

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
9 June 2005 (09.06.2005)

PCT

(10) International Publication Number
WO 2005/052811 A1

(51) International Patent Classification⁷: **G06F 17/30**

James [GB/GB]; 7 Aintree Drive, Waterlooville, Hampshire PO7 8NE (GB).

(21) International Application Number:
PCT/EP2004/053130

(74) Agent: **WILLIAMS, Julian, David**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Winchester, Hampshire SO21 2JN (GB).

(22) International Filing Date:
26 November 2004 (26.11.2004)

(25) Filing Language: English

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(26) Publication Language: English

(30) Priority Data:
0327589.8 27 November 2003 (27.11.2003) GB

(71) Applicant (*for all designated States except US*): **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York 10504 (US).

(71) Applicant (*for MG only*): **IBM UNITED KINGDOM LIMITED** [GB/GB]; PO Box 41 North Harbour, Portsmouth, Hampshire PO6 3AU (GB).

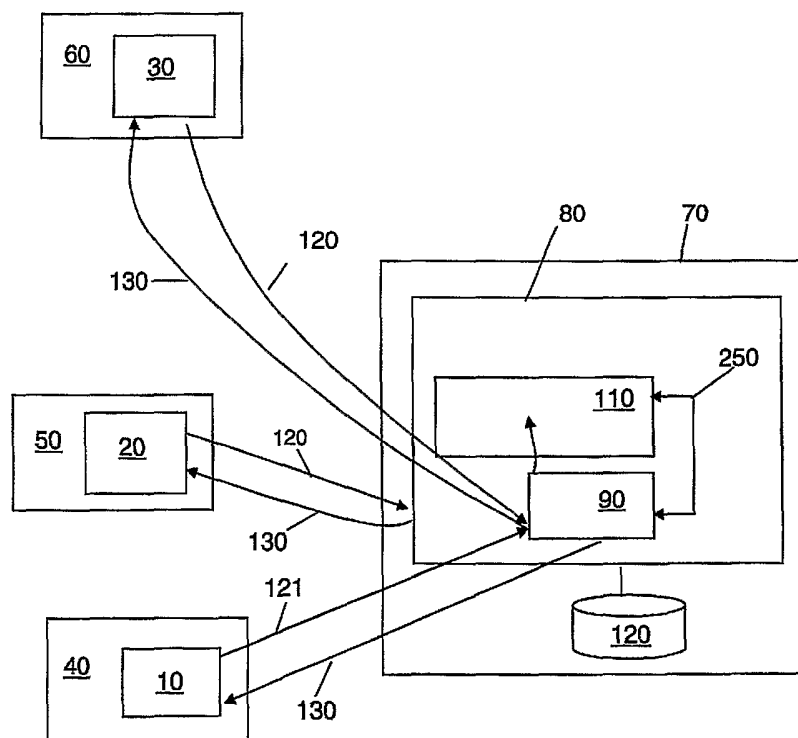
(72) Inventor; and

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE,

(75) Inventor/Applicant (*for US only*): **MIDGLEY, Nicholas**,

[Continued on next page]

(54) Title: **SEARCHING IN A COMPUTER NETWORK**



(57) Abstract: The invention allows a user to enter a search criteria in a search form once and search multiple web services, search engines and web pages without the need for repeatedly re-creating the search criteria for each web service or search engine because of different and varying supported formats. Once the user has defined their search query, the search query is passed to a query manager in the browser application for formatting the search criteria into a generic query format. The generic query format provides a unified and generic interface to all other query mechanisms supported by different web services, search engines and web pages.



SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *with international search report*

Description

SEARCHING IN A COMPUTER NETWORK

Technical Field

- [001] The present invention relates to a method, system and computer program for searching in a computer network and, in particular, defining a generic query format for use in a browser application to permit a single search query to be used for repeated searching within one or more websites or within a directory accessed via the Internet or an intranet.

Background Art

- [002] The advent of the Internet has brought a vast source of information to the finger tips of millions of people. Previously, this vast source of information had only been available in public and private libraries. As a result the Internet has become an invaluable tool for carrying out research through the means of tools such as search engines.
- [003] When performing searches over the Internet, or any other type of network such as a corporate Intranet, a user needs to search a number of repositories. A user performs a search by accessing their preferred search engine. The search engine requires the user to fill out a form with search terms and other criteria, or at least to type search terms into a search-specific entry field. A similar sequence of user interactions is required to initiate a Lightweight Directory Access Protocol (LDAP) directory search or a search via a search engine such as Google TM – firstly loading a web page, and then filling in a form to generate an LDAP directory search request or a Google TM web site search request.
- [004] The search engine takes the search criteria and formulates a search string. The search engine compares the search string to the search engine's database of keyword indices and any matches are returned in a 'hit list' to the user.
- [005] To create a database of keyword indices, search engines use software robots or software spiders to crawl the Internet. Each software robot has its own strategy for crawling the Internet, but generally each software robot starts from a predetermined list of historical Uniform Resource Locator (URL's) and from this list locates a document. The software robot may either parse the entire document, the title of the document or the first paragraph. The parsed information is indexed and stored in a database of keyword indices.
- [006] In some cases, a search is returned by a search engine displaying a list of successful

hits. In other cases, an unsuccessful hit may be returned and the user may be given the option of selecting another link to another web site in which a user may search further in order to delve deeper into the search engine's document repository. This step requires the user to re-enter their search query and perform a subsequent search.

