



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

(21) **PI0615323-2 A2**

(22) Data de Depósito: 30/08/2006
(43) Data da Publicação: 17/05/2011
(RPI 2106)



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

(54) Título: **MÉTODOS DE BUSCAS LOCAIS, DE GEOCODIFICAÇÃO E DE INDEXAÇÃO E SISTEMAS DE BUSCAS LOCAIS**

(30) Prioridade Unionista: 30/08/2005 US 60/712,146

(73) Titular(es): GOOGLE INC.

(72) Inventor(es): HONGJUN ZHU, HUICAN ZHU, KUN SHING LUK

(74) Procurador(es): HUGO SILVA , ROSA & MALDONADO - PROP. INT

(86) Pedido Internacional: PCT US2006033537 de 30/08/2006

(87) Publicação Internacional: WO 2007/027608 de 08/03/2007

(57) Resumo: Métodos de Buscas Locais, de Geocodificação e de Indexação e Sistemas de Buscas Locais
Resumo Um sistema recebe dados das páginas amarelas, dados de provedor de mapas e dados de documento em resposta a uma questão de busca local e geocodifica os dados de modo a atribuir um identificador geográfico e combinar pelo menos um endereço associado com a questão de busca local. O sistema também indexa os dados geocodificados para determinar informações de negócios e informações de localização associadas com a questão da busca local. O sistema ainda proporciona os resultados da busca local e um mapa com base nos dados indexados.

“Métodos de Buscas Locais, de Geocodificação e de Indexação e Sistemas de Buscas Locais”

Relatório Descritivo

Antecedentes

5 As modalidades aqui descritas relacionam-se em geral com a recuperação de informação e, mais particularmente, com a identificação de resultados de busca local.

 A World Wide Web (“Web”) contém uma vasta quantidade de informações. Localizar uma parte pretendida das informações pode, porém, ser desafiador. Este problema pode ser complexo, porque a
10 quantidade de informações sobre a Web e o número de novos usuários sem experiência em buscas na Web são rapidamente crescentes.

 Os sistemas de busca tentam retornar *hiperlinks* para páginas da Web em que o usuário está interessado. Geralmente, os
15 sistemas de busca baseiam a sua determinação do interesse do usuário em termos de busca (chamados de questão de busca) introduzidos pelo usuário. A meta do sistema de busca pode ser prover *links* para resultados relevantes, de alta qualidade (por exemplo, páginas da Web) para o usuário com base na questão de busca. Tipicamente, o sistema de
20 buscas realiza isto por comparação dos termos na questão de busca com um corpo de páginas da Web pré-armazenadas. As páginas da Web que contêm os termos de busca do usuário podem ser chamadas de “acertos” ou “resultados de busca” e podem ser retornadas ao usuário como *links*.

25 Os sistemas de buscas locais tentam retornar páginas relevantes da Web e/ou listagens de negócios dentro de uma área geográfica específica. Em alguns países, as informações detalhadas de endereços não estão disponíveis para alguns negócios. Como resultado, pode ser difícil realizar buscas locais.

De acordo com um aspecto, um método pode incluir receber dados das páginas amarelas, dados dos provedores de mapas de terceiros e dados de documentos em resposta a uma questão de busca local e geocodificar pelo menos um dos dados das páginas amarelas, os dados dos provedores de mapas de terceiros e os dados de documentos para atribuir um identificador geográfico e combinar pelo menos um endereço dentro da questão de busca local. O método pode também incluir a indexação dos dados geocodificados para identificar informações de negócios e informações de localização que correspondem à questão de busca local e fornecer os resultados da busca local e um mapa de terceiros com base nas informações de negócios e informações de localização identificadas.

De acordo com outro aspecto, um método de geocodificação com base numa questão de busca local pode incluir receber dados dos provedores de mapas de terceiros e dados das páginas amarelas, gerar um endereço com base na questão de busca local, analisar o endereço, localizar prefixos de coincidência mais longos no endereço, para identificar pelo menos uma parte do endereço, e localizar uma combinação no endereço para verificar o endereço.

De acordo ainda com outro aspecto, um método de indexação com base numa questão de busca local pode incluir pré-processar dados das páginas amarelas para um formato predeterminado, extrair informações de negócios a partir dos dados de documentos, armazenar as informações de negócios num repositório e indexar informações de endereços a partir de dados dos provedores de mapas de terceiros. O método pode também incluir agrupar os dados das páginas amarelas e os dados dos provedores de mapas de terceiros e destacar fragmentos nos dados de documentos. Para uma localização pontual, o método pode incluir ajustar uma distância de busca para a questão de busca local a uma distância predeterminada. Para uma localização limitada, o método pode ainda incluir ajustar a distância de busca a aproximadamente uma distância máxima de um centróide da localização limitada a

cantos da localização limitada.

De acordo com um aspecto adicional, um método pode incluir gerar uma lista de sinônimos relacionados com uma questão de busca local, reescrever a questão de busca local para expandir a questão, gerar os resultados da busca local com base na questão expandida, obter um mapa a partir de um provedor de mapas de terceiros com base nos resultados da busca local e gerar anúncios com base em informações geográficas relacionadas com os resultados da busca local.

De acordo com outro aspecto, um sistema pode incluir um indexador para receber dados dos provedores de mapas de terceiros, dados das páginas amarelas e dados de documentos, pré-processar os dados das páginas amarelas para determinar informações de negócios, extrair informações de negócios a partir dos dados de documentos, identificar as informações de localização numa questão de busca local e indexar dados de endereço dos dados dos provedores de mapas de terceiros. O sistema pode também incluir um geocodificador para receber informações a partir do indexador e atribuir identificadores geográficos e um servidor de *front end* para receber informações a partir do geocodificador, reescrever a questão de busca local, obter um mapa dos dados dos provedores de mapas de terceiros e gerar os resultados da busca local com base na questão de busca local.

De acordo ainda com outro aspecto, um sistema pode incluir meios para receber dados das páginas amarelas, dados de provedores de mapas de terceiros e dados de documentos em resposta a uma questão de busca local e meios para geocodificar pelo menos um dos dados das páginas amarelas, os dados dos provedores de mapas de terceiros e os dados de documentos para atribuir um identificador geográfico e comparar pelo menos um endereço dentro da questão de busca local. O sistema pode também incluir meios para indexar os dados geocodificados para identificar informações de negócios e informações de localização correspondendo à questão de busca local e meios

para fornecer os resultados da busca local e um mapa de terceiros com base nas informações de negócios e informações de localização identificadas.

De acordo com um aspecto adicional, um sistema pode
5 incluir uma memória para armazenar um grupo de instruções e um processador para executar instruções na memória. O processador pode identificar uma posição associada com uma questão de busca local, identificar os resultados da busca local relevantes para a questão de busca local e associados com a posição identificada, identificar um
10 identificador para cada um de um grupo de resultados da busca local e receber a partir de um provedor de mapas de terceiros um mapa associado com a posição identificada, em que o mapa identifica uma posição de pelo menos um resultado de busca local no grupo de resultados da busca local.

De acordo ainda com um aspecto adicional, um método
15 pode incluir receber uma questão de busca local, identificar uma localização associada com a questão de busca local, identificar um conjunto de resultados de busca relevantes para a questão de busca local e associados com a localização identificada e identificar um
20 identificador para cada um de um grupo de resultados de busca. O método pode também incluir fornecer o identificador para cada um dos grupos de resultados de busca para um provedor de mapas de terceiros e receber do provedor de mapas de terceiros um mapa associado com a posição identificada, em que o mapa identifica uma posição de pelo
25 menos um resultado de busca no grupo de resultados de busca.

De acordo com outro aspecto, um método pode incluir gerar uma lista de fichas, identificar um endereço potencial dentro de um documento da Web e analisar o endereço potencial a partir de um início para determinar se o endereço potencial inclui uma ficha associada com
30 uma cidade. O método pode também incluir analisar melhor o endereço potencial para determinar se o endereço potencial inclui uma ficha

associada com um distrito, identificar uma ficha de comparação mais longa no endereço potencial depois da ficha associada com a cidade ou a ficha associada com o distrito e determinar se o endereço potencial é um endereço real com base na ficha associada com a cidade, a ficha associada com o distrito e a ficha de comparação mais longa identificada.

Breve Descrição dos Desenhos

Os desenhos anexos, que são incorporados e constituem uma parte deste Relatório Descritivo, ilustram uma ou mais modalidades da invenção e, junto com a descrição, explicam aspectos da invenção. Nos desenhos:

a **Figura 1** é um diagrama de uma vista global de uma implementação exemplificativa aqui descrita;

a **Figura 2** é um diagrama de uma rede exemplificativa em que podem ser implementados os sistemas e métodos aqui descritos;

a **Figura 3** é um diagrama exemplificativo de um cliente ou servidor dentro da rede exemplificativa da Figura 2;

a **Figura 4** é um diagrama de blocos funcional de um sistema exemplificativo para identificar os resultados da busca local e prover um mapa associado com localizações identificadas;

a **Figura 5** é um diagrama exemplificativo de um repositório de índices/documentos do sistema exemplificativo da Figura 4;

a **Figura 6** é um diagrama exemplificativo de um geocodificador do sistema exemplificativo da Figura 4;

a **Figura 7** é um diagrama exemplificativo de um indexador do sistema exemplificativo da Figura 4;

a **Figura 8** é um diagrama exemplificativo de um servidor de *front end* do sistema exemplificativo de Figura 4;

a **Figura 9** é um diagrama de resultados de busca local exemplificativos e um mapa gerado pelo sistema exemplificativo da Figura 4; e

as **Figuras 10A-10D** são fluxogramas de um processo exemplificativo para identificar os resultados da busca local e prover um mapa associado com localizações identificadas.

Descrição Detalhada

A descrição detalhada seguinte da invenção refere-se aos desenhos anexos. Os mesmos números de referência em desenhos diferentes podem identificar os mesmos elementos ou semelhantes. Também a descrição detalhada seguinte não limita a invenção.

Em alguns países, como a China, por exemplo, os dados de mapas e dados das páginas amarelas podem não estar disponíveis a partir de um provedor único e devem ser obtidos a partir de vários provedores diferentes. Devido a restrições de exportação, pode não ser possível obter dados de mapas detalhados para prover o mapa de uma área ou conseguir a latitude e a longitude reais de endereços dentro da área. Como resultado, pode ser usada a aproximação de endereços para a geocodificação de endereços. Quando são enviados resultados a partir de uma busca local para o usuário, a página de resultados local pode incluir uma lista de resultados relevantes e um ponteiro para o servidor do provedor de mapas (terceiros). O provedor de mapas pode ser responsável por gerar o mapa exibido para o usuário.

Avaliação

As implementações aqui descritas podem identificar os resultados da busca local e gerar um mapa associado com localizações identificadas. Por exemplo, numa modalidade, conforme mostrado na Figura 1, um sistema pode receber uma entrada de questão de busca local por um usuário e pode identificar uma posição associada com a questão de busca local. O sistema pode identificar um conjunto de resultados da busca local (por exemplo, resultados de busca "A" até "H")

que podem estar relacionados com a questão de busca local e podem estar associados com a posição identificada. Os resultados da busca local podem incluir *links* para documentos que podem estar relacionados com a questão de busca local. O sistema pode identificar um

5 identificador para um grupo de resultados da busca local e pode prover o identificador para um provedor de mapas. O sistema pode receber um mapa associado com a localização identificada do provedor de mapas. O mapa pode identificar uma posição de pelo menos um resultado de busca (por exemplo, resultado de busca "A", como mostrado na Figura

10 1) no grupo de resultados da busca local.

Um "documento", como o termo é aqui usado, deve ser interpretado amplamente de modo a incluir qualquer produto de trabalho legível em máquina e armazenável em máquina. Um documento pode incluir, por exemplo, um correio eletrônico, um arquivo,

15 uma combinação de arquivos, um ou mais arquivos com *links* embutidos para outros arquivos, um grupo de notícias postando um blog, um anúncio de Web etc. No contexto da Internet, um documento comum é uma página da Web. As páginas da Web incluem freqüentemente informações textuais e podem incluir informações embutidas (tais como

20 meta-informações, imagens, *hyperlinks* etc.) e/ou instruções embutidas (tal como JavaScript etc.).

Um "*link*", como o termo é aqui usado, deve ser interpretado amplamente de modo a incluir qualquer referência a /a partir de um documento a partir de/para outro documento ou outra parte do mesmo

25 documento.

Configuração de Rede Exemplificativa

A Figura 2 é um diagrama exemplificativo de uma rede 200 em que podem ser implementados os sistemas e métodos aqui descritos. A rede 200 pode incluir clientes múltiplos 210 conectados a

30 servidores múltiplos 220-240 via uma rede 250. Dois clientes 210 e três servidores 220-240 foram ilustrados como conectados à rede 250

para simplicidade. Na prática, pode haver mais ou menos clientes e servidores. Também, em algumas instâncias, um cliente pode realizar uma ou mais funções de um servidor e/ou um servidor pode realizar uma ou mais funções de um cliente.

5 Os clientes 210 podem incluir entidades clientes. Uma entidade pode ser definida como um dispositivo, tal como um computador pessoal, um telefone sem fio, um assistente digital pessoal (PDA), um *laptop* ou outro tipo de dispositivo de computação ou comunicação, um *thread* ou processo que é executado num destes dispositivos e/ou
10 um objeto executável por um destes dispositivos. Os servidores 220-240 podem incluir: entidades servidoras que reúnem, processam, buscam e/ou mantêm documentos.

Numa implementação, o servidor 220 pode incluir um sistema de busca local 225 utilizável por clientes 210. O servidor 220
15 pode abarcar um corpo de documentos, indexar os documentos e armazenar as informações associadas com os documentos num repositório de documentos. Qualquer combinação de servidores 220-240 pode implementar o sistema de busca local 225 para identificar os resultados da busca local e fornecer um mapa associado com localiza-
20 ções identificadas.

Embora os servidores 220-240 sejam mostrados como entidades separadas, pode ser possível que um ou mais de servidores 220-240 realizem uma ou mais das funções de outro ou mais servidores 220-240. Por exemplo, pode ser possível que dois ou mais de servidores
25 220-240 sejam implementados como um servidor único. Pode ser também possível que um único dos servidores 220-240 seja implementado como dois ou mais dispositivos separados (e possivelmente distribuídos).

A rede 250 pode incluir uma rede local (LAN), uma rede remota (WAN), uma rede de telefones, tal como a Rede Pública de
30 Telefones Comutados (PSTN) ou uma rede celular, uma intranet, a

Internet ou uma combinação de redes. Os clientes 210 e servidores 220-240 podem conectar a rede 250 via conexões por fios, sem fios e/ou ópticas.

Arquitetura Exemplificativa

Cliente/Servidor

A Figura 3 é um diagrama exemplificativo de um cliente ou entidade de servidor (adiante chamada de “entidade cliente/servidor”), que pode corresponder a um ou mais clientes 210 e servidores 220-240. A entidade cliente/servidor pode incluir um barramento 310, um processador 320, uma memória principal 330, uma memória apenas de leitura (ROM) 340, um dispositivo de armazenamento 350, um dispositivo de entrada 360, um dispositivo de saída 370 e uma interface de comunicação 380. O barramento 310 pode incluir um caminho que permite a comunicação entre os elementos da entidade cliente/servidor.

O processador 320 pode incluir um processador, um microprocessador ou lógica de processamento que pode interpretar e executar instruções. A memória principal 330 pode incluir uma memória de acesso aleatório (RAM) ou outro tipo do dispositivo de memória dinâmica que pode armazenar informações e instruções para execução por processador 320. A ROM 340 pode incluir um dispositivo ROM ou outro tipo do dispositivo de armazenamento estático que pode armazenar informações e instruções estáticas para uso pelo processador 320. O dispositivo de armazenamento 350 pode incluir um meio de gravação magnética e/ou óptica e o seu *drive* correspondente.

O dispositivo de entrada 360 pode incluir um mecanismo que permite ao operador dar *input* a informações para a entidade cliente/servidor, tal como um teclado, um *mouse*, uma caneta, mecanismos de reconhecimento de som e/ou biométricos etc. O dispositivo de saída 370 pode incluir um mecanismo que dá a saída de informações para o operador, incluindo uma tela, uma impressora, um alto-falante etc. A interface de comunicação 380 pode incluir qualquer mecanismo

semelhante a um transceptor que ative a entidade cliente/servidor a comunicar-se com outros dispositivos e/ou sistemas. Por exemplo, a interface de comunicação 380 pode incluir mecanismos para comunicar com outro dispositivo ou sistema via rede, tal como a rede 250.

5 A entidade cliente/servidor pode desempenhar certas operações, como será descrito em detalhe abaixo. A entidade cliente/servidor pode realizar estas operações em resposta ao processador 320 executar instruções de software contidas num meio legível em computador, tal como a memória 330. Um meio legível em computador pode
10 ser definido como um dispositivo de memória física ou lógica e/ou onda portadora.

