

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6020517号
(P6020517)

(45) 発行日 平成28年11月2日 (2016. 11. 2)

(24) 登録日 平成28年10月14日 (2016. 10. 14)

(51) Int. Cl.	F I				
G06F 17/27 (2006.01)	G06F	17/27	665		
G06F 17/22 (2006.01)	G06F	17/22	635		
G06F 17/30 (2006.01)	G06F	17/30	170A		
G06Q 30/02 (2012.01)	G06F	17/30	220Z		
	G06F	17/30	419B		
請求項の数 16 (全 37 頁) 最終頁に続く					

(21) 出願番号	特願2014-122224 (P2014-122224)	(73) 特許権者	390002761 キヤノンマーケティングジャパン株式会社 東京都港区港南2丁目16番6号
(22) 出願日	平成26年6月13日 (2014. 6. 13)	(73) 特許権者	592135203 キヤノンITソリューションズ株式会社 東京都品川区東品川2丁目4番11号
(65) 公開番号	特開2015-64864 (P2015-64864A)	(74) 代理人	100189751 弁理士 木村 友輔
(43) 公開日	平成27年4月9日 (2015. 4. 9)	(74) 代理人	100188938 弁理士 榎葉 加奈子
審査請求日	平成26年9月30日 (2014. 9. 30)	(72) 発明者	田中 靖大 東京都品川区東品川2丁目4番11号 キヤノンITソリューションズ株式会社内
(31) 優先権主張番号	特願2013-175085 (P2013-175085)	審査官	成瀬 博之
(32) 優先日	平成25年8月26日 (2013. 8. 26)		最終頁に続く
(33) 優先権主張国	日本国(JP)		

(54) 【発明の名称】 情報処理装置、情報処理システム、評価分析方法、制御方法、及びプログラム

(57) 【特許請求の範囲】

【請求項1】

サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかると検索条件を用いて検索する検索手段と、

前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、

前記評価対象を特定するための特定サイトを記憶する記憶手段と、

前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、

前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、

を備えたことを特徴とする情報処理装置。

【請求項2】

前記取得手段で取得した文書に含まれるリンクのリンク先は外部サイトとなるリンクであることを特徴とする請求項1に記載の情報処理装置。

【請求項3】

前記重要度は、前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が

掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に対する、前記文書の数によって求められることを特徴とする請求項 1 又は 2 に記載の情報処理装置

【請求項 4】

前記特定手段によって前記リンク先のサイトに基づいて、前記リンク元の前記文書の評価対象を特定できない場合、前記取得手段は、前記リンク先のサイトの前記文書からリンクを取得することを特徴とする請求項 1 乃至 3 の何れか 1 項に記載の情報処理装置。

【請求項 5】

前記取得手段は、リンク先のサイトが前記記憶手段に記憶された特定サイトでない場合は、前記リンク先のサイトの前記文書からリンクを取得することを特徴とする請求項 1 乃至 4 の何れか 1 項に記載の情報処理装置。

10

【請求項 6】

前記特定手段は、前記取得手段によって取得したリンク先のサイトの前記文書が、リンク元の前記文書に対して外部のサイトである場合に、前記リンク元の前記文書の評価対象を特定することを特徴とする請求項 1 乃至 5 の何れか 1 項に記載の情報処理装置。

【請求項 7】

前記取得手段は、リンク先のサイトの文書におけるリンクを取得する回数が所定回数に至るまで、前記リンク先のサイトに基づいて、前記リンク元の前記文書の評価対象を特定できない場合は、前記リンク先のサイトの前記文書からリンクを取得することを特徴とする請求項 4 乃至 6 の何れか 1 項に記載の情報処理装置。

【請求項 8】

20

前記取得手段は、リンク先のサイトの文書におけるリンクを取得する回数が所定回数に至るまで、前記リンク先のサイトが前記記憶手段に記憶された特定サイトでない場合は、前記リンク先のサイトの前記文書からリンクを取得することを特徴とする請求項 4 乃至 6 の何れか 1 項に記載の情報処理装置。

【請求項 9】

前記特定手段は、前記取得手段によって取得するリンク先のサイトの前記文書とリンク元の前記文書とから求まる関連度が、規定値以上となる当該リンク先を用いて、前記リンク元の前記文書の評価対象を特定することを特徴とする請求項 1 乃至 8 の何れか 1 項に記載の情報処理装置。

【請求項 10】

30

前記特定手段は、前記取得手段によって取得するリンク先のサイトの文書内容と、リンク元の文書内容とから求まる関連度を用いて、前記リンク元の前記文書の評価対象を特定することを特徴とする請求項 9 に記載の情報処理装置。

【請求項 11】

前記特定手段は、前記取得手段によってリンクが取得できない場合、前記リンク元の文書のタイトルを用いて、前記リンク元の前記文書の評価対象を特定することを特徴とする請求項 6 または請求項 1 乃至 10 の何れか 1 項に記載の情報処理装置。

【請求項 12】

評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムであって、

40

前記情報処理装置は、

前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、

前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、

前記評価対象を特定するための特定サイトを記憶する記憶手段と、

前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、

50

前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、

前記特定手段によって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信手段と、

を備え、

前記クライアント端末は、

前記送信手段によって送信された特定情報を受信する受信手段と、

前記受信手段によって受信した特定情報を出力する出力手段と、

を備えたことを特徴とする情報処理システム。

10

【請求項 13】

情報処理装置における評価分析方法であって、

前記情報処理装置の検索手段は、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索ステップ、

前記情報処理装置の取得手段は、前記検索ステップによって検索した文書に含まれるリンクを取得する取得ステップ、

前記情報処理装置の判定手段は、前記取得ステップで取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定ステップ、

20

前記情報処理装置の特定手段は、前記判定ステップによって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定ステップ、

を含むことを特徴とする評価分析方法。

【請求項 14】

評価分析を行う情報処理装置において読取実行可能なプログラムであって、

前記情報処理装置を、

サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、

前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、

前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、

30

前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、

40

して機能させることを特徴とするプログラム。

【請求項 15】

評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムの制御方法であって、

前記情報処理装置は、

前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索ステップと、

前記検索ステップによって検索した文書に含まれるリンクを取得する取得ステップと、

50

前記取得ステップで取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定ステップと、

前記判定ステップによって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定ステップと、

前記特定ステップによって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信ステップと、

を実行し、

前記クライアント端末は、

前記送信ステップによって送信された特定情報を受信する受信ステップと、

前記受信ステップによって受信した特定情報を出力する出力ステップと、

を実行することを特徴とする情報処理システムの制御方法。

【請求項16】

評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムで読取実行可能なプログラムあって、

前記情報処理装置を、

前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、

前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、

前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、

前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、

前記特定手段によって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信手段と、

して機能させ、

前記クライアント端末を、

前記送信手段によって送信された特定情報を受信する受信手段と、

前記受信手段によって受信した特定情報を出力する出力手段と、

して機能させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、SNS上における企業のブランドや製品に対する意見や評価を分析する情報処理装置、情報処理システム、評価分析方法、制御方法、及びプログラムに関する。

【背景技術】

【0002】

以前より、インターネット上における個人の発言を定期的に収集し、ブランドや製品に対する意見や評価に関する記述を抽出し、言語的あるいは統計的に分析することで、マーケティング、需要予測、株価予測などに活用する様々な提案がなされていた。

【0003】

近年、スマートフォンの普及とSNSの利用者数増加に伴い、ネット上での評判を把握

10

20

30

40

50

することは企業にとってより重要な課題になってきている。

【0004】

このような課題を解決するにあたり、予め調査対象となる企業に関するキーワード（企業名など）や企業サイトへのリンク（URL）を登録し、ブログなどから抽出した文言に対し、前述したキーワードや企業サイトへのリンク（URL）があれば調査対象に関する言述と判断して、インターネット上における企業の評判を分析する手法が提示されている（例えば、特許文献1参照）。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特願2003-548126号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

しかしながら、特許文献1に記載の手法においては、評価の単位が企業全体であることが想定されており、製品や企業の施策（告知やイベント）ごとに評価することは想定されていない。

【0007】

例えば、特許文献1において、企業名や企業サイトのURLなど予め用意した辞書に基づいて言述の対象を判断しており、企業全体を分析の単位とする程度ならよいが、対象を拡張して製品などの分析に適用することを考えると、必要な辞書の事前準備だけでも非常に煩雑となり、更に新たな製品が出現することまで鑑みると、辞書を適正に維持することは非常に困難である。

【0008】

また、ニュースサイトの記事など、企業サイト以外で、調査対象について記載された文書へのリンク（URL）等については、情報ソースの提示または信頼度の判定に使うことしか想定されておらず、対象の特定に利用されていない。

【0009】

従って、企業のサイトにおいては、製品情報や施策（ニュースリリースやイベント告知等）が提示されるが、これに対し、ニュースサイトやブログにおいて関連する記事が作成されるため、SNSの投稿において、投稿者の情報の入手先により、同じ主題（製品や施策）に対する投稿でも、情報源として付与されるリンク（URL）が異なることにもなり得るが、このような事象に対応し得ることに關しては、記載や示唆はない。

【0010】

更に、このような態様であれば、SNSの投稿においては、文字数の制限や、モバイル環境からの投稿など、文章が短くなる要因が多く、このような企業外のサイトへのリンク以外から対象を特定することが困難な場合も多い。

【0011】

そこで、本発明においては、企業や製品等に関する意見や評価等を分析するにあたり、その意見や評価等の対象を適切に特定することで、妥当な評価を行うことの支援をすることが可能な情報処理装置、情報処理システム、評価分析方法、制御方法、及びプログラム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】

【0012】

上記目的を達成するための第1の発明は、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、前記評価対象を特定するための特定サイトを記憶する記憶手段と、前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得した

10

20

30

40

50

リンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、を備えたことを特徴とする情報処理装置である。

【0013】

上記目的を達成するための第2の発明は、評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムであって、前記情報処理装置は、前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、前記評価対象を特定するための特定サイトを記憶する記憶手段と、前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、前記特定手段によって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信手段と、を備え、前記クライアント端末は、前記送信手段によって送信された特定情報を受信する受信手段と、前記受信手段によって受信した特定情報を出力する出力手段と、を備えたことを特徴とする情報処理システムである。

【0014】

上記目的を達成するための第3の発明は、情報処理装置における評価分析方法であって、前記情報処理装置の検索手段は、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索ステップ、前記情報処理装置の取得手段は、前記検索ステップによって検索した文書に含まれるリンクを取得する取得ステップ、前記情報処理装置の判定手段は、前記取得ステップで取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定ステップ、前記情報処理装置の特定手段は、前記判定ステップによって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定ステップ、を含むことを特徴とする評価分析方法である。

【0015】

上記目的を達成するための第4の発明は、評価分析を行う情報処理装置において読取実行可能なプログラムであって、前記情報処理装置を、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、して機能させることを特徴とするプログラムである。

【0016】

上記目的を達成するための第5の発明は、評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムの制御方法であって、前記情報処理装置は、前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索ステップと、前記検索ステップによって検索した文書に含まれるリンクを取得する取得ステップと、前記取得ステップで取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定ステップと、前記判定ステップによって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定ステップと、前記特定ステップによって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信ステップと、を実行し、前記クライアント端末は、前記送信ステップによって送信された特定情報を受信する受信ステップと、前記受信ステップによって受信した特定情報を入力する出力ステップと、を実行することを特徴とする情報処理システムの制御方法である。