[007] An example of the above can be found by using a search engine such as Google™. If an unsuccessful (for example a HTTP 404 error) or successful hit, (for example a list of web pages) is returned, a user is given the option of 'searching within results'. Clicking on this link, a further search page is presented allowing a user to re-enter their search query and search within the returned web pages. Alternatively, a user can access another search engine and re-enter their search query into the search box to try and locate the results that they require.

[008] The above task when performed repeatedly over a number of search engines or within the same search engine can become repetitive and tedious, requiring repeated entry of a search query in a number of different formats. Often when performing an advanced search, a varying number of parameters are required to be entered, for example, case sensitive letters, keywords only, title only, body only, Boolean operations and specifying the maximum number of documents to be returned etc.

[009] Current web browsers provide a mechanism in which the web browser will 'remember' the last word that a user has typed into the search input box. As a user types in the first few letters of the word, the web browser will begin to match the letters with a word that was previously entered into the input box. For example, if a user had previously typed in the word 'toaster', the web browser would begin to pattern match (pre-fill the input box) as the user types 't', followed by 'o' followed by 'a', the user is then able to select the word 'toaster' if that is the word that they are looking for.

[010] A disadvantage with the above approach is that it does not allow for repeated searching across multiple web sites or web services.

[011] WO 02/091241 discloses a system and a method for distributed real-time search mechanism in a network. Network nodes operating as consumer or requesting nodes generate search requests. Nodes operating as hubs are configured to route the search requests in the network. Communication between nodes in the network may use a query routing protocol. The common query protocol is implemented as a server side protocol and is used to enable business to business services.

[012] A disadvantage of WO 02/091241 is that it does not allow a user freedom of action to create and manipulate their own searches and search mechanisms as the intelligence

and control of the search is totally under the control of the search engine.

[013] WO 02/091241 does not address the usability issues surrounding a user having to perform multiple, repeated searches across different web sites, search engines and web services. In this case, the user has to express a search query in an appropriate format for each and every web site, including the selection of various options for example the language that the search is to be performed in. At the same time, some systems will require authentication to be performed, for example, subscription based repositories, for which normal search services cannot obtain access, and therefore can not return the requested search results.

[014] Therefore, there is a need for a method, system and computer program for searching in a computer network and, in particular, defining a generic query format for use in a browser application to permit a search query to be used for repeated searching within one or more websites or within a directory accessed via the Internet or an intranet.

[015] For the remainder of this document the word web service will be used to encompass, but not limited to the following terms, a web page, a web site, a search engine or a web service, or any other means for obtaining search information or services over a computer network.

Disclosure of Invention

[016] Viewed from a first aspect, the present invention provides a method for processing a search query in a browser application, the browser application receiving a search query for submitting across a network and for receiving by a web service, the method comprising the steps of: formatting the received search query into a generic query format; translating the generic query format into a search format of the receiving web service; and transmitting the search format to the receiving web service for searching.

[017] Preferably, the present invention provides a method wherein the step of translating further comprises submitting a request to the receiving web service requesting identification of the receiving web service's search format.

[018] Preferably, the present invention provides a method wherein the step of translating further comprises, mapping each parameter within the generic query format to each corresponding search parameter of the search format of the web service.

[019] Preferably, the present invention provides a method further comprises the step of aggregating a response from the receiving web service.

[020] Preferably, the present invention provides a method wherein each parameters of the generic query format is updated on the trigger of an addition or a deletion of a search parameter of the search format.

- [021] The invention allows a user to enter search criteria in a search form once and to search multiple web services without the need for repeatedly re-creating the search criteria for each web service or search engine because of different and varying supported formats.
- [022] Once the user has defined their search query, the search query is passed to a query manager in the browser for formatting the search query into a generic query format. The generic query format is used to provide a unified and generic interface to all other query mechanisms supported by different web services.
- [023] The query manager initially sends a request to the web service or search engine that the user wants to search asking for information on its search capabilities. The request is sent in the form of a HyperText Transfer Protocol (HTTP) request or any other suitable protocol for communicating requests across a computer network. The target web service responds with a reply indicating its search capabilities and the search parameters that the web service utilises in order to search other web services.
- [024] The query manager maps the user's search query saved in the generic query format to the search capabilities and parameters of the target web service. The query manager continues the mapping process for each web service that is to be searched. A web service may pass on the formatted search criteria to other compliant web services. In this instance, a reply message is sent back to the query manager indicating that the search request has been sent onto other compliant web services. The present invention allows a search to be carried out on web services that require registration and authentication services, which would otherwise be out of bounds for a normal search engine. This is because authentication details may be submitted by a user to the query interface manager, such that when a web service challenges the query manager for authentication, the query manager is able to carry out negotiation on behalf of the user.
- [025] Once all the replies have been received by the query manager, the replies are aggregated and formatted into a response for passing onto the browser for displaying in the browser window.
- [026] The query manager maintains a data store of search queries in the generic query format that the user has elected to be stored for future use and reference. The saved search queries may be organised, for example, into categories of relevance or statistical data may be collated by the query manager to organise the saved queries by the ten most recent queries or by the most searched domain.
- [027] A button on the browser's tool bar, a menu option or a floating button within the browser's display window allows a saved search to be retrieved from the data store and

submitted to a web service. Associating a saved search query with, for example, a floating button allows the user to execute a search within a child browser window and not just the parent window as would be the case if a button or menu option was coded within a Hypertext Markup Language (html) web page.