 As instruções de software podem ser lidas para a memória 330 a partir de outro meio legível em computador, tal como o dispositivo de armazenamento de dados 350, ou a partir de outro dispositivo via a
15 interface de comunicação 380. As instruções de software contidas na memória 330 podem ocasionar que o processador 320 realize processos que serão descritos mais tarde. Alternativamente, podem ser usados circuitos físicos em lugar ou em combinação com instruções de software para implementar processos aqui descritos. Deste modo, as implemen-
20 tações aqui descritas não ficam limitadas a nenhuma combinação específica de circuitos de hardware e software.

Sistema Exemplificativo

de Busca Local

 A Figura 4 é um diagrama de blocos funcional de um
25 sistema exemplificativo (por exemplo, sistema de busca local 225) para identificar os resultados da busca local e fornecer um mapa associado com localizações identificadas. De acordo com uma implementação, uma ou mais das funções de sistema 225, como descritas abaixo, podem ser realizadas por um servidor (por exemplo, o servidor 220),
30 uma parte do servidor 220 ou uma combinação de servidores (por exemplo, servidores 220-240). De acordo com outra implementação,

uma ou mais destas funções podem ser realizadas por uma entidade separada a partir do servidor 220, tal como o cliente (por exemplo, o cliente 210), um computador associado com o servidor 220 ou um dentre os servidores 230 ou 240.

5 Conforme mostrado na Figura 4, o sistema 225 pode incluir dados de provedores de mapas 400, dados das páginas amarelas 405, dados de documentos da Web 410, um repositório de índices/docu-
mentos 415, informações de identificação (ID) de endereços e pontos de
interesse (POI) 420, a impressão digital de endereços (FP) e informações
10 de mapeamento de ID de POI 425, um geocodificador 430, um indexa-
dor 435, um servidor de *front end* 440, informações de Localizador de
Recurso Uniforme (URL) de mapas 445 e os resultados da busca local
450. Os dados dos provedores de mapas 400, os dados das páginas
amarelas 405 e os dados de documentos da Web 410 podem ser forne-
15 cidos ao indexador 435 e os dados de provedores de mapas 400 podem
ainda ser usados para derivar informações de endereços/ID de POI 420
e informações de mapeamento de ID de POI e endereços de FP 425. O
indexador 435 pode conectar-se a um repositório de índices/documen-
tos 415 e ao geocodificador 430 e as informações de endereços/ID de
20 POI 420 podem ser fornecidas ao geocodificador 430. O servidor de
front end 440 pode conectar-se ao geocodificador 430 e ao repositório de
índices/documentos 415 e pode receber informações de mapeamento de
ID de POI e endereços FP 425. O servidor de *front end* 440 pode gerar
informações de URL de mapa 445 e os resultados da busca local 450.
25 O sistema 225 pode incluir alternativamente outras conexões e/ou
inter-relações componentes não mostradas na Figura 4.

Em geral, o geocodificador 430 pode receber informações de
ID de POI/endereços e/ou informações a partir do indexador 435, pode
atribuir identificadores geográficos (por exemplo, localizações, coorde-
30 nadas etc.) para objetos e pode fornecer saídas para o servidor de *front*
end 440. O indexador 435 pode receber dados do provedor de mapas
400, dados das páginas amarelas 405 e/ou dados de documentos da

Web 410, podem pré-processar dados das páginas amarelas 405 para determinar informações de negócios, extrair informações de negócios a partir de dados de documentos da Web 410, identificar informações de localização numa questão de busca, indexar os dados de endereço dos dados de provedores de mapas 400 e podem prover saídas para repositório de índices/documentos 415 e/ou geocodificador 430. O servidor de *front end* 440 pode receber informações de ID de POI e FP de endereços 425, informações a partir do geocodificador 430 e/ou informações de repositório de índices/documentos 415, pode reescrever a questão de busca e pode gerar informações de URL de mapas 450, os resultados da busca local 450 e/ou informações geográficas para uso por um servidor de anúncios (anúncios) (não mostrado).

Dados dos Provedores de Mapas

Pode haver dois tipos de dados para informações de endereços chineses: dados de mapas e dados das páginas amarelas. Os dados dos provedores de mapas 400 podem incluir uma variedade de informações. Por exemplo, numa modalidade, um provedor de mapas de terceiros pode fornecer um conjunto de POIs (por exemplo, negócios e outros lugares de interesse, tais como museus, parques, hospitais, escolas etc.) e seus endereços para as cidades na China. Pode haver informações de ID de POI para cada POI. Devido a restrições legais, o provedor de mapas pode não fornecer informações de latitude e longitude para cada POI. Eles podem, porém, dividir um mapa em grades (por exemplo, grades de trezentos metros por trezentos metros) e podem fornecer um índice de grade para cada POI e um programa para computar a distância entre as grades. Os dados de provedores de mapas 400 podem ser periodicamente atualizados (por exemplo, diariamente, semanalmente, mensalmente etc.) e, durante cada atualização, o ID de POI pode mudar.

Noutra implementação, os dados de provedores de mapas 400 podem incluir as seguintes informações: (1) um POI normal que

pode conter um nome de negócios, endereço, número de telefone e índice de grade; (2) um POI de estrada que pode conter um nome de rua e um índice de grade do centro da rua; e/ou (3) um POI de código postal que pode conter códigos postais e um índice de grade do centro aproximado dos códigos postais.

Ainda noutra implementação, pode haver temas com dados de provedores de mapas 400. Por exemplo, os dados de provedores de mapas 400 podem conter uma parte de um endereço, mas, pode estar faltando outra parte do endereço. O número de rua, por exemplo, pode estar faltando nos dados de provedores de mapas 400 devido a restrições na China (por exemplo, só 20-30% dos POIs podem ter números de rua). Como resultado, os dados de provedores de mapas 400 podem incluir apenas o nome de rua do POI ou algum nível mais alto de abstração. Também ou alternativamente, pode estar faltando o número de telefone para alguns dos POIs de dados de provedores de mapas 400. Também ou alternativamente, os endereços fornecidos pelos dados de provedores de mapas 400 podem estar impropriamente formatados.

Dados das Páginas Amarelas

Os dados das páginas amarelas 405 podem incluir uma variedade de informações e podem ser recebidos a partir de múltiplas origens (por exemplo, provedores de terceiros). Por exemplo, numa implementação, os dados das páginas amarelas 405 podem incluir informações de endereços (por exemplo, município, cidade, província) e/ou número de telefone para POIs, nomes de negócios, endereços de correio eletrônico, números de fac-símile, endereços de sítios da Web, nomes de executivos, descrições de negócios, categorias do estilo SIC, tipos de negócios etc. Noutra modalidade, os dados das páginas amarelas 405 podem não conter nenhuma informação de latitude e longitude. Como resultado, informações completas (ou substancialmente completas) relativas ao POIs (por exemplo, informações de endereços, números de telefone e posições em mapas) podem ser deduzidas a partir

de dados de provedores de mapas 400 e dados de páginas amarelas 405.

Dados de Documentos da Web

Os dados de documentos da Web 410 podem incluir uma
 5 variedade de informações. Por exemplo, numa modalidade, os dados de documentos da Web 410 podem incluir um correio eletrónico, um arquivo, uma combinação de arquivos, um ou mais arquivos com *links* embutidos para outros arquivos, um grupo de notícias, um blog, um anúncio de Web, uma página da Web (que pode incluir informações
 10 textuais IO e pode incluir informações embutidas (como meta-informações, imagens, *hiperlinks* etc.) e/ou instruções embutidas (como JavaScript etc.), informações de negócios (por exemplo, endereço, número de telefone etc.) etc. Os dados de documentos da Web 410 podem também incluir documentos localizados com base numa questão
 15 de busca local.

Informações de ID de POI/Endereços

As informações de ID de POI 420 podem incluir uma variedade de informações. Por exemplo, numa implementação, as informações de ID de POI/endereços podem incluir endereços extraídos
 20 a partir de dados de provedores de mapas 400, IDs de POI extraídas a partir de dados de provedores de mapas 400 etc. Noutra implementação, as informações de ID de POI de endereços 420 podem incluir as informações seguintes relativas a endereços chineses: dados de endereços fornecidos por dados de provedores de mapas 400 (por exemplo,
 25 endereços de POIs, centros de ruas); POIs (por exemplo, escolas, parques, edifícios, hospitais etc.); centros de códigos postais (isto é, o centro de uma área coberta por um código postal); cidades adicionais que podem não ser incluídas na lista fornecida por dados de provedores de mapas 400 etc.

Informações de Mapeamento de FP de Endereços

e ID de POI

As informações de FP de endereços e ID de POI 425 podem incluir uma variedade de informações. Por exemplo, numa implementação, as informações de FP de endereços /ID de POI 425 podem incluir um mapeamento entre impressões digitais de endereços (FPs) e ID de POIs correspondentes etc. O mapeamento entre os FPs de endereços e ID de POIs correspondentes pode ser usado pelo servidor de *front end* 440 para pesquisar o ID de POI para cada resultado de busca local, conforme descrito em detalhe abaixo. Numa modalidade alternativa, a ID de POIs podem ser derivadas diretamente sem mapeamento entre FPs de endereços e ID de POIs correspondentes. Nessa alternativa, as informações 425 podem incluir a ID de POIs diretamente derivadas.

Embora informações exemplificativas incluídas em dados de provedores de mapas 400, dados das páginas amarelas 405, dados de documentos da Web 410, informações de ID de POI/endereços e informações de mapeamento FP de endereços/ID de POI 425 tenham sido descritas acima, em outras implementações, informações adicionais ou diferentes sobre endereços que podem ser úteis para identificar os resultados da busca local podem ser incluídas nas informações exemplificativas. As Figuras. 5-8 são diagramas exemplificativos de alguns componentes do sistema 225 da Figura 4.

Repositório de Índices/Documentos

O repositório de Índices/documentos 415 pode ser fornecido num único dispositivo de memória (por exemplo, memória principal 330, ROM 340 e/ou dispositivo de armazenamento 350). O repositório de índices/documentos 415, como mostrado na Figura 5, pode armazenar uma variedade de informações relacionadas a documentos, dados das páginas amarelas 405 e/ou dados de provedores de mapas 400. Por exemplo, numa implementação, o repositório de índices/documentos 415 pode armazenar as informações seguintes relativas a endereços chineses: dados de endereços fornecidos por dados de provedores de mapas 400 (por exemplo, endereços de POIs, centros de

ruas); POIs (por exemplo, escolas, parques, edifícios, hospitais etc.); centros de códigos postais (isto é, o centro de uma área coberta por um código postal); cidades adicionais que podem não ser incluídas na lista fornecida por dados de provedores de mapas 400; etc.

5 Noutra modalidade, como mostrado na Figura 5, o repositório de índices/documentos 415 pode incluir um campo [província] 500, um campo [cidade] 510, um campo [distrito] 520 e um campo [rua] 530 (que podem adicional ou alternativamente incluir um campo de endereço e/ou um campo de nome de POI). Os exemplos seguintes podem
10 corresponder a informações de endereços chineses armazenados em repositório de índices/documentos 415:

[Cidade de Beijing] [Distrito de XiCheng] [Estrada de BeiLiShi];

15 [Cidade de Beijing] [Distrito de HaiDian] [Estrada de YiHe Yuan Número 5];

[Cidade de Beijing] [Distrito de HaiDian] [Universidade de Beijing]; e [Província de JiangSu] [Cidade de NanJing] [Distrito de GuLou] [Estrada de HanZhong].

Ainda noutra modalidade, o repositório de índices/docu-
20 mentos 415 pode incluir um campo de impressão digital de endereço (FP) 540 e um campo de precisão de FP 550. Cada endereço pode ser tratado como um ponto pelo geocodificador 430 com uma exatidão de FP de endereços e de FP (por exemplo, os edifícios podem ter uma exatidão de FP mais alta do que os centros das ruas). Um FP de ende-
25 reços pode ser uma impressão digital gerada com base no endereço (por exemplo, um valor *hash* que pode ser gerado com base no endereço) que pode ser usado para pesquisar o ID de POI para exibir o mapa correto. Por exemplo, um gerador de FP 570 pode receber um endereço (por exemplo, campo [província] 500, campo [cidade] 510, campo [distrito]
30 520, campo [rua] 530 etc.) e pode gerar um FP de endereços (por exemplo, campo [FP] 540) e uma exatidão de FP (por exemplo, campo de

exatidão [FP] 550) com base no endereço. A exatidão de FP de endereços e de FP pode ser usada por uma tabela de pesquisa de ID de POI 580 para pesquisar o ID de POI para exibir o mapa correto (por exemplo, com os resultados da busca local). O gerador de FP 570 e/ou
 5 tabela de pesquisa de ID de POI 580 pode ser fornecido no geocodificador 430, no indexador 435 ou no servidor de *front end* 440.

Ainda numa modalidade adicional, o repositório de índices /documentos 415 pode incluir um campo [zcode] 560 associado ao endereço. Um zcode associado pode incluir um código como um código
 10 postal, que, no caso da China, pode ser computado a partir do código administrativo de seis dígitos para distritos na China. As localizações, tais como Província de JiangSu, Cidade de Beijing, Cidade de Beijing Distrito de XiCheng, podem ser computadas pelo geocodificador 430 como limites que contêm todas as localizações de pontos dentro deles.

As variantes dos endereços podem também ser determinadas e armazenadas no repositório de índices/documentos 415 para
 15 aumentar as chamadas. Por exemplo, podem ser identificados os sinônimos de caracteres únicos para as províncias e cidades (por exemplo, Cidade de Xangai → Lu; Província de HeBei → Qi). Adicional ou
 20 ou alternativamente, partes dos endereços podem ser omitidas (por exemplo, Cidade de Xangai → Xangai; Distrito de HaiDian → HaiDian; Estrada Central de HuaiHai → Estrada de HuaiHai). Adicional ou
 alternativamente, podem ser incluídos sinônimos para lugares famosos (por exemplo, Templo da Cidade de Xangai → Templo da Cidade Velha;
 25 Templo de LiuRong → LiuRong Ta).

Embora a Figura 5 mostre informações exemplificativas incluídas no repositório de índices/documentos 415, em outras imple-
 mentações, o repositório de índices/documentos 415 pode incluir
 quaisquer informações que possam ser úteis para identificar os resulta-
 30 dos da busca local.

O geocodificador 430, como mostrado na Figura 6, pode realizar uma variedade de tarefas para ajudar na identificação dos resultados da busca local. Numa modalidade, por exemplo, o geocodificador 430 pode atribuir identificadores geográficos (por exemplo, posições, coordenadas etc.) a objetos. Por exemplo, o geocodificador 430 pode comparar endereços com endereços no repositório de índices/documentos 415. A saída do geocodificador 430 pode ser usada com vários propósitos. Por exemplo, os dados das páginas amarelas 405 podem ser analisados para determinar informações de negócios (por exemplo, endereço, número de telefone, endereço de correio eletrônico, número de fac-símile, horas de operação etc.) para POIs, as informações de negócios podem ser extraídas a partir de dados de documentos da Web 410 e/ou podem ser identificadas informações de localização fornecidas numa questão de busca.

A fim de indexar dados de páginas amarelas 405, o indexador 435 pode precisar determinar a posição no mapa pesquisando endereços chineses. Isto pode implicar uma variedade de assuntos. Por exemplo, as informações de posição precisa e de endereço detalhado para a maior parte dos endereços na China não estão prontamente disponíveis. Os endereços chineses não têm um formato e hierarquia bem definida. Os componentes possíveis de um endereço chinês podem incluir cidade, distrito, cidade, aldeia, estrada, rua, número de rua e edifício. Os dados das páginas amarelas 405 podem estar na forma livre (isto é, não se conformando a nenhum formulário ou formato particular), especialmente em áreas rurais, e, para alguns POIs, não existem nenhuns componentes de endereço correspondentes disponíveis. Os endereços exemplificativos chineses podem incluir:

Cidade de ShenZhen Cidade de HeGang Aldeia de AnLiang Estrada de AnLiang Número 172 Edifício JingCheng;

Cidade de ShenZhen Cidade de LongHua Complexo de Escritórios MingZhiMingLe;

Cidade de ShenZhen Estrada de FuHong Praça ShiMao Bloco A 17º Andar; e

Cidade de Beijing Distrito de HaiDian SiJiQingXiao-Fu.

5 O geocodificador 430 pode comparar melhor os endereços nos dados das páginas amarelas 405 com endereços fornecidos em dados de provedores de mapas 400 e obter a posição do mapa mais precisa possível. O geocodificador 430 pode incluir um gerador de lista de endereços 600, um analisador 610, um localizador de prefixos de
10 coincidência mais longos 620, um localizador de combinações 630, uma unidade de geocodificação da questão 640 etc. Os componentes do geocodificador 430 podem realizar várias tarefas para cada endereço (por exemplo, Cidade de ShenZhen Cidade de HeGang Aldeia de AnLi-ang Estrada de AnLiang Número 172 Edifício JingCheng).