10

【0017】

上記目的を達成するための第6の発明は、評価の依頼を行うためのクライアント端末とネットワークを介して接続された情報処理装置からなる情報処理システムで読取実行可能なプログラムあって、前記情報処理装置を、前記クライアント端末から評価の依頼を受付けると、サイトに掲載された文書から評価対象に対して評価が掲載された文書を、当該評価対象にかかる検索条件を用いて検索する検索手段と、前記検索手段によって検索した文書に含まれるリンクを取得する取得手段と、前記取得手段で取得した文書に含まれるリンクと一致する、前記文書が掲載されたサイトと同一サイトに掲載された他の文書に含まれるリンクの数に基づいて求まる重要度が規定値を満たす場合、当該取得したリンクのリンク先のサイトが評価対象の特定にかかるサイトであると判定する判定手段と、前記判定手段によって評価対象の特定にかかるサイトであると判定されたリンク先のサイトが前記評価対象を特定するための特定サイトを記憶する記憶手段に記憶された特定サイトである場合に、当該特定サイトに対応する情報または当該リンク先の情報に基づいて、リンク元の前記文書の評価対象を特定する特定手段と、前記特定手段によって評価対象が特定された文書に関する特定情報をクライアント端末へ送信する送信手段と、して機能させ、前記クライアント端末を、前記送信手段によって送信された特定情報を受信する受信手段と、前記受信手段によって受信した特定情報を入力する出力手段と、して機能させることを特徴とするプログラムである。

20

30

【発明の効果】

【0018】

本発明に依れば、インターネット等で展開されるサービスに於いて行われる企業や製品等に関する意見や評価等の投稿に含まれるリンクに直接、あるいは間接問わず関連付けられる企業サイト（到達する企業サイトを用いて評価を行う）ことで、分析対象となる企業や製品等の範囲を幅広く抑えると共に、その意見や評価等の対象を適切に特定することで、妥当な評価を行うための支援をすることができる、という効果を奏する。

40

【図面の簡単な説明】

【0019】

【図1】本発明の実施形態に係る情報処理システムの構成の一例を示すシステム構成図である。

【図2】本発明の実施形態に係る情報処理装置、及び分析者端末のハードウェア構成を示すブロック図である。

【図3】本発明の実施形態において、各サイトへの情報収集及び解析処理を示すフローチ

50

ャートである。

【図4】本発明の実施形態において、SNS投稿収集処理の一例を示すフローチャートである。

【図5】本発明の実施形態において、設定保存領域の構成を示すブロック図である。

【図6】本発明の実施形態において、SNSデータ保存領域の構成を示すブロック図である。

【図7】本発明の実施形態において、設定保存領域の検索条件テーブルの構成を示す構成図である。

【図8】本発明の実施形態において、SNSデータ保存領域の投稿テーブルの構成を示す構成図である。

10

【図9】本発明の実施形態において、リンク解析処理の一例を示すフローチャートである。

【図10】本発明の実施形態において、リンク情報保存領域の構成を示すブロック図である。

【図11】本発明の実施形態において、文書内のリンク解析処理の一例を示すフローチャートである。

【図12】本発明の実施形態において、リンク情報追加処理の一例を示すフローチャートである。

【図13】本発明の実施形態において、情報処理装置における一時領域の一例を示す図である。

20

【図14】本発明の実施形態において、リンク先の文書の一例を示す図である。

【図15】本発明の実施形態において、設定保存領域の公式サイトURLパターンテーブルの構成を示す構成図である。

【図16】本発明の実施形態において、リンク情報保存領域のリンク情報テーブルの構成を示す構成図である。

【図17】本発明の実施形態において、リンク情報保存領域のタグ情報テーブルの構成を示す構成図である。

【図18】本発明の実施形態において、情報処理装置における一時領域の一例を示す図である。

【図19】本発明の実施形態において、リンク先の文書の一例を示す図である。

30

【図20】本発明の実施形態において、リンク先の文書の一例を示す図である。

【図21】本発明の実施形態において、リンク情報保存領域のリンク情報テーブルの構成を示す構成図である。

【図22】本発明の実施形態において、リンク情報保存領域のタグ情報テーブルの構成を示す構成図である。

【図23】本発明の実施形態において、リンク情報保存領域のリンク情報テーブルにおける到達URLによる投稿数の集計結果の一例を示す図である。

【図24】本発明の実施形態において、リンク情報保存領域のタグ情報テーブルにおけるタグによる投稿数の集計結果の一例を示す図である。

【図25】本発明の実施形態において、リンク情報保存領域の構成を示すブロック図である。

40

【図26】本発明の実施形態において、リンク解析処理の一例を示すフローチャートである。

【図27】本発明の実施形態において、文書内のリンク解析処理の一例を示すフローチャートである。

【図28】本発明の実施形態において、リンク元の文書の一例を示す図である。

【図29】本発明の実施形態において、関係情報テーブルの構成を示す構成図である。

【図30】本発明の実施形態において、リンク解析処理の一例を示すフローチャートである。

【図31】本発明の実施形態において、リンク元の文書の一例を示す図である。

50

【図32】本発明の実施形態において、リンク情報追加処理の一例を示すフローチャートである。

【図33】本発明の実施形態において、リンク情報保存領域のタグ情報テーブルの構成を示す構成図である。

【図34】本発明の実施形態において、リンク情報保存領域のタグ情報テーブルにおけるタグによる投稿数の集計結果の一例を示す図である。

【発明を実施するための形態】

【0020】

以下、図面を参照して、本発明の実施形態を詳細に説明する。

[第1の実施形態]

【0021】

図1は、本発明の情報処理システムの構成の一例を示す図である。

情報処理システム100は、サーバとしての情報処理装置110と、1又は複数のクライアント端末としての分析者端末120と、ローカルエリアネットワークを介して接続される構成となっている。また情報処理装置110は、外部ネットワークと接続されており、SNSサイト130やその他の外部サイト140に接続可能な構成となっている。

【0022】

情報処理装置110は、分析者端末120から入力された設定に基づき、SNSサイト130およびその他の外部サイト140から情報を取得し、分析を行う。

【0023】

分析者端末120は、情報処理装置110に対して情報の取得方法および分析方法についての設定を行う。また、情報処理装置110から得られた分析結果を表示する。

【0024】

情報処理装置110は、設定部111及び設定保存領域112を備えており、設定部111は、分析者端末120と設定保存領域112と接続されている。

【0025】

設定部111は、分析者端末120から入力された設定情報、例えば、後述する検索条件や公式サイトURLを受付けて、設定保存領域112へ出力し、設定保存領域112は、設定部111から出力された設定情報を自身に備えたテーブル等に記憶する。

【0026】

また、情報処理装置110は、SNS投稿収集部113及びSNSデータ保存領域114を備えており、SNS投稿収集部113は、設定保存領域112及びSNSデータ保存領域114と接続されている。

【0027】

SNS投稿収集部113は、情報処理装置110の外部に設置されたSNSサイト130やその他の外部サイト140から、設定保存領域112に記憶した検索条件を用いて企業や製品等に関する評価対象に対する評価等が掲載されたサイトに関する情報を収集し、SNSデータ保存領域114へ出力すると、SNSデータ保存領域114は、SNS投稿収集部113から出力されたサイトに関する情報を自身に備えたテーブル等に記憶する。

【0028】

また、情報処理装置110は、リンク解析部115及びリンク情報保存領域116を備えており、リンク解析部115は、設定保存領域112、SNSデータ保存領域114、及びリンク情報保存領域116と接続されている。

【0029】

リンク解析部115は、SNSデータ保存領域114に記憶した各サイトの情報を対象にして、設定保存領域112に記憶した公式サイトURLを用いて、その投稿に含まれるリンク(URL)を解析することで、何れの企業や製品等についての評価等が行われているのかを解明し、その解析結果をリンク情報保存領域116に出力し、リンク情報保存領域116は、リンク解析部115から出力された解析結果を記憶する。

【0030】

10

20

30

40

50

また、情報処理装置 110 は、分析部 117 を備えており、分析部 117 は、SNS データ保存領域 114、リンク情報保存領域 116、及び分析者端末 120 と接続されており、SNS データ保存領域 114 に記憶した各サイトの投稿における評価等が、どのような企業や製品等に関して行われているのかを、リンク情報保存領域 116 に記憶した解析結果を用いて、分析を行い、分析者端末 120 へその分析結果を送信部 118 を介して送信する。

【0031】

これらの各機能に関しては、その機能内容を概要として記載したが、詳細な内容に関しては、後述するフローチャート等を用いて説明を行う。

【0032】

以下、図 2 を用いて、図 1 に示した情報処理装置 110、分析者端末 120 に適用可能な情報処理装置のハードウェア構成について説明する。

【0033】

図 2 は、図 1 に示した情報処理装置 110、分析者端末 120 に適用可能な情報処理装置のハードウェア構成を示すブロック図である。

【0034】

図 2 において、201 は CPU で、システムバス 204 に接続される各デバイスやコントローラを統括的に制御する。また、ROM 202 あるいは外部メモリ 211 には、CPU 201 の制御プログラムである BIOS (Basic Input / Output System) やオペレーティングシステムプログラム (以下、OS) や、各サーバ或いは各 PC の実行する機能を実現するために必要な後述する各種プログラム等が記憶されている。

【0035】

203 は RAM で、CPU 201 の主メモリ、ワークエリア等として機能する。CPU 201 は、処理の実行に際して必要なプログラム等を ROM 202 あるいは外部メモリ 211 から RAM 203 にロードして、該ロードしたプログラムを実行することで各種動作を実現するものである。

【0036】

また、205 は入力コントローラで、キーボード (KB) 209 や不図示のマウス等のポインティングデバイス等からの入力を制御する。206 はビデオコントローラで、CRT ディスプレイ (CRT) 210 等の表示器への表示を制御する。

【0037】

なお、図 2 では、CRT 210 と記載しているが、表示器は CRT だけでなく、液晶ディスプレイ等の他の表示器であってもよい。これらは必要に応じて管理者が使用するものである。

【0038】

207 はメモリコントローラで、ブートプログラム、各種のアプリケーション、フォントデータ、ユーザファイル、編集ファイル、各種データ等を記憶する外部記憶装置 (ハードディスク (HD)) や、フレキシブルディスク (FD)、或いは PCMCIA カードスロットにアダプタを介して接続されるコンパクトフラッシュ (登録商標) メモリ等の外部メモリ 211 へのアクセスを制御する。

【0039】

208 は通信 I/F コントローラで、ネットワーク (例えば、図 1 に示した LAN 400) を介して外部機器と接続・通信するものであり、ネットワークでの通信制御処理を実行する。例えば、TCP/IP を用いた通信等が可能である。

【0040】

なお、CPU 201 は、例えば RAM 203 内の表示情報用領域へアウトラインフォントの展開 (ラスターライズ) 処理を実行することにより、CRT 210 上での表示を可能としている。また、CPU 201 は、CRT 210 上の不図示のマウスカーソル等でのユーザ指示を可能とする。

10

20

30

40

50

【 0 0 4 1 】

本発明を実現するための後述する各種プログラムは、外部メモリ 2 1 1 に記録されており、必要に応じて R A M 2 0 3 にロードされることにより C P U 2 0 1 によって実行されるものである。さらに、上記プログラムの実行時に用いられる定義ファイル及び各種情報テーブル等も、外部メモリ 2 1 1 に格納されており、これらについての詳細な説明も後述する。

【 0 0 4 2 】

次に、図 3 のフローチャートを参照して、本実施形態の情報処理装置 1 1 0 において定期的に行われる、各サイトに対する情報の収集および解析処理について説明する。

【 0 0 4 3 】

ステップ S 3 0 1 では、S N S 投稿収集部 1 1 3 は、分析者端末 1 2 0 から評価の依頼を受付けると、設定保存領域 1 1 2 に保存された検索条件に基づき S N S サイト 1 3 0 から投稿を取得し、S N S データ保存領域 1 1 4 に取得した投稿を保存する。