[028] Viewed from a second aspect, the present invention provides a system for processing a search query in a browser application, the browser application receiving a search query for submitting across a network and for receiving by a web service, the system comprising: a formatter component, for formatting the received search query into a generic query format; a translator component, for translating the generic query format into a search format of the receiving web service; and a transmitting component for transmitting the search format to the receiving web service for searching.

[029] Preferably, the present invention provides a system wherein the translator component further comprises submitting a request to the receiving web service requesting identification of the receiving web service's search format.

[030] Preferably, the present invention provides a system wherein the translator component further comprises, parsing each parameter within the generic query format to each corresponding search parameter of the search format of the web service.

[031] Preferably, the present invention provides a system further comprising an aggregator component for aggregating a response from the receiving web service.

[032] Preferably, the present invention provides a system wherein each parameter of the generic query format is updated on the trigger of an addition or a deletion of a search parameter of the search format.

[033] Viewed from a third aspect, the present invention provides a computer program product directly loadable into the internal memory of a digital computer, comprising software code portions for performing, when said product is run on a computer, to carry out the invention as described above.

Brief Description of the Drawings

[034] Embodiments of the invention are described below in detail, by way of example only, with reference to the accompanying drawings in which:

[035] Figure 1 is a block diagram showing interactions between a web browser and a number of servers within a computer network such as the Internet, according to an embodiment of the invention;

[036] Figure 2 is an example of the interface in which a use may build a search query;

[037] Figure 3 is a flow diagram showing a sequence of steps of a method for creating a search in accordance with a preferred embodiment of the present invention; and

[038] Figure 4 is a flow diagram showing a sequence of steps of a method for carrying out the search process in accordance with a preferred embodiment of the present invention.

Mode for the Invention

[039] Where reference is made in any one or more of the accompanying drawings to steps or features which have the same reference numerals, those steps or features have substantially the same function(s) or operation(s) for the purposes of this description.

[040] Figure 1 shows a number of components of a data processing network, including a number of web servers 10,20,30 running on mainframe servers 40,50,60 that are connected for communication with a user's computer 70. The user's computer 70 is running a web browser program 80 and a query manager 90. As is known in the art, a web browser is an application program which is capable of sending Hypertext Transfer Protocol (HTTP) requests to web servers to access information on the World Wide Web. Alternative embodiments of the invention include browsers or other client requestor programs which support the File Transfer Protocol (FTP), Lightweight Directory Access Protocol (LDAP) or other protocols for sending requests.

[041] The network may be the Internet, an Intranet, a WiFi network, a wireless network, or any other form of network or means for requesting and receiving data from one data processing device to another.

[042] Each of the user's computer 70 and the servers 40,50,60 may be remote from each other within the network, such as if the invention is used for searching for desired web pages on computers connected within the Internet and accessible via the World Wide Web. The user's computer 70 and the server 40,50,60 could equally be components of a local network or intranet.

[043] The user's computer 70 is not limited to a particular type of data processing apparatus, and may be a conventional desktop or lap-top personal computer, a personal digital assistant (PDA) or another specialized data processing device which is capable of running a client requestor program. The user's computer may connect to a network of data processing systems via wireless or hard-wired connections. Similarly, the servers 40,50,60 can be any data processing apparatuses which are capable of running a web server, directory server or similar server program. Software-implemented elements of the embodiment described in detail below are not limited to any specific operating system or programming language.

[044] In a first embodiment, the query manager 90 is implemented as a computer program module which extends and modifies the functions of a standard web browser 80. In

particular, this embodiment provides a “plug-in” program module for connecting to a standard connection interface of Netscape Corporation’s Netscape Communicator web browser program. As is known in the art, “plug-in” modules are programs that can be easily installed and used as part of a Web browser – once installed, “plug-in” modules are recognized automatically by the web browser and the web browser and plug-in modules call each other’s functions via a simple API.

[045] A number of “plug-in” components are already widely available for use with Microsoft Corporation’s Internet Explorer 5 or Netscape Corporation’s Netscape Communicator Web Browsers. At the time of writing this patent specification, an explanation of plug-ins for Netscape Communicator, their use and development is available from Netscape Corporation. Since the interfaces and development of “plug-in” components to add functions to an existing web browser are understood by those skilled in the art, the interfaces and development steps will not be described in detail herein.

[046] A data store 120 cooperates with the query manager 90 to store search queries. The data store 120 may comprise any type of storage and retrieval mechanism, such as, for example, IBM’s DB2 database system.

[047] The query manager 90 cooperates with the web browser 80 to process an entry of a search query. The search query is built by invoking the query manager interface 110 and entering a search string and a number of options to define the search query to retrieve accurate search results.

[048] The query manager interface 110 may be invoked in a number of ways, for example, by selecting a button from the web browser’s 80 toolbar, selecting an option from a menu function or selecting a floating button which is available to the user within a web browser window. It will be obvious to a person skilled in the art that the above methods for invoking the query manager interface 110 are not limited to the methods listed above but may comprise other methods of invocation as are known in the art.

[049] With reference to Figure 2, an example of a query manager interface is illustrated. The query manager interface 110 comprises a number of input fields 210, 215, 220, 225 and 230. The input fields allow a user to build and define a search query. The query manager interface 110 further comprises a tool bar 205 for accessing saved search queries, editing saved search queries and submitting search queries to a number of web services for searching.

[050] The query manager interface 110 provides a single search interface for defining a

search query and searching multiple web services without the need to repeatedly create the search query for each website or web service.