15 O gerador de listas de endereços 600 pode receber dados de provedores de mapas 400 e dados das páginas amarelas 405 e pode gerar uma lista de endereços (também chamados de fichas) com base nas províncias, cidades, distritos, nomes de rua, endereços, nomes de POI ou combinações dos mesmos a partir de dados de provedores de
20 mapas 400 e/ou dados das páginas amarelas 405. Para cada endereço, o analisador 610 pode tentar analisar a cidade a partir do início do endereço. Por exemplo, "Cidade de ShenZhen" pode ser analisada pelo analisador 610 do endereço. O analisador 610 pode também tentar analisar o distrito do endereço. Por exemplo, visto que um distrito está
25 ausente do endereço exemplificativo descrito acima, o analisador 610 pode não ser capaz de analisar o distrito a partir do endereço.

O localizador de prefixos de coincidência mais longos 620 do geocodificador 430 pode ser usado para localizar partes adicionais do endereço. Por exemplo, o localizador de prefixos de coincidência mais
30 longos 620 pode tentar localizar os prefixos de coincidência mais longos (por exemplo, fichas) a partir da lista de endereços. Isto pode falhar, se

não existir uma ficha específica na lista de endereços (por exemplo, se “Cidade de HeGang” não estivesse na lista de endereços). O localizador de prefixos de coincidência mais longos 620 pode avançar para o final da palavra “Cidade”, se estiver presente na lista de endereços. Se uma

5 ficha inteira (por exemplo, “Cidade de HeGang Aldeia de AnLiang Estrada de AnLiang”) estiver presente na lista de endereços, o localizador de prefixos de coincidência mais longos 620 pode identificar a ficha como uma boa coincidência (por exemplo, “Cidade de HeGang” pode ser identificada como uma boa coincidência). O localizador de prefixos de

10 coincidência mais longos 620 pode tentar novamente localizar o prefixo de coincidência mais longo (por exemplo, ficha) da lista de endereços. Se esta falhar, então, o localizador de prefixos de coincidência mais longos 620 pode avançar para a extremidade da palavra “Aldeia”, se estiver presente (por exemplo, “Aldeia de AnLiang”). O localizador de

15 prefixos de coincidência mais longos 620 pode repetir o processo novamente para localizar o prefixo de coincidência mais longo (por exemplo, ficha) a partir da lista de endereços. Desta vez, o localizador de prefixos de coincidência mais longos 620 pode tentar combinar nomes de rua (por exemplo, “Estrada” ou “Rua”). O localizador de

20 prefixos de coincidência mais longos 620 pode avançar para o final da(s) palavra(s) “Estrada” ou “Rua”, se elas estiverem presentes (por exemplo, “Estrada de AnLiang”). O localizador de prefixos de coincidência mais longos 620 pode também avançar para o final da palavra “Número”, se ela estiver presente (por exemplo, “Número 172”). Finalmente, o localizador de prefixos de coincidência mais longos 620 pode tentar localizar

25 o prefixo de coincidência mais longo (por exemplo, ficha) a partir do nome do POI. Isto pode proporcionar coincidências para nomes de edifícios, escolas, parques etc. (por exemplo, “Edifício JingCheng”). O localizador de combinações 630 pode localizar combinações particulares

30 de forma que os nomes de POI possam ser localizados primeiro, depois, possam ser localizados vários níveis dos endereços e, finalmente, possam ser localizados distritos ou cidades. Essa disposição de locali-

zação pode garantir que o endereço mais específico pode ser obtido pelo localizador de combinações 630. Por exemplo, o localizador de combinações 630 pode localizar as combinações exemplificativas seguintes:

cidade + distrito + nome de POI;

5 cidade+ distrito+ endereço localizado pelo localizador de prefixos de coincidência mais longos 620 (por exemplo, “Estrada” ou “Rua”);

10 cidade+ distrito+ endereço localizado pelo localizador de prefixos de coincidência mais longos 620 (por exemplo, “Aldeia”);

cidade+ distrito+ endereço localizado pelo localizador de prefixos de coincidência mais longos 620 (por exemplo, “Cidade”);

cidade+ distrito; e/ou

15 cidade.

Quando uma questão de busca local é introduzida por um usuário, a parte de localização da questão de busca pode ser enviada para um servidor de mapas possuído por um provedor de mapas. Se a questão de localização contiver caracteres chineses, japoneses ou coreanos (CJK), o servidor de mapas pode enviar isto para a unidade de geocodificação da questão 640 para geocodificação. Se a localização introduzida pelo usuário for ambígua (por exemplo, quando a pontuação do primeiro resultado do geocodificador for menor do que duas vezes a pontuação do segundo resultado), a unidade de geocodificação da questão 640 pode apresentar sugestões ao usuário. Por exemplo, se a questão de localização for “Museu de História”, a unidade de geocodificação da questão 640 pode apresentar as seguintes sugestões ao usuário: “Você quer procurar o Museu de História do Distrito de He-Dong Cidade de TianJian ou o Museu de História de Xangai do Distrito de PuDong Cidade de Xangai?”

A unidade de geocodificação da questão 640 pode computar a pontuação de um resultado de busca com base no número de pontos na localização do resultado de busca. Por exemplo, se o Distrito de ChaoYang Cidade de ChangChuan” pontuar menos do que o Distrito de ChaoYang Cidade de Beijing, então o Distrito de ChaoYang Cidade de Beijing pode ser exibido, quando a localização da questão do usuário for “Distrito de ChaoYang”.

Embora a Figura 6 mostre tarefas exemplificativas realizadas pelo geocodificador 430, em outras modalidades, o geocodificador 430 pode realizar tarefas adicionais ou diferentes que podem ser usadas para identificar os resultados da busca local. Além disso, embora a Figura 6 mostre os componentes do geocodificador 430 como interconectados, em outras implementações, os componentes do geocodificador 430 podem ser componentes separados, não interconectados.

Indexador

O indexador 435, conforme mostrado na Figura 7, pode realizar uma variedade de tarefas para ajudar na identificação dos resultados de busca local. Numa implementação, por exemplo, o indexador 435 pode pré-processar dados das páginas amarelas 405 para determinar informações de negócios (por exemplo, endereço, número de telefone, endereço de correio eletrônico, número de fac-símile, horas de operação etc.) para POIs, pode extrair informações de negócios a partir de dados de documentos da Web 410 e/ou pode identificar informações de localização numa questão de busca. O indexador 435 pode incluir um indexador de dados de mapa 700, um pré-processador de dados das páginas amarelas 710, um extrator de informações de negócios 720, um planificador de distâncias 730, um agrupador 740, um repositório de informações de negócios 750, um destacadador de fragmentos 760 etc.

O indexador de dados de mapa 700 pode indexar dados de endereços a partir de dados de provedores de mapas 400, que podem

incluir informações de posições de mapas. O geocodificador 430 pode tentar geocodificar o endereço novamente com base nos dados de endereços indexados e o agrupador 740 (descrito abaixo) pode ajustar a posição do grupo se a posição do grupo for diferente da posição dada, mas pode não ajustar a posição do grupo acima de um limite de distância predeterminado (por exemplo, dentro três quilômetros). Essa disposição pode ser usada para melhorar o agrupamento, como descrito em detalhe abaixo.

O pré-processador de dados das páginas amarelas 710 pode receber dados das páginas amarelas 405 a partir de diferentes provedores e pode pré-processar os dados das páginas amarelas 405 para um formato comum. Estes dados formatados podem ser fornecidos para o geocodificador 430 durante a indexação e o geocodificador 430 pode tentar geocodificar o endereço. Se o endereço puder ser geocodificado ao nível de edifício ou rua, o endereço pode ser indexado como uma entrada normal. Se o endereço puder ser geocodificado ao nível de cidade ou distrito, o endereço pode ser indexado como uma entrada com uma posição aproximada. Durante a pontuação, a entrada pode ser tratada como se estivesse pelo menos a vinte quilômetros do centróide (isto é, tendo essencialmente a sua pontuação degradada). Se o endereço não puder ser geocodificado, o endereço pode ser tratado como uma entrada sem uma posição. Durante a indexação, se a entrada sem uma posição puder ser agrupada com outra entrada (por exemplo, usando o seu número de telefone), a entrada pode ser retida. Caso contrário, a entrada pode ser descartada.

O extrator de informações de negócios 720 pode extrair informações de negócios (por exemplo, endereços e números de telefone) a partir de dados de documentos da Web 410 usando uma variedade de técnicas. As técnicas podem ser modificadas com base na diferente formatação de endereços e números de telefone na China e no Japão.

Por exemplo, numa primeira técnica, o extrator de informa-

ções de negócios 720 pode incluir um classificador que pode analisar documentos com endereços para determinar informações de negócios associadas aos endereços com base num modelo estatístico. Os documentos analisados pelo classificador podem incluir documentos com endereços para os quais não existe nenhum dado correspondente das páginas amarelas 405 e/ou documentos com endereços para os quais existem dados de páginas amarelas 405 possivelmente incorretos. As funções desempenhadas pelo classificador podem diferir com base sobre se as informações de negócios correspondem informações de nomes de negócios (título) ou informações de números de telefone. Ainda outras funções podem ser realizadas quando as informações de negócios incluem informações diferentes de nomes de negócios ou informações de números de telefone. Um nome de negócio (título) pode ser identificado analisando termos próximos do endereço e determinar a probabilidade de que cada prazo é parte do título. Uma pontuação de confiança pode ser atribuída para cada título de candidato que é identificado.

Um número de telefone pode ser associado a um endereço identificando um conjunto de números de telefones candidatos no documento. Pode ser determinada, com base no modelo estatístico, a probabilidade de que cada um dos números de telefone candidatos esteja associado ao endereço, dada a predição relativa ao número de telefone candidato precedente e dada uma janela de termos (por exemplo, examinando um número predeterminado de termos à esquerda e/ou direita) em torno do número de telefone candidato. Podem ser atribuídas pontuações de confiança aos números de telefone candidatos com base nas suas probabilidades determinadas. Opcionalmente, pode ser determinado o melhor número de telefone para o endereço. O número de telefone pode, então, ser associado ao endereço para formar ou suplementar uma listagem de negócios.

Numa segunda técnica exemplificativa, o extrator de informações de negócios 720 pode incluir um extrator de localização que pode ser incluído como parte do motor de busca. O extrator de

localização pode receber uma questão de busca e determinar se a questão de busca inclui uma referência geográfica. Quando a questão de busca incluir uma referência geográfica, o extrator de posição pode separar a referência geográfica a partir dos termos de busca na questão e enviá-los para um motor de busca local. Quando a questão de busca não incluir uma referência geográfica, o extrator de localização pode enviar os termos de busca para um motor de busca da Web que pode incluir um motor de busca tradicional da Web que retorne um conjunto de documentos relacionados com uma questão de busca. O motor de busca local pode incluir um motor de busca especializado, tal como um motor de busca de listagens de negócios. Em operação, o motor de busca local pode receber os termos de busca e a referência geográfica de uma questão de busca a partir do extrator de localização. O motor de busca local pode identificar um conjunto de documentos que combina a questão de busca (isto é, documentos que contêm o conjunto de termos de busca da questão de busca) comparando os termos de busca com documentos num corpo de documentos relativos à área geográfica associada à referência geográfica. O motor de busca local pode pontuar os documentos identificados, classificá-los com base nas suas pontuações e dar-lhes saída como uma lista de resultados de busca.

O extrator de localização pode determinar endereços não ambíguos (por exemplo, cidades) numa questão de busca configurando uma variável i igual a um e realizando uma busca para o nome de uma cidade para cada cidade(i) numa lista de cidades. O número de resultados de busca para esta busca pode ser contado como countcity. Uma busca pode também ser realizada para o nome da cidade com o nome da província correspondente. O número de resultados de busca para esta busca pode ser contado como countcity/province.

Pode, então, ser determinado se countcity/province é pelo menos $X\%$ (onde X é um número maior do que zero) de countcity. Quando countcity/province é pelo menos $X\%$ de countcity, então a cidade pode ser considerada uma "cidade não ambígua". Uma "cidade

não ambígua” pode referir-se a uma cidade cujo nome pode ser usado sozinho numa questão de busca e ficará entendido que o usuário pretendeu a cidade e não qualquer outra coisa mais. Quando countcity/province não é pelo menos X% de countcity, então pode ser determinado se existem mais cidades na lista. Se existirem mais cidades na lista, então a variável *i* pode ser incrementada de um e pode ser avaliada a próxima cidade na lista das cidades.

Os documentos dos resultados de busca podem ser analisados para identificar quaisquer códigos postais que eles contêm. Os códigos postais podem ser identificados usando uma técnica de comparação de padrões e verificados comparando-os com uma lista de códigos postais. Pode, então, ser determinado se os códigos postais correspondem a códigos postais associados com a cidade(*i*). O número de documentos que contêm códigos postais associados com a cidade(*i*) pode ser contado como countpostal. Pode ser determinado se countpostal é pelo menos X% (por exemplo, 5%) de countcity. Quando countpostal é pelo menos X% de countcity, então a cidade pode ser considerada uma cidade não ambígua.

Numa terceira técnica exemplificativa, o extrator de informações de negócios 720 pode identificar um documento geograficamente relevante. Um documento geograficamente relevante, como usado aqui, geralmente pode se referir a qualquer documento que, de alguma maneira, foi determinado ter relevância particular para uma localização geográfica. As listagens de negócios (por exemplo, listagens de páginas amarelas) podem ser consideradas um documento geograficamente relevante que é relevante para a região geográfica definida pelo endereço dos negócios. Outros documentos, tais como documentos da Web, podem também ter relevância geográfica particular. Por exemplo, um negócio pode ter uma *home page*, pode ser o assunto de um documento que comenta ou repassa os negócios ou pode ser mencionado por uma página da Web que, de algum outro modo, se relaciona com os negócios. A localização geográfica particular a que um documento está associado

pode ser determinada de vários modos, tal como a partir do endereço postal ou de outros sinais geográficos.

A região geográfica associada ao documento geograficamente relevante pode ser mapeada para um identificador de localização correspondente. Os identificadores de localização adicionais podem ser determinados para o documento. Em particular, podem também ser determinados os identificadores de localização correspondentes às regiões circundantes dentro de uma faixa predeterminada. Cada documento geograficamente relevante pode ser indexado como se o documento incluísse os identificadores de localização associados à região do documento, assim como também às regiões circundantes identificadas.

Embora a Figura 7 mostre o extrator de informações de negócios 720 como parte do indexador 435, em outras modalidades, o extrator de informações de negócios 720 pode ser separado do indexador 435. Para assegurar que o conjunto de documentos de combinação pode estar disponível na indexação, as informações extraídas pelo extrator de informações de negócios 720 podem ser fornecidas no repositório de informações de negócios 750. O repositório de informações de negócios 750 pode incluir uma variedade de informações, por exemplo, os documentos a partir dos quais as informações de negócios foram extraídas pelo extrator de informações de negócios 720. O repositório de informações de negócios 750, junto com as informações de negócios extraídas, pode ser fornecido dentro do indexador 435. Embora a Figura 7 mostre o repositório de informações de negócios 750 como parte do indexador 435, em outras implementações, o repositório de informações de negócios 750 pode ser separado do indexador 435.

O planificador de distâncias 730 pode configurar um raio ou distância de busca para uma questão de busca local. Por exemplo, numa modalidade, cada questão de busca local pode ser geocodificada pelo geocodificador 430 para uma localização particular. Cada localiza-

ção pode ser uma localização de ponto (por exemplo, edifícios, lugares turísticos famosos, escolas, centros das ruas etc.) ou uma localização limitada (por exemplo, distritos, cidades, províncias etc.). Para localizações de ponto, o planejador de distâncias 730 pode configurar o raio de busca para uma distância predeterminada (por exemplo, aproximadamente cinco quilômetros em torno do ponto). Para localizações limitadas, o planejador de distâncias 730 pode configurar o raio de busca para aproximadamente a distância máxima a partir de um centróide da posição para os cantos. As pontuações dos resultados de busca num conjunto zcode associado à localização (isto é, o conjunto de zcodes que compõem a posição) podem ser promovidas. Deste modo, quando um usuário busca próximo a um nome de distrito, os resultados superiores podem estar dentro desse distrito.