【 0 0 4 4 】

そして、リンク解析部 1 1 5 は、ステップ S 3 0 1 で取得した投稿に対し、ステップ S 3 0 2 に関して、繰り返し処理を開始する。

【 0 0 4 5 】

ステップ S 3 0 2 では、リンク解析部 1 1 5 は、処理対象である投稿に含まれるリンク (U R L) を解析し、解析結果をリンク情報保存領域 1 1 6 に保存する。

【 0 0 4 6 】

そして、リンク解析部 1 1 5 は、次の投稿がある場合、ステップ S 3 0 2 からの処理を実施する。次の投稿がない場合、処理を終了する。

【 0 0 4 7 】

本説明において、S N S 投稿の取得とリンク解析処理を同期的に行うように記述しているが、非同期に行うように構成しても構わない。

【 0 0 4 8 】

次に図 4 に示すフローチャートを用いて S N S において投稿された情報を収集する処理について詳細について説明する。

【 0 0 4 9 】

ステップ S 4 0 1 では、S N S 投稿収集部 1 1 3 は、分析者端末 1 2 0 から評価の依頼を受付けると、設定保存領域 1 1 2 における検索条件テーブル 5 0 1 から検索条件を取得する。

【 0 0 5 0 】

尚、設定保存領域 1 1 2 には、図 5 に示すように S N S サイト 1 3 0 に投稿された文書類等 (S N S 投稿) を検索するための検索条件を記憶する検索条件テーブル 5 0 1 (図 7 参照)、各企業が提供するサイトを特定するための公式サイト情報を記憶する公式サイト U R L パターンテーブル 5 0 2 (図 1 5 参照、詳細後述) を備えている。

【 0 0 5 1 】

図 7 に示すように、検索条件テーブル 5 0 1 は、S N S サイト 1 3 0 に対して投稿された文書類等 (S N S 投稿) を取得するための条件を備えており、検索条件を一意に識別するための i d、検索を行う際のキーワード (検索内容) や投稿がなされた時刻以降の S N S 投稿を取得する (最終取得時刻) ための条件等の項目を含んで構成されている。

【 0 0 5 2 】

S N S 投稿収集部 1 1 3 は、ステップ S 4 0 1 で取得した検索条件に対し、ステップ S 4 0 2 及びステップ S 4 0 3 の処理を繰り返し実行する。

【 0 0 5 3 】

ステップ S 4 0 2 では、S N S 投稿収集部 1 1 3 は、S N S サイト 1 3 0 に対し、処理対象である検索条件を用いて問い合わせを行い、S N S 投稿を取得する。

【 0 0 5 4 】

ステップ S 4 0 3 では、S N S 投稿収集部 1 1 3 は、ステップ S 4 0 2 で取得した S N

10

20

30

40

50

S 投稿を SNS データ保存領域 1 1 4 における投稿テーブル 6 0 1 に保存する。

【 0 0 5 5 】

尚、 SNS データ保存領域 1 1 4 は、図 6 に示すように、 SNS サイト 1 3 0 から取得した SNS 投稿を記憶する投稿テーブル 6 0 1 を備えており、投稿テーブル 6 0 1 は、図 8 に示すように、取得した SNS 投稿を一意に識別するための id、 SNS 投稿を投稿したユーザを識別するためのユーザ名、投稿された文書やリンク等を示す投稿文、投稿された日時を示す保存時刻等の項目を含んで構成されている。

【 0 0 5 6 】

ステップ S 4 0 3 の処理を終了すると、 SNS 投稿収集部 1 1 3 は、次の検索条件がある場合、ステップ S 4 0 2 及びステップ S 4 0 3 の処理を実施する。次の検索条件がない場合、処理を終了する。

次に、具体的な例を用いて、本フローチャートの説明を行う。

【 0 0 5 7 】

ステップ S 4 0 1 では、 SNS 投稿収集部 1 1 3 は、検索条件テーブル 5 0 1 から検索条件を取得する。

【 0 0 5 8 】

SNS 投稿収集部 1 1 3 は、ステップ S 4 0 1 で取得した検索条件に対し、ステップ S 4 0 2 及びステップ S 4 0 3 の処理を繰り返し実行する。

【 0 0 5 9 】

まず、最初に、 SNS 投稿収集部 1 1 3 は、ステップ S 4 0 1 で取得した検索条件の最初の検索条件 7 0 1 に対し、ステップ S 4 0 2 及びステップ S 4 0 3 の処理を実行する。

【 0 0 6 0 】

ステップ S 4 0 2 では、 SNS 投稿収集部 1 1 3 は、 SNS サイト 1 3 0 に対し、検索条件 7 0 1 を用いて問い合わせを行い、最終取得時刻 (2013-07-12 01:02:30) 以降に投稿され「KANNON」「kannon」「カンノン」のいずれかの文字列を含む SNS 投稿を取得する。

【 0 0 6 1 】

ステップ S 4 0 3 では、 SNS 投稿収集部 1 1 3 は、ステップ S 4 0 2 で取得した投稿 8 0 1 および 8 0 3 ~ 8 0 6 を投稿テーブル 6 0 1 に保存する。

【 0 0 6 2 】

ステップ S 4 0 3 の処理を終了すると、 SNS 投稿収集部 1 1 3 は、次の検索条件 7 0 2 があるので、検索条件 7 0 2 に対し、ステップ S 4 0 2 及びステップ S 4 0 3 の処理を実施する。

【 0 0 6 3 】

ステップ S 4 0 2 では、 SNS 投稿収集部 1 1 3 は、 SNS サイト 1 3 0 に対し、検索条件 7 0 2 を用いて問い合わせを行い、最終取得時刻 (2013-07-12 01:05:43) 以降に投稿され「NIKKOU」「nikkou」「ニッコウ」のいずれかの文字列を含む SNS 投稿を取得する。

【 0 0 6 4 】

ステップ S 4 0 3 では、 SNS 投稿収集部 1 1 3 は、ステップ S 4 0 2 で取得した投稿 8 0 2 を投稿テーブル 6 0 1 に保存する。

【 0 0 6 5 】

ステップ S 4 0 3 の処理を終了すると、 SNS 投稿収集部 1 1 3 は、次の検索条件がないので、処理を終了する。

【 0 0 6 6 】

次に、図 9 に示すフローチャートを用いて SNS にて投稿された文書内のリンクを解析する処理の詳細について説明する。

【 0 0 6 7 】

ステップ S 9 0 1 では、リンク解析部 1 1 5 は、投稿テーブル 6 0 1 に記憶した、処理対象の SNS 投稿における投稿文からリンク (URL) を取得する。

10

20

30

40

50

【 0 0 6 8 】

投稿文からリンク（URL）を取得するには正規表現などを用いてもよい。本実施形態では、投稿文からリンク（URL）を取得しているが、SNSの投稿が構造化されており、別途リンク（URL）に関する情報が取得できる場合はそちらを用いてもよい。

【 0 0 6 9 】

そして、リンク解析部 1 1 5 は、ステップ S 9 0 1 において取得したリンク（URL）に対し、ステップ S 9 1 2 までの繰り返し処理を開始する。

【 0 0 7 0 】

ステップ S 9 0 2 では、リンク解析部 1 1 5 は、一時領域をクリアし、処理対象のリンク（URL）をリンク先とした次数 0 の関係情報（詳細後述）を一時領域に追加する。

10

【 0 0 7 1 】

ステップ S 9 0 3 では、リンク解析部 1 1 5 は、終了フラグを未設定にセットし、処理次数を 0 にセットする。

【 0 0 7 2 】

そして、リンク解析部 1 1 5 は、処理次数が規定値以下かつ終了フラグが未設定の間、ステップ S 9 0 4 からステップ S 9 1 0 までの繰り返し処理を開始する。

【 0 0 7 3 】

ここで規定値を設ける理由としては、無駄な処理を行わないことがあげられるが、特にあまりにも辿るリンク、つまり公式サイトへ到達するリンクが遠い場合、投稿者が当該公式サイトを主題としていることが推定し難いことがあげられるため、このような値を設けている。

20

【 0 0 7 4 】

ステップ S 9 0 4 では、リンク解析部 1 1 5 は、一時領域から次数が処理次数に等しい関係情報を取得する。

【 0 0 7 5 】

そして、リンク解析部 1 1 5 は、ステップ S 9 0 4 で取得した関係情報に対し、ステップ S 9 0 9 までの繰り返し処理を開始する。

【 0 0 7 6 】

ステップ S 9 0 5 では、リンク解析部 1 1 5 は、処理中の関係情報のリンク先が示す文書を外部サイト 1 4 0 から取得する。

30

【 0 0 7 7 】

ステップ S 9 0 6 では、リンク解析部 1 1 5 は、リンク先の URL が公式サイトへのリンクであるか否かを判定し、公式サイトへのリンクであれば、ステップ S 9 0 7 に処理を移し、公式サイトへのリンクでなければステップ S 9 0 9 に処理を移す。

【 0 0 7 8 】

尚、URL が示す文書が公式サイトに属するか否かを判定するには、設定保存領域 1 1 2 における公式サイト URL パターンテーブル 5 0 2 に予め設定されたパターンにより判定を行う。

【 0 0 7 9 】

本実施形態では、URL のオーソリティ部に対し、公式サイト URL パターンテーブル 5 0 2 に予め設定されたパターンが後方一致する場合に公式サイトであると判定し、対応するブランド（企業名等）に関連すると判定する。

40

【 0 0 8 0 】

ステップ S 9 0 7 においてリンク解析部 1 1 5 は、ステップ S 9 0 5 で取得した文書と処理対象の関係情報およびステップ S 9 0 6 で判定したブランドに対し、リンク情報追加処理を実施する。詳細については、図 1 2 を用いて後述する。

ステップ S 9 0 8 では、リンク解析部 1 1 5 は、終了フラグを設定する。

【 0 0 8 1 】

ステップ S 9 0 9 では、リンク解析部 1 1 5 は、ステップ S 9 0 5 で取得した文書に対し、文書内のリンク解析処理を実施する。詳細については、図 1 1 を用いて後述する。

50

【 0 0 8 2 】

ステップ S 9 0 8 あるいはステップ S 9 0 9 の処理を終了すると、リンク解析部 1 1 5 は、次の関係情報がある場合、ステップ S 9 0 8 からの処理を実施する。次の関係情報がない場合、ステップ S 9 1 0 に処理を移す。

ステップ S 9 1 0 では、リンク解析部 1 1 5 は、処理回数を 1 増やす。

【 0 0 8 3 】

ステップ S 9 1 0 の処理が終了すると、リンク解析部 1 1 5 は、処理回数が規定値以下でありかつ終了フラグが未設定である場合、ステップ S 9 0 4 からの処理を実施する。処理回数が規定値より大きいかまたは終了フラグが設定されている場合、ステップ S 9 1 1 に処理を移す。

10

【 0 0 8 4 】

ステップ S 9 1 1 では、リンク解析部 1 1 5 は、終了フラグが設定されているか否かを判定し、終了フラグが設定されていない場合、ステップ S 9 1 2 に処理を移し、終了フラグが設定されており、次のリンクがある場合、ステップ S 9 0 2 からの処理を実施し、次のリンクがない場合、処理を終了する。

【 0 0 8 5 】

ステップ S 9 1 2 では、リンク解析部 1 1 5 は、ステップ S 9 0 5 で取得した文書に対し、関係情報がない状態で、リンク情報追加処理を実施する。詳細については、図 1 2 を用いて後述する。

