US 20050216434A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2005/0216434 A1**

Haveliwala et al. (43) **Pub. Date: Sep. 29, 2005**

(54) **VARIABLE PERSONALIZATION OF SEARCH RESULTS IN A SEARCH ENGINE**

(76) Inventors: **Taher H. Haveliwala**, Mountain View, CA (US); **Glen M. Jeh**, San Francisco, CA (US); **Sepandar D. Kamvar**, San Francisco, CA (US)

Correspondence Address:
FENWICK & WEST LLP
SILICON VALLEY CENTER
801 CALIFORNIA STREET
MOUNTAIN VIEW, CA 94041 (US)

(21) Appl. No.: **11/002,474**

(22) Filed: **Dec. 1, 2004**

Related U.S. Application Data

(60) Provisional application No. 60/557,783, filed on Mar. 29, 2004.

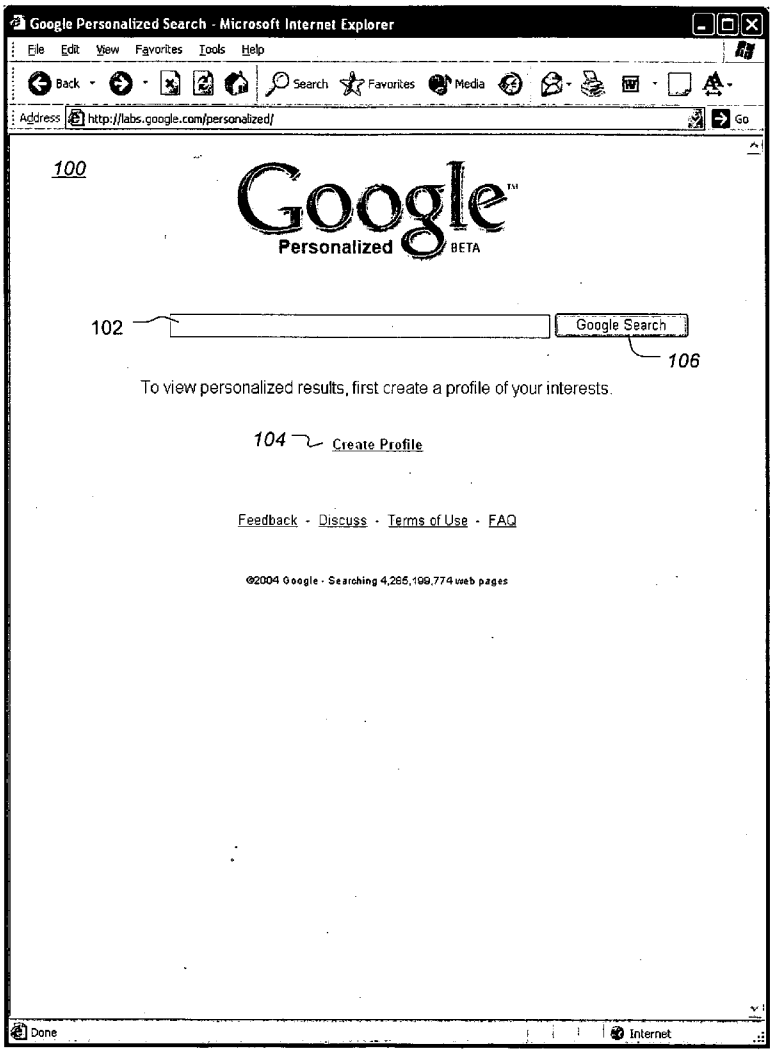
Publication Classification

(51) **Int. Cl.**[7] **G06F 7/00**

(52) **U.S. Cl.** **707/1**

(57) **ABSTRACT**

A search engine provides personalized rankings of search results. A user interest profile identifies topics of interest to a user. Each topic is associated with one or more sites, and a boost value, which can be used to augment an information retrieval score of any document from the site. Search results from any search are provided to the user, with a variable control of the ranking of the results. The results can be ranked by their unboosted information retrieval score, thus reflecting no personalization, or by their fully or partially boosted information retrieval scores. This allows the user to selectively control how their interests affect the ranking of the documents.