O agrupador 740 pode agrupar dados de provedores de mapas 400 e dados das páginas amarelas 405 em conjunto. Muitos dos dados de provedores de mapas 400 podem incluir informações de posição precisa de forma que pode ser exibido com precisão num mapa, mas pode não incluir informações detalhadas de endereço ou número de telefone (por exemplo, poderia incluir um endereço sem um número de rua e/ou pode estar faltando um número de telefone). Por outro lado, os dados das páginas amarelas 405 podem incluir informações detalhadas de número de endereço e telefone, mas podem não incluir informações precisas da posição. Por exemplo, dados de provedores de mapas 400 podem incluir uma origem (por exemplo, provedor de mapas), um título (por exemplo, “Universidade de Beijing”), um endereço (por exemplo, “Cidade de Beijing Distrito de HaiDian Estrada de YiHe Yuan”) e/ou um ID de POI (para exibição de mapa) (por exemplo, “A1234567”). Entretanto, os dados das páginas amarelas 405 podem incluir uma origem (por exemplo, provedor de dados das páginas amarelas), um título (por exemplo, “Universidade de Beijing”), um endereço (por exemplo, “Cidade de Beijing Distrito de HaiDian Estrada de YiHe Yuan Número 5”) e/ou um número de telefone (por exemplo,

“010-62752114”). Se estas entradas forem agrupadas em conjunto pelo agrupador 740, então, o servidor de *front end* 440 pode ser capaz de prover o usuário com informações detalhadas de endereço e número de telefone, assim como também uma posição precisa num mapa.

5 A posição obtida a partir da geocodificação (por exemplo, com o geocodificador 430) do endereço de dados das páginas amarelas 405 pode ser uma aproximação e pode estar longe da posição precisa fornecida por dados de provedores de mapas 400. Neste caso, o mesmo negócio a partir de dois provedores pode ficar localizado em muitos
10 bairros diferentes e, deste modo, podem não ser agrupados em conjunto pelo agrupador 740. A solução para isto pode incluir a geocodificação (por exemplo, com o geocodificador 430) do endereço a partir de dados de provedores de mapas 400 até uma posição de grupo. A posição do grupo pode ser usada para geração de vizinhança, assim como também
15 para agrupar pelo agrupador 740. A posição real fornecida pelos dados dos provedores de mapas 400 pode, então, ser usada para exibir o mapa.

O destacador de fragmentos 760 pode destacar fragmentos em documentos (por exemplo, documentos da Web). O destaque do
20 fragmento de documento da Web pode ser tipicamente realizado por desvios de termos nos documentos. Visto que um documento CJK pode ter os termos contados de modo diferente a partir de outras linguagens (por exemplo, usar um espaço como delimitador, um parágrafo longo de texto pode precisar de ser segmentado para obter os termos), a fim de
25 destacar termos específicos, o documento inteiro pode ser segmentado para obter os termos correspondentes, o que pode ser ineficiente. Contudo, o destacador de fragmentos 760 pode armazenar desvios de baítes em vez de desvios de termos para identificar um endereço ou número de telefone (ou algumas outras informações de negócios) num
30 documento da Web durante a indexação. O destacador de fragmentos 760 pode usar o deslocamento de baítes para realizar o destaque e pode não ser exigida nenhuma segmentação. Embora a Figura 7 mostre o

destacador de fragmentos 760 como parte do indexador 435, em outras implementações, o destacador de fragmentos 760 pode ser separado do indexador e/ou estar incluído noutro componente (por exemplo, dentro servidor de *front end* 440).

5 Embora a Figura 7 mostre tarefas exemplificativas realizadas pelo indexador 435, em outras modalidades, o indexador 435 pode realizar tarefas adicionais ou diferentes que podem ser usadas para identificar os resultados da busca local. Além disso, embora a Figura 7 mostre os componentes de indexador 435 como interconectados, em
10 outras modalidades, os componentes de indexador 435 podem ser componentes separados, não interconectados.

Servidor de *Front End*

O servidor de *front end* 440, conforme mostrado na Figura 8, pode realizar uma variedade de tarefas para ajudar na identificação
15 dos resultados da busca local. Numa modalidade, por exemplo, o servidor de *front end* 440 pode incluir um regravador de questão 800, um gerador de busca local 810, um gerador de mapas 820, um gerador de informações geográficas 830 etc. Embora Figura 8 mostre tarefas exemplificativas realizadas pelo servidor de *front end* 440, em outras
20 implementações, o servidor de *front end* 440 pode realizar tarefas adicionais ou diferentes que podem ser usadas para ajudar na identificação dos resultados da busca local.

Numa modalidade, o regravador de questão 800 pode realizar uma variedade de tarefas. Por exemplo, visto que o número de
25 agrupamentos de documento de Web pode ser muito menor para dados chineses em comparação com os dados ingleses, pode ser usada uma coincidência num título ou categoria para retornar resultados válidos. Também, em chinês, pode haver vários sinônimos que se referem ao mesmo termo. Por exemplo, o termo “restaurante” pode ter vários
30 sinônimos comuns em chinês. Além disso, os diferentes provedores de dados das páginas amarelas 400 podem usar diferentes nomes para

representar a mesma categoria. A fim de proporcionar a maior chamada para a questão de busca introduzida pelo usuário, o regravador de questão 800 pode gerar uma lista de sinônimos para cada uma das categorias. O regravador de questão 800 pode também regravar cada

5 questão de busca local, de modo a expandir a questão para um par de termos de busca que podem ser juntos por um operador “OU”.

O gerador de busca local 810 pode gerar os resultados da busca local 450. Por exemplo, o gerador de busca local 810 pode gerar resultados correspondendo a páginas relevantes da Web e/ou listagens

10 de negócios dentro de uma área geográfica específica, com base nas informações de endereços obtidas e processadas pelo sistema de busca local 225. Os resultados da busca local 450 podem ser exibidos (por exemplo, para o usuário que introduz a questão de busca local) num monitor (por exemplo, dispositivo de saída 370).

O gerador de mapas 820 pode gerar um mapa relativo aos

15 resultados da busca local 450. Por exemplo, a fim de exibir o mapa para os resultados, o gerador de mapa 820 pode construir uma moldura i (isto é, uma moldura flutuante inserida numa página da Web) no resultado de busca da página e pode executar uma ação posterior para

20 um URL fornecido pelo provedor de mapas. Os parâmetros no posto podem conter os campos seguintes para cada resultado da busca: um título; um endereço; um número de telefone; ID de POI (para exibir o ponto no mapa); e exatidão. As localizações que podem ser geocodificadas a um nível de edifício podem ser marcadas como precisas no mapa

25 pelo gerador de mapas 820. As localizações que podem ser geocodificadas a um nível de rua (ou algum nível mais alto de abstração) podem ser marcadas como estimadas no mapa pelo gerador de mapas 820. A Figura 9 ilustra como estes dois níveis de exatidão podem ser distinguidos num mapa 900 gerado com base nos resultados da busca local 910.

30 Por exemplo, as localizações precisas 920 podem ser identificadas por uma cor de marcador (por exemplo, verde) e as posições estimadas 930 podem ser identificadas por outra cor de marcador (por exemplo,

vermelho). Embora sejam mostrados na Figura 9 dois níveis de exatidão (precisa e estimada), os níveis adicionais de exatidão podem ser usados em outras modalidades consistentes com os princípios da invenção.

5 Durante a indexação pelo indexador 435, a impressão digital do endereço (por exemplo, FP de endereços 540) pode ser armazenada junto com a sua exatidão associada (por exemplo, exatidão de FP 550). O mapeamento entre a impressão digital de endereço e o ID de POI (por exemplo, informações de mapeamento de FP de endereços/ID de POI 425) pode ser usado pelo gerador de mapa 820 para pesquisar o
10 ID de POI para cada resultado de busca. O ID de POI pode ser usado pelo gerador de mapa 820 para identificar uma posição no mapa (por exemplo, mapa 900) para o provedor de mapas, de forma que o provedor de mapas pode mostrar a posição no mapa fornecido dentro da página de resultado.

15 A fim de exibir anúncios locais para a China, o gerador de informações geográficas 830 pode fornecer informações geográficas para o servidor de anúncios 840 numa variedade de modos. Por exemplo, numa implementação, cada província ou cidade auto-administrada pode ter um código de dois dígitos, que pode ser a segunda parte do
20 código da região definido por ISO-3166-2. O geocodificador 430 pode indexar este código (por exemplo, "CN-dd") com cada endereço e passá-lo para o servidor de *front end* 440 para cada endereço geocodificado com sucesso. O gerador de informações geográficas 830 pode enviar o código para o servidor de anúncios 840 como um geocódigo de região.

25 Noutra implementação, os anunciantes podem ofertar sobre pares de palavra-chave de <palavra-chave, posição>, em vez de usar *geotargeting*, porque, para buscas na Web, o *geotargeting* pode não trabalhar tão bem em países asiáticos como ocorre nos Estados Unidos. O inventário para anúncios pode ser mais para esses pares de palavra-
30 chave. Como resultado, o gerador de informações geográficas 830 pode concatenar palavras-chaves com as posições entradas pela questão de

busca e pode usar estas concatenações como palavras-chaves enviadas para o servidor de anúncios 840. Por exemplo, quando o usuário busca “restaurante” próximo de “Beijing”, a palavra-chave enviada para o servidor de anúncios 840 pelo gerador de informações geográficas 830 pode ser “restaurante Beijing”. Essa disposição pode ser fornecida tanto para o Japão como a China.

Ainda noutra implementação, o gerador de informações geográficas 830 pode determinar se a questão de busca inclui uma referência geográfica. Se a questão de busca não incluir uma referência geográfica, então, podem ser apresentados anúncios regulares pelo servidor de anúncios 840. Todavia, pode ser determinado se um indicador da posição do usuário, tal como o endereço de IP do usuário, está disponível. Quando um indicador da posição do usuário está disponível, então, podem ser apresentados anúncios locais com base na localização do usuário.

Se, por outro lado, a questão de busca incluir uma referência geográfica, o gerador de informações geográficas 830 pode determinar se a referência geográfica corresponde apenas a um nome de cidade (isto é, sem quaisquer outras informações geográficas, tal como nenhuma informação de província). Se a questão de busca incluir uma referência geográfica diferente de um nome de cidade apenas, então podem ser apresentados anúncios locais.

Se a questão de busca incluir uma referência geográfica correspondendo apenas a um nome de cidade, então o gerador de informações geográficas 830 pode determinar se a cidade corresponde a uma cidade não ambígua. Se a cidade não corresponder a uma cidade não ambígua, então, podem ser apresentados anúncios regulares. Se a cidade corresponder a uma cidade não ambígua, então, o gerador de informações geográficas 830 pode determinar se o nome da cidade com um ou mais outros termos de busca da questão aparece numa lista negra. Uma lista negra pode ser mantida para nomes de cidades não

ambíguas que, quando combinados com uma ou mais palavras, signifiquem algo diferente das respectivas cidades. Se o nome da cidade com um ou mais outros termos de busca da questão aparecer na lista negra, então, podem ser apresentados anúncios regulares. Se o nome da cidade com um ou mais outros termos de busca da questão não aparecer na lista negra, então podem ser apresentados anúncios locais com base na referência geográfica da questão.

Se deverem ser apresentados anúncios locais (“alvejados”), podem ser usadas informações relativas à localização do usuário (por exemplo, o endereço de IP do usuário) pelo gerador de informações geográficas 830 para determinar se a localização está dentro de uma distância predeterminada da localização correspondente à referência geográfica. Se a localização do usuário estiver dentro da distância predeterminada, então, podem ser apresentados anúncios locais. Se a posição do usuário estiver fora da distância predeterminada, porém, então, podem ser apresentados anúncios regulares.

O servidor de *front end* 440 (ou componentes do servidor de *front end* 440) pode realizar algumas tarefas que podem ser específicas para a China. Por exemplo, o servidor de *front end* 440 pode esconder direções de direção, fornecer uma unidade de exibição (quilômetro contra milha) que possa ser dependente do país, realizar filtração para palavras-chaves sensíveis, quando o usuário estiver na China, fornecer formatação específica de endereços e números de telefone chineses para exibição, mostrar a localização geocodificada na parte superior do mapa e/ou arredondar a distância para 0,5 quilômetros em vez de 0,1 quilômetros e remover a direção.

Processo Exemplificativo

As Figuras 10A-10D são fluxogramas de um processo exemplificativo para identificar os resultados da busca local e fornecer um mapa associado a posições identificadas.

Resultado de Busca Local

e Processo de Geração de Mapas Exemplificativos

Conforme mostrado na Figura 10A, um processo 1000 para identificar os resultados da busca local e fornecer um mapa associado com posições identificadas pode começar com o recibo das páginas amarelas, o provedor de mapas e dados de documentos (bloco 1005). Por exemplo, numa modalidade descrita acima com relação à Figura 4, os dados do provedor de mapas 400, os dados das páginas amarelas 405 e os dados de documentos da Web 410 podem ser fornecidos para o indexador 435 e os dados de provedores de mapas 400 podem ainda ser usados para derivar informações de ID de POI/endereços e informações de mapeamento de FP de endereços/ID de POI 425.

O processo 1000 pode realizar a geocodificação sobre os dados (bloco 1010). Por exemplo, numa implementação descrita acima com relação à Figura 6, o gerador de listas de endereços 600 do geocodificador 430 pode receber dados de provedores de mapas 400 e dados das páginas amarelas 405 e pode gerar uma lista de endereços com base nas províncias, cidades, distritos, nomes de rua, endereços, nomes de POI ou combinações dos mesmos a partir de dados de provedores de mapas 400 e/ou dados das páginas amarelas 405. O analisador 610 do geocodificador 430 pode tentar analisar a cidade e/ou distrito a partir do início de cada endereço. O localizador de prefixos de coincidência mais longos 620 do geocodificador 430 pode ser usado para localizar partes adicionais de cada endereço. O localizador de combinações 630 do geocodificador pode localizar combinações particulares de forma que os nomes de POI possam ser localizados primeiro, então vários níveis dos endereços possam ser localizados e, finalmente, possam ser localizados distritos ou cidades. A unidade de geocodificação da questão 640 do geocodificador 430 pode computar a pontuação do resultado de uma busca com base no número de pontos na localização do resultado de busca.

Como melhor mostrado na Figura 10A, o processo 1000

pode realizar a indexação sobre os dados (bloco 1015). Por exemplo, numa implementação descrita acima com relação à Figura 7, o indexador de dados de mapa 700 do indexador 435 pode indexar dados de endereço a partir dos dados de provedores de mapas 400, o que pode

5 incluir informações de posições de mapas. O pré-processador de dados das páginas amarelas 710 do indexador 435 pode receber dados das páginas amarelas 405 a partir dos diferentes provedores e pode pré-processar dados das páginas amarelas 405 para um formato comum. O extrator de informações de negócios 720 do indexador 435 pode extrair

10 informações de negócios (por exemplo, endereços e números de telefone) a partir de dados de documentos da Web 410 usando uma variedade de técnicas. As técnicas podem ser modificadas com base na formatação diferente de endereços e números de telefone na China e no Japão. As informações de negócios extraídas podem ser fornecidas no repositório

15 de informações de negócios 750. O planificador de distâncias 730 do indexador 435 pode configurar um raio de busca para uma questão de busca local. O agrupador 740 do indexador 435 pode agrupar dados de provedores de mapas 400 e dados das páginas amarelas 405 em conjunto. O destacador de fragmentos 760 do indexador 435 pode destacar

20 fragmentos em documentos (por exemplo, documentos da Web).

O processo 1000 pode gerar os resultados da busca local e pode fornecer um URL de mapas (bloco 1020). Por exemplo, numa implementação descrita acima com relação à Figura 8, o gerador de busca local 810 do servidor de *front end* 440 pode gerar resultados de

25 busca local 450 (por exemplo, páginas relevantes da Web e/ou listagens de negócios dentro de uma área geográfica específica, com base nas informações de endereços obtidas e processadas pelo sistema de busca local 225). O gerador de mapas 820 do servidor de *front end* 440 pode gerar um mapa relativo aos resultados da busca local 450. Por exemplo,

30 para a fim de exibir o mapa para os resultados, o gerador de mapas 820 pode construir uma moldura *i* (isto é, uma moldura flutuante dentro de uma página da Web) sobre a página dos resultados de busca e pode

executar uma ação posterior para um URL fornecido pelo provedor de mapas.

Processo de Geocodificação Exemplificativo

O bloco do processo 1010 (Figura 10A) do processo 1000
 5 pode incluir os blocos mostrados na Figura 10B. Deste modo, o bloco do processo 1010 pode gerar uma lista de endereços (bloco 1025). Por exemplo, numa modalidade descrita acima com relação à Figura 6, o gerador de listas de endereços 600 do geocodificador 430 pode receber dados dos provedores de mapas 400 e dados das páginas amarelas 405
 10 e pode gerar uma lista de endereços (fichas) com base nas províncias, cidades, distritos, nomes de rua, endereços, nomes de POI ou combinações dos mesmos a partir de dados dos provedores de mapas 400 e/ou dados das páginas amarelas 405. O bloco do processo 1010 pode analisar cada endereço na lista de endereços (bloco 1030). Por exemplo,
 15 numa implementação descrita acima com relação à Figura 6, para cada endereço, o analisador 610 do geocodificador 430 pode tentar analisar a cidade e o distrito do endereço.

