

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0613079-8 A2**



* B R P I 0 6 1 3 0 7 9 A 2 *

(22) Data de Depósito: 28/06/2006
(43) Data da Publicação: 21/12/2010
(RPI 2085)

(51) *Int.Cl.:*
G06F 17/30

(54) Título: **MÉTODO IMPLEMENTADO POR COMPUTADOR, APARELHO E MEIO LEGÍVEL POR COMPUTADOR**

(30) Prioridade Unionista: 29/06/2005 US 11/170.725

(73) Titular(es): GOOGLE INC

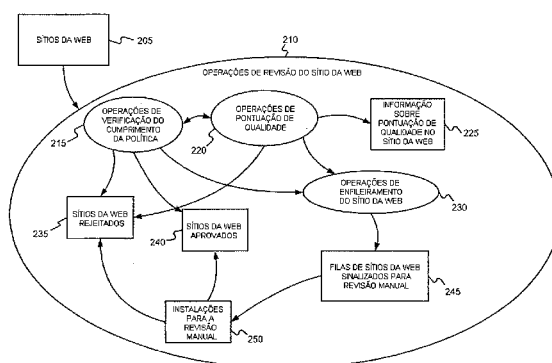
(72) Inventor(es): DMITRIY PORTNOV, GOKUL RAJARAM, LAWRENCE IP, PAVAN KUMAR DESIKAN, PRASENJIT PHUKAN, SANJEEV KULKARNI, TIMOTHY JAMES

(74) Procurador(es): Orlando de Souza

(86) Pedido Internacional: PCT US2006025244 de 28/06/2006

(87) Publicação Internacional: WO 2007/002785de 04/01/2007

(57) **Resumo:** MÉTODO IMPLEMENTADO POR COMPUTADOR, APARELHO E MEIO LEGÍVEL POR COMPUTADOR. A maneira pela qual os sites da Web são revistos para utilização em uma rede de propaganda poderá ser aprimorada pela (a) aceitação de uma coleção incluindo um ou mais documentos, (b) determinar se a coleção cumpre ou não com políticas de uma rede de propaganda, e (c) aprovar a coleção se foi determinado que a coleção cumpre com uma ou mais políticas tal que (por exemplo, direcionada por conteúdo) os anúncios poderão ser servidos em associação com os documentos na coleção. A política poderá abranger (A) conteúdo de um ou mais documentos, (B) capacidade de uso do site da Web e/ou (C) fraude ou decepção possíveis na rede de propaganda ou nos participantes da rede de propaganda.



MÉTODO IMPLEMENTADO POR COMPUTADOR, APARELHO E MEIO LEGÍVEL
POR COMPUTADOR

1-HISTÓRICO DA INVENÇÃO

1.1-CAMPO DA INVENÇÃO

5 A presente invenção abrange ajudar a assegurar uma
experiência de qualidade em uma rede de propaganda de
anúncios. Em particular, a presente invenção abrange
determinar se um sítio da Web cumpre ou não com políticas,
e/ou satisfaz certos padrões de qualidade de uma rede de
10 propaganda.

1.2-HISTÓRICO DA INVENÇÃO

 A propaganda com a utilização da mídia tradicional,
como a televisão, o rádio, jornais e revistas, é bem
conhecida. Infelizmente, mesmo quando armado de estudos
15 demográficos e suposições inteiramente razoáveis sobre a
audiência típica de várias saídas da mídia, os anunciantes
reconhecem que boa parte de seu orçamento de propaganda é
simplesmente desperdiçada. Além disso, é muito difícil
identificar e eliminar esse desperdício.

20 Recentemente, tornou-se popular anunciar através da
mídia mais interativa. Por exemplo, na medida em que o
número de pessoas que utilizam a Internet explodiu, os
anunciantes passaram a apreciar a mídia e os serviços
ofertados pela Internet como uma maneira potencialmente
25 poderosa de anunciar.

 A propaganda interativa fornece oportunidades para os
anunciantes direcionar seus anúncios para uma audiência
receptiva. Isto é, os anúncios dirigidos têm maior
probabilidade de serem úteis para os usuários finais, pois
30 os anúncios poderão ser relevantes a uma necessidade

inferida de alguma atividade do usuário (por exemplo, relevante para a busca de uma consulta a um motor de busca, relevante para o conteúdo em um documento solicitado pelo usuário, etc.). O direcionamento de palavras chaves em 5 consultas tem sido utilizado pelas ferramentas de busca para entregar anúncios relevantes. Por exemplo, o sistema de propaganda AdWords de Google de Mountain View, CA, entrega anúncios direcionados a palavras chaves de consultas de busca. De modo similar, sistemas de entrega de 10 anúncios direcionados por conteúdo foram propostos. Por exemplo, os pedidos de Patentes dos Estados Unidos de número de série: 10/314.427 (aqui incorporado por referência e referido como "a aplicação 427") intitulada "METHODS AND APPARATUS FOR SERVING RELEVANT - 15 ADVERTISEMENTS", requerida em 6 de dezembro de 2002, e que relaciona Jeffrey A Dean, Georges R. Harik e Paul Buchheit como inventores; e 10/375.900 (aqui incorporada por referência e referida como "a aplicação 900") intitulada "SERVING ADVERTISEMENTS BASED ON CONTENT", requerida em 26 20 de fevereiro de 2003 e que relaciona Darrell Anderson, Paul Buchheit, Alex Carobus, Claire Cui, Jeffrey A. Dean, Georges R. Harik, Deepak Jindal e Narayanan Shivakumar como inventores, descrevem métodos e aparelho para servir anúncios relevantes para o conteúdo de um documento, como 25 uma página da Web, por exemplo. Os sistemas de entrega de anúncios dirigidos por conteúdo, como o sistema de propaganda AdSense de Google, por exemplo, foram utilizados para servir anúncios em páginas da Web.

Como pode ser apreciado pelo que antecede, servir 30 anúncios relevante aos conceitos do texto em um texto é

útil, pois esses anúncios supostamente abrangem um interesse atual do usuário. Conseqüentemente, essa propaganda dirigida ao conteúdo tornou-se crescentemente popular. No entanto, esses sistemas de propaganda ainda têm
5 espaço para melhorias.

O sucesso das redes de anúncio dirigidas pelo conteúdo, como o programa Google AdSense, por exemplo, depende de três constituintes importantes - anunciantes, editores do sítio da Web (referidos abaixo como "editoras
10 da Rede"), e os usuários finais. Muitos anunciantes querem ser assegurados de que os sítios da Web em que seus anúncios são exibidos são de alta qualidade, irão gerar um retorno positivo sobre o investimento (ROI), e não causarão qualquer controvérsia por eles. As editoras da Rede
- 15 precisam continuar a atrair usuários finais através de conteúdo relevante e de alta qualidade. Por último, os usuários finais geralmente desejam visitar apenas aqueles sítios da Web que fornecem uma boa experiência. Ainda, os usuários finais geralmente selecionam apenas aqueles
20 anúncios que são relevantes ao contexto da página da Web e que os ajudam a encontrar o produto ou serviço que satisfaça suas necessidades. Uma rede de alta qualidade, assim, é crítica para assegurar que seja criado um ciclo virtuoso que ajude a expandir a rede de propaganda ao
25 atrair novas editoras, publicitários, e usuários finais.

As maneiras atuais de assegurar a qualidade de um sítio da Web que participa de uma rede de propaganda incluem, por exemplo, a revisão dos sítios da Web manualmente (por exemplo, para assegurar que eles cumprem
30 as orientações de políticas, que eles não criem uma fraca

experiência para o usuário final, e que o conteúdo permite que anúncios relevantes sejam dirigidos). Infelizmente, a revisão manual dos sítios da Web está mal equipada para lidar com o crescimento explosivo experimentado pelas redes de propaganda online.

Tendo em vista as dificuldades da revisão manual, de muitos novos sítios da Web de maneira eficaz de tempo e de custo, algumas redes de propaganda limitaram essa revisão manual para apenas aqueles sítios da Web sobre os quais os anunciantes e/ou os usuários finais reclamaram. Infelizmente, os usuários finais e/ou os anunciantes poderão exigir uma abordagem mais proativa. Portanto, ter um processo de aprovação de sítio da Web de propaganda mais proativo seria útil.

15 2-SINOPSE DA INVENÇÃO

Modalidades consistentes com a presente invenção melhorar a maneira pela qual os sítios da Web são revistos para utilização em uma rede de propaganda. Pelo menos algumas modalidades consistentes com a presente invenção poderão (a) aceitar uma coleção incluindo um ou mais documentos, (b) determinar se uma coleção cumpre ou não com as políticas de uma rede de propaganda, e (c) aprovar a coleção se foi determinado que a coleção cumpre com as políticas.

25 Em pelo menos algumas modalidades consistentes com a presente invenção, a coleção é acrescentada à rede de propaganda se a coleção é aprovada tal que (por exemplo, dirigida por conteúdo) anúncios poderão ser servidos em associação com renderizações de documentos incluídos na
30 coleção.

Em pelo menos algumas modalidades consistentes com a presente invenção, a coleção que inclui um ou mais documentos é um sítio da Web que inclui uma ou mais páginas da Web.

5 Em pelo menos algumas modalidades consistentes com a presente invenção, a política abrange (A) conteúdo do um ou mais documentos da coleção, (B) utilidade do sítio da Web, em que a coleção de um ou mais documentos é um sítio da Web que inclui uma ou mais páginas da Web, e/ou (C) fraude ou
10 decepção possíveis na rede de propaganda ou dos participantes da rede de propaganda pela coleção.

3-DESCRIÇÃO SUCINTA DOS DESENHOS

A Figura 1 é um diagrama que mostra as partes ou entidades que podem interagir com uma rede de propaganda, como uma rede de propaganda dirigida por conteúdo, por
15 exemplo.

A Figura 2 é um diagrama de bolhas das operações exemplares para verificar automaticamente os sítios da Web pelo cumprimento das políticas e/ou da pontuação dos sítios da Web, que poderá ser efetuado de maneira consistente com
20 a presente invenção, bem como informação que poderá ser utilizada e/ou gerada por tais operações.

A Figura 3 é um diagrama de fluxo de um método exemplar que poderá ser utilizado para determinar se os
25 sítios da Web cumprem com as políticas de uma maneira consistente com a presente invenção.

A Figura 4 é um diagrama de fluxo de um método exemplar que poderá ser utilizado para determinar uma pontuação de qualidade para os sítios da Web de uma maneira
30 consistente com a presente invenção.

A Figura 5 é um diagrama de fluxo de um método exemplar que poderá ser utilizado para sinalizar e enfileirar os sítios da Web para uma revisão mais detalhada (por exemplo, manual) de maneira consistente com a presente invenção.

A Figura 6 é um diagrama de fluxo de um método exemplar que poderá ser utilizado para processar sítios da Web revistos manualmente de uma maneira consistente com a presente invenção.

10 A Figura 7 é um diagrama de blocos de um aparelho exemplar que poderá efetuar várias operações de uma maneira consistente com a presente invenção.

4-DESCRIÇÃO DETALHADA

A presente invenção poderá envolver métodos, aparelho, . 15 formatos de mensagens, e/ou estruturas de dados para verificar sítios da Web quanto a violações de política, e/ou determinar pontuações de qualidade para os sítios da Web, que poderão ser utilizados para ajudar uma rede de propaganda a assegurar que ela contém apenas sítios da Web
20 de alta qualidade. A descrição seguinte é apresentada para permitir que alguém habilitado na tecnologia faça e utilize a invenção, e é fornecida no contexto de aplicações particulares e de seus requisitos. Assim, a descrição seguinte de modalidades consistentes com a presente
25 invenção fornece ilustração e descrição, mas não pretende ser abrangente ou limitar a presente invenção à forma precisa revelada. Várias modificações nas modalidades reveladas ficarão aparentes para aqueles habilitados na tecnologia, e os princípios gerais aqui apresentados
30 poderão ser aplicados a outras modalidades e aplicações.

Por exemplo, embora uma série de atos poderá ser descrita com referência a um diagrama de fluxo, a ordem dos atos poderá diferir em outras implementações quando o desempenho de um ato não estiver dependente do término de outro ato.