[051] The query manager interface 110 provides additional input fields which accept user names and passwords for web services that require registration and authentication, which would otherwise be out of bounds for a search engine.

[052] The query manager interface 110 may be implemented as an http input form, or other mechanisms providing input of data, for example, a Java program or any other suitable programming language for implementing applications over a computer network (Java is a trademark of Sun Microsystems Inc. in the U.S. and other countries).

[053] In one embodiment communication over the computer network 121, 130 is carried by using a request/reply protocol such as the Hypertext Transfer Protocol (HTTP).

[054] In another embodiment, the query manager 90 may issue request/reply messages directly to multiple target web sites via a web service call utilising SOAP messages. SOAP provides the definition of XML based information which may be used for exchanging structured and information in a distributed environment.

[055] The query manager interface 110 as illustrated in Figure 2 is for illustration purposes only, but it will be apparent to a person skilled in the art that the query manager interface may comprise a number of different fields in response to the parameters supplied to the generic query format from XML data feeds containing information on the latest query parameters sent from compliant web services.

[056] The user begins by building their search query by defining a search query in one or more of the search entry fields 100 as explained with reference to Figure 2.

[057] The query manager 90 receives the entered search query from the query manager interface 110. The entered search query is stored in the browser's data store 120 in a generic query format. The generic query format may be saved in the data store 120 as an XML schema or in a textual format as dictated by a relational database.

[058] The generic query format maintains a list of parameters used to define a search query. The following is an example list of such parameters, but is not limited to the following:

[059] Search Term

[060] The search term defines the word or words that comprise the search criteria which is to be used for the search.

[061] Exclude Term

[062] An example of an excluded word may be 'burnt', if a user is searching for recipes

on how to create the perfect piece of toast.

[063] Exact phrase term

[064] An exact phrase may be used if a user is searching for a specific term, for example, Châteaux Midgley in the Loire region of the South West of France. This enables the search engine to only return the search results that meet the search criteria above.

[065] Boolean term

[066] A Boolean expression allows a search query to be expressed using notation such as 'and', 'or' and 'not'. For example searching for the terms Babbage 'and' analytical engine, enables the search engine to return search results that comprise the words 'Babbage' and 'analytical engine' in the same web page.

[067] Mode of search

[068] Whether the search term is a Boolean search.

[069] Scope of search

[070] Enables a search to be carried out on a particular domain, for example country specific i.e. the UK or domain specific www.ibm.com.

[071] Language

[072] A search may be limited to a specific language, for example, only search only websites with the extension.

[073] Time period

[074] If searching for articles concerning recent new events, articles older than 1 month ago may be of no use and hence it is possible to limit the search to articles which are 1 week old. Alternatively if a search is trying to locate articles concerning a subject that took place a year ago, it is further possible to limit the search to articles with dates of 12 months etc. It will also be appreciated by a person skilled in the art that any date settings may be configured for use by the user.

[075] Timeframe

[076] Alternatively it is possible to specify a time frame in which to search documents in, for example date range from/date range to.

[077] File type

[078] Different types of file types can be defined in the search criteria, for example only search for .pdf files. Other file types include but not limited to doc, html, exe, Zip, ppt,xls,lwp,123 etc.

[079] Occurrences

[080] A user may only be interested in terms that appear in specific parts of a document, for example, a search is requested to return only documents with the words laser

technology in the 21st century in the title. Another example may comprise only returning documents with Pythagoras theorem in the body of the document etc.

[081] Location

[082] To enable searching within a specific domain such as www.ibm.com

[083] Results per page

[084] The search engine will return a designated number of searched per page, for example, the top 10 nearest hits.

[085] Chain

[086] Designating a chain allows a search query to be routed to other search providers.

[087] In a preferred embodiment, the above parameters are defined in the generic query format as an XML schema for use in a web services environment, for example, with reference to Example 1, it can be seen that the parameters, days, case, occurrences and language are contained within the coding of the XML format.

[088] Example 1

[089] <?—xml version ="1.0" encoding="UTF-8"?>

[090] <!--The definition of the search entered by the user comprising a search phrase and an associated search option)>

[091] <define search query>

[092] <search term></search term>

[093] <time period></time period>

[094] <case (Y|N)> </case>

[095] </define search query>

[096] With reference to Figure 3 the steps that a user may perform to enter, save and submit a search query are explained.

[097] At step 305, the user launches the query manager interface 110 and specifies a search criteria at step 310.

[098] At step 315, the search query is stored in the generic query format and stored in the data store 120 of the web browser 80 for retrieval by the user.

[099] The query manager 90 maintains the data store 120 of the stored search queries and retrieves the stored search queries according to user defined criteria. For example, the user may define 'the most performed active queries', or by a 'most searched domain', for example the domain http://www.ibm.com being a most frequent searched domain.

[100] To enable the search query to be used repeatedly without a need for manually re-entering of the search query into the query manager interface 110, the search query is associated with a button, menu function or floating button of the web browser 80. Each

time the user wishes to execute a saved search query, the user retrieves from the list of saved search queries, the relevant search query. A menu function or floating button is selected to execute the search. An active search query is available when a new instance of the web browser 80 is created i.e. when a child window is created, for example, in Internet Explorer through invoking File/New/Window or Ctrl+ N from a keyboard connected to the user's computer.