Como ainda mostrado na Figura 10B, o bloco do processo 1010 pode localizar os prefixos de coincidência mais longos a partir de
 20 cada endereço na lista de endereços (bloco 1035). Por exemplo, numa modalidade descrita acima com relação à Figura 6, pode ser usado o localizador de prefixos de coincidência mais longos 620 do geocodificador 430 para localizar partes adicionais do endereço (por exemplo, uma cidade, uma aldeia, uma estrada ou rua, número, nomes de POI etc.).
 25 Se os prefixos mais longos de comparação não são localizados no endereço (bloco 1040 - NO), então, o bloco do processo 1010 pode continuar a localizar partes adicionais do endereço (bloco 1035).

Se os prefixos mais longos de comparação são localizados no endereço (bloco 1040 - SIM), então, o bloco do processo 1010 pode
 30 localizar combinações em cada endereço (bloco 1045). Por exemplo, numa implementação descrita acima com relação à Figura 6, o localiza-

dor de combinações 630 do geocodificador 430 pode localizar combinações particulares de forma que possam ser localizados nomes de POI primeiro, depois, possam ser localizados vários níveis dos endereços e, finalmente, possam ser localizados distritos ou cidades. Essa disposição de um localizar pode garantir ou verificar que o endereço mais específico possível pode ser obtido pelo localizador de combinações 630.

Processo de Indexação Exemplicativo

O bloco do processo 1015 (Figura 10A) do processo 1000 pode incluir os blocos mostrados na Figura 10C. Deste modo, o bloco do processo 1015 pode pré-processar os dados das páginas amarelas (bloco 1050). Por exemplo, numa implementação descrita acima com relação à Figura 7, o pré-processador de dados das páginas amarelas 710 do indexador 435 pode receber dados das páginas amarelas 405 a partir dos diferentes provedores e pode pré-processar dados das páginas de amarelas 405 para um formato comum (ou predeterminado). Estes dados formatados podem ser fornecidos para o geocodificador 430 durante a indexação e o geocodificador 430 pode tentar geocodificar o endereço. Se o endereço puder ser geocodificado ao nível de edifício ou rua, o endereço pode ser indexado como uma entrada normal. Se o endereço puder ser geocodificado ao nível de cidade ou distrito, o endereço pode ser indexado como uma entrada com uma posição aproximada.

O bloco do processo 1015 pode extrair informações de negócios a partir dos documentos (bloco 1055) e pode armazenar as informações de negócios (bloco 1060). Por exemplo, numa implementação descrita acima com relação à Figura 7, o extrator de informações de negócios 720 do indexador 435 pode extrair informações de negócios (por exemplo, endereços e números de telefone) a partir de dados de documentos da Web 410 usando uma variedade de técnicas. As técnicas podem ser modificadas com base na formatação diferente de endereços e números de telefone na China e no Japão. Noutra implementa-

ção descrita acima com relação à Figura 7, as informações de negócios extraídas podem ser providas no repositório de informações de negócios 750. O repositório de informações de negócios 750 pode incluir uma variedade de informações, por exemplo, os documentos a partir dos
5 quais foram extraídas as informações de negócios pelo extrator de informações de negócios 720.

O bloco do processo 1015 pode indexar dados de endereço recebidos de um provedor de mapas (bloco 1065). Por exemplo, numa modalidade descrita acima com relação à Figura 7, o indexador de
10 dados de mapa 700 do indexador 435 pode indexar dados de endereço a partir dos dados de provedores de mapas 400, que podem incluir informações de posições de mapas. O geocodificador 430 pode tentar geocodificar o endereço novamente com base nos dados de endereços indexados e o agrupador 740 pode configurar a posição do grupo, se a
15 posição do grupo for diferente da posição dada, mas pode não configurar a posição do grupo acima de um limite de distância predeterminada (por exemplo, dentro três quilômetros).

Como ainda mostrado na Figura 10C, o bloco do processo 1015 pode agrupar dados de páginas amarelas e dados de provedores
20 de mapas (bloco 1070). Por exemplo, numa modalidade descrita acima com relação à Figura 7, o agrupador 740 do indexador 435 pode agrupar dados de provedores de mapas 400 e dados das páginas amarelas 405 em conjunto. Se estas entradas são agrupadas em conjunto pelo agrupador 740, então o servidor de *front end* 440 pode ser capaz de
25 prover o usuário com informações detalhadas de endereço e de número de telefone, assim como também uma posição precisa sobre um mapa.

O bloco do processo 1015 pode destacar fragmentos fornecidos em documentos (bloco 1075). Por exemplo, numa implementação descrita acima com relação à Figura 7, o destacador de fragmentos 760
30 do indexador 435 pode destacar fragmentos em documentos (por exemplo, documentos da Web). O destacador de fragmentos 760 pode

armazenar deslocamentos de baites em vez de deslocamentos de termos para identificar um endereço ou número de telefone (ou algumas outras informações de negócios) num documento da Web durante a indexação. O destacador de fragmentos 760 pode usar o deslocamento de baites para realizar os destaques e nenhuma segmentação pode ser exigida.

Como ainda mostrado na Figura 10C, o bloco do processo 1015 pode configurar uma distância de busca (bloco 1080). Por exemplo, numa modalidade descrita acima com relação à Figura 7, o planificador de distâncias 730 do indexador 435 pode configurar um raio de busca para uma questão de busca local geocodificada pelo geocodificador 430 para uma localização particular. Cada localização pode ser uma posição de ponto (por exemplo, edifícios, lugares turísticos famosos, escolas, centros das ruas etc.) ou uma localização limitada (por exemplo, distritos, cidades, províncias etc.). Para localizações de ponto, o planificador de distâncias 730 pode configurar o raio de busca para uma distância predeterminada (por exemplo, aproximadamente cinco quilômetros em torno do ponto). Para localizações limitadas, o planificador de distâncias 730 pode configurar o raio de busca para aproximadamente a distância máxima a partir de um centróide da localização para os cantos.

Processo Exemplificativo

do Servidor de *Front End*

O bloco de processo 1020 (Figura 10A) do processo 1000 pode incluir os blocos mostrados na Figura 10D. Deste modo, o bloco do processo 1020 pode regravar uma questão de busca (bloco 1085). Por exemplo, numa implementação descrita acima com relação à Figura 8, pode haver vários sinônimos que se referem ao mesmo termo em chinês. Além disso, os diferentes provedores de dados das páginas amarelas 400 podem usar diferentes nomes para representar a mesma categoria. A fim de proporcionar a maior chamada para a questão de busca introduzida pelo usuário, o regravador de questão 800 do servi-

dor de *front end* 440 pode gerar uma lista de sinônimos para cada uma das categorias. O regravador de questão 800 pode também regravar cada questão de busca local que pode expandir a questão para um par de termos de busca que podem ser juntos por um operador "OU".

5 O bloco do processo 1020 pode gerar os resultados da busca local com base na questão de busca e pode gerar um mapa que mostra localizações dos resultados de busca (bloco 1090). Por exemplo, numa implementação descrita acima com relação à Figura 8, o gerador de busca local 810 do servidor de *front end* 440 pode gerar os resulta-
 10 dos da busca local (por exemplo, páginas da Web e/ou listagens de negócios relevantes dentro de uma área geográfica específica, com base nas informações de endereço obtidas e processadas pelo sistema de busca local 225). Noutra modalidade descrita em relação com a Figura 8, o gerador de mapas 820 do servidor de *front end* 440 pode gerar um
 15 mapa relativo aos resultados da busca local 450. Por exemplo, a fim de exibir o mapa para os resultados, o gerador de mapas 820 pode construir uma moldura i (isto é, uma moldura flutuante dentro de uma página da Web) sobre a página de resultados de busca e pode executar uma ação posterior para um URL fornecido pelo provedor de mapas. O
 20 ID de POI pode ser usado pelo gerador de mapas 820 para identificar uma posição sobre o mapa (por exemplo, mapa 900) para o provedor de mapas de forma que o provedor de mapas pode mostrar a posição no mapa fornecido dentro da página de resultados.

Como ainda mostrado na Figura 10D, o bloco do processo
 25 1020 pode gerar informações geográficas para um servidor de anúncios (bloco 1095). Por exemplo, numa modalidade descrita acima com relação à Figura 8, a fim de exibir anúncios locais para a China, o gerador de informações geográficas 830 do servidor de *front end* 440 podem fornecer informações geográficas para o servidor de anúncios
 30 840. Por exemplo, o gerador de informações geográficas 830 pode enviar o código da região para o servidor de anúncios 840 como um geocódigo de região. Noutro exemplo, o gerador de informações geográ-

ficas 830 pode concatenar palavras-chaves com as posições introduzidas pela entrada da questão de busca do usuário e pode usar estas concatenações como as palavras-chaves enviadas para o servidor de anúncios 840.

5

Conclusão

As implementações aqui descritas podem proporcionar sistemas e métodos para identificar os resultados da busca local e gerar um mapa associado com as localizações identificadas. O sistema pode receber uma entrada de questão de busca local pelo usuário e pode
10 identificar uma posição associada à questão da busca local. O sistema pode identificar um conjunto de resultados da busca local que pode estar relacionado com a questão de busca local e pode estar associado à localização identificada. Os resultados da busca local podem incluir *links* para documentos que podem estar relacionados com a questão de
15 busca local. O sistema pode identificar um identificador para um grupo de resultados da busca local e pode fornecer o identificador para um provedor de mapas. O sistema pode receber um mapa associado com a posição identificada do provedor de mapas. O mapa pode identificar uma posição de pelo menos um resultado de busca no grupo de resul-
20 tados da busca local.

As modalidades descritas proporcionam uma ou mais das vantagens seguintes. Por exemplo, dados de mapas e dados das páginas amarelas podem ser utilizados a partir de diferentes provedores para identificar os resultados da busca local e gerar um mapa associado
25 com localizações identificadas. O mapa pode ser convenientemente exibido com os resultados da busca local. Essa disposição evita a geração de resultados da busca local e um ponteiro para um servidor do provedor de mapas de terceiros.

Noutro exemplo, o mapa pode proporcionar dados de mapa
30 detalhados com base nos dados das páginas amarelas. Isto pode tornar possível gerar um mapa que inclui dados de mapa detalhados em países

onde as restrições de exportação podem limitar a disponibilidade de dados de mapa detalhados para entregar o mapa ou podem limitar a disponibilidade da latitude e longitude reais de endereços dentro da área.

5 A descrição precedente de modalidades da presente invenção proporciona a ilustração e descrição, mas não se pretende que seja exaustiva nem que limite a invenção à forma precisa revelada. São possíveis modificações e variações levando em conta os ensinamentos acima ou podem ser adquiridas a partir da prática da invenção. Por
10 exemplo, embora tenham sido descritas séries de atos no que se relaciona com as Figuras 10A-10D, a ordem dos atos pode ser modificada em outras implementações consistentes com os princípios da invenção. Além disso, atos não dependentes podem ser realizados em paralelo. Numa implementação, o servidor 220 pode realizar a maioria, se não
15 todos, os atos descritos no que se relaciona com o processamento das Figuras 10A-10D. Noutra implementação, um ou mais ou todos os atos podem ser realizados por outra entidade, tal como outro servidor 230 e/ou 240 ou cliente 210.

Noutro exemplo, podem ser utilizadas abordagens alternati-
20 vas para endereços de geocodificação que não são providos pelo provedor de mapas. Numa abordagem alternativa, o geocodificador 430 pode tentar localizar o ponto mais próximo de um endereço a ser geocodificado. Por exemplo, suponha-se que os pontos "1 Rua ABC" e "10 Rua ABC" são identificados pelo provedor de mapas. Quando se tenta
25 geocodificar o endereço "3 Rua ABC", o geocodificador 430 pode retornar a localização "1 Rua ABC", que é o ponto mais próximo de "3 Rua ABC". Noutra abordagem alternativa, o geocodificador 430 pode tentar interpolar um ponto. Por exemplo, se "1 Rua ABC" estiver no índice de grade (0, 0) e "10 Rua ABC" estiver no índice de grade (10, 20), o geoco-
30 dificador 430 pode determinar que o endereço de "3 Rua ABC" está no índice de grade (3, 6), com base na interpolação.

Ainda noutro exemplo, podem ser utilizadas abordagens alternativas para obter o ID de POIs para desenhar o mapa. Numa abordagem alternativa, o ID de POI pode ser armazenado com os dados de localização a partir do provedor de mapas. Os IDs de POIs podem ser retornados pelo servidor de *front end* 440 durante o tempo serviço. Os IDs de POIs podem mudar em versões diferentes dos dados de provedores de mapas. Armazenar o ID de POI no índice torna o índice dependente dos dados a partir do provedor de mapas. Noutra abordagem alternativa, os endereços dos resultados de busca podem ser geocodificados durante o tempo de serviço e pode ser pedido que o geocodificador 430 forneça os pontos de comparação mais próximos. O geocodificador 430 pode retornar os IDs de POIs dos pontos. Os pedidos dos pontos de comparação mais próximos podem ser enviados em lotes (por exemplo, lotes de dez) de pedidos de geocodificação de forma que o impacto de desempenho pode ser pequeno.

Num exemplo adicional, se bem que os sistemas e métodos fossem descritos em termos de uma busca local chinesa, noutra implementação, algumas das técnicas aqui descritas podem aplicar-se igualmente a buscas locais em outros países.

Será evidente para uma pessoa de capacidade ordinária na técnica que aspectos da invenção, como descrita acima, podem ser implementados de muitas maneiras diferentes de software, *firmware* e hardware nas modalidades ilustradas nas Figuras. O código de software real ou hardware de controle especializado usado para implementar aspectos consistentes com os princípios da invenção não são limitativos da invenção. Deste modo, a operação e o comportamento dos aspectos foram descritos sem referência ao código de software específico - ficando entendido que uma pessoa de capacidade ordinária na técnica seria capaz de projetar software e hardware de controle para implementar os aspectos com base na presente descrição.

Nenhum elemento, ato ou instrução usada no presente

pedido deve ser interpretado como crítico ou essencial para a invenção, a menos que explicitamente descrito como tal. Também, conforme aqui usado, pretende-se que o artigo “um” inclua um ou mais itens. Nos casos em que apenas um item é pretendido, é usado o termo “um” ou
5 linguagem semelhante. Além disso, pretende-se que a frase “com base em” signifique “com base, pelo menos em parte, em”, a menos que explicitamente declarado de outra forma.

“Métodos de Buscas Locais, de Geocodificação e de Indexação e Sistemas de Buscas Locais”

Reivindicações

1 - Método de Buscas Locais, caracterizado por que compreende:

5 receber dados das páginas amarelas, dados do provedor de mapas de terceiros e dados de documento em resposta a uma questão de busca local;

 geocodificar pelo menos um dos dados das páginas amarelas, os dados do provedor de mapas de terceiros e os dados de documento para atribuir um identificador geográfico e combinar pelo menos
10 um endereço dentro da questão de busca local;

 indexar os dados geocodificados para identificar informações de negócios e informações de localização correspondendo à questão de busca local; e

15 proporcionar os resultados da busca local e um mapa de terceiros com base nas informações de negócios e informações de localização identificadas.

2 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que os dados das páginas amarelas compreendem
20 pelo menos uma de informações de endereço para pontos de interesse (POIs), informações de número de telefone para POIs, um nome de negócios, um endereço de e-mail, um número de facsímile, um endereço de site da Rede, um nome de CEO, uma descrição de negócios, uma categoria de estilo SIC ou um tipo de negócios.

25 **3 - Método de Buscas Locais**, de acordo com a Reivindicação 1, **caracterizado** por que os dados do provedor de mapas de terceiros compreendem um ponto de interesse (POI) e um endereço do POI.

4 - Método de Buscas Locais, de acordo com a Reivindicação 3, **caracterizado** por que os dados do provedor de mapas de terceiros
30 compreendem ainda uma grade para o POL.

5 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que os dados do provedor de mapas de terceiros compreendem pelo menos um dentre:

um ponto normal de interesse (POI) que inclui um nome de negócios, um endereço, um número de telefone e um índice de grade,

uma estrada POI que inclui um nome de rua e um índice de grade de um centro da rua ou

um código postal POI que inclui um código postal e um índice de grade de um centro aproximado do código postal.

6 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que a geocodificação compreende:

gerar uma lista de endereços com base na questão de busca local;

analisar cada endereço na lista de endereços;

localizar os prefixos de coincidência mais longos em cada endereço para identificar partes de cada endereço; e

localizar combinações em cada endereço para verificar cada endereço.