5 Ainda, atos não dependentes poderão ser efetuados em paralelo. Nenhum elemento, ato ou instrução utilizado na descrição deve ser considerado como crítico ou essencial para a presente invenção a menos que seja explicitamente descrito como tal. Também, como é aqui utilizado, o artigo
10 "um/uma" pretende incluir um ou mais itens. Quando for pretendido apenas um item, o termo "um (1)" ou linguagem similar é utilizado. Assim, a presente invenção não pretende ser limitada às modalidades mostradas e os inventores consideram que sua invenção inclui qualquer
- 15 matéria patenteável descrita.

No que se segue, definições dos termos que poderão ser utilizados na especificação estão fornecidos sob o título "4.1". Depois, os ambientes em que, ou com os quais, a presente invenção poderá operar, estão descritos sob o
20 título "4.2". Modalidades exemplares da presente invenção são descritas sob o título "4.3". Finalmente, algumas conclusões sobre a presente invenção são apresentadas sob o título "4.4".

4.1-DEFINIÇÕES

25 Anúncios online possuem várias características intrínsecas. Essas características poderão ser especificadas por uma aplicação e/ou um anunciante. Essas características são referidas abaixo como "características do anúncio". Por exemplo, no caso de um anúncio de texto,
30 as características do anúncio poderão incluir uma linha de

título, texto do anúncio, e um enlace embutido. No caso de um anúncio com imagem, as características do anúncio poderão incluir imagens, código executável, e um enlace embutido. Dependendo do tipo de anúncio online, as

5 características do anúncio poderão incluir um ou mais do seguinte: texto, um enlace, um arquivo de áudio, um arquivo de vídeo, um arquivo de imagem, código executável, informação embutida, etc.

Quando um anúncio online é servido, um ou mais

10 parâmetros poderão ser utilizados para descrever como, quando, e/ou onde o anúncio foi servido. Esses parâmetros são referidos abaixo como "parâmetros de servir". Parâmetros de servir, por exemplo, poderão incluir um ou mais do seguinte: características (incluindo informação)

15 sobre o documento em que, ou com o qual, o anúncio foi servido, uma consulta de busca ou resultados de busca associados ao servir do anúncio, uma característica do usuário (por exemplo, sua localização geográfica, o idioma utilizado pelo usuário, o tipo de navegador utilizado,

20 visões de páginas anteriores, comportamento anterior, conta do usuário, qualquer cookie da Rede utilizado pelo sistema, características do dispositivo do usuário, etc.), o sítio hospedeiro ou afiliado (por exemplo, América Online, Google, Yahoo) que iniciou a solicitação, a posição

25 absoluta do anúncio na página em que ele foi servido, a posição (espacial ou temporal) do anúncio em relação a outros anúncios servidos, o tamanho absoluto do anúncio, o tamanho do anúncio em relação aos demais anúncios, a cor do anúncio, o número de outros anúncios servidos, tipos de

30 outros anúncios servidos, hora do dia em que foi servido,

hora da semana servida, hora do ano servido, etc. Naturalmente, há outros parâmetros de servir que poderão ser utilizados no contexto da invenção.

Embora os parâmetros de servir poderão ser extrínsecos
5 às características do anúncio, eles poderão ser associados ao anúncio como condições ou restrições de servir. Quando utilizados como condições ou restrições de servir, esses parâmetros de servir são simplesmente referidos como "restrições de servir" (ou "critérios de direcionamento").
10 Por exemplo, em alguns sistemas, o anunciante poderá ser capaz de direcionar o servir de seu anúncio ao especificar que ele só deve ser serviços nos dias úteis, não abaixo de certa posição, apenas para os usuários em determinada localização, etc. Como outro exemplo, em alguns sistemas, o
15 anunciante poderá especificar que seu anúncio só deverá ser servido se a página ou consulta de busca incluir certas palavras chaves ou frases. Como ainda outro exemplo, em alguns sistemas, o anunciante poderá especificar que seu anúncio deverá ser servido apenas se o documento em que, ou
20 com o qual, o anúncio deverá ser servido, inclui certos tópicos ou conceitos, ou se enquadra sob um grupo ou agrupamentos particulares, ou alguma outra classificação ou classificações (por exemplo, verticais). Em alguns sistemas, o anunciante poderá especificar que seu anúncio
25 deverá ser servido apenas para (ou não deverá ser servido para) dispositivos do usuário dotados de certas características. Finalmente, em alguns sistemas o anúncio pode ser direcionado de modo a ser servido em resposta a uma solicitação originada de uma localidade particular, ou
30 em resposta a uma solicitação sobre uma localidade

particular.

"Informação do anúncio" poderá incluir qualquer combinação de características do anúncio, restrições de servir o anúncio, informação derivada das características do anúncio ou das restrições de servir o anúncio (referidas como "informação derivada do anúncio"), e/ou informação relacionada ao anúncio (referida como "informação relacionada ao anúncio"), bem como uma extensão dessa informação (por exemplo, informação derivada de informação relacionada ao anúncio).

A proporção do número de seleções (por exemplo, cliques) do anúncio ao número de impressões do anúncio (isto é, o número de vezes que o anúncio é renderizado) é definido como a "taxa de seleção" (ou "taxa de cliques") do anúncio.

Diz-se que ocorreu uma "conversão" quando o usuário consoma uma transação relacionada a um anúncio anteriormente servido. O que constitui a conversão poderá variar de um caso para outro e pode ser determinado de uma variedade de formas. Por exemplo, poderá ser o caso que a conversão ocorre quando o usuário clica no anúncio, é referido à página da Web do anunciante, e consoma a compra lá antes de deixar aquela página da Web. Alternativamente, a conversão poderá ser definida como o usuário ter um anúncio a ele exibido, e fazer a compra na página da Web do anunciante dentro de um tempo predeterminado (por exemplo, sete dias). Em ainda outra alternativa, a conversão poderá ser definida pelo anunciante como sendo qualquer ação do usuário mensurável/observável como, por exemplo, baixar um trabalho, navegar por pelo menos uma profundidade dada de

um sítio da Web, visualizar pelo menos um certo número de páginas da Web, gastar pelo menos uma quantidade de tempo predeterminada em um sítio da Web ou uma página da Web, registrar em um sítio da Web, etc. Muitas vezes, se a ação do usuário não indicarem uma compra consumada, elas poderão indicar uma dica de venda, embora ações do usuário que constituem uma conversão não estão limitadas a essas. Na verdade, muitas outras definições daquilo que constitui uma conversão são possíveis.

10 A proporção entre o número de conversões e o número de impressões do anúncio (isto é, o número de vezes que o anúncio é renderizado) e a proporção entre o número de conversões e o número de seleções (ou o número de algum outro evento anterior) são ambos referidos como a "taxa de conversão". O tipo de taxa de conversão será aparente do contexto em que ela é utilizada. Se a conversão é definida como sendo capaz de ocorrer dentro de um tempo predeterminado desde o servir do anúncio, uma definição possível da taxa de conversão pode considerar apenas os
15 anúncios que foram servidos mais do que o tempo predeterminado no passado.

Uma "propriedade" é algo em que os anúncios podem ser apresentados. A propriedade poderá incluir conteúdo online (por exemplo, um sítio da Web, um programa de áudio MP3, jogos online, etc.), conteúdo offline (por exemplo, um
25 jornal, uma revista, uma produção teatral, um concerto, um evento esportivo, etc.), e/ou objetos offline (por exemplo, um cartaz, placar em estádios, parede em estádios, a lateral de caminhões, etc.). Propriedades com conteúdo (por exemplo, revistas, jornais, sítios da Web, mensagens de
30

correspondência eletrônica, etc.) poderão ser referidas como "propriedades de mídia". Embora as propriedades poderão elas próprias estar offline, a informação pertinente a respeito de uma propriedade (por exemplo, atributos, tópicos, conceitos, categorias, palavras chaves, informação de relevância, tipos de anúncios suportados, etc.) poderá estar disponível online. Por exemplo, um festival de música de jazz ao ar livre poderá ter entrado os tópicos "musica" e "jazz", a localização dos concertos, a hora dos concertos, artistas programados para aparecer no festival, e tipos de espaços de propaganda disponíveis (por exemplo, pontos em um programa impresso, pontos no palco, pontos em forros de assentos, anúncios de áudio dos patrocinadores, etc.).

O "documento" é para ser interpretado amplamente como incluindo qual produto de trabalho legível ou armazenado por máquina. Um documento poderá ser um arquivo, uma combinação de arquivos, um ou mais arquivos com enlaces embutidos para outros arquivos, etc. Os arquivos poderão ser de qualquer tipo, como de texto, de áudio, de imagem, de vídeo, etc. Partes de um documento a ser renderizado para o usuário final podem ser consideradas como "conteúdo" do documento. Um documento poderá incluir "dados estruturados" contendo tanto conteúdo (palavras, fotos, etc.) e alguma indicação do significado daquele conteúdo (por exemplo, campos de correspondência eletrônica e dados associados, etiquetas HTML e dados associados, etc.). Pontos de anúncios no documento poderão ser definidos por informação ou instruções embutidas. No contexto da Internet, um documento comum é a página da Web. Páginas da

Web muitas vezes incluem conteúdo e poderão incluir
informação embutida (como meta informação, hyperlinks, etc)
e/ou instruções embutidas (como JavaScript, etc.). Em
muitos casos, o documento tem um local de armazenamento
5 endereçável e pode, portanto, ser identificado
singularmente por este local endereçável. Um locador de
recurso universal (URL) é o endereço utilizado para acessar
informação na Internet.

Um "documento da Rede" inclui qualquer documento
10 publicado na Rede. Exemplos de documentos da Rede incluem,
por exemplo, um sítio da Web ou uma página da Web.

A "informação do documento" poderá incluir qualquer
informação incluída no documento, informação derivada da
informação incluída no documento (referida como "informação
15 derivada do documento") e/ou informação relacionada com o
documento (referida como "informação relacionada ao
documento"), bem como extensões de tal informação (por
exemplo, informação derivada de informação relacionada). Um
exemplo de informação derivada do documento é uma
20 classificação com base no conteúdo textual de um documento.
Exemplos de informação relacionada ao documento informação
de documento de outros documentos com enlaces para o
documento em tela, bem como informação de documento de
outros documentos aos quais o documento em tela se enlaça.

25 O conteúdo de um documento poderá ser renderizado em
"aplicação ou dispositivo de renderização de conteúdo".
Exemplos de aplicações de renderização de conteúdo incluem
um navegador da Internet (por exemplo, Explorer, Netscape,
Opera, Firefox, etc.), um reprodutor de mídia (por exemplo,
30 um reprodutor MP3, um reprodutor de arquivo de áudio

streaming Realnetworks, etc.), um visualizador (por exemplo, uma leitora pdf da Adobe Acrobat), etc.

Um "proprietário do conteúdo" é a pessoa ou entidade que tem algum direito de propriedade no conteúdo do documento. Um proprietário do conteúdo poderá ser o autor do conteúdo. Além disso, ou alternativamente, o proprietário do conteúdo poderá ter o direito para reproduzir o conteúdo, direito para preparar obras derivativas do conteúdo, direito de exibir ou efetuar o conteúdo publicamente, e/ou outros direitos interditivos no conteúdo. Embora o servidor do conteúdo pode ser um proprietário do conteúdo no conteúdo dos documentos que ele serve, isto não é necessário. O "editor da Rede" é um exemplo de um proprietário do conteúdo.

A "política" de uma rede de propaganda é uma regra que é determinativo de se a página da Web ou o sítio da Web (ou algum outro documento ou coleção de documentos) poderá participar na rede de propaganda. Isto é, se uma página da Web ou um sítio da Web violar uma política, a rede de propaganda a proibirá de participar na rede de propaganda (embora o sítio da Web poderá tentar participar na rede de propaganda utilizando meios fraudulentos ou ilegítimos). Como as políticas poderão ser definidas pela rede de propaganda, diferentes redes de propaganda poderão ter políticas diferentes tal que o sítio da Web que viola uma política em uma rede de propaganda pode não violar qualquer política em outra rede de propaganda.

A "pontuação de qualidade" é tipicamente um indicador quantitativo do valor ou do valor relativo da participação de um sítio da Web ou de uma página da Web (ou algum outro

documento ou coleção de documentos) em uma rede de propaganda. Uma pontuação de qualidade poderá ser aplicada a um limite para determinar um indicador qualitativo.

