セキュリティを確保したりプライバシーを保護したりすることができる。また、以上のようにして行動マップデータを作成することにより、対象のトピック

10

20

30

40

50

に相応する行動メタデータが存在していなくても、類似のトピックに基づいて関連行動を選択して、関連行動情報を出力することができる。これにより、例えば、対象のトピックがマイナーな話題などであり、Webページから十分な行動メタデータが得られない場合であっても、ユーザになんらかの関連行動を提示することができる。

【0031】

また、行動マップデータにおいてリンクに重み係数を設定することで、リンクを辿ったりリンクの重み係数を参酌したりすることにより、規則性のある行動や関連行動を容易に検出することができる。従って、ユーザがしたいと思われる行動を効果的に選択して提示することができる。

【0032】

また、Webページに記載されている文章における行動の連続性として、当該文章から抽出した行動メタデータの出現順序を考慮して、クラスとして表現される行動間の繋がりを表すリンクに重み係数を設定することで、役に立つがありがちな行動と、意外性のある行動とを識別可能にする。このため、前者の行動をユーザに提示することにより、実用性をユーザに提示できると共に、後者の行動をユーザに提示することにより、面白さをユーザに提示することができる。

【0033】

即ち、上述の実施の形態においては、リンクに設定された重み係数が所定値以上であるノードを関連行動として選択したが、逆に、リンクに設定された重み係数が所定値より小さいノード、即ち、関連性が弱い行動を選択することもできる。そして、この行動を示す情報を出力することにより、ユーザが少し意外に感じる行動を提示することができる。

【0034】

尚、ユーザ個人に拠らない意外性のある行動の提示は、上述したようにユーザ自身のみならず不特定多数のユーザが記載したブログなどのWebページから行動メタデータを抽出することによっても実現される。

【0035】

[変形例]

なお、本発明は前記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、前記実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。また、以下に例示するような種々の変形が可能である。

【0036】

上述した実施の形態において、データ作成装置100で実行される各種プログラムを、インターネット等のネットワークに接続されたコンピュータ上に格納し、ネットワーク経由でダウンロードさせることにより提供するように構成しても良い。また、当該各種プログラムを、インストール可能な形式又は実行可能な形式のファイルでCD-ROM、フレキシブルディスク(FD)、CD-R、DVD(Digital Versatile Disk)等のコンピュータで読み取り可能な記録媒体に記録して提供するように構成しても良い。

【0037】

上述した実施の形態においては、Webページを収集する方法や行動メタデータを抽出する方法は特に限定されない。

【0038】

上述した実施の形態においては、トピックは、複数であっても良く、複数のトピック毎に行動メタデータを抽出したり、行動マップを作成したりしても良い。

【0039】

上述した実施の形態においては、行動オントロジー及び商品オントロジーは上述したものに限らない。また、その他様々なものについての言語表現の概念をクラスとクラス間の相互関係をグラフにより表現するオントロジーを用いて、行動マップデータを作成する

10

20

30

40

50

ようにしても良い。

【0040】

上述した実施の形態においては、行動リンク重み付け部104は、新たに作成したリンクに対して、更に、重み係数を設定するようにしても良い。

【0041】

上述した実施の形態において、データ作成装置100がステップS13で出力する関連行動情報は、関連行動自体を示すものに限らず、例えば、関連行動に関連する情報であっても良い。データ作成装置100は、例えば、上述の例において「寝る」という関連行動に関連する情報として、癒し系グッズや快眠グッズに関する情報を出力するようにしても良いし、当該情報を例えばWebページとして提供する情報提供サイトとなる外部装置へネットワークを介して接続するようにしても良い。