7 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que a indexação compreende:

pré-processar os dados das páginas amarelas para um formato predeterminado;

extrair as informações de negócios a partir dos dados de documento;

indexar as informações de endereço a partir dos dados do provedor de mapas de terceiros; e

agrupar os dados das páginas amarelas e os dados do provedor de mapas de terceiros.

8 - Método de Buscas Locais, de acordo com a Reivindicação 7,

caracterizado por que a indexação compreende ainda:

- armazenar as informações de negócios num repositório;
 - destacar fragmentos nos dados de documento; e
 - configurar uma distância de busca para a questão da busca
- 5 local.

9 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que o fornecimento compreende:

gerar uma lista de sinônimos relacionados com a questão da busca local;

10 reescrever a questão de busca local para expandir a questão;

gerar os resultados da busca local com base na questão expandida;

15 obter o mapa de terceiros com base nos resultados da busca local; e

gerar anúncios com base em informações geográficas relacionadas com os resultados da busca local.

10 - Método de Buscas Locais, de acordo com a Reivindicação 1, **caracterizado** por que as informações de localização são determinadas a partir dos dados do provedor de mapas de terceiros e os dados das

20 páginas amarelas.

11 - Método de Geocodificação, com base numa questão de busca local, **caracterizado** por que compreende:

25 receber dados do provedor de mapas de terceiros e dados das páginas amarelas;

gerar um endereço com base na questão de busca local;

analisar o endereço;

localizar prefixos de coincidência mais longos no endereço para identificar pelo menos uma parte do endereço; e

localizar uma combinação no endereço para verificar o endereço.

12 - Método de Geocodificação, de acordo com a Reivindicação 11, **caracterizado** por que compreende ainda:

5 gerar o endereço com base em informações incluídas nos dados do provedor de mapas de terceiros ou nos dados das páginas amarelas.

13 - Método de Geocodificação, de acordo com a Reivindicação 11, **caracterizado** por que compreende ainda:

10 analisar uma cidade e um distrito do endereço.

14 - Método de Geocodificação, de acordo com a Reivindicação 11, **caracterizado** por que a parte do endereço compreende pelo menos um dentre uma cidade, uma aldeia, uma estrada, uma rua, um número ou um nome de um ponto de interesse.

15 **15 - Método de Geocodificação**, de acordo com a Reivindicação 11, **caracterizado** por que compreende ainda:

 localizar a combinação no endereço de forma que um nome de um ponto de interesse seja localizado em primeiro lugar, outras partes do endereço sejam localizadas em segundo lugar e um distrito ou
20 uma cidade seja localizada em terceiro lugar.

16 - Método de Geocodificação, de acordo com a Reivindicação 11, **caracterizado** por que compreende ainda:

 indexar o endereço como uma posição exata, se o endereço incluir um edifício ou uma parte de rua.

25 **17 - Método de Geocodificação**, de acordo com a Reivindicação 11, **caracterizado** por que compreende ainda:

 Indexar o endereço para uma posição aproximada, se o endereço incluir uma parte da cidade ou uma parte de distrito.

18 - Método de Indexação, com base numa questão de busca local,

caracterizado por que compreende:

pré-processar dados das páginas amarelas para um determinado formato;

extrair informações de negócios a partir de dados de documento;

armazenar as informações de negócios num repositório;

indexar as informações de endereço a partir de dados do provedor de mapas de terceiros;

agrupar os dados das páginas amarelas e os dados do provedor de mapas de terceiros;

destacar fragmentos nos dados de documento;

para uma localização de ponto, configurar uma distância de busca para a questão de busca local para uma distância predeterminada; e

para uma localização limitada, configurar a distância de busca para aproximadamente uma distância máxima a partir de um centróide da localização limitada aos cantos da localização limitada.

19 – Método de Indexação, de acordo com a Reivindicação 18, **caracterizado** por que compreende ainda:

pré-processar os dados das páginas amarelas para um formato predeterminado.

20 – Método de Indexação, de acordo com a Reivindicação 18, **caracterizado** por que as informações de negócios compreendem pelo menos um de um endereço de negócios ou um número de telefone de negócios.

21 – Método de Indexação, de acordo com a Reivindicação 20, **caracterizado** por que compreende ainda:

modificar a extração das informações de negócios com base em diferenças de formato do endereço de negócios ou o número de telefone de negócios.

22 – Método de Indexação, de acordo com a Reivindicação 18, **caracterizado** por que compreende ainda:

armazenar documentos associados com as informações de negócios no repositório.

5 **23 – Método de Indexação**, de acordo com a Reivindicação 18, **caracterizado** por que compreende ainda:

geocodificar um endereço com base nas informações de endereço indexadas; e

determinar uma nova posição de agrupamento com base no
10 endereço geocodificado, se a nova posição de agrupamento for diferente de uma posição de agrupamento prévia.

24 – Método de Indexação, de acordo com a Reivindicação 18, **caracterizado** por que compreende ainda:

destacar os fragmentos nos dados de documento com base
15 em desvios de bautes que são armazenados para identificar as informações de negócios nos dados de documento.

25 - Método de Buscas Locais, caracterizado por que compreende:

gerar uma lista de sinônimos relacionados com uma questão de busca local;

20 reescrever a questão de busca local de modo a expandir a questão;

gerar resultados de busca local com base na questão expandida;

obter um mapa a partir de um provedor de mapas de
25 terceiros com base nos resultados da busca local; e

gerar anúncios com base em informações geográficas relacionadas com os resultados da busca local.

26 - Método de Buscas Locais, de acordo com a Reivindicação 25, **caracterizado** por que compreende ainda:

reescrever a questão de busca local por junção de termos de busca com um operador “OU”.

27 - Método de Buscas Locais, de acordo com a Reivindicação 25, **caracterizado** por que obter o mapa compreende:

5 construir uma moldura de mapa com os resultados da busca local;

 executar uma ação de posto para um URL fornecido pelo provedor de mapas de terceiros; e

10 identificar uma posição no mapa de maneira a ativar o provedor de mapas de terceiros para mostrar a posição sobre o mapa.

28 - Método de Buscas Locais, de acordo com a Reivindicação 25, **caracterizado** por que compreende ainda:

 gerar os anúncios com base num código de região relacionado com os resultados da busca local.

15 **29 - Método de Buscas Locais**, de acordo com a Reivindicação 25, **caracterizado** por que compreende ainda:

 concatenar uma palavra-chave com uma posição relacionada com a questão de busca local; e

 gerar os anúncios com base na palavra-chave concatenada.

20 **30 - Sistema de Buscas Locais**, **caracterizado** por que compreende:

 um indexador para receber dados do provedor de mapas de terceiros, dados das páginas amarelas e dados de documento, pré-processar os dados das páginas amarelas para determinar informações de negócios, extrair informações de negócios a partir dos dados de documento, identificar as informações de localização numa questão de busca local e indexar dados de endereços dos dados do provedor de mapas de terceiros;

 um geocodificador para receber informações a partir do indexador e atribuir identificadores geográficos; e

um servidor de frente para receber informações a partir do geocodificador, reescrever a questão de busca local, obter um mapa a partir dos dados do provedor de mapas de terceiros e gerar os resultados da busca local com base na questão de busca local.

- 5 **31 - Sistema de Buscas Locais**, de acordo com a Reivindicação 30, **caracterizado** por que compreende ainda:

um repositório de índices/documentos que armazena pelo menos um dos dados de endereço fornecidos pelos dados do provedor de mapas de terceiros, informações de pontos de interesse, centros de
10 códigos postais ou cidades não incluídas nos dados do provedor de mapas de terceiros.

32 - Sistema de Buscas Locais, de acordo com a Reivindicação 30, **caracterizado** por que o indexador pré-processa os dados das páginas amarelas para um formato predeterminado.

- 15 **33 - Sistema de Buscas Locais**, de acordo com a Reivindicação 30, **caracterizado** por que o indexador fixa uma distância de busca para a questão de busca local.

34 - Sistema de Buscas Locais, de acordo com a Reivindicação 30, **caracterizado** por que o indexador agrupa os dados do provedor de
20 mapas de terceiros e os dados das páginas amarelas.

35 - Sistema de Buscas Locais, de acordo com a Reivindicação 30, **caracterizado** por que o indexador destaca fragmentos nos dados de documento.

36 - Sistema de Buscas Locais, de acordo com a Reivindicação 30,
25 **caracterizado** por que o geocodificador recebe os dados do provedor de mapas de terceiros e os dados das páginas amarelas e gera uma lista de endereços a partir dos dados do provedor de mapas de terceiros e os dados das páginas amarelas.

37 - Sistema de Buscas Locais, de acordo com a Reivindicação 30,
30 **caracterizado** por que o geocodificador localiza partes de um endereço.

38 - Sistema de Buscas Locais, de acordo com a Reivindicação 37, **caracterizado** por que o geocodificador localiza combinações dentro do endereço.

5 **39 - Sistema de Buscas Locais**, de acordo com a Reivindicação 30, **caracterizado** por que o geocodificador computa as pontuações de cada resultado de busca local com base em vários pontos numa localização de resultados de busca.

10 **40 - Sistema de Buscas Locais**, de acordo com a Reivindicação 30, **caracterizado** por que o servidor de frente gera uma lista de sinônimos relacionados com a questão de busca local.

41 - Sistema de Buscas Locais, de acordo com a Reivindicação 30, **caracterizado** por que o servidor de frente gera uma localização estimada e uma localização exata sobre o mapa para pelo menos um resultado de busca local.

15 **42 - Sistema de Buscas Locais**, de acordo com a Reivindicação 30, **caracterizado** por que o servidor de frente gera anúncios com base em informações geográficas relacionadas com os resultados da busca local.

43 - Sistema de Buscas Locais, caracterizado por que compreende:

20 meios para receber dados das páginas amarelas, dados do provedor de mapas de terceiros e dados de documentos em resposta a uma questão de busca local;

meios de geocodificação de pelo menos um dos dados das páginas amarelas, os dados do provedor de mapas de terceiros e os dados de documento para atribuir um identificador geográfico e combi-
25 nar pelo menos um endereço dentro da questão de busca local;

meios de indexação dos dados geocodificados para identificar informações de negócios e informações de localização correspondendo à questão de busca local; e

30 meios para proporcionar os resultados da busca local e um mapa de terceiros com base nas informações de negócios e informações

de localização identificadas.

44 - Sistema de Buscas Locais, caracterizado por que compreende:

uma memória para armazenar uma pluralidade de instruções; e

5 um processador para executar instruções na memória para:

identificar uma localização associada com uma questão de busca local,

identificar os resultados da busca local relevantes para a questão de busca local e associados com a posição identificada,
10 da,

identificar um identificador para cada um de um grupo dos resultados da busca local e

receber a partir de um provedor de mapas de terceiros um mapa associado com a posição identificada, em que o
15 mapa identifica uma posição de pelo menos um resultado de busca local no grupo de resultados da busca local.

45 - Método de Buscas Locais, caracterizado por que compreende:

receber uma questão de busca local;

identificar uma posição associada com a questão de busca
20 local;

identificar um conjunto de resultados de busca relevantes para a questão de busca local e associados com a localização identificada;

identificar um identificador para cada um de um grupo dos
25 resultados de busca;

proporcionar um identificador para cada um do grupo de resultados de busca para um provedor de mapas de terceiros; e

receber a partir do provedor de mapas de terceiros um mapa associado com a localização identificada, em que o mapa identifi-

ca uma posição de pelo menos um resultado de busca no grupo de resultados de busca.

46 - Método de Buscas Locais, caracterizado por que compreende:

gerar uma lista de fichas;

5 identificar um endereço potencial num documento da Rede;
 analisar o endereço potencial a partir de um início para
determinar se o endereço potencial inclui uma ficha associada com uma
cidade;

 analisar ainda o endereço potencial para determinar se o
10 endereço potencial inclui uma ficha associada com um distrito;

 identificar uma ficha de coincidência mais longa no endereço
potencial depois da ficha associada com a cidade ou a ficha associada
com o distrito; e

 determinar se o endereço potencial é um endereço real com
15 base na ficha associada com a cidade, a ficha associada com o distrito e
a ficha de coincidência mais longa identificada.