5 "Informação do usuário" poderá informação do comportamento do usuário e/ou informação do perfil do usuário.

10 "Informação de correspondência eletrônica" poderá incluir qualquer informação incluída em uma correspondência eletrônica (também referida como "informação de correspondência eletrônica interna"), informação derivada da informação incluída na correspondência eletrônica e/ou informação relacionada com a correspondência eletrônica, bem como extensões de tal informação (por exemplo, informação derivada da informação relacionada). Um exemplo
15 da informação derivada da informação de correspondência eletrônica é a informação extraída ou de outra forma derivada de resultados de busca devolvidos em resposta a uma consulta de busca composta de termos extraídos de uma linha de assunto da correspondência eletrônica. Exemplos de
20 informação relacionada à informação de correspondência eletrônica incluem a informação de correspondência eletrônica sobre uma ou mais outras correspondências eletrônicas enviadas pelo mesmo remetente de uma dada correspondência eletrônica, ou informação do usuário a
25 respeito de um receptor de correspondência eletrônica. A informação derivada ou relacionada à informação de correspondência eletrônica poderá ser referida como "informação de correspondência eletrônica externa".

30 4.2-AMBIENTES EXEMPLARES NOS QUAIS, OU COM OS QUAIS, A PRESENTE INVENÇÃO PODERÁ OPERAR

A Figura 1 é um diagrama que mostra as partes ou as entidades que podem interagir com uma rede de propaganda 120 em um ambiente exemplar 100. O ambiente 100 poderá incluir uma rede de propaganda 120 que serve anúncios, ou está de outra forma em associação, com um número de sítios da Web de editoras da Rede 110 e são visualizados pelos usuários finais 140. As editoras da Rede 110 poderão solicitar para participar da rede de propaganda 120 para obter receita ao anúncios renderizados em seus sítios da Web. Os anunciantes 130 poderão entrar, acompanhar e manter seus anúncios na rede de propaganda 120. Os anúncios poderão ser na forma de anúncios gráficos como os chamados anúncios de bandeiras, anúncios apenas de texto, anúncios de imagem, anúncios de áudio, anúncios de vídeo, anúncios que combinam uma ou mais de qualquer um desses componentes, etc. Os anúncios também poderão incluir informação embutida, como um enlace, e/ou instruções executadas por máquina. Esses anúncios poderão ser renderizados, ou de outra forma estarem em associação, com sítios da Web de editoras da Rede 110 quando os usuários finais 140 acessam esses sítios da Web.

Como resultado, pode ser compreendido que a qualidade dos sítios da Web que participam de uma rede de propaganda 120 é importante para a satisfação dos anunciantes 130 e dos usuários finais 140. Para reiterar, os anunciantes 130 muitas vezes querem estar assegurados de que os sítios da Web onde seus anúncios são exibidos são de alta qualidade, gerarão um retorno positivo sobre o investimento, e não serão fonte de controvérsia para o anunciante. As editoras da Rede 110 querem continuar a atrair os usuários finais

140 através de conteúdo relevante e de alta qualidade. Isto ocorre porque os usuários finais 140 têm maior probabilidade de querer visitar sítios da Web que fornecem uma experiência boa, e têm maior probabilidade de
5 selecionar anúncios que são relevantes ao contexto da página da Web e os ajudem a encontrar um produto ou serviço que satisfaça as suas necessidades.

Do que antecede, é aparente que um sistema capaz de avaliar automaticamente o cumprimento de políticas, e/ou a
10 qualidade dos sítios da Web seria útil. Pelo menos algumas modalidades consistentes com a presente invenção poderão avaliar automaticamente sítios da Web quanto ao cumprimento de políticas de uma rede de propaganda 120. Tal avaliação poderá ser feita antes do sítio da Web poder participar na
15 rede de propaganda. Tal avaliação poderá ser repetida periodicamente. Essas reavaliações periódicas poderão ser muito úteis para o sistema de propaganda em que o conteúdo de um sítio da Web aprovado poderá ser mudado pela editora da Rede. Essas avaliações de cumprimento de políticas
20 poderão ser úteis para ajudar a encontrar sítios da Web que foram capazes de participar na rede de propaganda utilizando meios não autorizados. Sob avaliação do sítio da Web, essas modalidades poderão aprovar automaticamente ou desaprovar o sítio da Web, ou poderá procurar por auxílio
25 através da revisão manual.

Pelo menos algumas modalidades consistentes com a presente invenção poderão analisar automaticamente sítios da Web para determinar uma ou mais métrica de qualidade (pontuações). Essas pontuações de qualidade poderão ser
30 utilizadas para determinar incentivos ou desincentivos (ou

evento para remover) sítios da Web que participam em uma rede de propaganda. Alternativamente, ou em acréscimo, essas pontuações de qualidade poderão ser utilizadas para prospectar por sítios da Web de alta qualidade que não participam da rede de propaganda.

Várias modalidades exemplares da presente invenção são descritas agora na 4.3.

4.3-MODALIDADES EXEMPLARES

A Figura 2 é um diagrama de bolhas de operações exemplares para revisar sítios da Web de maneira consistente com a presente invenção, bem como informação que poderá ser utilizada e/ou gerada por essas operações. As operações de revisão do sítio da Web 210 poderão aceitar sítios da Web 205 (como as editoras da Rede que procuram ingressar na rede de propaganda, por exemplo). Essas operações 210 poderão avaliar se esses sítios da Web 205 cumprem com políticas e/ou determinam um ou mais pontuações de qualidade para os sítios da Web.

Os sítios da Web que são de alta qualidade e que cumprem com políticas da rede de propaganda poderão ser servidos e promovidos, enquanto os sítios da Web que violam as políticas da rede de propaganda poderão ser excluídos da rede de propaganda e os sítios da Web que são de baixa qualidade poderão ser penalizados (que poderá incluir serem excluídos da rede de propaganda). As operações de revisão do sítio da Web 210 podem revisar os sítios da Web que poderão estar participando da rede de propaganda, ou sítios da Web fora da rede de propaganda. Os sítios da Web revisados pelas operações de revisão do sítio da Web 210 que estão na rede de propaganda poderão ter sido

anteriormente examinados e aprovados. Assim, as operações de revisão do sítio da Web 210 podem ser utilizados para reavaliar esses sítios da Web como parte da rotina de acompanhamento. Isto é útil nos casos em que as editoras da Rede mudam seus sítios da Web de maneira que possa fazer com que os sítios da Web violem políticas e/ou degradem a sua qualidade. Também é útil para reavaliar sítios da Web em vista de novas políticas. Além disso, alguns sítios da Web que estão na rede de propaganda poderão ser desautorizados. Um sítio da Web autorizado é um sítio da Web que passou pelos procedimentos apropriados para ingressar na rede de propaganda (por exemplo, solicitado formalmente para participar na rede de propaganda) e que tinha sido aprovado pela rede de propaganda. Inversamente, um sítio da Web não autorizado é um sítio da Web que participa de modo inapropriado na rede de propaganda. Esses sítios da Web podem não haver solicitado para ingressar na rede de propaganda, ou poderão haver solicitado para ingressar na rede de propaganda mas não foram aprovados. Assim, uma utilização das operações de revisão do sítio da Web 210 poderá ser a de examinar sítios da Web não autorizados na rede de propaganda para determinar se eles devem continuar a ser servidos (por exemplo, ratificado como participante legítimo da rede de propaganda), ou excluído da rede de propaganda.

Apesar disso, é esperado que as operações de revisão do sítio da Web 210 estarão revisando vários sítios da Web 205 que não participam da rede de propaganda, mas estão buscando ingressar na rede de propaganda. Em pelo menos algumas modalidades consistentes com a presente invenção,

cada sítio da Web 205 que solicitar seu ingresso na rede de propaganda passará pelas operações de revisão do sítio da Web 210 para serem conferidos quanto ao cumprimento das políticas. Uma vez que as operações de revisão do sítio da Web 210 aprovelem o sítio da Web, o sítio da Web poderá ser admitido na rede de propaganda como um participante legítimo. Verificações de acompanhamento periódicas podem ser conduzidas para ajudar a assegurar que a rede de propaganda contém sítios da Web de alta qualidade.

10 Além disso, outros sítios da Web que não participam da rede de propaganda e não solicitaram seu ingresso na rede de propaganda ainda poderão ser revisados pelas operações 210 para ajudar a encontrar sítios da Web de alta qualidade com perspectivas para solicitar seu ingresso na rede de propaganda e ajudar a manter ou elevar sua qualidade.

15 As operações de revisão de sítios da Web 210 poderão incluir um número de operações efetuando várias tarefas para avaliar o sítio da Web 205. Especificamente, as operações de revisão do sítio da Web 210 poderão incluir 20 operações de verificação do cumprimento de políticas 215, operações de pontuação de qualidade 220, operações de enfileiramento 230, e recursos para revisão manual 250. As operações de verificação do cumprimento de políticas 215 poderão efetuar verificações iniciais e de acompanhamento 25 para determinar se os sítios da Web 205 têm qualquer violação de política. Os sítios da Web aprovados pelas operações de verificação do cumprimento das políticas 215 poderão ser armazenados como sítios da Web aprovados 240, enquanto os sítios da Web desaprovados poderão ser 30 armazenados como sítios da Web rejeitados 235. Os sítios da

Web que não foram aprovados, mas que não podem ser desaprovados com incerteza poderão ser sinalizados e enviados para uma fila de revisão manual 245 apropriada pelas operações de enfileiramento 230. Ainda, as operações de verificação 215 do cumprimento das políticas poderá encaminhar os sítios da Web para as operações de pontuação de qualidade 220.

As operações de pontuação de qualidade 220 poderão ser utilizadas para determinar pontuações de qualidade 225 para os sítios da Web. Em casos graves, as operações de pontuação de qualidade 220 poderão remover o sítio da Web da rede de propaganda. Em outros casos, as operações de pontuação de qualidade 220 poderão sinalizar os sítios da Web com baixa qualidade e enviá-los para as operações de enfileiramento 230 onde eles poderão ser enfileirados para uma revisão manual. A informação de pontuação de qualidade do sítio da Web 225 poderá ser utilizada por seres humanos e/ou por outras operações automatizadas (não mostradas) para maior análise ou para outros usos.

As operações de enfileiramento do sítio da Web 230 poderão enfileirar sítios da Web com base no tipo e/ou na gravidade das violações de políticas conforme determinado pelas operações de verificação de cumprimento das políticas 215, e/ou com base nas baixas pontuações de qualidade conforme determinado pelas operações de pontuação de qualidade 220. Assim, sítios da Web sinalizados poderão ser colocados em uma ou mais filas 245 para revisão manual. Nas instalações de revisão manual 250, seres humanos poderão obter os sítios sinalizados enfileirados e decidir se aceitam ou rejeitam o sítio da Web. O sítio da Web poderá

ser subseqüentemente processado de acordo.

As políticas a serem verificadas pelas operações de verificação do cumprimento das políticas 215 poderão ser definidas pela rede de propaganda particular. No contexto da presente invenção, as políticas são definidas tal que, se um sítio da Web violar qualquer política, ele poderá não participar (legitimamente) da rede de propaganda. Assim, políticas são aquelas regras cuja violação não é tolerada. Portanto, as operações de verificação do cumprimento das políticas 215 poderá ou aprovar ou desaprovar os sítios da Web 205. Violações das políticas poderão ser geralmente classificadas como (a) violações relacionadas ao conteúdo do sítio da Web (por exemplo, conteúdo genérico demais para fornecer direcionamento específico, não há conteúdo suficiente, conteúdo ruim ou controverso, etc.), (b) violações relacionadas à editora ou origem do sítio da Web, (c) violações relacionadas com o uso do sítio da Web (por exemplo, sob construção, enlaces quebrados, carregamento de página lento, utilização inadequada de quadros, etc.) e (d) violações relacionadas a fraude (por exemplo, tentativa de defraudar anunciantes e/ou a rede de propaganda).