10

【図面の簡単な説明】

【0042】

【図1】一実施の形態にかかるデータ作成装置100の機能的構成を例示する図である。

【図2】同実施の形態にかかる行動オントロジーのデータ構成を例示する図である。

【図3】同実施の形態にかかる商品オントロジーのデータ構成を例示する図である。

【図4】同実施の形態にかかる地名オントロジーのデータ構成を例示する図である。

【図5】同実施の形態にかかる行動マップデータのデータ構成を概念的に示す図である。

【図6】同実施の形態にかかる行動マップデータのデータ構成を概念的に示す図である。

【図7】同実施の形態にかかるデータ作成装置100の行う行動マップデータ作成処理の手順を示すフローチャートである。

20

【図8】同実施の形態にかかるデータ作成装置100が行動マップデータを用いて関連行動情報を出力する処理の手順を示すフローチャートである。

【図9】同実施の形態にかかる行動マップデータにおける関連行動について説明するための図である。

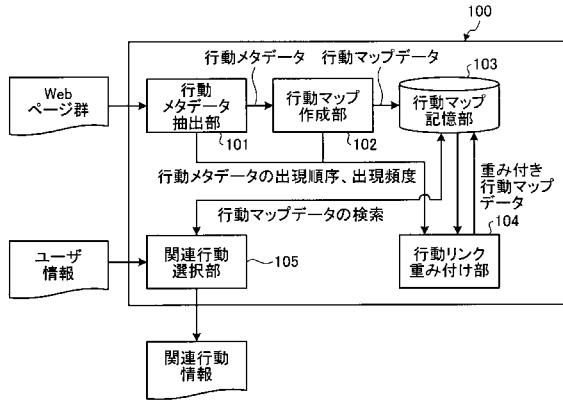
【符号の説明】

【0043】

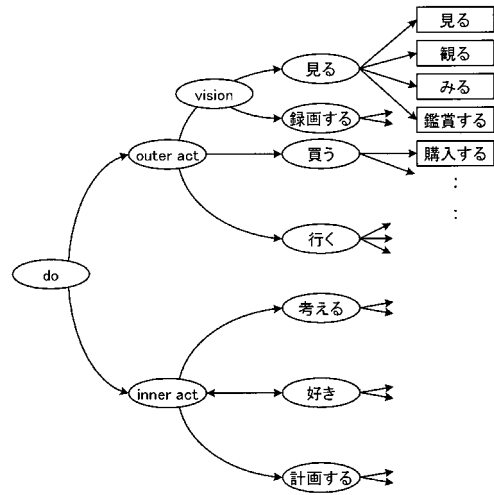
- 100 データ作成装置
- 101 行動メタデータ抽出部
- 102 行動マップ作成部
- 103 行動マップ記憶部
- 104 行動リンク重み付け部
- 105 関連行動選択部