Figura 1

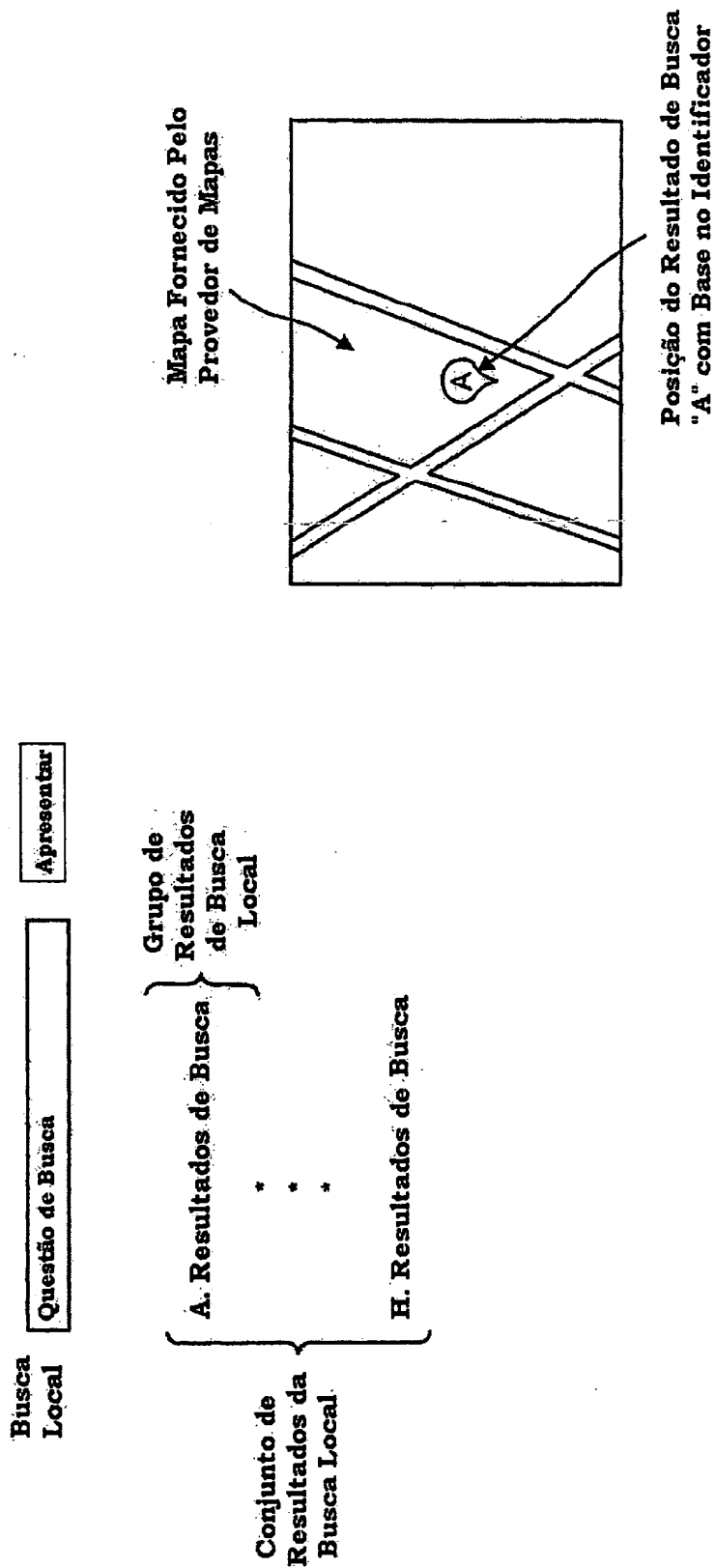
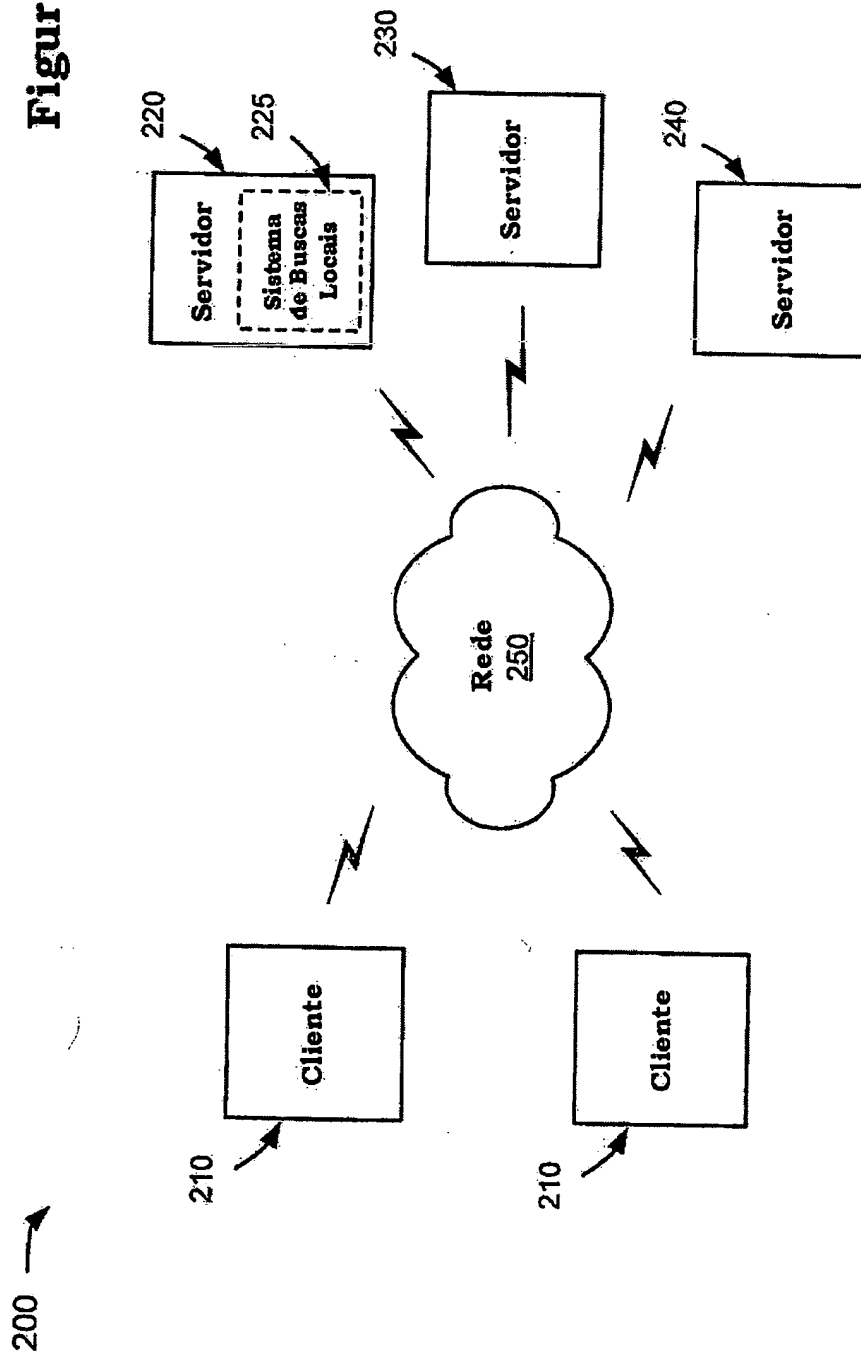
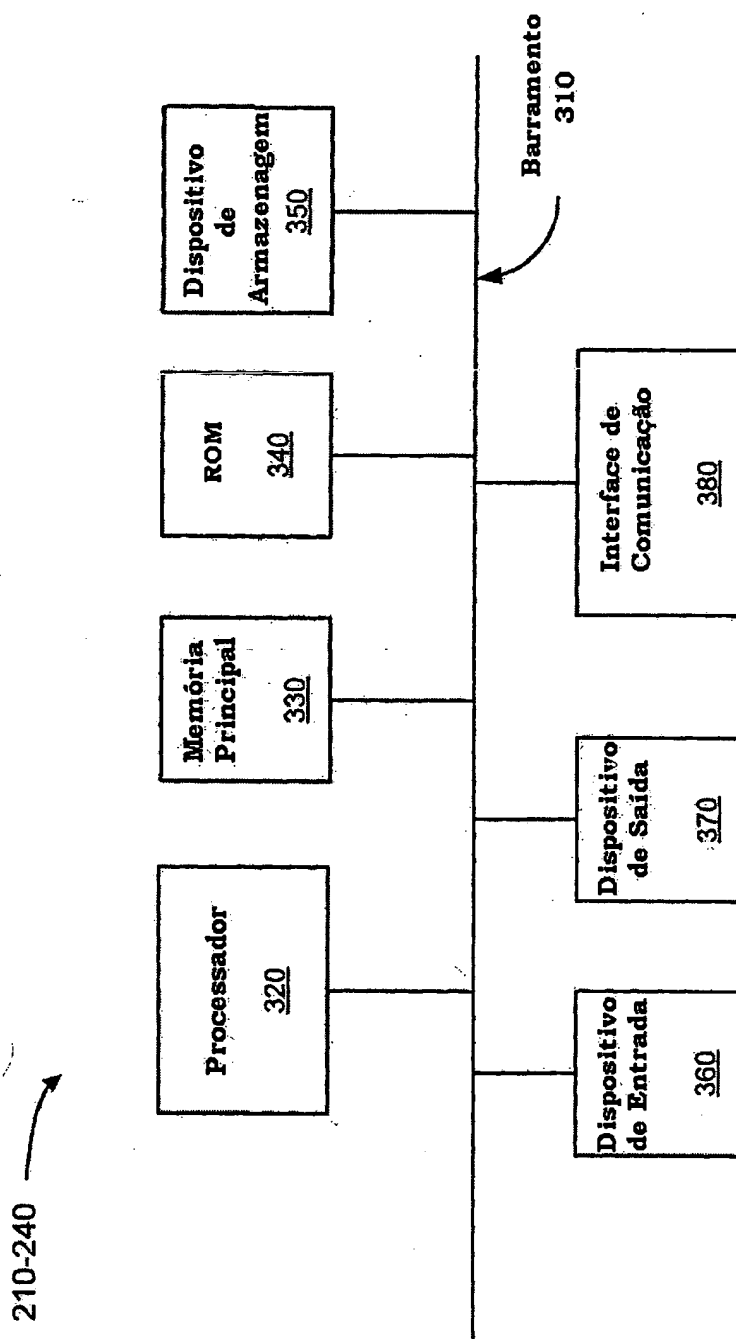