Alguns exemplos de possíveis violações de políticas de conteúdo do sítio da Web incluem sítios da Web ódio-cêntricas (por exemplo, racista, anti-semítica, antigay, antiimigrantes, etc.), sítios da Web essencialmente relacionado a campanhas de ódio individual (por exemplo, Hate Bush, Microsoftsucks.com., etc.), sítios da Web que advogam ou facilitam a violência ou o terror (por exemplo, Al Qaeda, etc.), sítios da Web que oferecem aconselhamento malicioso ou que advogam ou facilitam atividades maliciosas

(por exemplo, bombas caseiras, pregar peças, etc.), sítios da Web que oferecem ou facilitam a venda de drogas ilícitas ou parafernália para o uso de drogas, sítios da Web que oferecem ou facilitam a venda ilegal de produtos farmacêuticos (uma exceção poderá ser feita para as drogas vendidas no varejo ou drogas herbais se não forem prescritivas), sítios da Web que facilitam ou advogam invasores de software ou de rede (por exemplo, sítios da Web relacionados ou que forneçam dicas sobre como invadir uma estrutura de IT), sítios da Web que facilitam ou advogam quebra (cracking) de software ou de redes (por exemplo, sítios da Web relacionados ou fornecendo dicas sobre como quebrar aplicações de software), sítios da Web que oferecem ou facilitam a venda de armas de fogo (ou de outras armas) ou a venda ilegal de armas de fogo (ou outras armas), sítios da Web que incluem pornografia, pornografia hardcore e/ou conteúdo orientado para adultos (por exemplo, sítios da Web que contém fotos orientadas para adultos, texto, serviços de consultório sentimental, histórias eróticas, ou enlaces), sítios da Web com profanidade ou profanidade em excesso, sítios da Web que oferecem ou facilitam ou promovem a venda de produtos de ponta-de-estoque ou mercadorias falsificadas, sítios da Web que promovem ou facilitam a perpetração de enganações ou de embustes (por exemplo, esquemas de corrente), sítios da Web que dependem de erros tipográficos de marcas registradas (por exemplo, sítios da Web intromissão em erros de digitação cibernéticos), sítios da Web que oferecem ou facilitam a venda de documentos falsificados (por exemplo, passaportes falsificados, carteiras de motorista

falsificadas, certificados de nascimento falsos, etc.), sítios da Web que oferecem ou facilitam a venda de teses (por exemplo, "fábricas de teses"), sítios da Web que oferecem ou facilitam a baixa não autorizada ou a
5 renderização de conteúdo com direitos autorais, sítios da Web que oferecem a venda de álcool, sítios da Web que oferecem a venda de álcool que não seja vinho, sítios da Web que oferecem a venda de álcool a menores, sítios da Web que oferecem ou facilitam a venda de fumo, sítios da Web
10 que oferecem ou facilitam a venda de fumo a menores, sítios da Web que facilitam a jogatina online, sítios da Web que facilitam a jogatina online para menores, etc.

Pelo menos algumas das violações de políticas de conteúdo anteriores podem, por exemplo, ser descobertas ao
15 determinar se o sítio da Web inclui certas palavras, frases, ou coleções de palavras e/ou frases, se o sítio da Web inclui mais do que uma porcentagem predeterminada e/ou contagem de certas palavras e/ou frases, se o sítio da Web tem certas imagens (por exemplo, suásticas, quadros de
20 apostas, nus, armas, marcas de cigarro, marcas de bebidas alcoólicas, etc.), etc. Ainda, uma lista manual de sítios da Web com um dado tipo ou classe de violação de política poderá ser utilizado para treinar o sistema especializado (por exemplo, redes neurais, redes bayesianas, máquinas
25 vetoras de suporte, etc.) a classificar outros sítios da Web como tendo ou não a violação de política.

Alguns exemplos de violações de políticas relacionados à editora ou origem do sítio da Web incluem sítios da Web suportados por países na lista negra (por exemplo, Cuba),
30 editoras da Rede com um histórico de fraude, editoras da

Rede anteriormente eliminada da rede de propaganda, etc. Pelo menos algumas dessas violações de políticas podem ser determinadas utilizando a localização do servidor da Rede. Outra informação disponível para auxiliar na determinação de violações relacionadas à editora poderão incluir o endereço, o número de seguro social, a correspondência eletrônica, o endereço IP do servidor da Rede, o nome do sítio da Web, etc.

Alguns exemplos de violações da política de uso do sítio da Web incluem sítios da Web com erros de servidor de nome de domínio (DNS) (como, por exemplo, URL que não existe, URL desconectada, etc.), sítios da Web com enlaces quebrados, sítios da Web "offline", sítios da Web "em construção", sítios da Web com pop-ups excessivos (por exemplo, mais de N (por exemplo, um) anúncio de pop-up em qualquer carregamento de página da Web), sítios da Web de chats, sítios da Web não HTML (por exemplo, índices de diretórios, flash, WAP), sítios da Web com instrumentos de espionagem, sítios da Web para tomar páginas da Web, etc.

Outros exemplos poderiam incluir sítios da Web que tentam instalar software malicioso no computador do usuário, ou sítios da Web que afetam o uso ao desativar o botão liga/desliga, pop-ups excessivos, etc. Pelo menos algumas dessas violações de políticas podem ser determinadas ao identificar códigos de erro, e/ou código executável suspeito. Isto poderá ser efetuado ao tentar automaticamente carregar o sítio da Web em um navegador e capturar os eventos "ruins" ao nível do navegador, ou ao nível do sistema operacional.

Finalmente, alguns exemplos de violações de sítios da

Web relacionados a fraudes incluem os chamados sítios da Web "pay-to-click" (pagar para clicar) (por exemplo aqueles cuja finalidade principal é fazer com que as pessoas selecionem anúncios de pagamento por clique) etc. Por exemplo, sítios da Web do tipo clique-mensagem em massa (spam) possuem tipicamente conteúdo gerado utilizando um modelo e/ou uma alta proporção anúncios/texto. Ainda, há técnicas existentes para detectar indicadores de mensagem em massa (isto é, sítios da Web que utilizam meios ilegítimos para classificar mais alto nos resultados de busca). Acredita-se que poderá haver uma correlação entre aqueles que engajam em indicar mensagem em massa com aqueles que engajam em clique de mensagem em massa. Tal correlação poderá ser utilizada para ajudar a encontrar clique de mensagem em massa ao encontrar os indicadores de mensagem em massa. Qualquer uma da informação anterior pode ser utilizada para ajudar a descobrir esses tipos de violações. Em outra abordagem, os tipos citados acima de violações poderão ser descobertas ao medir um desvio da efetiva taxa de cliques no sítio da Web examinado da taxa de clique média para todos os sítios da Web. Além disso, poderá ser empregado a aprendizagem de algoritmos (naive Bayes, SVMs, etc.). Por exemplo, esses algoritmos poderão ser utilizados para treinar as redes utilizando a estatística de alguns sítios da Web conhecidos como "pagar para clicar". Essas redes treinadas poderão então ser utilizadas para identificar novos sítios da Web.

A pontuação de qualidade do sítio da Web poderá ser determinada utilizando critérios de qualidade definidos. Alguns exemplos de critérios de qualidade de um sítio da

Web incluem: dados de utilização da rede de propaganda ou de outras fontes (por exemplo, impressões, seleções, geolocalização do usuário, conversões, e derivativos dessa informação); popularidade do sítio da Web (seja na rede de 5 propaganda, ou não) (conforme medido pela barra de ferramenta do Google, por exemplo); mensagem em massa do sítio da Web (por exemplo, os chamados "link farms" (campos de enlace), onde os sítios da Web são montados tendo pouco conteúdo de relevância, mas muitos enlaces uns para os 10 outros para elevar as pontuações dos ferramentas de busca, o conteúdo inválido estando em uma página como o sítio da Web de pornografia tendo um segundo plano de preto e texto branco relacionado a pornografia, mas texto preto (e com isso oculto) que se refere a câmeras digitais ou outras 15 consultas populares, para elevar as pontuações dos ferramentas de busca, etc.); etc. Alguns exemplos de mensagem em massa de sítios da Web estão descritos no Pedido de Patente dos Estados Unidos número de série 10/748.664 (aqui incorporado por referência e referido como 20 "a aplicação 664"), intitulado "INFORMATION RETRIEVAL BASED ON HISTORICAL DATA", requerido em 31 de dezembro de 2003 e que lista Anurag Acharya, Matt Cutts, Jeffrey Dean, Paul Haahr, Monika Henzinger, Urs Hoelze, Steve Lawrence, Karl Pflieger, Olcan Sercinoglu e Simon Tong como inventores.

25 4.3.1-MÉTODOS EXEMPLARES

A Figura 3 é um diagrama de fluxo de um método exemplar 300 que poderá ser utilizado para verificar automaticamente sítios da Web quanto a violações de políticas de uma maneira consistente com a presente invenção. Como é 30 indicado pelo bloco de evento 305, os principais atos do

método 300 poderão ser efetuados em resposta ao receber uma solicitação para verificar um sítio da Web. Em resposta ao recebimento de uma solicitação de verificação, o método 300 poderá aceitar o sítio da Web (bloco 310) e examiná-lo quanto a possíveis violações de políticas (bloco 315). Se nenhuma violação de políticas for encontrada (bloco de decisão 320), então o método 300 poderá simplesmente aprovar o sítio da Web (bloco 325) e programar uma verificação de políticas de acompanhamento (bloco 330) antes do método 300 ser deixado (nó 360).

Com referência de volta ao bloco de decisão 320, se, por outro lado, violações de políticas forem encontradas, o método 300 poderá desaprovar o sítio da Web (bloco 335). O método 300 poderá efetuar outros atos com base na confiança da determinação da violação de política (bloco 340). Mais especificamente, se a confiança da determinação da violação de política é baixa, então o método 300 poderá sinalizar (e/ou enfileirar) o sítio da Web para uma revisão manual (bloco 345) (relembrar, por exemplo, as operações de enfileiramento 230 da Figura 2) e poderá notificar a editora do sítio da Web (bloco 355) antes do método 300 ser deixado (nó 360). Isto poderá permitir que possíveis violações de políticas sejam confirmadas ou apagadas manualmente. Também poderá permitir que as editoras da Rede argumentem seu caso de porque seu sítio da Web, na verdade, não viola uma política, ou tomar medidas corretivas. Com referência de volta ao bloco de decisão 340, se a confiança da determinação de violação é alta, então o método 300 poderá sinalizar o sítio da Web para indicar que ele foi rejeitado (por exemplo, nenhuma revisão manual é

necessária) (bloco 350) antes do método 300 ser deixado (nó 360).

Com referência de volta ao bloco de evento 305, as solicitações para um exame poderão originar-se de fontes diferentes. Especificamente, o método 300 poderá receber solicitações de editoras da Rede que desejem participar da rede de propaganda, solicitações como parte da rotina de acompanhamento, etc. A solicitação poderá originar-se de uma fonte de desenvolvimento de empresas. Por exemplo, como será descrito em maior detalhe abaixo, as operações de pontuação de qualidade do sítio da Web (relembrar, por exemplo, o elemento 220 da Figura 2) poderão ser utilizadas para ajudar a rede de propaganda a procurar por sítios da Web que já não estão participando da rede de propaganda, mas que são bons candidatos potenciais com base em sua pontuação de qualidade. A fonte de desenvolvimento de empresas poderá querer pré-selecionar sítios da Web quanto a violações de políticas para estreitar a sua busca de sítios da Web em perspectiva para solicitar.

Com referência de volta ao bloco 315, o método 300 examina os sítios da Web recebidos para determinar se eles violam qualquer uma das políticas colocadas pela rede de propaganda. Exemplos das políticas foram listadas no 4.3 acima.

A Figura 4 é um diagrama de fluxo de um método exemplar 400 para a pontuação da qualidade do sítio da Web de maneira consistente com a presente invenção. O sítio da Web é aceito (bloco 402). Como é indicado pelo bloco 404, no método exemplar 400, tipos diferentes de sítios da Web poderão ser processados de modos diferentes. Por exemplo, o

sítio da Web poderá ter sido "aprovado anteriormente" (por exemplo, passou anteriormente por uma verificação de política). Como outro exemplo, o sítio da Web poderá estar na rede de propaganda mas poderá ter ganho acesso por meio
5 ilegítimo ou não autorizado. Como outro exemplo, o sítio da Web poderá ser um que foi rejeitado anteriormente por uma ou mais violações de política. Como ainda outro exemplo, o sítio da Web poderá ser um que não foi verificado anteriormente (por exemplo, um sítio da Web que nunca
10 procurou participar na rede de propaganda).