【 0 0 8 6 】

次に、図 1 1 に示すフローチャートを用いて文書内のリンクを解析する処理の詳細について説明する。

20

【 0 0 8 7 】

ステップ S 1 1 0 1 では、リンク解析部 1 1 5 は、ステップ S 9 0 5 で取得した文書からリンク (U R L) を取得する。

【 0 0 8 8 】

そして、リンク解析部 1 1 5 は、ステップ S 1 1 0 1 において取得したリンク (U R L) に対し、ステップ S 1 1 0 5 までの繰り返し処理を開始する。

【 0 0 8 9 】

ステップ S 1 1 0 2 では、リンク解析部 1 1 5 は、処理中のリンク (U R L) が外部サイトへのリンクか否かを判定し、外部サイトへのリンクであれば、ステップ S 1 1 0 3 に処理を移し、外部サイトへのリンクでなければ、次のリンク (U R L) があれば、ステップ S 1 1 0 2 からの処理を実施し、次のリンク (U R L) がない場合、処理を終了する。

30

【 0 0 9 0 】

本実施の例においては、処理中のリンクが外部サイトに属するか否かを判定するには、処理中の文書へのリンク (U R L) および処理中のリンク (U R L) のオーソリティ部が等しいか否かにより判定し、等しければ内部リンクと判定し、異なれば外部リンクと判定する。

【 0 0 9 1 】

ステップ S 1 1 0 3 では、リンク解析部 1 1 5 は、処理回数に 1 を加算した値を回数、処理対象である文書の U R L をリンク元とし、処理中のリンク (U R L) をリンク先とする関係情報を生成する。

40

【 0 0 9 2 】

ステップ S 1 1 0 4 では、リンク解析部 1 1 5 は、ステップ S 1 1 0 3 において生成した関係情報が一時領域に登録されているか否かを判定し、登録されていない場合は、ステップ S 1 1 0 5 に処理を移し、登録されている場合は、次のリンク (U R L) があれば、ステップ S 1 1 0 2 からの処理を実施し、次のリンク (U R L) がない場合、処理を終了する。

【 0 0 9 3 】

ステップ S 1 1 0 5 では、リンク解析部 1 1 5 は、ステップ S 1 1 0 3 で生成された関

50

係情報を一時領域に登録する。

【0094】

本処理では、内部リンクを後続の処理の対象から外しているが、これも、前述したことと同様に、あまりにも辿るリンク、つまり公式サイトへ到達するリンクが遠い場合、投稿者が当該公式サイトを主題としていることが推定し難いことがあげられるため、このような処理を行っている。

【0095】

内部リンクに関しては、例えば、ニュースサイト内の文書から同じニュースサイトの別のページにリンクがあるような場合が想定され、この場合、後方のページは、主題から遠ざかる可能性が十分に有り得るため、後続の処理の対象から外している。

10

【0096】

次に、図12に示すフローチャートを用いて、リンク情報の追加を行う処理の詳細について説明する。

【0097】

ステップS1201では、リンク解析部115は、リンク情報(詳細後述)を生成し、処理対象のリンクを直接リンクとして設定する。

【0098】

ステップS1202では、リンク解析部115は、処理対象の関係情報が指定されているか否かを判定し、関係情報が指定されている場合、ステップS1203に処理を移し、関係情報が指定されていない場合、ステップS1206に処理を移す。

20

【0099】

ステップS1203では、リンク解析部115は、処理対象の関係情報のリンク先をリンク情報の到達リンクに設定する。

【0100】

ステップS1204では、リンク解析部115は、処理対象のブランドをリンク情報のブランドに設定する。

【0101】

ステップS1205では、リンク解析部115は、処理対象の文書におけるタイトルからタグを抽出し、処理対象の投稿と関連付けて、リンク情報保存領域116におけるタグ情報テーブル1002(図17参照)に登録する。

30

【0102】

本実施の例において、タグの抽出は、正規表現等を用い、英数字および記号が連続する箇所を抽出するものとする。

【0103】

タグは予め定義した辞書を用いて特定してもよいし、tf/idfのような単語の出現頻度などの統計情報に基づいて抽出の基準を定義して特定してもよい。

【0104】

リンク情報保存領域116には、図10に示すように、投稿文のリンクから到達する公式サイトとの関係を記憶するリンク情報テーブル1001、SNS投稿の特徴とSNS投稿によって評価されているブランドとの関係を記憶するタグ情報テーブル1002を備えている。

40

【0105】

ステップS1206では、リンク解析部115は、直接リンクをリンク情報の到達リンクに設定する。

【0106】

ステップS1207では、リンク解析部115は、リンク情報を処理対象の投稿と関連付けて、リンク情報保存領域116におけるリンク情報テーブル1001に登録する。

【0107】

次に、具体例として、図8に示す投稿テーブル601に記憶した投稿801に対して、処理次数の規定値を2とし、図9、図11および図12の処理を実施した場合について説

50

明する。

【0108】

設定保存領域112における公式サイトURLパターンテーブル502の具体例について、図15に示す。

【0109】

公式サイトURLパターンテーブル502は、各企業が提供するサイトのURLに含まれる文字列を示す公式サイトURLパターン文字列、及び当該企業のブランド等の項目を含んで構成されている。

【0110】

ステップS901では、リンク解析部115は、投稿801の投稿文からリンク807

10

【0111】

そして、リンク解析部115は、リンク807に対してステップS912までの繰り返し処理を開始する。

【0112】

ステップS902では、リンク解析部115は、一時領域をクリアし、リンク807をリンク先とした次数0の関係情報1301を一時領域に追加する(図13参照)。

【0113】

ステップS903では、リンク解析部115は、終了フラグを未設定にセットし、処理

20

【0114】

ステップS904では、リンク解析部115は、処理次数0が規定値2以下かつ終了フラグが未設定なので、ステップS910までの繰り返し処理を開始する。

【0115】

ステップS905では、リンク解析部115は、一時領域から次数が処理次数0に等しい関係情報1301を取得する。

【0116】

そして、リンク解析部115は、ステップS904において取得した関係情報1301に対し、ステップS908までの繰り返し処理を開始する。

【0117】

ステップS906では、リンク解析部115は、処理中の関係情報1301のリンク先「<http://www.kannon.jp/news20130711.html>」が示す文書1401(図14参照)を外部

30

【0118】

ステップS906では、リンク解析部115は、関係情報1301のリンク先「<http://www.kannon.jp/news20130711.html>」のオーソリティ部「www.kannon.jp」が公式サイトURLパターン文字列1501に後方一致するので、ブランド「kannon」の公式サイトと判定し、関係情報1301のリンク先が公式サイトへのリンクであるのでステップS907に処理を移す。

【0119】

ステップS907では、リンク解析部115は、ステップS905で取得した文書1401と処理対象の関係情報1301およびステップS906において判定したブランド「kannon」に対し、リンク情報追加処理(図12参照)を実施する。

40

【0120】

ステップS1201では、リンク解析部115は、リンク情報1601(図16参照)を生成し、処理対象のリンク「<http://www.kannon.jp/news20130711.html>」を直接リンクとして設定する。

【0121】

ステップS1202では、リンク解析部115は、関係情報1301が指定されているので、ステップS1203に処理を移す。

50

【 0 1 2 2 】

ステップ S 1 2 0 3 では、リンク解析部 1 1 5 は、処理対象の関係情報 1 3 0 1 のリンク先「<http://www.kannon.jp/news20130711.html>」をリンク情報 1 6 0 1 の到達リンクに設定する。

【 0 1 2 3 】

ステップ S 1 2 0 4 では、リンク解析部 1 1 5 は、処理対象のブランド「kannon」をリンク情報 1 6 0 1 のブランドに設定する。

【 0 1 2 4 】

ステップ S 1 2 0 5 では、リンク解析部 1 1 5 は、処理対象の文書 1 4 0 1 のタイトルからタグ「32D mk II」を抽出し、投稿 8 0 1 (i d = 1) と関連付けて、リンク情報保存領域 1 1 6 におけるタグ情報テーブル 1 0 0 2 に登録する。

10

【 0 1 2 5 】

ステップ S 1 2 0 7 では、リンク解析部 1 1 5 は、リンク情報を処理対象の投稿 8 0 1 (i d = 1) と関連付けて、リンク情報保存領域 1 1 6 におけるリンク情報テーブル 1 0 0 1 に登録して、ステップ S 9 0 8 に処理を戻す。

ステップ S 9 0 8 では、リンク解析部 1 1 5 は、終了フラグを設定する。

【 0 1 2 6 】

そして、リンク解析部 1 1 5 は、次の関係情報がないので、ステップ S 9 1 0 に処理を移す。

ステップ S 9 1 0 では、リンク解析部 1 1 5 は、処理回数を 1 増やす。

20

【 0 1 2 7 】

そして、リンク解析部 1 1 5 は、終了フラグが設定されているので、ステップ S 9 1 1 に処理を移す。

【 0 1 2 8 】

ステップ S 9 1 1 では、リンク解析部 1 1 5 は、終了フラグが設定されており、次のリンクがないので、処理を終了する。

次に投稿 8 0 2 に対する具体例について説明する。

【 0 1 2 9 】

ステップ S 9 0 1 では、リンク解析部 1 1 5 は、投稿 8 0 2 の投稿文からリンク 8 0 8 を取得する。

30

【 0 1 3 0 】

そして、リンク解析部 1 1 5 は、リンク 8 0 8 に対し、ステップ S 9 1 2 までの繰り返し処理を開始する。

【 0 1 3 1 】

ステップ S 9 0 2 では、リンク解析部 1 1 5 は、一時領域をクリアし、処理対象のリンク 8 0 8 をリンク先とした回数 0 の関係情報 1 8 0 1 (図 1 8 参照) を一時領域に追加する。

【 0 1 3 2 】

ステップ S 9 0 3 では、リンク解析部 1 1 5 は、終了フラグを未設定にセットし、処理回数を 0 にセットする。

40

【 0 1 3 3 】

そして、リンク解析部 1 1 5 は、処理回数 0 が規定値 2 以下かつ終了フラグが未設定なので、ステップ S 9 1 0 までの繰り返し処理を開始する。

【 0 1 3 4 】

ステップ S 9 0 4 では、リンク解析部 1 1 5 は、一時領域から回数が処理回数 0 に等しい関係情報 1 8 0 1 を取得する。

【 0 1 3 5 】

そして、リンク解析部 1 1 5 は、ステップ S 9 0 4 で取得した関係情報 1 8 0 1 に対し、ステップ S 9 1 1 までの繰り返し処理を開始する。

【 0 1 3 6 】

50

ステップS 9 0 5では、リンク解析部 1 1 5は、処理中の関係情報 1 8 0 1のリンク先「<http://nblog.ne.jp/nagatomo/201307112211.html>」が示す文書 1 9 0 1 (図 1 9 参照) を外部サイト 1 4 0 から取得する。

【 0 1 3 7 】

ステップS 9 0 6では、リンク解析部 1 1 5は、関係情報 1 8 0 1のリンク先「<http://nblog.ne.jp/nagatomo/201307112211.html>」のオーソリティ部「nblog.ne.jp」が公式サイトURLパターン文字列のいずれにも一致しないので、関係情報 1 8 0 1のリンク先が公式サイトへのリンクでないと判断し、ステップS 9 0 9に処理を移す。

【 0 1 3 8 】

ステップS 9 0 9では、リンク解析部 1 1 5は、ステップS 9 0 5で取得した文書 1 9 0 1 10 0 1に対し、文書内のリンク解析処理 (図 1 1 参照) を実施する。

【 0 1 3 9 】

ステップS 1 1 0 1では、リンク解析部 1 1 5は、ステップS 9 0 5において取得した文書 1 9 0 1 からリンク 1 9 0 2 ~ 1 9 0 4 を取得する。