- [101] At step 320, the query manager 90 receives the search query from the query manager interface 110 and begins to process the search query 400. At any time, processing of the search query may be stopped at the instruction of the user at step 340.
- [102] With reference to Figure 4, starting at step 401, the steps the query manager 90 performs in order to process a search query are explained in further detail.
- [103] At step 400 the query manager 90 receives the search query from the query manager interface and begins to process the search request. The query manager 90 begins by identifying the search capabilities of the web service to be searched at step 405.
- [104] Each target web service the user wishes to search may have different search capabilities influenced by the way in which the web service was designed and the programming language or markup language used to implement the design.
- [105] To illustrate the differences that may occur between a plurality of web services. Example 2 illustrates a web page implemented in HTML. As is shown in Example 2 the web page may comprises a plurality of HTML tags. The HTML defines how information should be viewed in a browser.
- [106] Example 2
- [107] <html>
- [108] <head>
- [109] <meta http-equiv="content-type" content="text/html; charset=iso-8859-1">
- [110] <title>XML and Search: SearchTools Report</title>
- [111] </head>
- [112] <body bgcolor="white">
- [113] <h1>This is a heading</h1>
- [114] <P>This is a paragraph</P>
- [115] </body>
- [116] </html>
- [117] Each HTML page begins with an <html> tag, which informs the browser the

markup language being used is HTML. Proceeding the HTML tag, a head tag <head> is used to specify information to the browser about the web page for example, meta information such as an HTML URL encoding reference which allows hexadecimal values to display non standard letters and characters in browsers and plug-in applications.

[118] A title tag enables the web page to be given a title which is displayed in the top most title bar of the browser window. Once the above tags are defined, the body of the HTML page can be developed by including body tags which may define the colour of the HTML page (<body bgcolor="white">), heading tags (<h1> This is a heading </h1>), which define how large the font of the heading should be and a paragraph tag (<P> This is a paragraph</P>), to enter the text of the HTML, therefore specifying how the text of the web page should be used and displayed.

[119] A disadvantage associated with using HTML is that HTML does not describe the contents of the text. For example, if the text within a paragraph tag comprised a catalogue of different types of door handles, HTML does not provide any functionality which allows the catalogue to be sorted by, for example, size and colour.

[120] Conversely, XML provides web developers with the aforementioned functionality allowing more control over the data. In order to have more control over the data, further tags and fields can be defined in the XML format which adequately describes the data. Instead of a search engine searching the whole text of an XML page (which a search engine would have to do in an HTML page), search engines may use the XML tags to specify which parts of the XML page to search. Hence there are many differences when performing a search on an HTML page opposed to an XML page.

[121] Further, because different webservices accessed by a web browser 80 use different tags to identify entry fields, check boxes and drop down menus etc, only limited pre-filling can take place.

[122] With reference to step 410 of Figure 4 the query manager 90 begins by issuing a request to the web service, for example, an HTTP request or SOAP message, requesting details of the web services search capabilities. The web service responds with a reply message comprising MIME content type details about the query mechanism that is currently being utilised by the web service. In another embodiment, the web service may respond using a web services protocol such as SOAP.

[123] Each web service may support different parameters in order to execute a search. Table 1 illustrates these difference between two well known search engines – Alta Vista ® and Google ™. The parameter column defines the parameters maintained by

the query manager 90.

[124] Table 1

Table 001

Parameter	Altavista® Mapping	Google™ Mapping
Search Term	aqa = "text"	as_q = "text"
Exclude Term	aqn = "text"	as_eq = "text"
Exact phrase term	aqp = "text"	as_epq = "text"
Atleast	<no mapping>	as_oq = "text"
Boolean term	aqb = "text"	as_q = "text"
Mode of search	aqmode=bls	<no mapping>
Scope of search	kgs = 0 1	<no mapping>
Language	kl=<language code>	lr = <language code>
Time period	dt = tmperiod dtrange	as_qdr=m3 m6 y all
Timeframe	d2 = 1 2 3 4 5	<no mapping>
Date start	d0 = mm/dd/yy	<no mapping>
Date end	d1 = mm/dd/yy	<no mapping>
Filetype	filetype = pdf html <blank>	as_filetype = pdf ps doc xls ppt rtf <blank>
Occurrences	<no mapping>	as_occt = title body url link
Location	rc = url dmn	as_sitesearch="text"
Results per page	nbq = <numeric>	num = <numeric>
Chain	<no mapping>	<no mapping>

[125] The parameters as listed in the first column of Table 1 provide the generic search parameters for the generic query format. The parameters provide a generic reference in which to map onto other query parameters as specified by a compliant web service.

[126] The second and third columns detail the respective parameters for the search engines Alta Vista ® and Google ™. It can be seen that in some cases, there is no equivalent parameters. For example, the Alta Vista ® parameter, mode of search, has no equivalent parameter in Google ™.