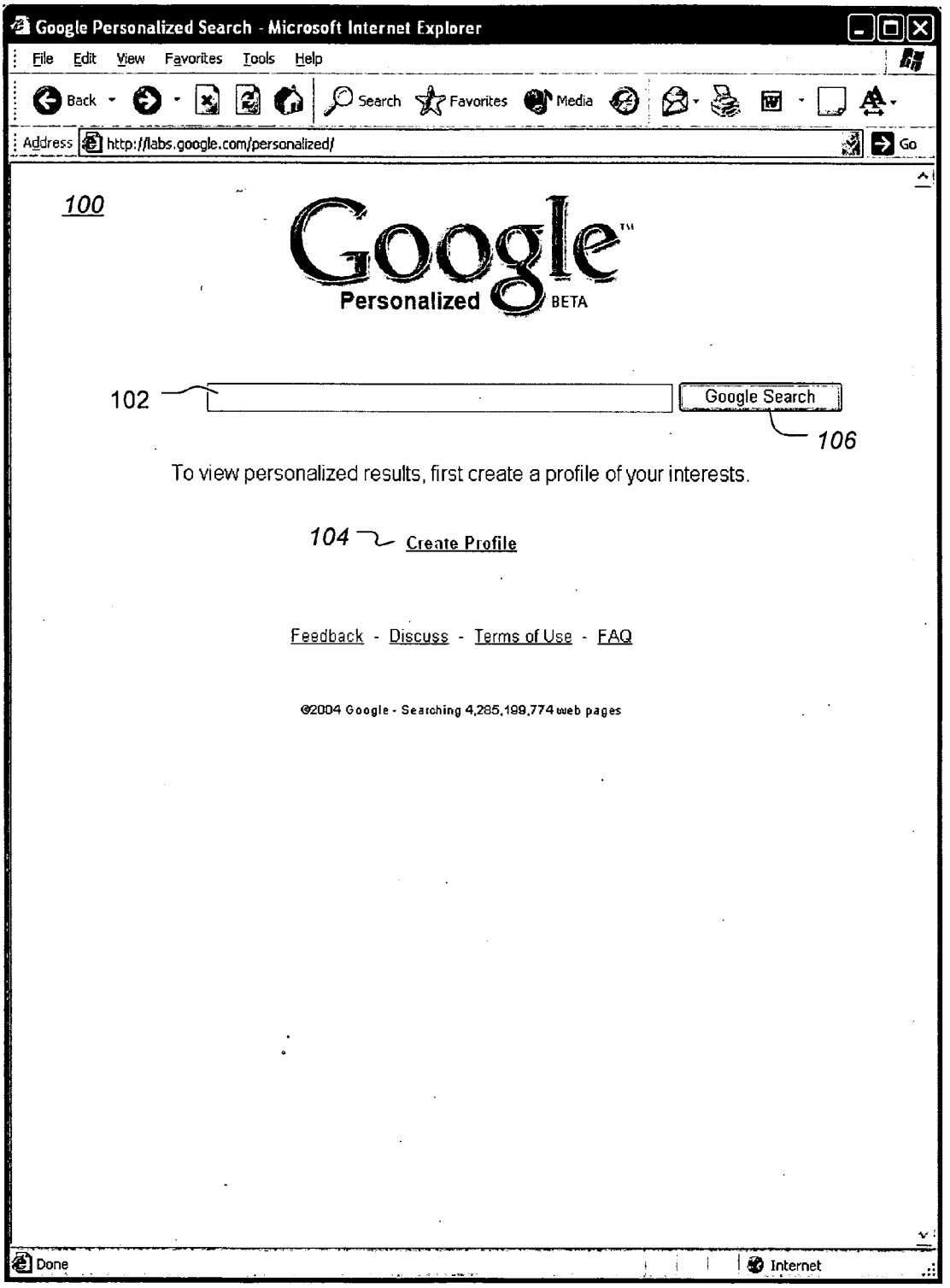

Google Personalized Search - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/
Go
100
Google
Personalized BETA
102
Google Search
106
To view personalized results, first create a profile of your interests.
104
Create Profile
Feedback - Discuss - Terms of Use - FAQ
©2004 Google - Searching 4,285,199,774 web pages
Done
Internet