【 0 1 4 0 】

そして、リンク解析部 1 1 5は、ステップS 1 1 0 1において取得したリンク 1 9 0 2 「<http://www.jnews.co.jp/20120127.html>」に対し、ステップS 1 1 0 5までの繰り返し処理を開始する。

【 0 1 4 1 】

ステップS 1 1 0 2では、リンク解析部 1 1 5は、処理中のリンク情報 1 6 0 2 が外部 20 サイトへのリンクか否かを判定する。

【 0 1 4 2 】

リンク 1 9 0 2 のオーソリティ部は「www.jnews.co.jp」であり、処理中の文書 1 9 0 1 へのリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」のオーソリティ部「nblog.ne.jp」と異なるので外部サイトへのリンクと判断し、ステップS 1 1 0 3に処理を移す。

【 0 1 4 3 】

ステップS 1 1 0 3では、リンク解析部 1 1 5は、回数 1 (= 処理回数 0 + 1)、文書 1 9 0 1 へのリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」をリンク元とし、処理中のリンク 1 9 0 2 「<http://www.jnews.co.jp/20120127.html>」をリンク先とする関係情報 1 8 0 2 を生成する。 30

【 0 1 4 4 】

ステップS 1 1 0 4では、リンク解析部 1 1 5は、ステップS 1 1 0 3において生成した関係情報 1 8 0 2 が一時領域に登録されていないので、ステップS 1 1 0 5に処理を移す。

【 0 1 4 5 】

ステップS 1 1 0 5では、リンク解析部 1 1 5は、ステップS 1 1 0 4で生成された関係情報 1 8 0 2 を一時領域に登録する。

【 0 1 4 6 】

そして、リンク解析部 1 1 5は、次のリンク 1 9 0 3 があるので、ステップS 1 1 0 2 40 からの処理を実施する。

【 0 1 4 7 】

ステップS 1 1 0 2では、リンク解析部 1 1 5は、リンク 1 9 0 3 で示される「<http://nblog.ne.jp/nagatomo/201307092032.html>」に対し、ステップS 1 1 0 5までの繰り返し処理を開始する。

【 0 1 4 8 】

ステップS 1 1 0 3では、リンク解析部 1 1 5は、処理中のリンク 1 9 0 3 が外部サイトへのリンクか否かを判定し、リンク 1 9 0 3 のオーソリティ部は「nblog.ne.jp」であり、処理中の文書 1 9 0 1 へのリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」のオーソリティ部「nblog.ne.jp」と等しいので、内部サイトへのリンクと判断し、ス 50

ステップ S 1 1 0 5 の後に処理を移す。

【 0 1 4 9 】

そして、リンク解析部 1 1 5 は、次のリンク 1 9 0 4 があるので、ステップ S 1 1 0 2 からの処理を実施する。

【 0 1 5 0 】

ステップ S 1 1 0 2 では、リンク解析部 1 1 5 は、リンク 1 9 0 4 で示される「<http://www.kshop.co.jp/kannon32d.html>」に対し、ステップ S 1 1 0 5 までの繰り返し処理を開始する。

【 0 1 5 1 】

ステップ S 1 1 0 3 では、リンク解析部 1 1 5 は、処理中のリンク 1 9 0 4 が外部サイトへのリンクか否かを判定する。

10

【 0 1 5 2 】

リンク 1 9 0 4 のオーソリティ部は「www.kshop.co.jp」であり、処理中の文書 1 9 0 1 へのリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」のオーソリティ部「nblog.ne.jp」と異なるので外部サイトへのリンクと判断し、ステップ S 1 1 0 3 に処理を移す。

【 0 1 5 3 】

ステップ S 1 1 0 4 では、リンク解析部 1 1 5 は、次数 1 (= 処理次数 + 1)、文書 1 9 0 1 へのリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」をリンク元とし、処理中のリンク 1 9 0 4 「<http://www.kshop.co.jp/kannon32d.html>」をリンク先とする関係情報 1 8 0 3 を生成する。

20

【 0 1 5 4 】

ステップ S 1 1 0 4 では、リンク解析部 1 1 5 は、ステップ S 1 1 0 3 で生成した関係情報 1 8 0 3 が一時領域に登録されていないので、ステップ S 1 1 0 5 に処理を移す。

【 0 1 5 5 】

ステップ S 1 1 0 5 では、リンク解析部 1 1 5 は、ステップ S 1 1 0 4 で生成された関係情報 1 8 0 3 を一時領域に登録する。

【 0 1 5 6 】

そして、リンク解析部 1 1 5 は、次のリンクがないので、処理を終了し、次の関係情報がないので、ステップ S 9 1 0 に処理を移す。

30

【 0 1 5 7 】

ステップ S 9 1 0 では、リンク解析部 1 1 5 は、処理次数を 1 増やし、処理次数を 1 とする。

【 0 1 5 8 】

そして、リンク解析部 1 1 5 は、処理次数 1 が規定値 2 以下でありかつ終了フラグが未設定であるので、ステップ S 9 0 4 からステップ S 9 1 0 までの繰り返し処理を開始する。

【 0 1 5 9 】

ステップ S 9 0 4 では、リンク解析部 1 1 5 は、一時領域から次数が処理次数 1 に等しい関係情報 1 8 0 2 および関係情報 1 8 0 3 を取得する。

40

【 0 1 6 0 】

そして、リンク解析部 1 1 5 は、ステップ S 9 0 4 において取得した関係情報 1 8 0 2 に対し、ステップ S 9 1 0 までの繰り返し処理を開始する。

【 0 1 6 1 】

ステップ S 9 0 5 では、リンク解析部 1 1 5 は、処理中の関係情報 1 8 0 2 のリンク先「<http://www.jnews.co.jp/20120127.html>」が示す文書 2 0 0 1 (図 2 0) を外部サイト 1 4 0 から取得する。

【 0 1 6 2 】

ステップ S 9 0 6 では、リンク解析部 1 1 5 は、関係情報 1 8 0 2 のリンク先「<http://www.jnews.co.jp/20120127.html>」のオーソリティ部「www.jnews.co.jp」が公式サイト

50

URLパターン文字列のいずれにも一致しないので、関係情報1802のリンク先が公式サイトへのリンクでないと判断し、ステップS909に処理を移す。

【0163】

ステップS909では、リンク解析部115は、ステップS906において取得した文書2001に対し、文書内のリンク解析処理(図11参照)を実施する。

【0164】

ステップS1101では、リンク解析部115は、ステップS905で取得した文書2001からリンク2002および2003を取得する。

【0165】

そして、リンク解析部115は、ステップS1101で取得したリンク2002「<http://www.kannon.jp/news20130711.html>」に対し、ステップS1105までの繰り返し処理を開始する。

10

【0166】

ステップS1102では、リンク解析部115は、処理中のリンク2002が外部サイトへのリンクか否かを判定し、リンク2002のオーソリティ部は「www.kannon.jp」であり、処理中の文書2001へのリンク「<http://www.jnews.co.jp/20120127.html>」のオーソリティ部「www.jnews.co.jp」と異なるので外部サイトへのリンクと判断し、ステップS1103に処理を移す。

【0167】

ステップS1103では、リンク解析部115は、回数2(=処理回数1+1)、文書2001へのリンク「<http://www.jnews.co.jp/20120127.html>」をリンク元とし、処理中のリンク2002「<http://www.kannon.jp/news20130711.html>」をリンク先とする関係情報1804を生成する。

20

【0168】

ステップS1105では、リンク解析部115は、ステップS1103で生成した関係情報1804が一時領域に登録されていないので、ステップS1105に処理を移す。

【0169】

ステップS1105では、リンク解析部115は、ステップS1104で生成された関係情報1804を一時領域に登録する。

【0170】

そして、リンク解析部115は、次のリンク2003があるので、ステップS1102からの処理を実施する。

30

【0171】

ステップS1102では、リンク解析部115は、リンク2103で示される「<http://www.jnews.co.jp/2013-0527-kannon.html>」に対し、ステップS1105までの繰り返し処理を開始する。

【0172】

ステップS1102では、リンク解析部115は、処理中のリンク2003が外部サイトへのリンクか否かを判定し、リンク2003のオーソリティ部は「www.jnews.co.jp」であり、処理中の文書2001へのリンク「<http://www.jnews.co.jp/20120127.html>」のオーソリティ部「www.jnews.co.jp」と等しいので、内部サイトへのリンクと判断し、次のリンクがないので、処理を終了する。

40

【0173】

そして、リンク解析部115は、次の関係情報1803があるので、ステップS905に処理を移す。

【0174】

そして、リンク解析部115は、関係情報1803に対し、ステップS910までの繰り返し処理を開始する。

【0175】

ステップS905では、リンク解析部115は、処理中の関係情報1803のリンク先

50

「<http://www.kshop.co.jp/kannon32d.html>」が示す文書を外部サイト140から取得する。ここで「<http://www.kshop.co.jp/kannon32d.html>」が示す文書はリンクを持っていないものとする。

【0176】

ステップS906では、リンク解析部115は、関係情報1803のリンク先「<http://www.kshop.co.jp/kannon32d.html>」のオーソリティ部「www.kshop.co.jp」が公式サイトURLパターン文字列のいずれにも一致しないので、関係情報1803のリンク先が公式サイトへのリンクでないと判断し、ステップS909に処理を移す。

【0177】

ステップS909では、リンク解析部115は、ステップS905で取得した文書に対し、文書内のリンク解析処理（図11参照）を実施するが処理対象である文書はリンクを持たないので、リンク解析処理が終了しても状態は変化しないので、詳細は省略する。

10

【0178】

そして、リンク解析部115は、次の関係情報がないので、ステップS910に処理を移す。

【0179】

ステップS910では、リンク解析部115は、処理次数を1増やし、処理次数を2とする。

【0180】

そして、リンク解析部115は、処理次数2が規定値2以下でありかつ終了フラグが未設定であるので、ステップS904からステップS910までの繰り返し処理を開始する。

20

【0181】

ステップS904では、リンク解析部115は、一時領域から次数が処理次数2に等しい関係情報1804を取得する。

【0182】

そして、リンク解析部115は、ステップS904において取得した関係情報1804に対し、ステップS912までの繰り返し処理を開始する。

【0183】

ステップS905では、リンク解析部115は、処理中の関係情報1804のリンク先「<http://www.kannon.jp/news20130711.html>」が示す文書1401（図14参照）を外部サイト140から取得する。

30

【0184】

ステップS906では、リンク解析部115は、関係情報1804のリンク先「<http://www.kannon.jp/news20130711.html>」のオーソリティ部「www.kannon.jp」が公式サイトURLパターン文字列1501に後方一致するので、ブランド「kannon」の公式サイトと判定し、関係情報1804のリンク先が公式サイトへのリンクであるのでステップS907に処理を移す。

【0185】

ステップS907では、リンク解析部115は、ステップS905において取得した文書1401と処理対象の関係情報1804およびステップS906で判定したブランド「kannon」に対し、リンク情報追加処理（図12参照）を実施する。

40

【0186】

ステップS1201では、リンク解析部115は、リンク情報1602（図16参照）を生成し、処理対象のリンク「<http://nblog.ne.jp/nagatomo/201307112211.html>」を直接リンクとして設定する。

【0187】

ステップS1202では、リンク解析部115は、関係情報1804が指定されているので、ステップS1203に処理を移す。

【0188】

50

ステップS 1 2 0 3では、リンク解析部 1 1 5は、処理対象の関係情報 1 8 0 4のリンク先「<http://www.kannon.jp/news20130711.html>」をリンク情報 1 6 0 2の到達リンクに設定する。