[127] As is known to a person skilled in the art, the above query parameters are not an

- exhaustive list, as other query parameters may be available from both search engines and service providers.
- [128] To enable parameters defined by the generic query format to be dynamically updated, the query manager 90 may provide an auto update function which allows the download of parameters from compliant web services which in turn allows the generic query format to be constantly updated in line with parameters of target websites, web service or search engines (column two and three of Table 1).
- [129] Alternatively, the query manager 90 may access a specific website to download the latest parameter file which contains the most up to date mappings for popular engines. The data contained in the latest downloaded parameter file is extracted and stored in the data store 120 in an appropriate format along with other query parameters.
- [130] The generic query format provides a generic format in which all of the parameters from different web services can be linked together to facilitate searching across multiple sites. If no mapping exists i.e. the mapping is not recognised within the query manager, the default position is to ignore the absent parameter.
- [131] The query parameters can be sent to the query manager in a number of different ways. For example, in a web services environment, the target web service may send the query parameters in an XML Document Type Definition to the query manager.
- [132] Once the query manager 90 receives the search capabilities of the web service, the query manager 90 begins to map the search query stored in the generic query format onto the query parameters as requested by the web service to construct a payload at step 415. The query manager parses each of the parameters defined in the generic query format locating the metadata defined in the generic query format's tags (in an XML implementation). A lookup is performed by the query manager 90 to determine the equivalent query parameter of the web service, search engine or web page to be searched.
- [133] The following is an example of a search query received by the query manager 90 from the query manager interface 110 and mapped onto the search parameters as defined by a target web service.
- [134] Search item: Charles Babbage
- [135] Exclude term: Ada Lovelace
- [136] Time period: 12/12/2002 to 11/11/2003
- [137] The interface manager 90 receives the search query from the query manager interface 110. In this example the search query is saved as an XML schema.
- [138] The query manager 90 sends an http request, or SOAP request depending on the im-

plementation requirements of the web service(s) to be searched, requesting details of the web services search capabilities.

[139] In this example, the reply message defines that one of the web services to be searched requires a DTD format and the other is a web page requiring an HTML format. The reply message further defines the naming convention of the query parameters to be used.

[140] In response to the data supplied in the reply message the query manager 90 performs a lookup in the data store and determines the equivalent parameter to the generic query parameter. The parameter of the generic query format is mapped onto the query parameter of the web service to be searched. The query manager 90 continues parsing the generic query format until each of the parameters defined in the generic query format have been mapped on to the equivalent query parameters. The resulting output can be seen in Example 3 as an XML search query as defined by the web service to be searched.

[141] Example 3

[142] POST /search HTTP/1.1

[143] Host: www.infobase.com

[144] Content-Type: application/soap+xml; charset=utf-8

[145] Content-Length: nnn

[146] <?xml version="1.0"?>

[147] <soap:Envelope xmlns:soap="http://www.w3.org/2001/12/soap-envelope"

[148] soap:encodingStyle="http://www.w3.org/2001/12/soap-encoding">

[149] <soap:Body xmlns:q="http://www.infobase.com/search">

[150] <q:query>

[151] <aqa:searchphrase>

[152] <aqa:searchitem>Babbage</q:searchitem>

[153] </aqa:searchphrase>

[154] <aqa:searchoptions author=""

[155] case="N" chain="Y" time period="12/12/2002 to 11/11/2003" domain=""

format="ANY" Exclude term ="Ada Lovelace"

[156] geography="" language=""

[157] occurrences="ANYWHERE" type="ALLWORDS">searchoptions

[158] </aqa:searchoptions>

[159] </q:query>

[160] </soap:Body>

[161] </soap:Envelope>

[162] The payload will be constructed depending on the requirements of the web service, search engine or web page to be searched. In one embodiment the search site may support a web services interface and supply the payload as a DTD as in Example 3, or the payload may be constructed as a standard HTTP request. How the payload is constructed depends on the requirements of the web service to be searched.

[163] When performing any search over a computer network it is often the case that multiple searches will occur. In this instance, the query manager 90 will be required to map the search query stored in the generic query format onto many different types of query parameters in order to carry out multiple concurrent searches.

[164] Referring back to Figure 4, once the search query is in a format accepted by the web service, the query manager 90 submits the payload to the web service for searching at step 430.

[165] The web service searches the content of the website (in accordance with the defined search query) and identifies occurrences of the search terms within the designated web service or web pages, providing its output as a list of identified web pages or other content within the web site.

[166] The query manager 90 intercepts the received search results and aggregates the results, generating a consolidated result which the query manager 90 forwards to the web browser 80 for display at step 440. This intercepting step 435 may involve collating received search responses from different web services before presenting the results back to the user via the web browser 80.

[167] In another embodiment the search query may be repeated a number of times by locating, specifying or retrieving details of other complaint search engines which may also include the use of directories such as the Universal Description, Discovery and Integration Service (UDDI), which defines a set of standard interfaces for accessing a database of web services. Search engine details may be registered in a UDDI directory and can be classified by particular data categories for example web services, medical, or information technology. When the web browser 80 receives a response from a search engine the response may indicate that the search query has been passed on to another cooperating entity that the search engine knows about, that may be able to provide a more accurate answer to the request than the search engine that was originally sent the query.