Se o sítio da Web aceito é um sítio da Web "aprovado" (por exemplo, aprovado pelas operações de verificação de cumprimento de política 215), o método 400 poderá determinar um ou mais pontuações de qualidade para o sítio
15 da Web (bloco 406). O método 400 poderá então efetuar atos diferentes dependendo de se a pontuação determinada é ou não aceitável (bloco 408). Se a pontuação é aceitável, o método 400 poderá fornecer incentivos ou desincentivos (por exemplo, com relação ao preço cobrado para o anunciante,
20 partilhamento da receita com a editora da Rede, emolumentos para a editora da Rede, etc.) dependendo do valor da pontuação, ou simplesmente nada fazer (bloco 410) antes do método 400 ser deixado (nó 490). Se a pontuação é baixa demais, o método 400 poderá avisar a editora da Rede sobre
25 a pontuação de qualidade baixa, e/ou remover o sítio da Web de participação na rede de propaganda (blocos 412 e 414) antes do método 400 ser deixado (nó 490).

Com referência de volta ao bloco 404, se o sítio da Web recebido for um sítio da Web não autorizado na rede de
30 propaganda, o método 400 poderá inicialmente determinar se

o sítio da Web não autorizado viola ou não qualquer política (bloco 416). Com referência ao bloco 418, se o sítio da Web não autorizado fracassar na verificação de políticas, o método 400 poderá desativar o sítio da Web de participação na rede de propaganda (bloco 424), e talvez atualizar o depósito de fraudes para incluir o sítio da Web rejeitado (bloco 426) antes do método 400 ser deixado (nó 490). Com referência de volta ao bloco 418, se o sítio da Web não autorizado passar na verificação de políticas, o método 400 poderá prosseguir para determinar uma pontuação de qualidade (bloco 420). Com referência ao bloco 422, se a pontuação de qualidade determinada for aceitável, o método 400 poderá ratificar ou autorizar a participação do sítio da Web na rede de propaganda (bloco 428) antes do método 400 ser deixado (nó 490). Com referência novamente ao bloco 422, se, por outro lado, a pontuação de qualidade determinada for baixa demais (não aceitável), o método 400 poderá desativar o sítio da Web de participação na rede de propaganda (bloco 424), e talvez atualizar o depósito de fraudes para incluir o sítio da Web rejeitado (bloco 426) antes do método 400 ser deixado (nó 490).

Com referência de volta ao bloco 404, se o sítio da Web recebido for um que foi anteriormente rejeitado por violação de políticas, o método 400 poderá simplesmente desconsiderar o sítio da Web e não tomar qualquer outra medida. Isto é porque poderá ser um desgaste dos recursos verificar a pontuação de qualidade de um sítio da Web que não pode participar da rede de propaganda de qualquer forma devido à violação de políticas.

Com referência de volta ao bloco 404, quando do

recebimento de um sítio da Web que não é parte da rede de propaganda e que não solicitou anteriormente ingressar na rede de propaganda, o método 400 poderá inicialmente verificar o sítio da Web por qualquer violação de políticas (bloco 430). Com referência ao bloco 432, se o sítio da Web fracassar no exame de políticas, o método 400 poderá simplesmente ignorar o sítio da Web. Se, por outro lado, o sítio da Web passar no exame de políticas, o método 400 poderá prosseguir para determinar uma pontuação de qualidade (bloco 434). Com referência ao bloco 436, se a pontuação de qualidade determinada for alta, o método 400 poderá marcar o sítio da Web como um membro em perspectiva da rede de propaganda para uma solicitação de seguimento antes do método 400 ser deixado (nó 490). (Bloco 438) Com referência novamente ao bloco 436, se, por outro lado, a pontuação de qualidade determinada for baixa demais (não aceitável), o método 400 poderá simplesmente ignorar o sítio da Web.

Com referência de volta aos blocos 406, 420 e 434, o método 400 poderá utilizar uma ou mais funções para determinar um ou mais pontuações de qualidade. Como foi mencionado anteriormente, a pontuação de qualidade poderá refletir alguma medida de valor do sítio da Web para a rede de propaganda.

Ainda com referência aos blocos 406, 420 e 434, após os valores de pontuação terem sido determinados, o método 400 poderá utilizar valores limites (por exemplo, predeterminados) para decidir quais sítios da Web deverão ser aceitos e quais deverão ser rejeitados.

Com referência de volta ao bloco 404, o método 400

poderá receber vários tipos de sítios da Web, alguns dos quais não passaram por uma verificação de políticas. (Relembre, por exemplo, as operações de verificação de cumprimento de políticas 215 da Figura 2.) Por exemplo, 5 sítios da Web que não são autorizados na rede de propaganda e os sítios da Web que não solicitaram anteriormente participação na rede podem não terem sido verificados quanto a violações de políticas. Assim, o método 400 poderá enviar solicitações para as operações de verificação de 10 cumprimento de política 215, e receber os resultados da verificação, como é mostrado pelos blocos 416 e 430.

Com referência de volta ao bloco 410, o método 400 poderá fornecer incentivos ou desincentivos dependendo do valor da pontuação. Por exemplo, se o valor da pontuação do 15 sítio da Web é alto, a rede de propaganda poderá fornecer a editora da Rede com incentivos (por exemplo, uma quantidade acrescida por seleção, uma porcentagem aumentada de receita de propaganda, custos de participação diminuídos, etc.), pois ela fornece um sítio da Web de alta qualidade. Por 20 outro lado, se o valor da pontuação do sítio da Web for baixo (mas aceitável), então a rede de propaganda poderá fornecer desincentivos à editora da Rede (por exemplo, quantidade menor por seleção, uma porcentagem diminuída da receita de propaganda, custos adicionados, etc.) por não 25 fornecer um sítio da Web de alta qualidade. Isto poderá estimular as editoras da Rede a melhorarem seus sítios da Web.

A Figura 5 é um diagrama de fluxo de um método exemplar 500 que poderá ser utilizado para enfileirar sítios da Web 30 para a revisão manual de uma maneira consistente com a

presente invenção. Especificamente, o método 500 poderá obter sítios da Web com pontuações de qualidade baixa ou sítios da Web contendo possíveis violações de políticas (relembrar, por exemplo, o bloco 345 da Figura 3) (bloco 5 520). A seguir, os sítios da Web sinalizados poderão ser acrescentados a uma fila apropriada compreendida de sítios da Web sinalizados prontos para serem apresentados para a revisão manual (bloco 540).

Com referência de volta ao bloco 520, o método 500 10 poderá obter sítios da Web com pontuações de baixa qualidade (por exemplo, das operações de pontuação de qualidade de sítios da Web 220 da Figura 2) e/ou sítios da Web contendo possíveis violações de política (por exemplo, das operações de verificação do cumprimento de políticas 15 215 da Figura 2). Por exemplo, os sítios da Web que contêm possíveis violações poderão ser aqueles sítios da Web que foram desaprovados com baixa confiança pelas operações de verificação do cumprimento de políticas 215. Relembre que essas operações 215 poderão simplesmente sinalizar esses 20 sítios da Web para a revisão manual ao encaminhá-las para as operações de enfileiramento do sítio da Web 230. Os sítios da Web com pontuação de baixa qualidade poderão ser aqueles sítios da Web aprovados após terem sido examinados por violações de políticas, mas que atingiram valores de 25 pontuação que são baixos (mas talvez não baixos o suficiente para a terminação automática). Portanto, o método 500 poderá programar esses sítios da Web para uma revisão manual em que seres humanos poderão tomar a decisão final sobre se a pontuação de qualidade é ou não aceitável.

30 Com referência de volta ao bloco 540, o método 500

poderá acrescentar os sítios da Web sinalizados para uma ou mais filas de sítios da Web sinalizados. Filas diferentes poderão refletir prioridades e/ou tipos diferentes de prioridade. Como exemplo mais específico, filas poderão ser 5 priorizadas utilizando uma combinação de gravidade da violação, tipo da violação, número de violações, etc. Isto permite que sítios da Web com uma única violação de política menos grave (que têm uma probabilidade mais alta de terem sua rejeição anulada manualmente) sejam revisados 10 mais rapidamente do que os sítios da Web com múltiplas violações de política graves (que têm menor probabilidade de serem aprovadas).

Como outro exemplo, as filas poderão ser fornecidas com base no tipo de violação de política. Isto permite a 15 revisão manual de sítios da Web com violações de política a serem especializadas. Por exemplo, uma ou mais pessoas poderão ser responsáveis pela revisão dos sítios da Web que se acredita sejam ódio-cêntricos, enquanto uma ou mais outras pessoas poderão ser responsáveis pela revisão de 20 sítios da Web que se acredita venderem álcool. Essa especialização permite uma aplicação mais uniforme e apropriada das políticas.

A Figura 6 é um diagrama de fluxo de um método exemplar 600 que poderá ser utilizado para processar manualmente os 25 sítios da Web revistos de uma maneira consistente com a presente invenção. Em particular, o método 600 poderá aceitar a decisão (por exemplo, de aprovar ou desaprovar um sítio da Web sinalizado, ou talvez redesignar a revisão para outro revisor/departamento) de um revisor humano 30 (bloco 610). Com referência ao bloco 620, se o sítio da Web

é aprovado manualmente, então o método 600 poderá marcar o sítio da Web como aprovado (ou efetuar os atos que forem necessários para ter o sítio da Web acrescentado à rede de propaganda), e uma verificação de política de acompanhamento poderá ser programada (blocos 630 e 640) antes do método 600 ser deixado (nó 690). Ainda, o método 600 poderá tomar as medidas necessárias para determinar a pontuação de qualidade (bloco 650).

Com referência de volta ao bloco 620, se o sítio da Web é desaprovado manualmente, então o método 600 poderá remover o sítio da Web da participação na rede, se ela estiver ativa (bloco 660) antes do método 600 ser deixado (nó 690). Nesse caso, o método 600 poderá também registrar a informação a respeito do sítio da Web rejeitado, e/ou informar a editora do sítio da Web (blocos 670 e 680).

Como um aditamento ao acima, as decisões do operador humano poderão ser utilizadas em conjunto com algoritmos de aprendizagem (naïve Bayes, SVMs, etc.) de modo a treinar (por exemplo, com base em decisões de operador humano anteriores) e sintonizar as operações de pontuação da página da Web (relembrar, por exemplo, o método 400 da Figura 4). Fazê-lo poderá resultar em um procedimento de pontuação mais eficaz e preciso, assim tornando o sistema mais eficiente enquanto reduz a intervenção humana (revisões manuais).

4.3.2-APARELHO EXEMPLAR

A Figura 7 é um diagrama de blocos de alto nível de uma máquina 700 que poderá efetuar uma ou mais das operações discutidas acima, e armazenar várias informações discutidas acima. A máquina 700 inclui um ou mais processadores 710,

uma ou mais unidades de interface de entrada/saída 730, um ou mais dispositivos de armazenamento 720, e um ou mais barramentos do sistema e/ou de redes 740 para facilitar a comunicação da informação entre os elementos acoplados. Um
5 ou mais dispositivos de entrada 732 e um ou mais dispositivos de saída 734 poderão ser acoplados com a uma ou mais interfaces de entrada/saída 730.

Um ou mais processadores 710 poderão executar instruções executadas por máquinas (por exemplo, C ou C++
10 processando no sistema operacional Solaris disponível de Sun Microsystems Inc. de Palo Alto, Califórnia, EUA, ou o sistema operacional Linux amplamente disponível de um número de fornecedores, como Red Hat, Inc., de Durham, North Carolina, EUA) para efetuar um ou mais aspectos da
15 presente invenção. Pelo menos uma parte das instruções executadas por máquina poderá ser armazenada (temporária ou mais permanentemente) no um ou mais dispositivos de armazenamento 720 e/ou poderão ser recebidos de uma fonte externa através de uma ou mais unidades de interface de
20 entrada 730.