【 0 1 8 9 】

ステップS 1 2 0 4では、リンク解析部 1 1 5は、処理対象のブランド「kannon」をリンク情報 1 6 0 2のブランドに設定する。

【 0 1 9 0 】

ステップS 1 2 0 5では、リンク解析部 1 1 5は、処理対象の文書 1 4 0 1からタグ「32D mk II」を抽出し、投稿 8 0 2 (i d = 2)と関連付けて、リンク情報保存領域 1 1 6におけるタグ情報テーブル 1 0 0 2に登録する。

10

【 0 1 9 1 】

ステップS 1 2 0 7では、リンク解析部 1 1 5は、リンク情報 1 6 0 2を処理対象の投稿 8 0 2 (i d = 2)と関連付けて、リンク情報保存領域 1 1 6におけるリンク情報テーブル 1 0 0 1に登録して、ステップS 9 0 8に処理を戻す。

ステップS 9 0 8では、リンク解析部 1 1 5は、終了フラグを設定する。

【 0 1 9 2 】

そして、リンク解析部 1 1 5は、次の関係情報がないので、ステップS 9 1 0に処理を移す。

【 0 1 9 3 】

ステップS 9 1 0では、リンク解析部 1 1 5は、処理回数を 1 増やし、処理回数を 3 とする。

20

【 0 1 9 4 】

そして、リンク解析部 1 1 5は、処理回数 3 が規定値 2 以下でありかつ終了フラグが設定されているので、ステップS 9 1 1に処理を移す。

【 0 1 9 5 】

ステップS 9 1 1では、リンク解析部 1 1 5は、終了フラグが設定されており、次のリンクがないので、処理を終了する。

【 0 1 9 6 】

以下、投稿 8 0 3 ~ 8 0 6 までに対し、同様の処理を繰り返した場合に得られる、リンク情報保存領域 1 1 6 におけるリンク情報テーブル 1 0 0 1 およびタグ情報テーブル 1 0 0 2 の状態を図 2 1 および図 2 2 に示す。

30

【 0 1 9 7 】

文書「<http://www.kannon-shop.co.jp/kannon32dmk2.html>」のタイトルは「KANNON ダイレクトショップ - 32D mk II」とする。

【 0 1 9 8 】

尚、分析部 1 1 7 は、分析者端末 1 2 0 からの指示に応じて、リンク情報保存領域 1 1 6 におけるリンク情報テーブル 1 0 0 1 およびタグ情報テーブル 1 0 0 2 に保存された情報に基づく分析を行い、分析結果を送信部 1 1 8 を介して分析者端末 1 2 0 の受信部 1 2 1 へ送信し、出力部 1 2 2 は、受信部 1 2 1 で受信した分析結果を CRT 2 1 0 へ表示する。

40

【 0 1 9 9 】

公式サイトの各ページや抽出したタグ、ブランドに関する投稿の集計・ランキングや投稿の一覧表示、ネガポジ分析など投稿に対する言語処理を用いた分析などが様々な分析手法があるが詳細は省略する。

【 0 2 0 0 】

図 2 3 は、図 2 1 に示すリンク情報テーブル 1 0 0 1 に対し、到達リンクで集計した結果（分析結果）を示す表である。

【 0 2 0 1 】

投稿上のリンク先が異なるが、投稿 8 0 1 ~ 8 0 3 および 8 0 6 (i d 1 ~ 3 および 6) は同一の文書 (<http://www.kannon.jp/news20130711.html>) に関する投稿として分析す

50

ることが可能となっている。

【0202】

図24は、図22に示すタグ情報テーブル1002に対しタグで集計した結果（分析結果）を示す表である。投稿801～804および806（id1～4および6）は同一のタグ（32D mk 11）に関する投稿として分析することが可能となっている。

【0203】

本実施形態によれば、投稿802、803、806のように企業のサイトへのリンクでなくても、主題が、企業のサイトの文書（<http://www.kannon.jp/news20130711.html>）と同じであり、投稿801と同様に評価の対象に含めた分析を実現することが可能となる。

【0204】

更に、企業サイトと関連付けられるリンクを持つ投稿801～804および806に対し、企業サイトのページのタイトルから抽出した有用性の高いキーワード（「32D mk 11」）に基づいた分析を実現することも可能となる。

[第2の実施形態]

【0205】

第2の実施形態では、第1の実施形態の構成及び処理に基づくものであるが、第2の実施形態において、リンク情報保存領域116は図25に示すように関係情報テーブル2501を拡張する。

【0206】

次に図26および図27に示すフローチャートを用いて第2の実施形態において、SNSにて投稿された文書内のリンクを解析する処理について説明する。

【0207】

図26に示すフローチャートは図9に示すフローチャートに対し、ステップS2601の処理が追加されている。

【0208】

ステップS2601においてリンク解析部115は、一時領域に保存された関係情報を関係情報テーブル2501に追加する。このとき、リンクのサイトを判定する情報として、関係情報におけるリンク元URLのオーソリティ部を一緒に追加する。一時領域に保存されている関係情報のうちリンク元URLがない関係情報は追加しない。

【0209】

図27に示すフローチャートは図11に示すフローチャートに対し、ステップS2701およびステップS2702の処理が追加されている。

【0210】

ステップS2701においてリンク解析部115は、処理対象であるリンクの重要度を算出する。リンクの重要度は、リンクに対する関係情報から、対象となるリンクがリンク元の文書において特有のリンクであるか、同一サイトにおいて共通に用いられているかを数値化したものである。リンクの重要度は、対象のリンクが、リンク元文書と同一サイト（同じオーソリティ部を持つ）にある文書において共通に用いられている場合に低く、リンク元文書において特有のリンクである場合に高くなるように定義されている。以下にリンクの重要度の判定式の一例を示す。

【0211】

$$(\text{リンクの重要度}) = \log_2 \frac{N_{\text{src}} + C + 1}{N_{\text{link}} + 1}$$

Nsrc: リンク元のオーソリティ部と一致する関係情報のリンク元URLの異なり数

Nlink: リンクをリンク先URLに持つ関係情報の異なり数

C: 定数

【0212】

ステップS2702においてリンク解析部115は、ステップS2701において算出したリンクの重要度が規定値以上の場合、ステップS1105に処理を移す。リンクの重

10

20

30

40

50

要度が規定値未満である場合、次のリンクに対する繰り返し処理を実行する。

【0213】

(具体例)

図28に示すリンク元の文書に対し、第2の実施形態におけるリンク解析処理が実施された場合について具体的に説明する。本実施形態の例におけるリンクの重要度に対する規定値は1.0とする。

【0214】

図28に示すリンク元文書2801「<http://nblog.ne.jp/kiyotake/2013070101.html>」に対し、図26に示すリンク解析処理を実施する。

【0215】

詳細は省略するが、リンク元文書2801におけるリンク2802～2806のうち、リンク2803はリンク元文書と同一オーソリティであるため、外部リンクと判断されず、一時領域に保存されない。

【0216】

結果としてS2601においてリンク2802およびリンク2804～2806が関係情報テーブル2501に追加され、図29に示すように関係情報2901～2904が登録される。

【0217】

同様にリンク2804～2806を持つ同サイト(オーソリティ部が一致する)の4つの文書が更に処理されたものとする。詳細は省略するが、ここまでの処理において、ステップS2701で算出されるリンクの重要度はいずれも規定値以上となる。結果として関係情報テーブル2501は図29の状態となる。尚、<http://nblog.ne.jp/kiyotake/2013070103.html>のリンク元URL、<http://nblog.ne.jp/kiyotake/2013070104.html>のリンク元のURLに関しては、関係情報として、<http://www.camera-sokuhou.ne.jp/camera/>を備える。

【0218】

図28に示すリンク元文書2807「<http://nblog.ne.jp/kiyotake/2013070106.html>」に対し、図26に示すリンク解析処理を実施する。

【0219】

文書2807におけるリンク2808「<http://www.jnews.co.jp/20120129.html>」に対し処理が行われ、ステップS2701でリンクの重要度が算出される。

【0220】

リンク元である文書2807のオーソリティ部「<http://nblog.ne.jp/>」を持つリンク元のURLの異なり数は5であるので、Nsrc=5となる、リンク2808をリンク先に持つ関係情報はないのでNlink=0となる。定数Cの値を5とすると、リンク2808に対する重要度は $\log_2((5+5+1)/(1+1)) = 2.459431619$ となる。

【0221】

ステップS2702において、リンク2808の重要度は規定値1.0以上なのでステップS1105の処理が実施され、以降の処理でも有効な関係情報として処理される。

【0222】

文書2807におけるリンク2810「<http://www.camera-sokuhou.ne.jp/camera/>」に対しても、同様にステップS2701においてリンクの重要度が算出される。

【0223】

リンク元である文書2807のオーソリティ部「<http://nblog.ne.jp/>」を持つリンク元のURLの異なり数は5であるので、Nsrc=5となる、リンク2810をリンク先に持つ関係情報の異なり数も5であるのでNlink=5となり、リンク2810に対する重要度は $\log_2((5+5+1)/(5+1)) = 0.874469118$ となる。

【0224】

ステップS2702において、リンク2810の重要度は規定値1.0未満なのでステップS1105の処理は実施されず、リンク2810は以降の処理において対象とならない

10

20

30

40

50

。

【0225】

上述したように、リンク先文書のうち、リンク元文書とは関係が薄く、リンク元文書の属するオーソリティ（ドメイン）において定型的に記載されているリンクを解析の対象から除外することが可能となる。結果として、不要なリンクを辿る処理を削減でき、処理効率および処理精度ともに向上させることが可能となる。

【0226】

[第3の実施形態]

第3の実施形態では、第1及び第2の実施形態の構成及び処理に基づくものであるが、まず、図30に示すフローチャートを用いて第3の実施形態において、SNSにて投稿された文書内のリンクを解析する処理の詳細について説明する。

10

【0227】

図30に示すフローチャートは図9に示すフローチャートに対し、ステップS3001およびステップS3002の処理が追加されている。

【0228】

ステップS3001においてリンク解析部115は、リンク元文書とリンク先文書を比較して、関連度を算出する。関連度は、リンク元文書とリンク先文書の主題との同一性を数値化したものであり、文書そのものの類似性から算出してもよいし、タイトル等の一部分から判定してもよい。本実施形態の例においてはタイトル文字列の類似性を評価する指標として用いる。以下に関連度を算出する式の一例を示す。

20

【0229】

$$\text{(関連度)} = \frac{(\text{共通単語の数}) \times 2}{(\text{出現単語の総和})}$$

ステップS3001においてリンク解析部115は、ステップS3001で求めた関連度が規定値以上であれば、ステップS906に処理を移す。関連度が規定値未満であれば、次の関係情報に対する繰り返し処理を開始する。

(具体例)

【0230】

図31に示すリンク元の文書に対し、第3の実施形態におけるリンク解析処理が実施された場合について具体的に説明する。本実施形態の例において関連度の規定値は0.5とする。

30

【0231】

ステップS901においてリンク解析部115は、リンク元の文書3101からリンク3102および3103を取得する。

【0232】

リンク3102に対する繰り返し処理が開始され、ステップS905まで処理が進む。

【0233】

ステップS905においてリンク解析部115は、リンク3102が示す文書を取得する。このときリンク3102が示す文書のタイトルが「カンノン：5月24日より32D mk II 発売開始」であったとする。

40

【0234】

ステップS3001においてリンク解析部115は、リンク元である文書3101のタイトル「JNews - カンノンから 32D mk II 24日より発売」とステップS905で取得したリンク先文書のタイトルとから関連度を算出する。本実施形態の例において、形態素解析の結果から記号類を除いたものを単語として抽出する。

【0235】

リンク元文書のタイトルからは「JNews」「カンノン」「から」「32D mk II」「24日」「より」「発売」の7つの単語が得られる。

【0236】

50

リンク先文書のタイトルからは「カンノン」「5月」「24日」「より」「32D mk II」「発売」「開始」の7つの単語が得られる。

【0237】

一致する単語は「カンノン」「32D mk II」「24日」「より」「発売」の5つであり、関連度は $(5 \times 2) / (7+7) = 10/14 \quad 0.71429$ となる。