Claims

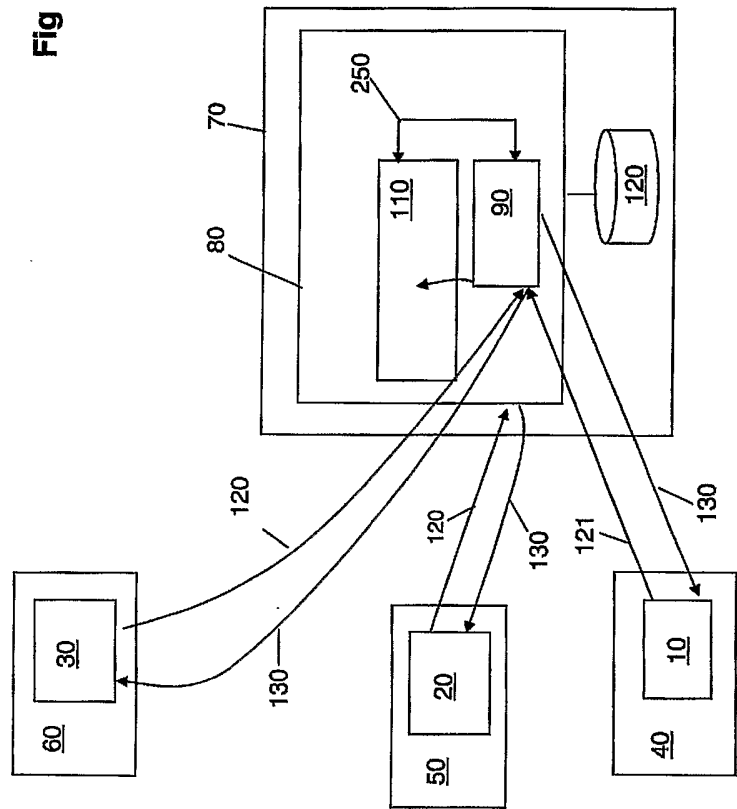
- [001] A method for processing a search query in a browser application, the browser application receiving a search query for submitting to a network and for receiving by a web service, the method comprising the steps of: formatting the received search query into a generic query format; translating the generic query format into a search format of the receiving web service; and transmitting the search format to the receiving web service for searching.
- [002] A method as claimed in claim 1 wherein the step of translating further comprises submitting a request to the receiving web service requesting identification of the receiving web service's search format.
- [003] A method as claimed in claim 1 wherein the step of translating further comprises, mapping each parameter within the generic query format to each corresponding search parameter of the search format of the web service.
- [004] A method as claimed in claim 1 further comprises the step of collecting a response from the receiving web service.
- [005] A method as claimed in claim 3 wherein each of the parameters of the generic query format are updated on the trigger of an addition or a deletion of a search parameter of the search format.
- [006] A system for processing a search query in a browser application, the browser application receiving a search query for submitting across a network and for receiving by a web service, the system comprising: a formatter component, for formatting the received search query into a generic query format; a translator component, for translating the generic query format into a search format of the receiving web service; and a transmitting component for transmitting the search format to the receiving web service for searching.
- [007] A system as claimed in claim 6 wherein the translator component further comprises submitting a request to the receiving web service requesting identification of the receiving web service's search format.
- [008] A system as claimed in claim 6 wherein the translator component further comprises, parsing each of the parameters within the generic query format to each corresponding search parameter of the search format of the web service.
- [009] A system as claimed in claim 6 further comprises an aggregator component for collecting a response from the receiving web service.
- [010] A system as claimed in claim 8 wherein each parameter of the generic query

format is updated on the trigger of an addition or a deletion of a search parameter of the search format.

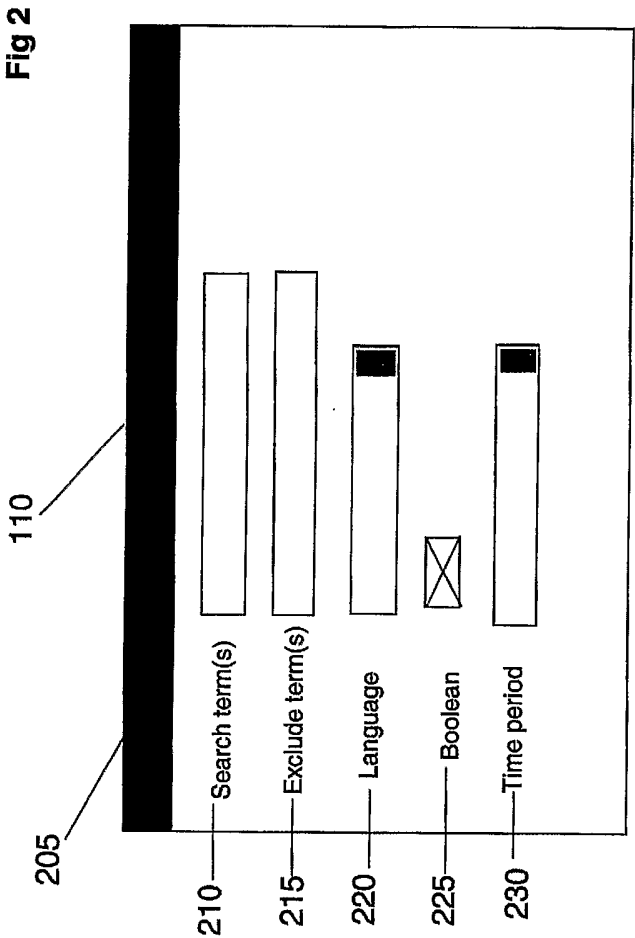
[011] A computer program product directly loadable into the internal memory of a digital computer, comprising software code portions for performing, when said product is run on a computer, to carry out the invention of claims 1 to 5.

[Fig.]