FIG. 1

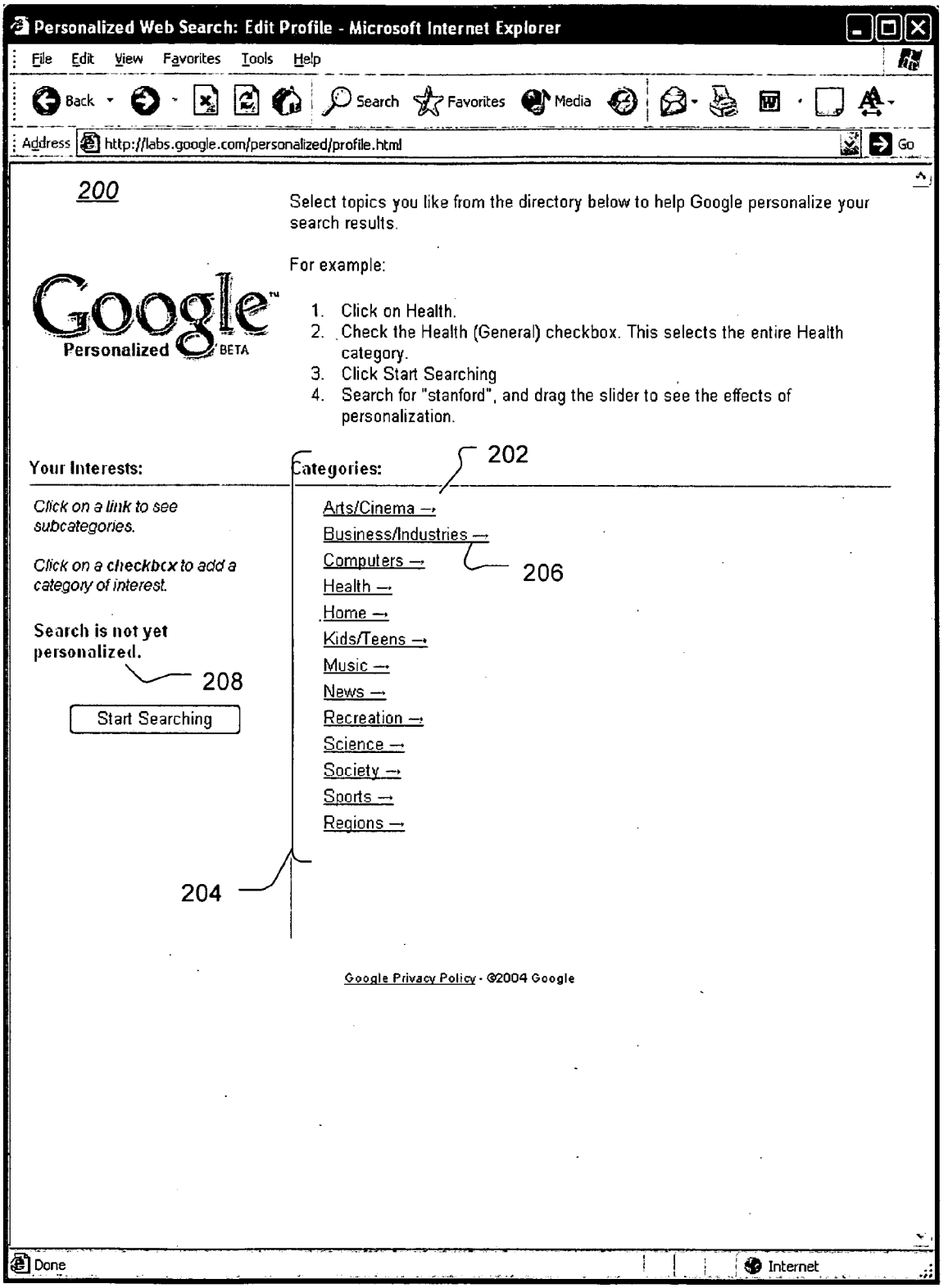

Personalized Web Search: Edit Profile - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/profile.html
Go
200
Select topics you like from the directory below to help Google personalize your search results.
For example:
1. Click on Health.
2. Check the Health (General) checkbox. This selects the entire Health category.
3. Click Start Searching
4. Search for "stanford", and drag the slider to see the effects of personalization.
Google Personalized BETA
Your Interests:
Click on a link to see subcategories.
Click on a checkbox to add a category of interest.
Search is not yet personalized.
208
Start Searching
Categories:
202
Arts/Cinema →
Business/Industries →
Computers →
206
Health →
Home →
Kids/Teens →
Music →
News →
Recreation →
Science →
Society →
Sports →
Regions →
204
Google Privacy Policy - ©2004 Google
Done
Internet