Em uma versão, a máquina 700 poderá ser um ou mais computadores pessoais convencionais. Neste caso, as unidades de processamento 710 poderão ser um ou mais microprocessadores. O barramento 740 poderá incluir um
25 barramento do sistema. Os dispositivos de armazenamento 720 poderão incluir memória do sistema, como memória de apenas leitura (ROM) e/ou memória de acesso aleatório (RAM). Os dispositivos de armazenamento 720 também poderão incluir uma unidade de disco rígido para ler e gravar em um disco
30 rígido, uma unidade de disco magnético para ler ou gravar

em disco magnético (por exemplo, removível), e uma unidade de disco óptico para ler ou gravar para um disco (magneto) óptico removível como um disco compacto ou outra mídia (magneto) óptica.

5 Um usuário poderá entrar com comandos e informação dentro do computador pessoal através de dispositivos de entrada 732, como, por exemplo, um teclado e dispositivo apontador (por exemplo, um mouse). Outros dispositivos de entrada como um microfone, um manche, uma almofada de
10 jogos, uma antena de satélite, um digitalizador, ou assemelhados, também (ou alternativamente) poderão ser incluídos. Esses e outros dispositivos de entrada são muitas vezes conectados às unidades de processamento 710 através de uma interface apropriada 730 acoplada ao
15 barramento de sistema 740. Os dispositivos de saída 734 poderão incluir um monitor ou outro tipo de dispositivo de exibição, que também poderá estar conectado ao barramento de sistema 740 através de uma interface apropriada. Além (ou em vez) de o monitor, o computador pessoal poderá
20 incluir outros dispositivos de saída (periféricos) (não mostrados) como, por exemplo, alto falantes e impressoras.

4.3.3-REFINAMENTOS E ALTERNATIVAS

A presente invenção não é limitada às violações de política particulares descritas acima, nem é a presente
25 invenção limitada às pontuações de qualidade particulares descritas acima.

Embora tanto a verificação de cumprimento de políticas dos sítios da Web como a pontuação dos sítios da Web foram descritas, não é necessário efetuar os dois.

30 A conferência do cumprimento de políticas poderá ser

solicitada por várias partes, e poderá ocorrer em várias ocasiões. Por exemplo, a editora da Rede poderá solicitar que seu sítio da Web seja acrescentado a uma rede de propaganda. A rede de propaganda, ou um agente para a rede de propaganda, poderá verificar o sítio da Web quanto ao cumprimento de políticas antes dele ser acrescentado à rede de propaganda. Como outro exemplo, como o sítio da Web poderá ser modificado após ele ter sido aceito na rede de propaganda, a rede de propaganda ou um agente para a rede de propaganda (referido coletivamente como "a rede de propaganda") poderá reverificar o sítio da Web quanto ao cumprimento de políticas. Como ainda outro exemplo, como as políticas poderão ser modificadas (por exemplo, relaxadas, tornadas mais rígidas, expandidas, reduzidas, etc.) poderá ser desejável para a rede de propaganda reavaliar um sítio da Web (anteriormente aceito e/ou anteriormente rejeitado) em vista da mudança na política. Como ainda outro exemplo, uma rede de propaganda poderá verificar os participantes em sua rede de propaganda quanto ao cumprimento de políticas para ajudar a encontrar sítios da Web não autorizados. Como ainda mais outro exemplo, o desenvolvimento empresarial para a rede de propaganda poderá querer atrair editoras da Rede para participarem da rede de propaganda. O desenvolvimento de empresas poderá utilizar a presente invenção para filtrar candidatos em potencial que violariam uma ou mais políticas da rede de propaganda.

Se uma pontuação de qualidade é ou não baixo poderá ser determinado utilizando heurísticas predeterminadas, limites predeterminados ou dinamicamente determinados (absoluto e/ou relativo), etc.

A pontuação de qualidade poderá ser solicitada por várias partes, e poderá ocorrer em várias ocasiões. Por exemplo, a rede de propaganda poderá medir a qualidade dos sítios da Web que participam de sua rede (por exemplo, para 5 satisfazer anunciantes e/ou usuários finais). A rede de propaganda poderá desejar fornecer incentivos e/ou desincentivos com a meta de melhorar a qualidade dos sítios da Web na rede de propaganda, ou pelo menos ajustar os pagamentos do anunciante e/ou a compensação da editora da 10 Rede conseqüentemente.

Embora muitas das modalidades exemplares foram descritas no contexto de sítios da Web, elas poderão ser aplicadas a páginas da Web e outros tipos de documentos ou coleções de documentos. Como um exemplo, anúncios poderão 15 ser servidos com correspondência eletrônica, como com G-Mail oferecida por Google, Inc. O conteúdo da correspondência eletrônica pode ser verificado (por exemplo, em tempo real) quanto a violações de políticas.

4.4-CONCLUSÕES

20 Como pode ser apreciado do que antecede, modalidades consistentes com a presente invenção ajudam as redes de propaganda a (i) tornar-se cidadãos empresariais, (ii) assegurar que os sítios da Web em sua rede não serão ofensivos ou controversos, (iii) evitar ilegalidades, (iv) melhorar a 25 qualidade dos sítios da Web em sua rede, (v) proteger e aprimorar sua marca, e/ou (vi) ajustar os pagamentos do anunciante e/ou a compensação da editora da Rede e/ou os emolumentos com base na qualidade do sítio da Web. Essas modalidades ainda permitem que as redes de propaganda realizem 30 tais meios com maior eficiência e mais rapidamente.

REIVINDICAÇÕES

1. Método implementado por computador, caracterizado pelo fato de compreender:

a) aceitar uma coleção incluindo um ou mais documentos;

5 b) determinar se a coleção cumpre ou não com as políticas de uma rede de propaganda; e

c) aprovar a coleção se foi determinado que a coleção cumpre com as políticas.

2. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato de ainda compreender:

15 d) acrescentar a coleção à rede de propaganda se a coleção for aprovada tal que anúncios poderão ser servidos em associação com renderizações dos documentos incluídos na coleção.

3. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato de ainda compreender:

20 d) acrescentar a coleção à rede de propaganda se a coleção for aprovada tal que anúncios de conteúdo direcionado poderão ser servidos em associação com renderizações dos documentos incluídos na coleção.

4. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato de ainda compreender:

25 d) programar a determinação de acompanhamento de se a coleção cumpre ou não com as políticas da rede de propaganda.

5. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato da coleção

incluindo um ou mais documentos ser um sítio da Web incluindo uma ou mais páginas da Web.

6. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato da política
5 abranger o conteúdo de um ou mais documentos da coleção.

7. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato da política incluir pelo menos uma política selecionada de um grupo de políticas que consiste de (A) nenhuma coleção de documentos
10 ódio-cêntricos, (B) nenhuma coleção de documentos abrangendo essencialmente campanhas de ódio individual, (C) nenhuma coleção de documentos que advogam ou facilitam a violência ou o terror, (D) nenhuma coleção de documentos que oferece aconselhamento malicioso ou advoga ou facilita
15 atividades maliciosas, (E) nenhuma coleção de documentos que oferece ou facilita a venda de drogas ilegais ou de parafernália para drogas, (F) nenhuma coleção de documentos que oferece ou facilita a venda ilegal de produtos farmacêuticos, (G) nenhuma coleção de documentos que
20 facilite ou advogue a invasão de software ou de rede, (H) nenhuma coleção de documentos que facilite ou advogue a quebra de software ou de rede, (I) nenhuma coleção de documentos que oferece ou facilita a venda de armas de fogo, (J) nenhuma coleção de documentos com pornografia,
25 (K) nenhuma coleção de documentos com pornografia pesada, (L) nenhuma coleção de documentos com profanidade, (M) nenhuma coleção de documentos com excesso de profanidade, (N) nenhuma coleção de documentos que oferece ou facilita ou promove a venda de produtos de ponta de estoque ou de
30 bens falsificados, (O) nenhuma coleção de documentos que

promovem ou facilitam a perpetração de enganações ou de embustes, (P) nenhuma coleção de documentos que dependem de erros tipográficos de marcas registradas, (Q) nenhuma coleção de documentos que ofereça ou facilite a venda de documentos falsificados, (R) nenhuma coleção de documentos que ofereça ou que facilite a venda de apresentações de trabalhos acadêmicos, (S) nenhuma coleção de documentos que ofereça ou que facilite a transferência (download) não autorizada ou a renderização de conteúdo coberto por direitos autorais, (T) nenhuma coleção de documentos suportada por editoras indesejadas, (U) nenhuma coleção de documentos que ofereça a venda de álcool, (V) nenhuma coleção de documentos que ofereça a venda de álcool exceto o vinho, (W) nenhuma coleção de documentos que ofereça a venda de álcool para menores, (X) nenhuma coleção de documentos que ofereça ou facilita a venda de fumo, (Y) nenhuma coleção de documentos que ofereça ou que facilite a venda de fumo a menores, (Z) nenhuma coleção de documentos que facilite a jogatina online, e (AA) nenhuma coleção de documentos que facilita a jogatina online para menores.

8. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato da coleção incluindo um ou mais documentos ser um sítio da Web que inclui uma ou mais páginas da Web,

25 e em que a política abrange a usabilidade do sítio da Web.

9. Método implementado por computador, de acordo com a reivindicação 1, caracterizado pelo fato da coleção ser um sítio da Web que inclui uma ou mais páginas da Web, e

30 em que a política inclui pelo menos uma política

selecionado de um grupo de políticas que consistem de (A) sítios da Web com erros de Domain Name System (DNS - Sistema de Nome de Domínio), (B) sítios da Web offline, (C) sítios da Web em construção, (D) sítios da Web com 5 excessivos pop-ups, (E) sítios da Web com excessivos pop-under, (F) sítios da Web de conversação, (G) sítios da Web não HTML, (H) sítios da Web spyware, e (I) sítios da Web de "assalto" a páginas da Web.

10. Método implementado por computador, de acordo com a 10 reivindicação 1, caracterizado pelo fato da política abranger uma possível fraude ou decepção na rede de propaganda ou dos participantes da rede de propaganda pela coleção.

11. Método implementado por computador, de acordo com a 15 reivindicação 1, caracterizado pelo fato de ainda compreender:

d) armazenar, se é determinado que a coleção não cumpre com a política, uma indicação que os anúncios da rede de propaganda não devem ser servidos com documentos incluídos 20 na coleção.

12. Método implementado por computador, de acordo com a reivindicação 11, caracterizado pelo fato de ainda compreender:

e) determinar um nível de confiança na determinação de 25 que a coleção não cumpre com a política; e

f) gravar a coleção para uma determinação do cumprimento de política manual se o nível de confiança determinado estiver abaixo de um limite predeterminado.

13. Método implementado por computador, de acordo com a 30 reivindicação 12, caracterizado pelo fato de ainda

compreender:

g) aceitar uma determinação manual do cumprimento da política da coleção; e

h1) aprovar a coleção se foi determinado manualmente que a coleção cumpre com as políticas.

14. Método implementado por computador, de acordo com a reivindicação 13, caracterizado pelo fato de ainda compreender:

h2) acrescentar a coleção à rede de propaganda se a coleção for aprovada tal que anúncios poderão ser servidos em associação com renderizações dos documentos incluídos na coleção.

15. Método implementado por computador, de acordo com a reivindicação 13, caracterizado pelo fato de ainda compreender:

h2) armazenar, se é determinado manualmente que a coleção não cumpre com a política, uma indicação de que os anúncios da rede de propaganda não serão servidos com documentos incluídos na coleção.

16. Aparelho, caracterizado pelo fato de compreender:

a) meio para aceitar uma coleção incluindo um ou mais documentos;

b) meio para determinar se a coleção cumpre ou não com políticas de uma rede de propaganda; e

c) meio para aprovar a coleção se foi determinado que a coleção cumpre com as políticas.

17. Meio legível por computador caracterizado por ter armazenado nele instruções legíveis por computador que, quando executadas por um computador, efetuam um método compreendendo:

a) aceitar a coleção incluindo um ou mais documentos;

b) determinar se a coleção cumpre ou não com políticas de uma rede de propaganda; e

c) aprovar a coleção se foi determinado que a coleção
5 cumpre com as políticas.