Fig 1

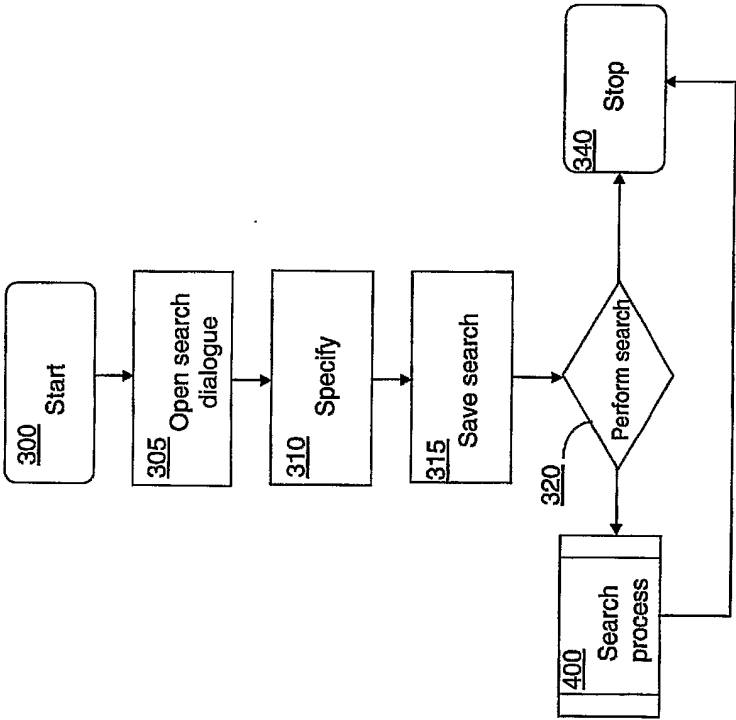


[Fig.]



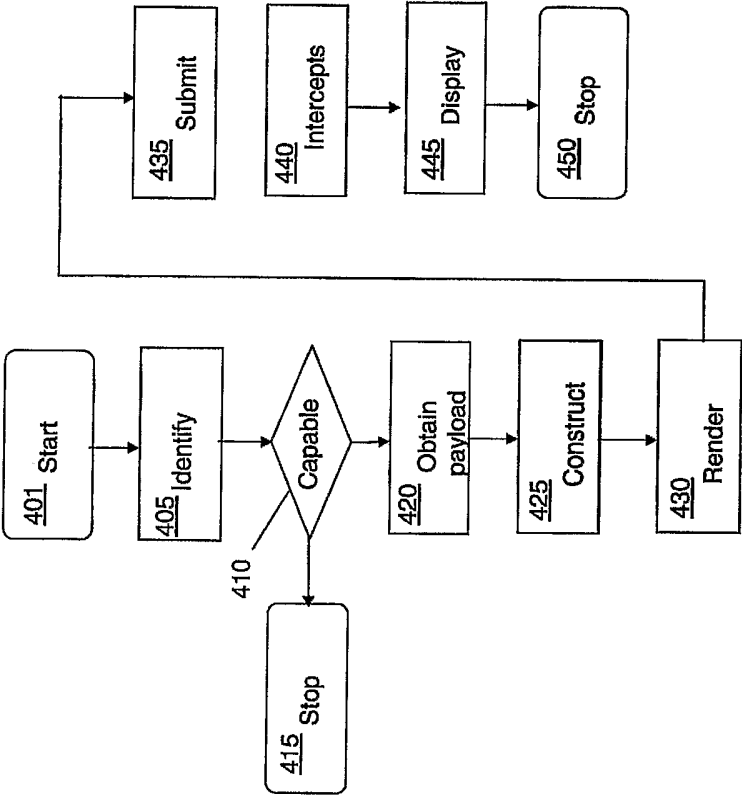
[Fig.]

Fig 3



[Fig.]

Fig 4



INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/053130

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/049756 A1 (CHUA GHIM-SIM ET AL) 25 April 2002 (2002-04-25) abstract paragraph '0006! - paragraph '0009! paragraph '0021! - paragraph '0056! figures 2-5	1-11
X	US 6 085 186 A (CHRISTIANSON ET AL) 4 July 2000 (2000-07-04) abstract column 2, line 39 - column 3, line 50 column 9, line 25 - column 10, line 41	1-11
X	US 2003/204497 A1 (KALOGERAKI VASILIKI ET AL) 30 October 2003 (2003-10-30) abstract paragraph '0010! - paragraph '0013! paragraph '0039! - paragraph '0040!	1,6,11
	----- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

10 March 2005

Date of mailing of the international search report

16/03/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Bernardi, L

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/053130

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/055818 A1 (FAYBISHENKO YAROSLAV ET AL) 20 March 2003 (2003-03-20) abstract paragraph '0008! - paragraph '0013! -----	1,6,11
A	POWELL J ET AL: "Multilingual Federated Searching Across Heterogeneous Collections" ., September 1998 (1998-09), XP002255043 the whole document -----	1,6,11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2004/053130

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002049756	A1	25-04-2002	NONE
US 6085186	A	04-07-2000	US 6102969 A 15-08-2000 AT 255246 T 15-12-2003 AU 731420 B2 29-03-2001 AU 4499497 A 14-04-1998 CA 2269227 A1 26-03-1998 CN 1249042 A 29-03-2000 EP 1015964 A2 05-07-2000 IL 129066 A 04-01-2004 JP 2001501003 T 23-01-2001 KR 2000048514 A 25-07-2000 NZ 335247 A 28-04-2000 WO 9812881 A2 26-03-1998
US 2003204497	A1	30-10-2003	NONE
US 2003055818	A1	20-03-2003	US 2003088544 A1 08-05-2003 WO 02091243 A2 14-11-2002 WO 02091238 A2 14-11-2002 WO 02091239 A2 14-11-2002 WO 02091107 A2 14-11-2002 WO 02091241 A2 14-11-2002 WO 02091242 A2 14-11-2002 US 2003158839 A1 21-08-2003 US 2003050959 A1 13-03-2003 US 2003050924 A1 13-03-2003 US 2003126120 A1 03-07-2003