【0238】

ステップS3002においてリンク解析部115は、関連度が規定値0.5以上なので、リンク3102に対し以降の処理が実施される。

【0239】

同様にリンク3103に対する繰り返し処理が開始され、ステップS905まで処理が進む。 10

【0240】

ステップS905においてリンク解析部115は、リンク3103が示す文書を取得する。このときリンク3103が示す文書のタイトルが「カンノン：ホームページ」であったとする。

【0241】

ステップS3001においてリンク解析部115は、リンク元である文書3101のタイトルとステップS905で取得したリンク先文書のタイトルとから関連度を算出する。

【0242】

リンク先文書のタイトルからは「カンノン」「ホームページ」の2つの単語が得られる。 20

【0243】

一致する単語は「カンノン」のみであり、関連度は $(1 \times 2) / (7+2) = 2/9 \quad 0.22222$ となる。

【0244】

ステップS3002においてリンク解析部115は、関連度が規定値0.5未満なので、リンク3103に対し以降の処理が実施されない。

【0245】

上述したように、リンク先の文書に対し、主題が同一であると判断された文書へのリンクのみを処理対象とすることが可能となり、処理効率および処理精度とも更に向上させることが可能となる。 30

【0246】

[第4の実施形態]

第4の実施形態では、第1乃至第3の実施形態の構成及び処理に基づくものであるが、まず、図32に示すフローチャートを用いて第4の実施形態において、SNSにて投稿された文書内のリンクを解析する処理の詳細について説明する。

【0247】

図32に示すフローチャートは図12に示すフローチャートに対し、ステップS3201の処理が追加されている。

【0248】

ステップS3201においてリンク解析部115は、直接リンク先の文書におけるタイトルからタグを抽出し、処理対象の投稿と関連付けて、リンク情報保存領域116におけるタグ情報テーブル1002に登録する。 40

【0249】

(具体例)

第1の実施形態の例における図8の投稿に対し、第4の実施形態におけるリンク処理を実施すると、投稿805に含まれているリンク「<http://www.kshop.co.jp/kannon32d.htm>」に対しては、ステップS1206まで処理が進む。

【0250】

ステップS3201においてリンク解析部115は、処理対象である直接リンク「[http](http://www.kshop.co.jp/kannon32d.htm)」 50

://www.kshop.co.jp/kannon32d.html」からタグを取得する。直接リンクが示す文書のタイトルが「カンノン32D最安値！」であったとすると、「32D」をタグとして取得し、タグ情報テーブル1002に登録する。

【0251】

結果としてタグ情報テーブル1002は図33に示した状態となり、図34に示した集計結果を表示することが可能となる。

【0252】

上述したように、ノイズとなる可能性もあるが、公式サイトまで辿れない文書も分析の対象とすることが可能となる。

【0253】

以上、本実施形態によれば、インターネット等で展開されるサービスに於いて行われる企業や製品等に関する意見や評価等の投稿に含まれるリンクに直接、あるいは間接問わず関連付けられる企業サイト（到達する企業サイトを用いて評価を行う）ことで、分析対象となる企業や製品等の範囲を幅広く抑えると共に、その意見や評価等の対象を適切に特定することで、妥当な評価を行うことができる。

【0254】

以上のように、前述した実施形態の機能を実現するプログラムを記録した記録媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記録媒体に格納されたプログラムを読み取り実行することによっても、本発明の目的が達成されることは言うまでもない。

【0255】

この場合、記録媒体から読み出されたプログラム自体が本発明の新規な機能を実現することになり、そのプログラムを記憶した記録媒体は本発明を構成することになる。プログラムを供給するための記録媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、DVD-ROM、磁気テープ、不揮発性のメモリカード、ROM、EEPROM、シリコンディスク等を用いることができる。

【0256】

また、コンピュータが読み出したプログラムを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムの指示に基づき、コンピュータで稼働しているOS等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0257】

さらに、記録媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPU等が実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0258】

また、本発明は、システムあるいは装置にプログラムを供給することによって達成される場合にも適応できることは言うまでもない。この場合、本発明を達成するためのプログラムを格納した記録媒体を該システムあるいは装置に読み出すことによって、そのシステムあるいは装置が、本発明の効果を享受することが可能となる。

【0259】

さらに、本発明を達成するためのプログラムをネットワーク上のサーバ、データベース等から通信プログラムによりダウンロードして読み出すことによって、そのシステム、あるいは装置が、本発明の効果を享受することが可能となる。なお、上述した各実施形態およびその変形例を組み合わせた構成も全て本発明に含まれるものである。

【符号の説明】

【0260】

10

20

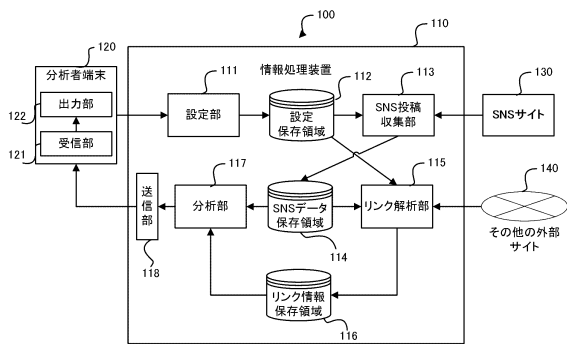
30

40

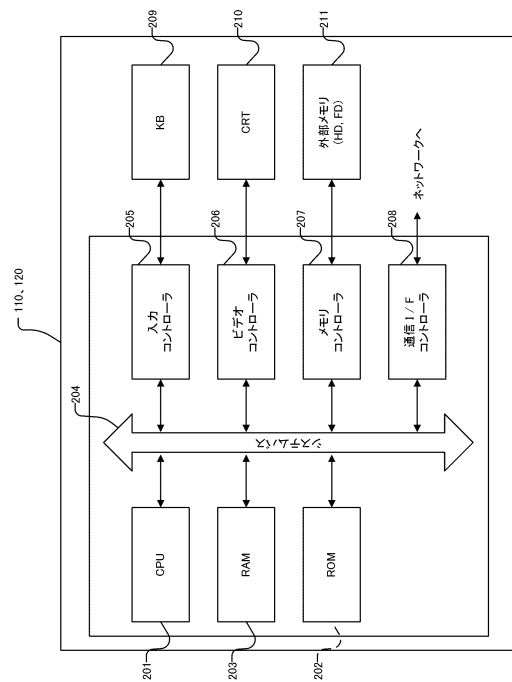
50

- 1 0 0 情報処理装置
- 1 2 0 分析者端末
- 1 3 0 SNSサイト
- 1 4 0 その他の外部ネットワーク
- 2 0 1 CPU
- 2 0 2 RAM
- 2 0 3 ROM
- 2 0 4 システムバス
- 2 0 5 入力コントローラ
- 2 0 6 ビデオコントローラ
- 2 0 7 メモリコントローラ
- 2 0 8 通信I/F(インターフェース)コントローラ
- 2 0 9 キーボード(KB)
- 2 1 0 ディスプレイ装置(CRT)
- 2 1 1 外部メモリ

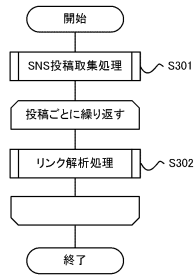
【図1】



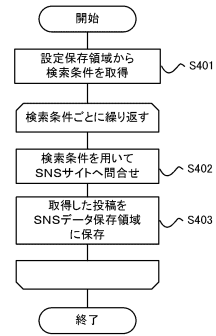
【図2】



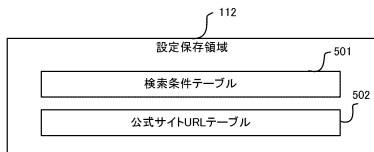
【図3】



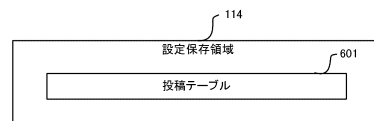
【図4】



【図5】



【図6】



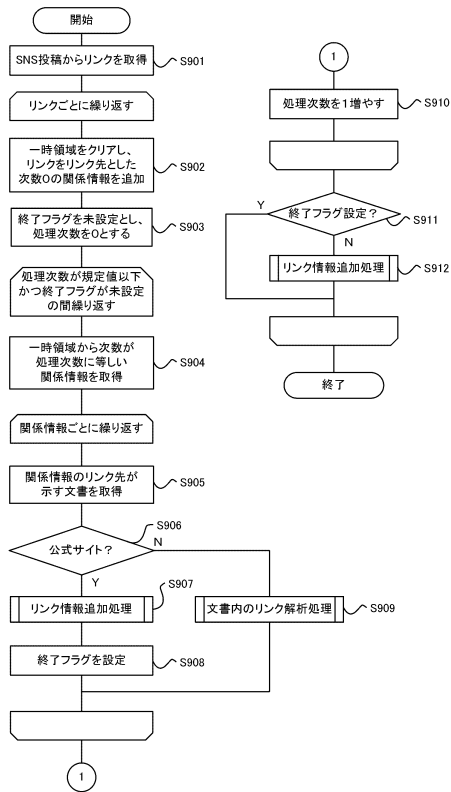
【図7】

id	検索内容	最終取得時刻
1	KANNON OR kannon OR カノン	2013-07-12 01:02:30
2	NIKKOU OR nikkou OR ニッコウ	2013-07-12 01:05:43
.	.	.

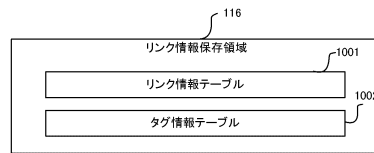
【図8】

id	ユーザ	検索内容	最終取得時刻
1	karawa	セミナー「カノン、3D mk 11 発表のお知らせ」 http://www.kannon.jp/news/2013/07/11.html	2013-07-12 03:00:20
2	negatomo	来ましたね。Nikkouのヤツと比較してみただ、いいな。 http://nbl.loc.ne.jp/nakatomo/2013/07/11/2211.html	2013-07-12 03:00:20
3	uchiida	報告。「Jnews - カノンから 3D mk 11 24日より発売」 http://www.jnews.co.jp/2013-07/11-kannon.html	2013-07-12 03:00:20
4	okazaki	予約した。 http://www.kannon-shop.co.jp/kannon3dmk2.html	2013-07-12 03:00:20
5	kshop	カノン 3D発表会場 http://www.kshop.co.jp/kannon3d.html	2013-07-12 03:00:20
6	honda	これ欲しいな。「Jnews - カノンから 3D mk 11 24日より発売」 http://www.jnews.co.jp/2013-07/11-kannon.html	2013-07-12 03:00:40

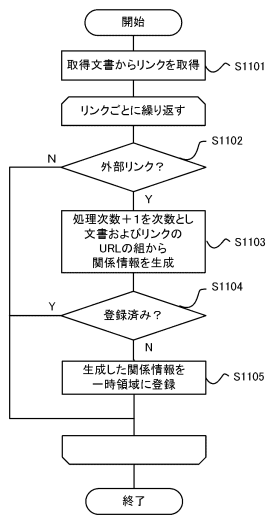
【図9】



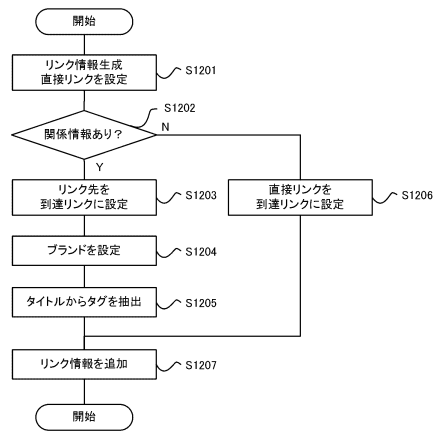
【図10】



【図 1 1】



【図 1 2】



【図 1 3】

リンク元URL	リンク先URL	次数
	http://www.kannon.jp/news20130711.html	0

1301

【図 1 4】

```

<html>
<title>カンノン : 32D mk II 発売のお知らせ</title>
<body>
<div>
:
<a ref="http://www.kannon.jp/index.html">トップ</a>
<a ref="http://www.kannon.jp/32dmk2/32dmk2top.html">製品詳細</a>
</div>
</body>
</html>
  