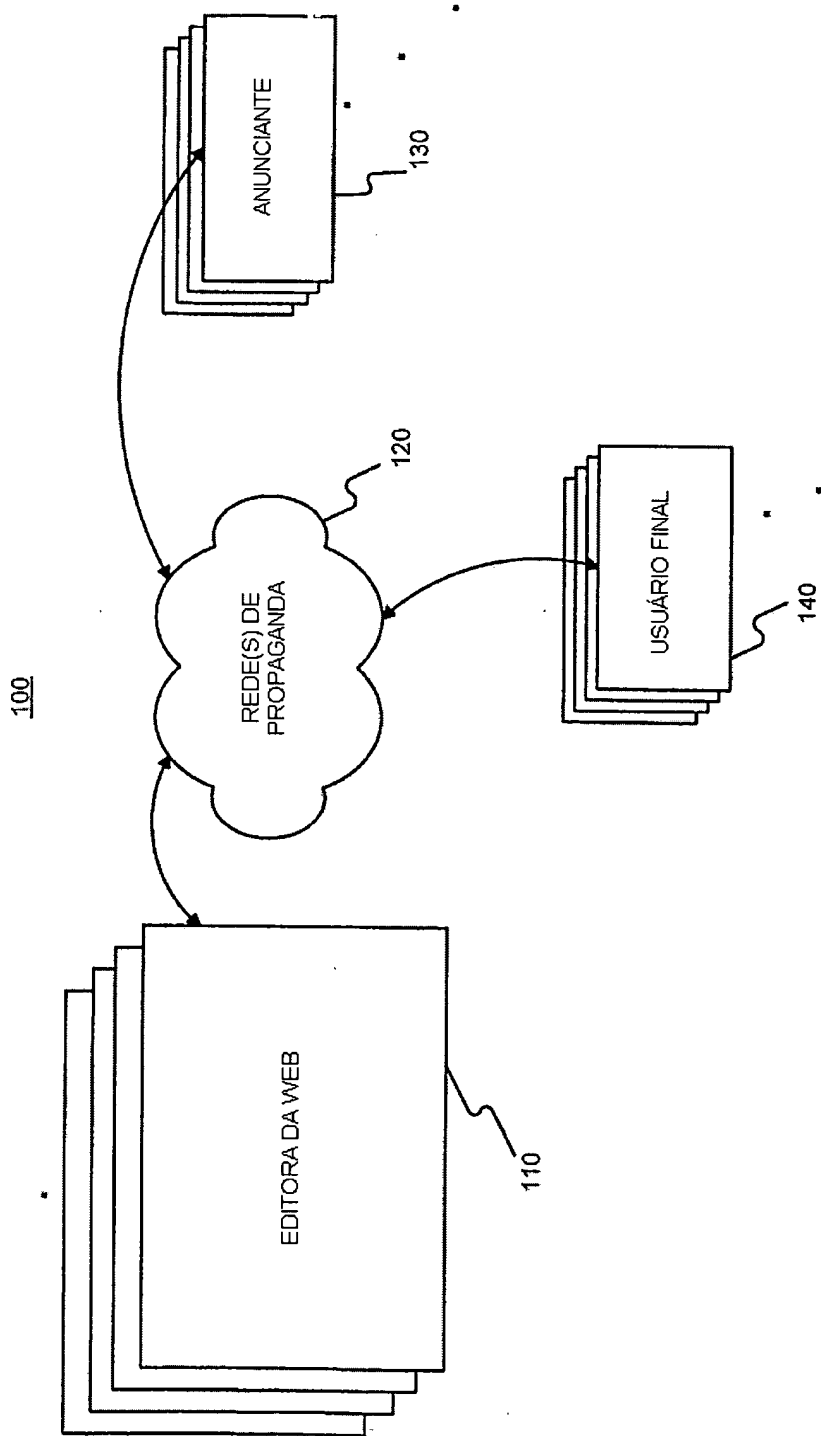
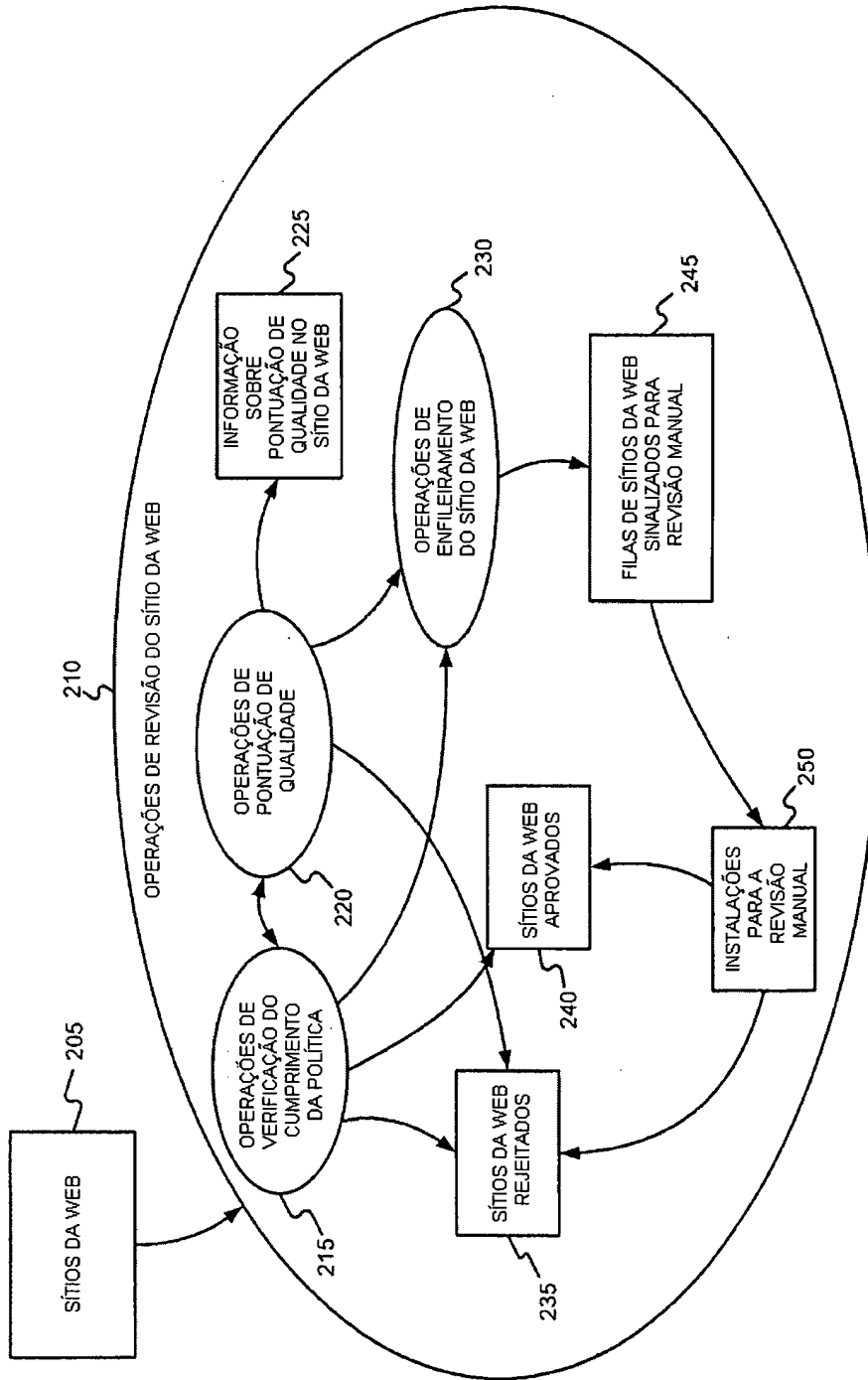