Figura 2

**Figura 3**

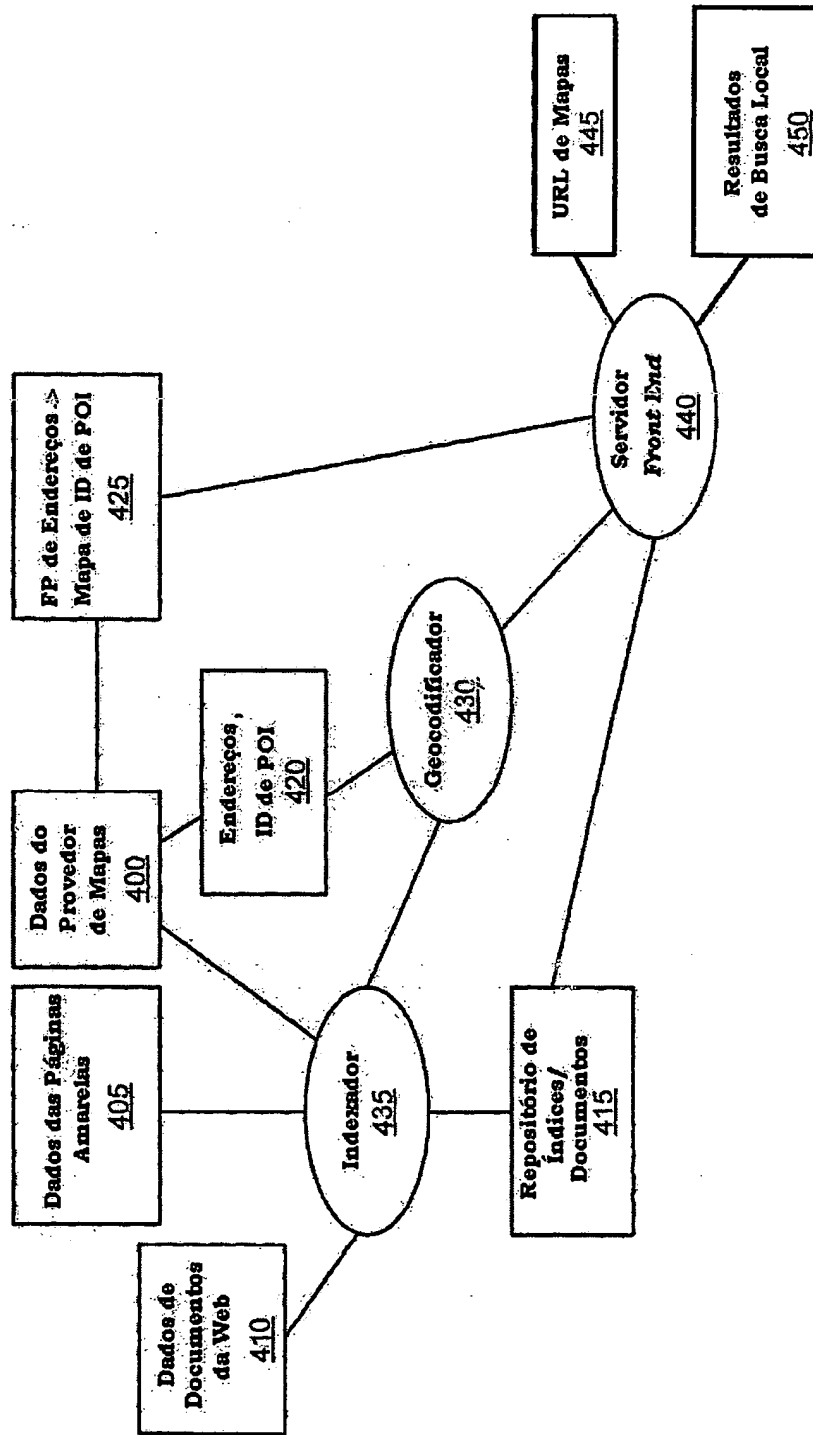


Figura 4

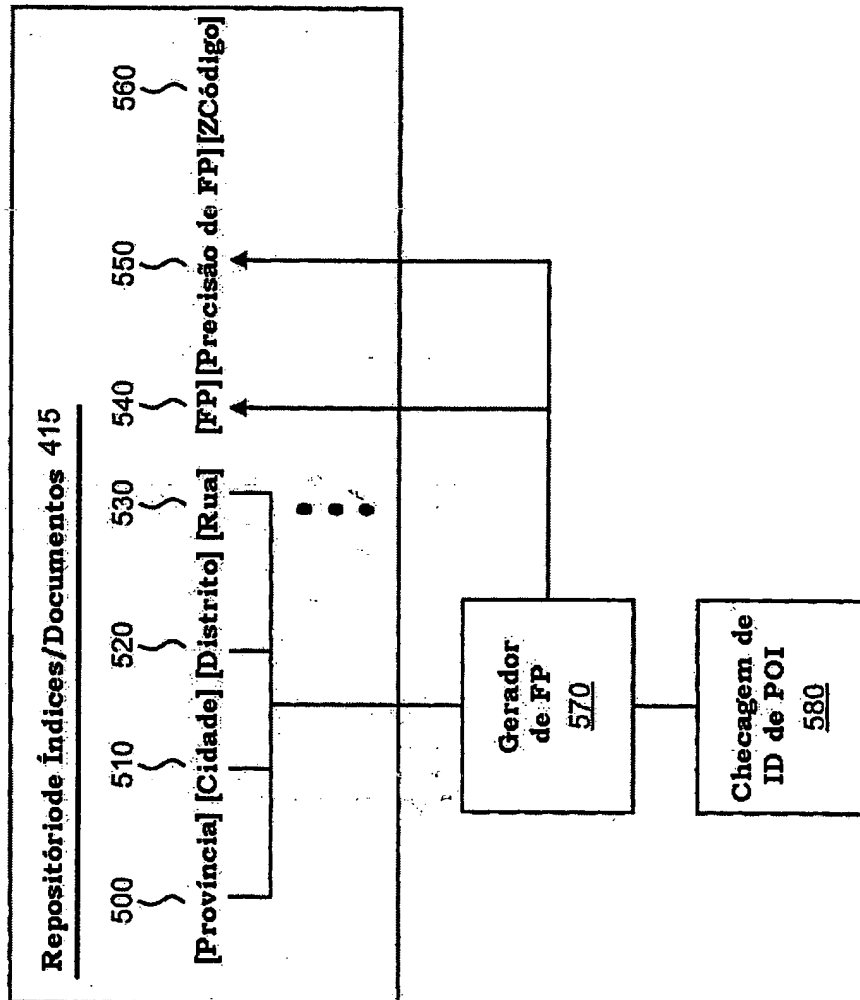


Figura 5

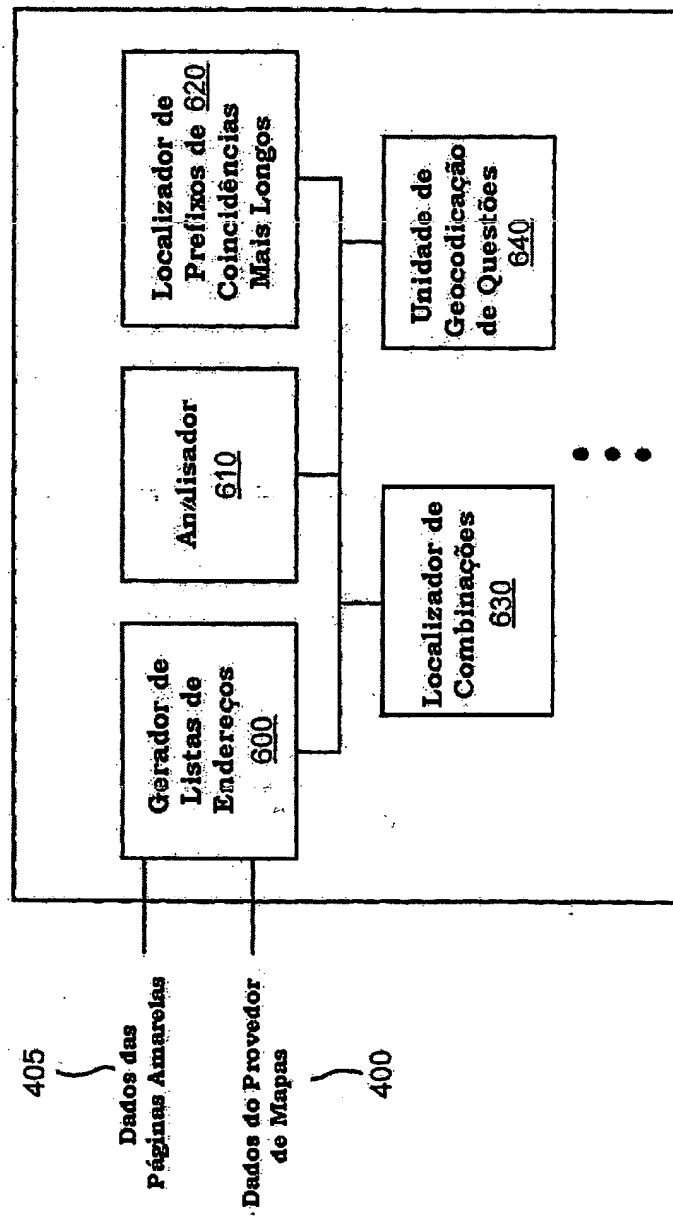
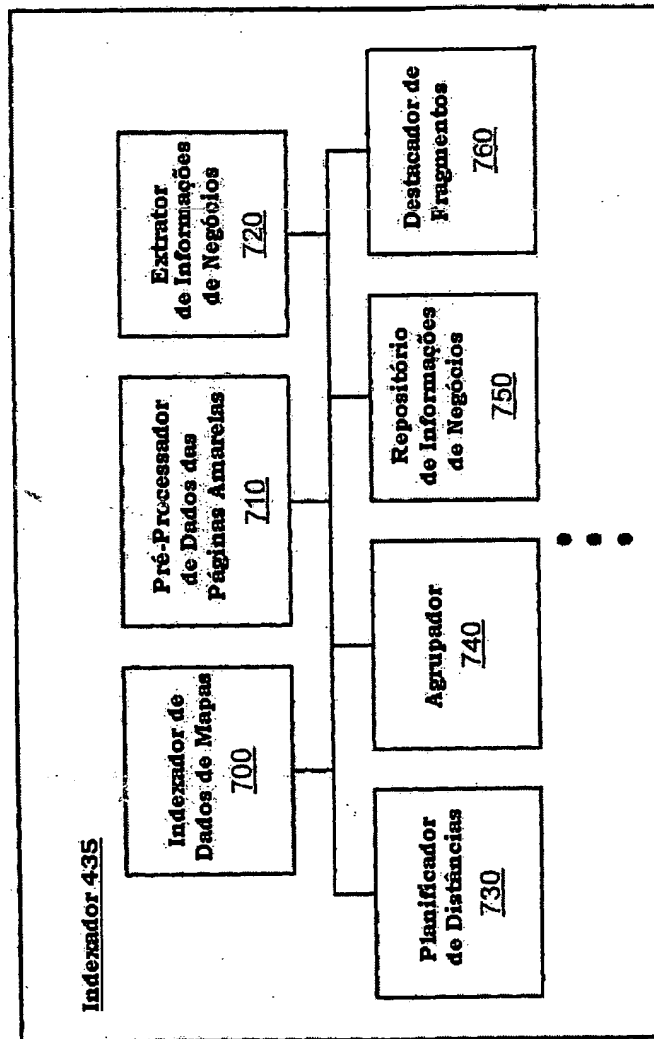
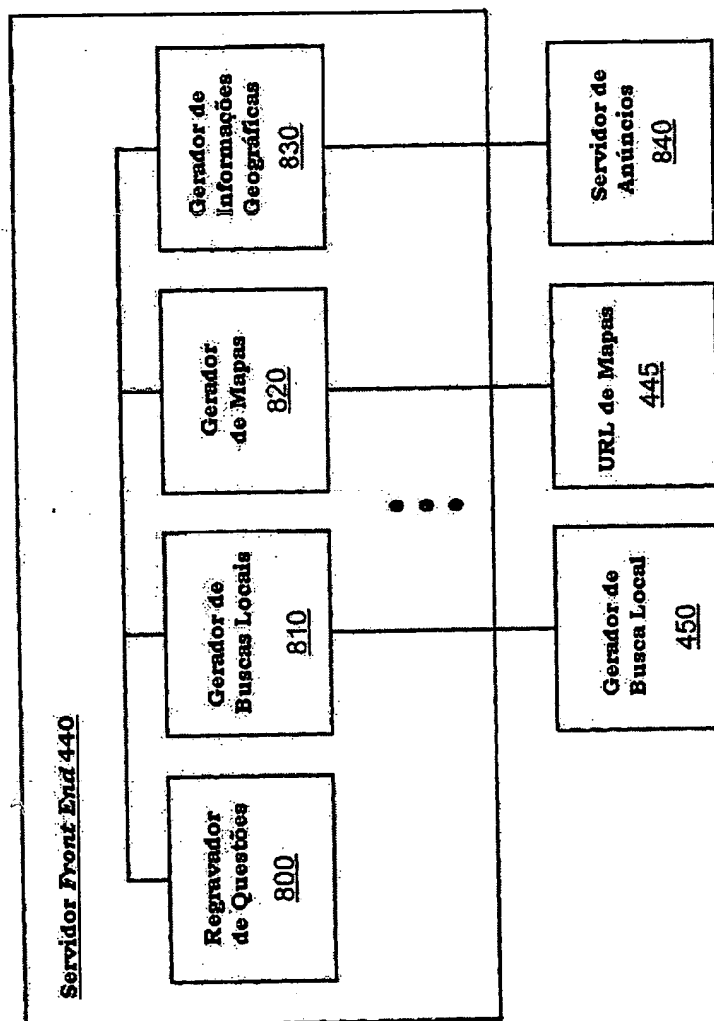
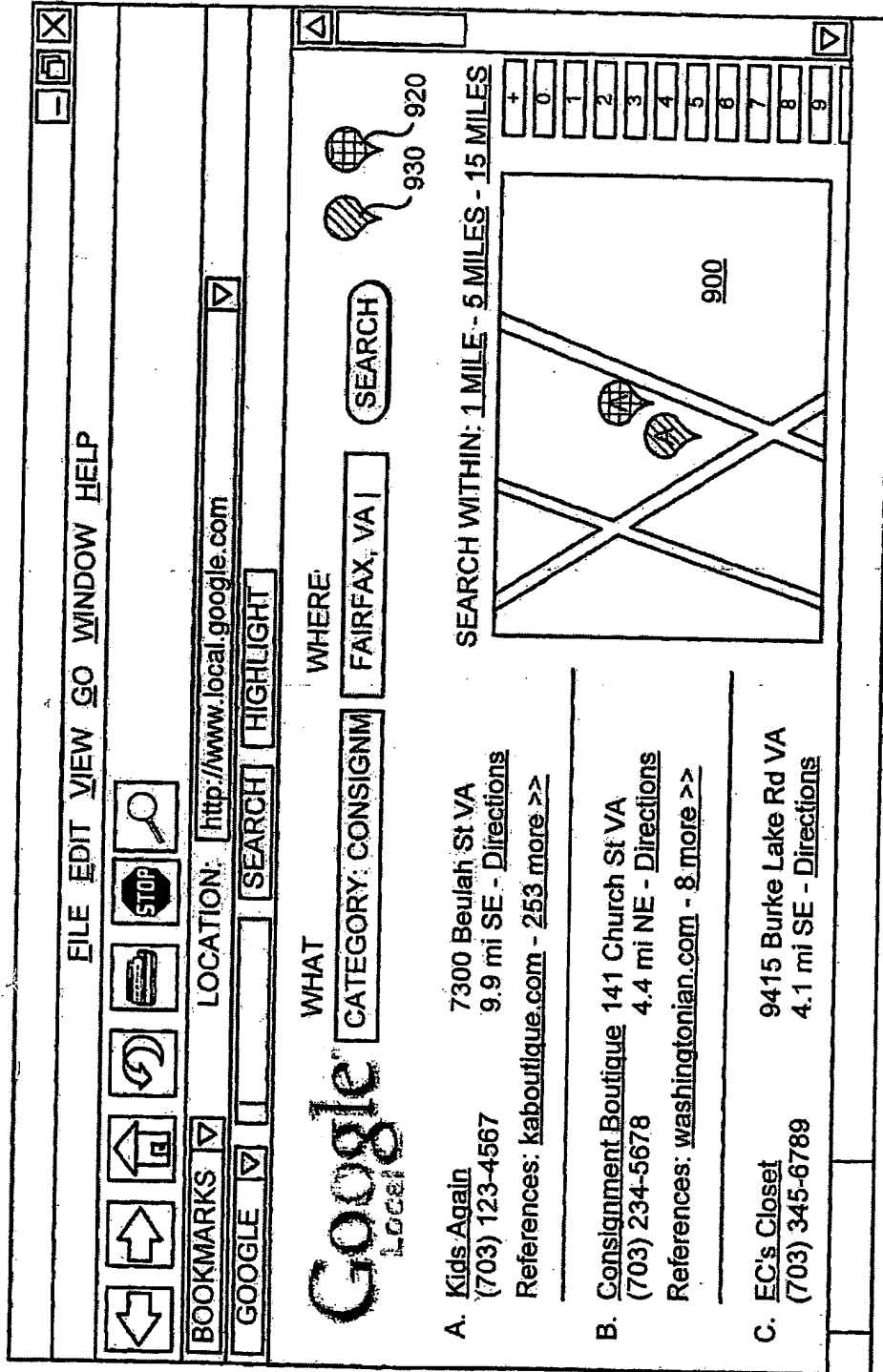


Figura 6

**Figura 7**





910

Figura 9

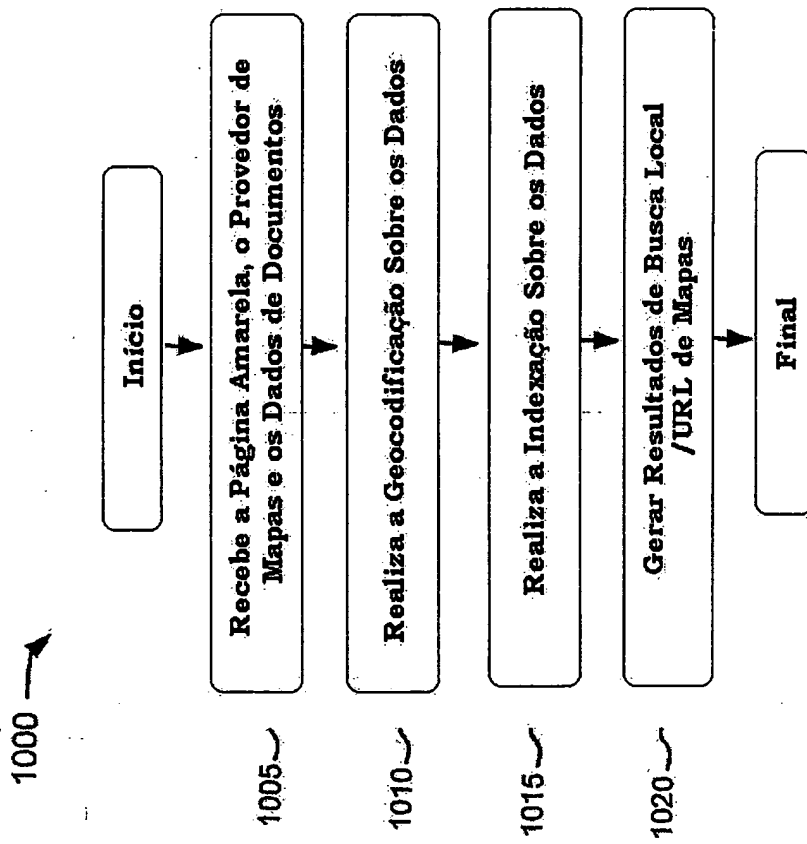
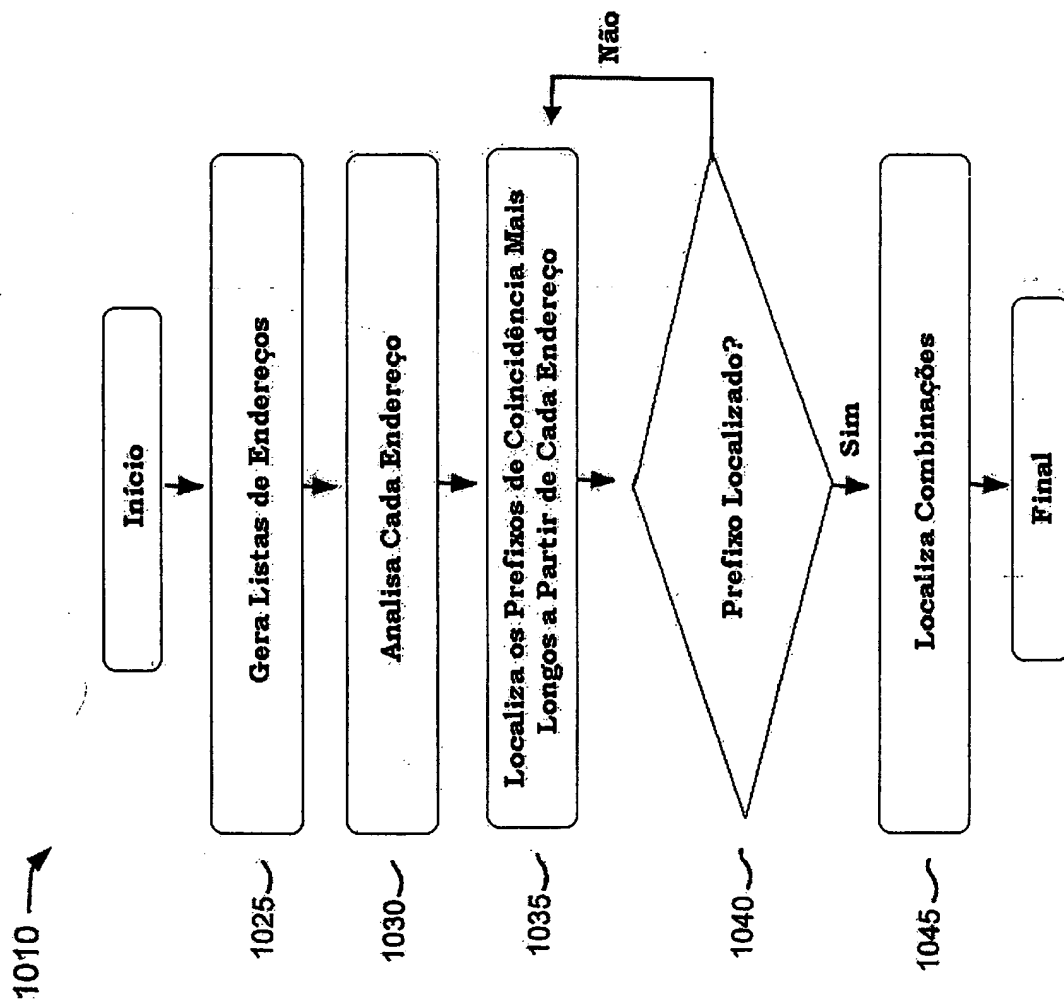


Figura 10A

Figura 10B



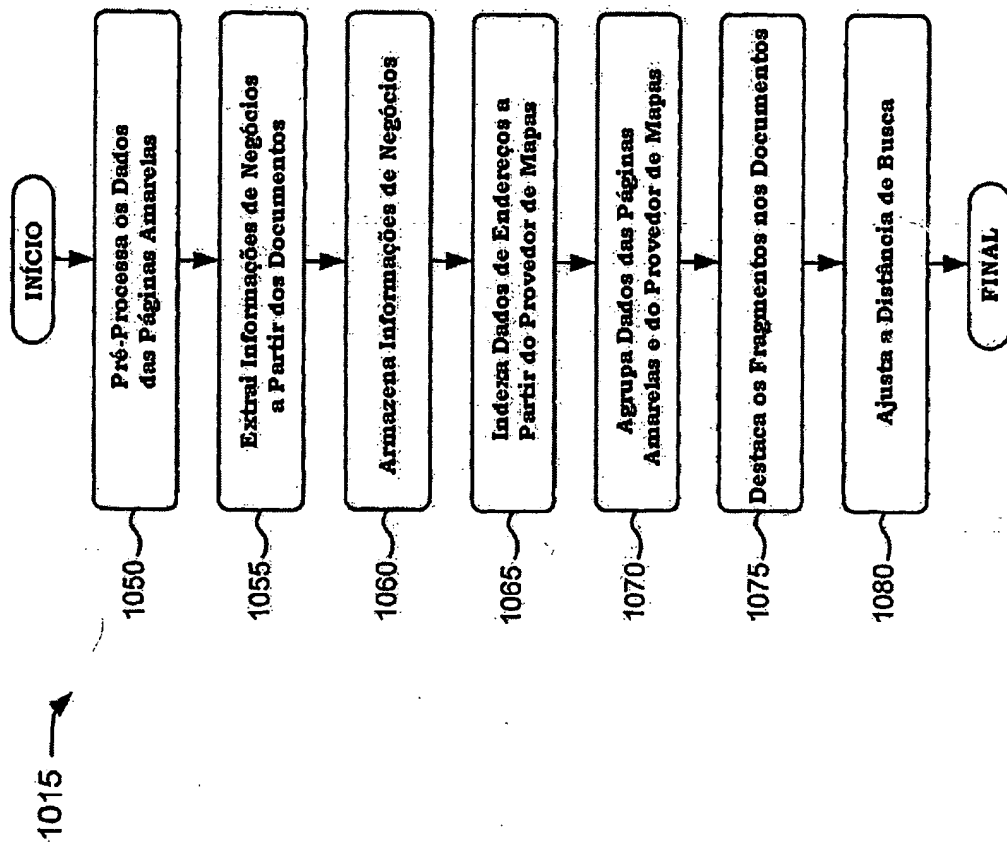


Figura 10C

1020 →

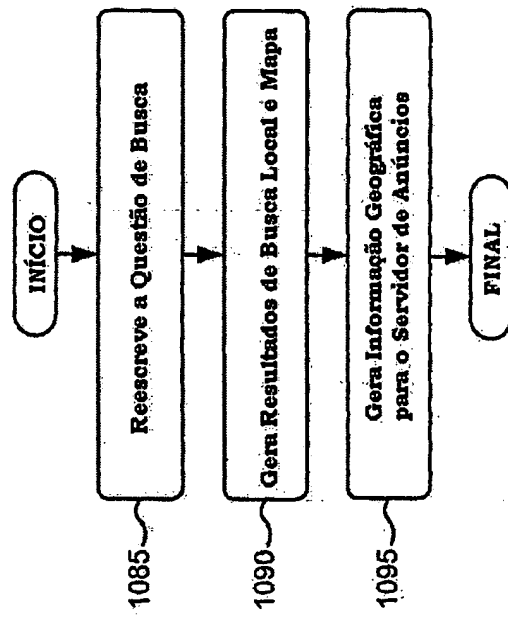


Figura 10D

**“Métodos de Buscas Locais, de Geocodificação
e de Indexação e Sistemas de Buscas Locais”**

Resumo

Um sistema recebe dados das páginas amarelas, dados de
5 provedor de mapas e dados de documento em resposta a uma questão
de busca local e geocodifica os dados de modo a atribuir um identifica-
dor geográfico e combinar pelo menos um endereço associado com a
questão de busca local. O sistema também indexa os dados geocodifi-
cados para determinar informações de negócios e informações de
10 localização associadas com a questão da busca local. O sistema ainda
proporciona os resultados da busca local e um mapa com base nos
dados indexados.