```

1401

【図15】

公式サイトURLパターン文字列	ブランド
kannon.jp	kannon
kannon.co.jp	kannon
kannon-shop.co.jp	kannon
nikkou.com	nikkou
nikkou.jp	nikkou

【図16】

投稿id	直接リンクURL	到達URL	ブランド
1	http://www.kannon.jp/news/20130711.html	http://www.kannon.jp/news/20130711.html	kannon
2	http://nb.log.ne.jp/nagatomo/201307112211.html	http://www.kannon.jp/news/20130711.html	kannon

【図17】

投稿ID	タグ	ブランド
1	"32D mk l"	kannon
2	"32D mk l"	kannon

【図18】

リンク元URL	リンク先URL	回数
http://nb.log.ne.jp/nagatomo/201307112211.html	http://nb.log.ne.jp/nagatomo/201307112211.html	0
http://nb.log.ne.jp/nagatomo/201307112211.html	http://www.jnews.co.jp/20120127.html	1
http://nb.log.ne.jp/nagatomo/201307112211.html	http://www.ashop.co.jp/kannon32d.html	1
http://www.jnews.co.jp/20120127.html	http://www.kannon.jp/news/20130711.html	2

【 図 1 9 】

```

<html>
<title> MAGATOMIC BLOG | KANNON 32D mk2 になって何が変わるのか？</title>
<body>
<div>
<div>
<div>
<a href="http://www.jnews.co.jp/2013/07/27.html">「JNews - カノンから 32D mk 11 24日
より発売」</a>
1902
<a href="http://blog.ne.jp/nagatomo/201307202032.html">前の記事</a>
</div>
</div>
</div>
</div>
</body>
</html>

```

【 図 2 0 】

```

<html>
<title> JNews - カノンから 32D mk 11 24日より発売</title>
<body>
<div>
25日、カノン株式会社より、32Dの後継機として24日より32D mk 11の発売を開始する…
<a href="http://www.kannon.jp/news/20130711.html">リリースノート</a>
<div>関連記事</div>
<a href="http://www.jnews.co.jp/2013-05/27-kannon.html">カノン 32D mk11 β機使用レポート</a>
2003
</div>
</body>
</html>

```

【 図 2 1 】

投稿ID	直接リンクURL	到達URL	ブランド
1	http://www.kannon.jp/news/20130711.html	http://www.kannon.jp/news/20130711.html	kannon
2	http://blog.ne.jp/nagatomo/201307112211.html	http://www.kannon.jp/news/20130711.html	kannon
3	http://www.jnews.co.jp/2013-07/11-kannon.html	http://www.kannon.jp/news/20130711.html	kannon
4	http://www.kannon-shop.co.jp/kannon32mk2.html	http://www.kannon-shop.co.jp/kannon32mk2.html	kannon
5	http://www.eshop.co.jp/kannon32d.html	http://www.eshop.co.jp/kannon32d.html	-
6	http://www.jnews.co.jp/2013-07/11-kannon.html	http://www.kannon.jp/news/20130711.html	kannon

【 図 2 2 】

投稿ID	タグ	ブランド
1	32D mk 11	kannon
2	32D mk 11	kannon
3	32D mk 11	kannon
4	32D mk 11	kannon
6	KANNON	kannon
6	32D mk 11	kannon

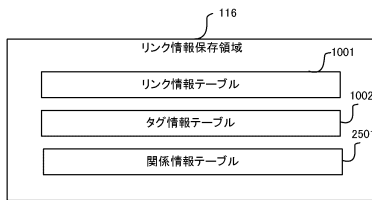
【図23】

到達URL	投稿数
http://www.kannon.jp/news20130711.html	4
http://www.kannon-shop.co.jp/kannon32dmk2.html	1
http://www.kshop.co.jp/kannon32d.html	1

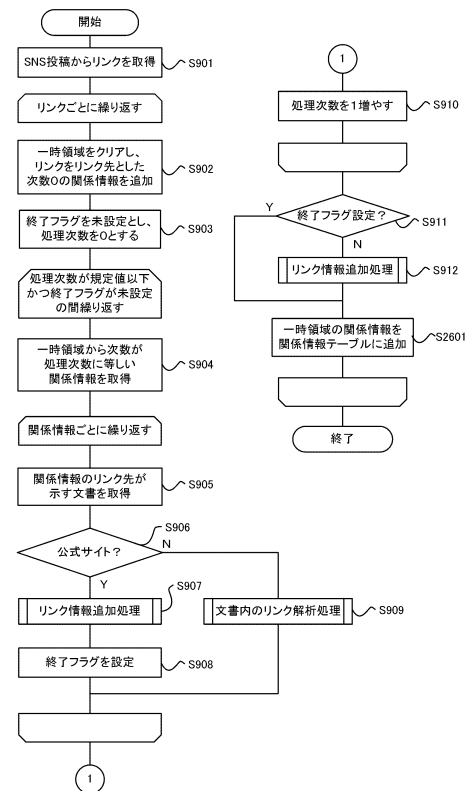
【図24】

タグ	投稿数
32D mk	5
KANNON	1

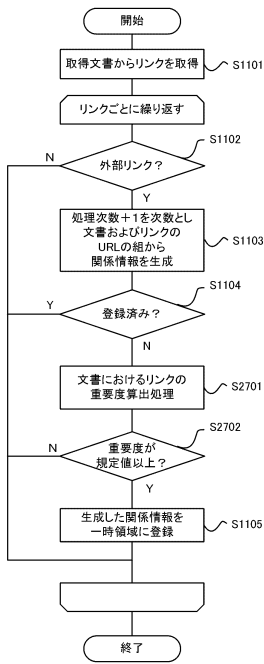
【図25】



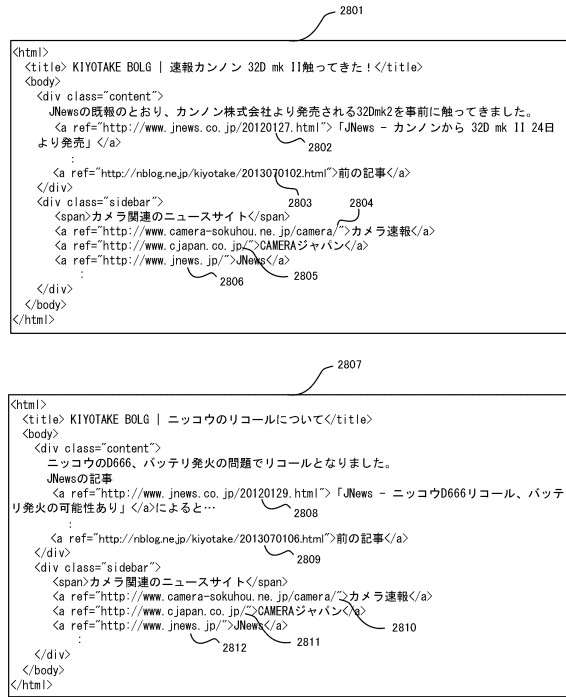
【図26】



【図 27】



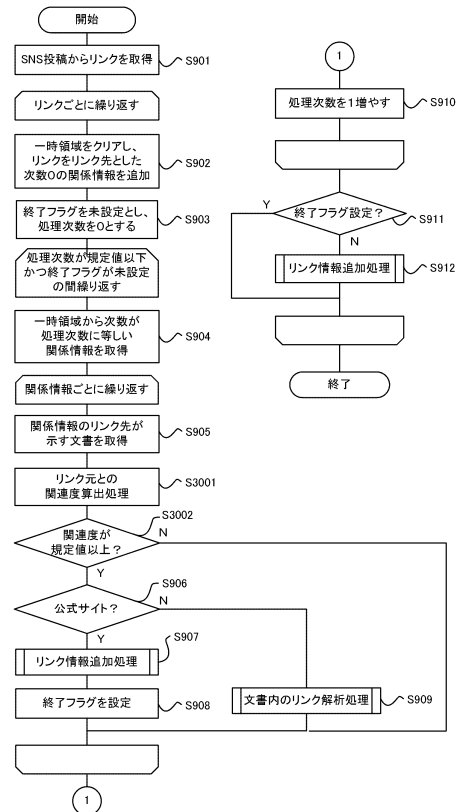
【図 28】



【図 29】

リンク元URL	リンク先URL	リンク元オーリディ
...
http://nblog.ne.jp/kiyotake/2013070101.html	http://www.jnews.co.jp/20120127.html	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070101.html	http://www.camera-sokuhou.ne.jp/camera/	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070101.html	http://www.cjapan.co.jp	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070101.html	http://www.jnews.jp	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070102.html	http://www.cjapan.co.jp/article/20120127.html	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070102.html	http://www.camera-sokuhou.ne.jp/camera/	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070102.html	http://www.cjapan.co.jp	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070102.html	http://www.jnews.jp	http://nblog.ne.jp/
...
http://nblog.ne.jp/kiyotake/2013070105.html	http://www.cjapan.co.jp/article/20120128.html	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070105.html	http://www.camera-sokuhou.ne.jp/camera/	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070105.html	http://www.cjapan.co.jp	http://nblog.ne.jp/
http://nblog.ne.jp/kiyotake/2013070105.html	http://www.jnews.jp	http://nblog.ne.jp/
...

【図 30】



【図 3 1】

```

<html>
<title> JNews - カンノンから 32D mk II 24日より発売</title>
<body>
<div class="content">
.
.
.
<a ref="http://www.kannon.jp/news/20150711.html">リリースノート</a>
</div>
</body>
</html>
    
```

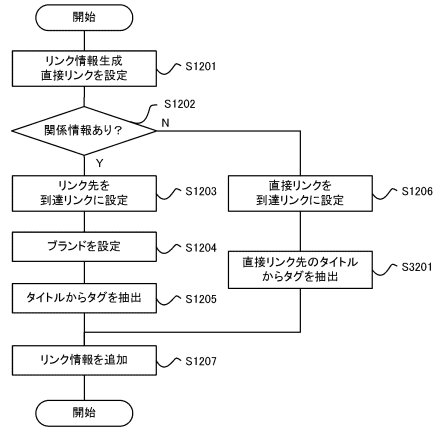
3101

3102

3103

23日、カンノン株式会社より、32Dの後継機として24日より32D mk IIの発売を開始する...

【図 3 2】



【図 3 3】

1002

投稿ID	タグ	ブランド
1	32D mk II	kannon
2	32D mk II	kannon
3	32D mk II	kannon
4	32D mk II	kannon
5	32D	
6	KANNON	kannon
6	32D mk II	kannon

【図 3 4】

タグ	投稿数
32D mk II	5
KANNON	1
32D	1

フロントページの続き

(51)Int.Cl. F I
G 0 6 Q 30/02 3 1 2

(56)参考文献 特開2013-092851(JP,A)
特開2011-108053(JP,A)

(58)調査した分野(Int.Cl., DB名)
G 0 6 F 1 7 / 2 0 - 1 7 / 3 0
G 0 6 Q 3 0 / 0 2