30

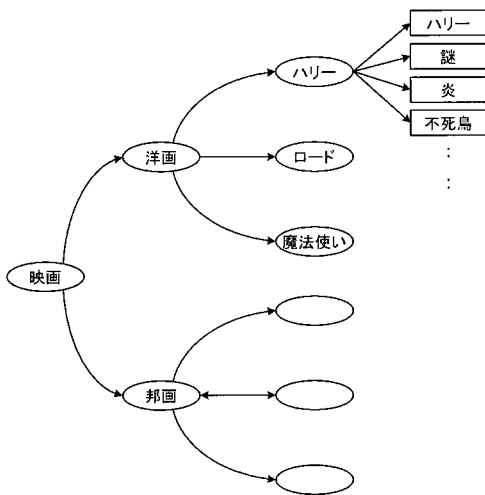
【 図 1 】



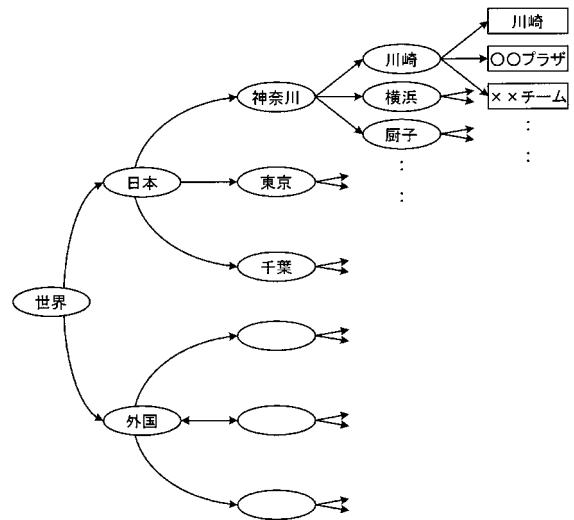
【 図 2 】



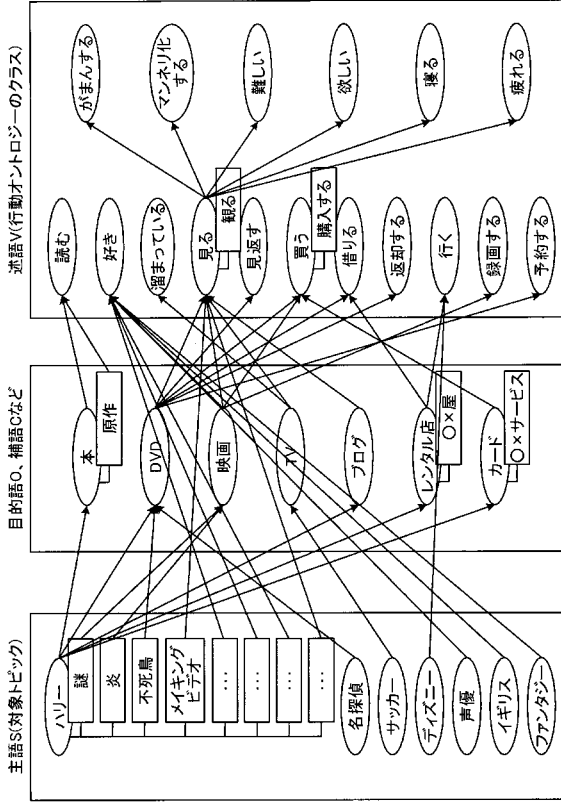
【 図 3 】



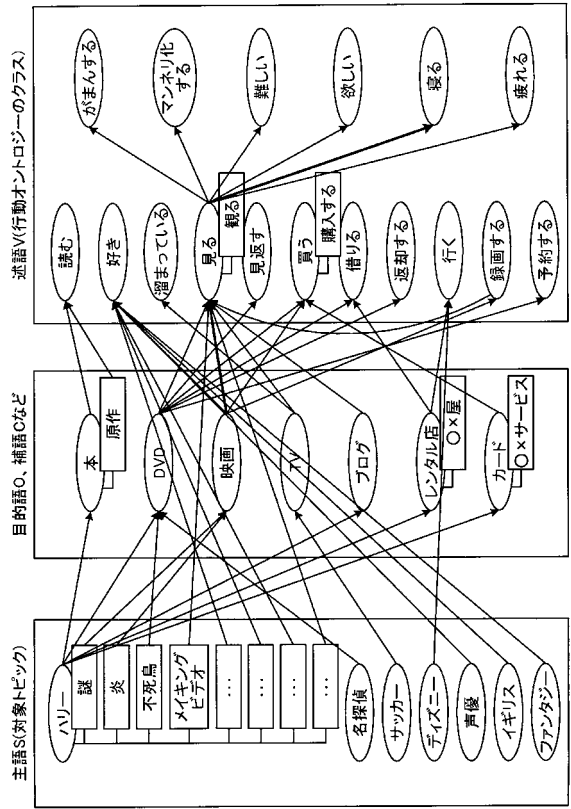
【 図 4 】



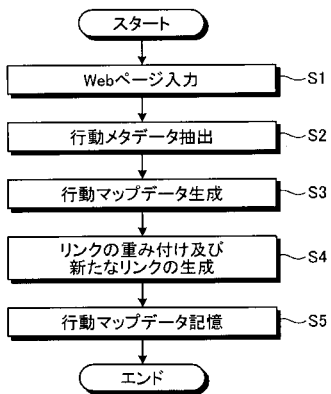
【 図 5 】



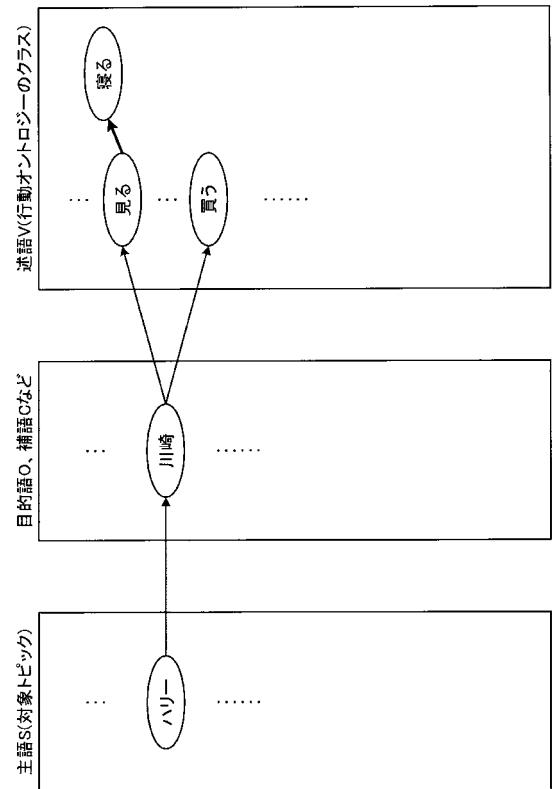
【 図 6 】



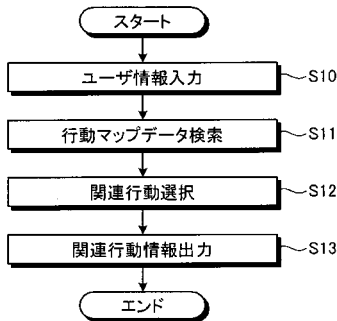
【 図 7 】



【 図 9 】



【 図 8 】



フロントページの続き

- (72)発明者 下郡 祐美子
東京都港区芝浦一丁目1番1号 株式会社東芝内
- (72)発明者 飯田 貴之
東京都港区芝浦一丁目1番1号 株式会社東芝内
- Fターム(参考) 5B075 NK44 NS10 PQ02 PQ13 QP01