FIGURA 1



200
FIGURA 2

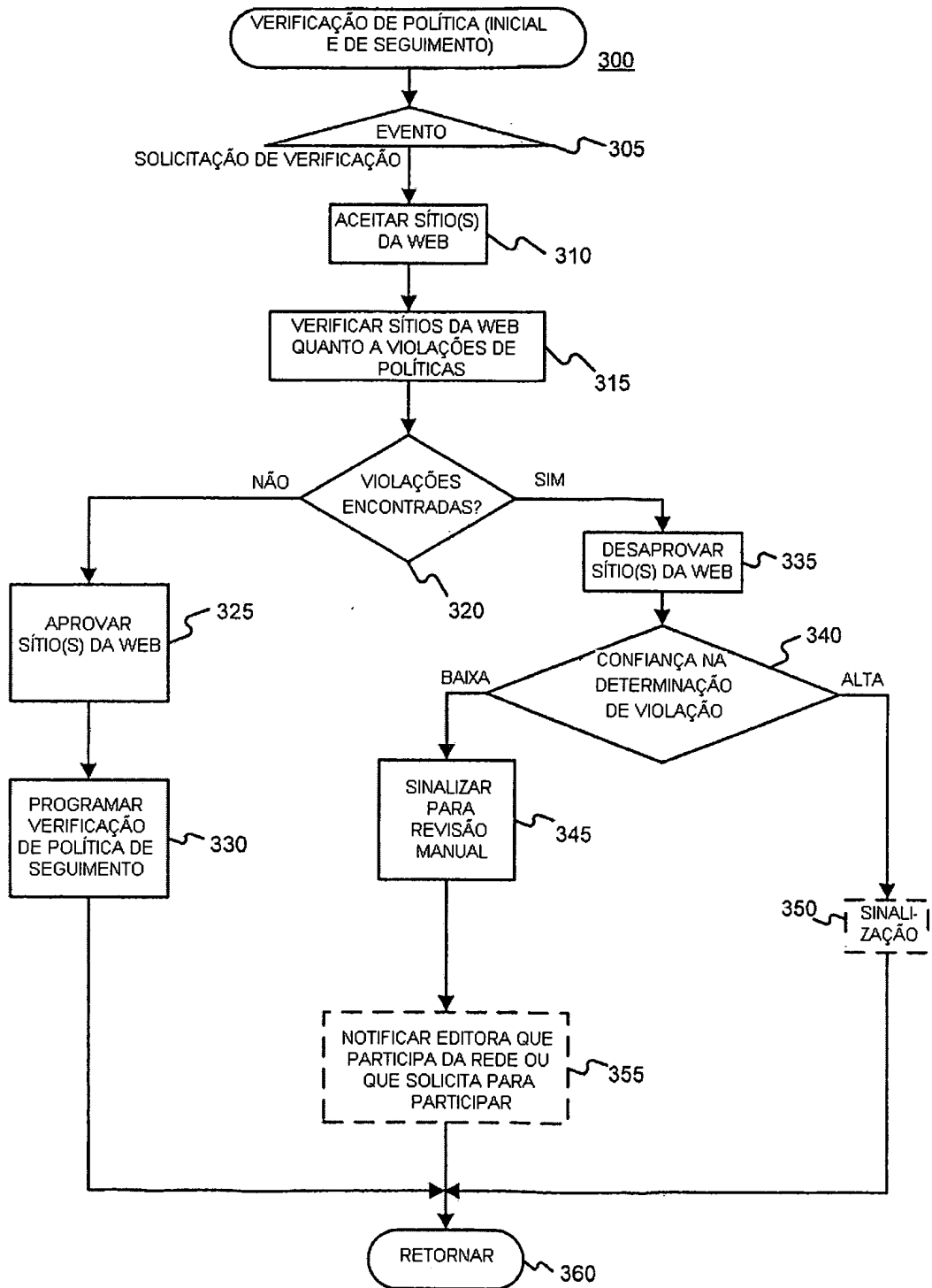


FIGURA 3

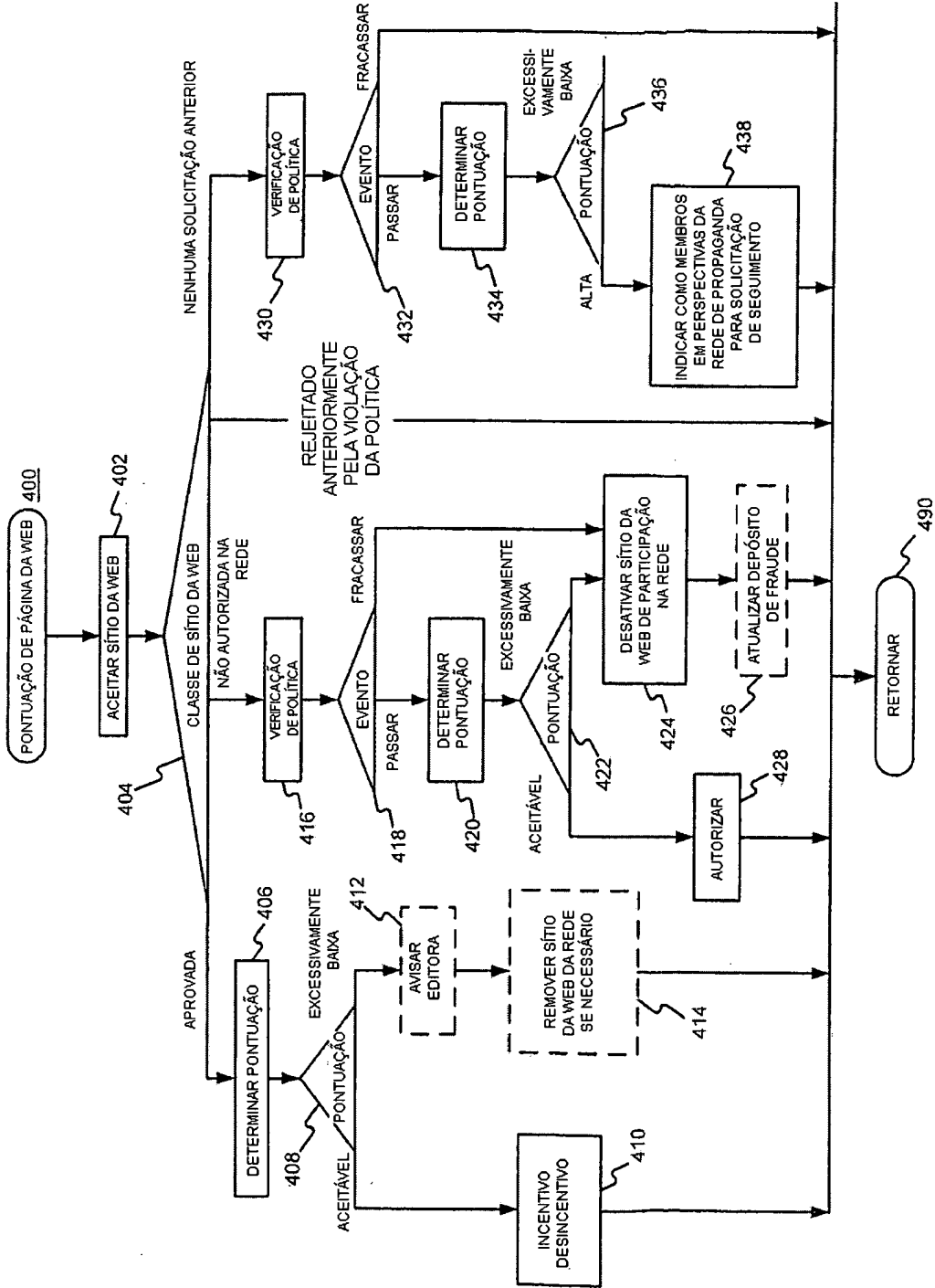


FIGURA 4

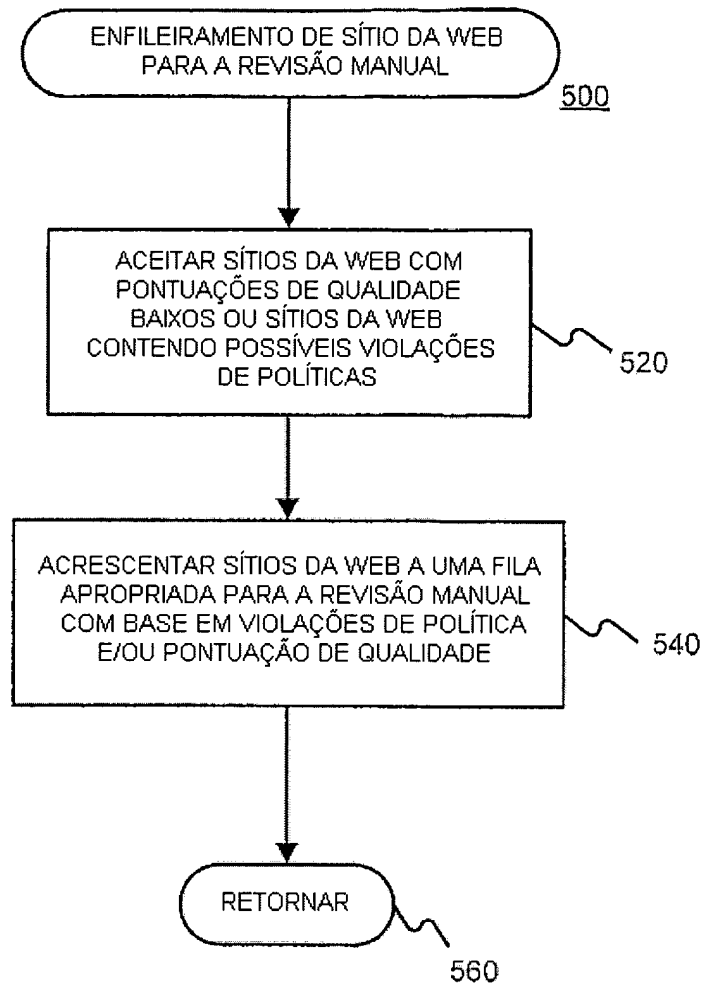


FIGURA 5

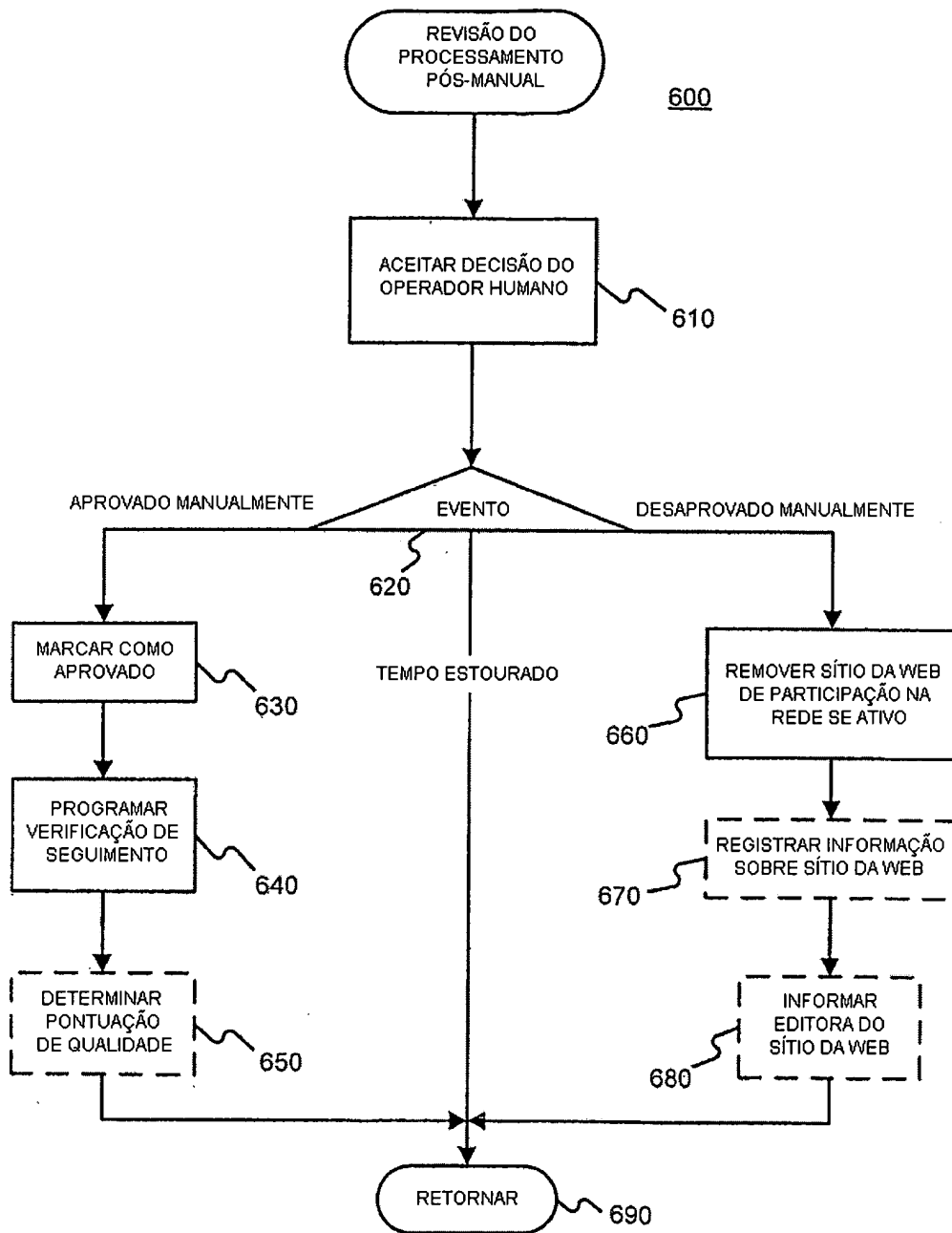


FIGURA 6

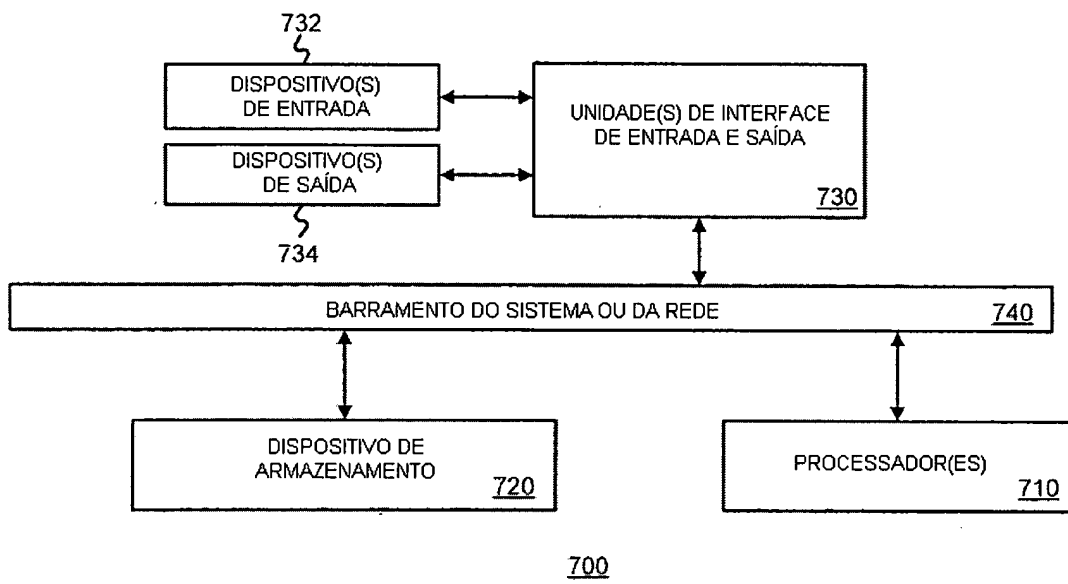


FIGURA 7

**MÉTODO IMPLEMENTADO POR COMPUTADOR, APARELHO E MEIO LEGÍVEL
POR COMPUTADOR**

A maneira pela qual os sítios da Web são revistos para
utilização em uma rede de propaganda poderá ser aprimorada
5 pela (a) aceitação de uma coleção incluindo um ou mais
documentos, (b) determinar se a coleção cumpre ou não com
políticas de uma rede de propaganda, e (c) aprovar a
coleção se foi determinado que a coleção cumpre com uma ou
mais políticas tal que (por exemplo, direcionada por
10 conteúdo) os anúncios poderão ser servidos em associação
com os documentos na coleção. A política poderá abranger
(A) conteúdo de um ou mais documentos, (B) capacidade de
uso do sítio da Web e/ou (C) fraude ou decepção possíveis
na rede de propaganda ou nos participantes da rede de
15 propaganda.