FIG. 2

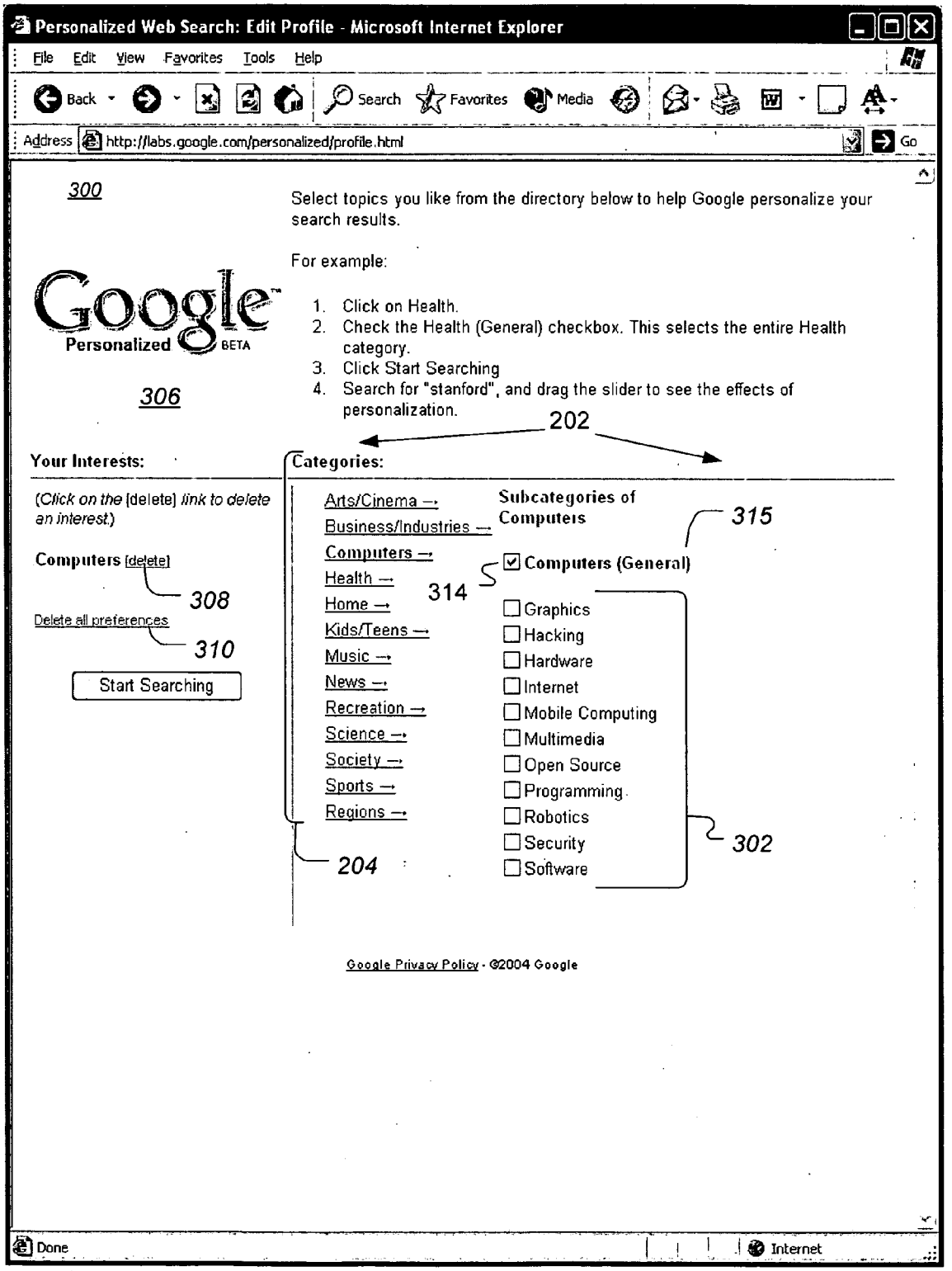

Personalized Web Search: Edit Profile - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/profile.html
Go
300
Select topics you like from the directory below to help Google personalize your search results.
For example:
1. Click on Health.
2. Check the Health (General) checkbox. This selects the entire Health category.
3. Click Start Searching
4. Search for "stanford", and drag the slider to see the effects of personalization.
Google Personalized BETA
306
202
Your Interests:
Categories:
(Click on the [delete] link to delete an interest.)
Computers [delete]
308
Delete all preferences
310
Start Searching
Arts/Cinema →
Business/Industries →
Computers →
Health →
Home →
Kids/Teens →
Music →
News →
Recreation →
Science →
Society →
Sports →
Regions →
204
Subcategories of Computers
315
Computers (General)
314
Graphics
Hacking
Hardware
Internet
Mobile Computing
Multimedia
Open Source
Programming
Robotics
Security
Software
302
Google Privacy Policy - ©2004 Google
Done
Internet

FIG. 3

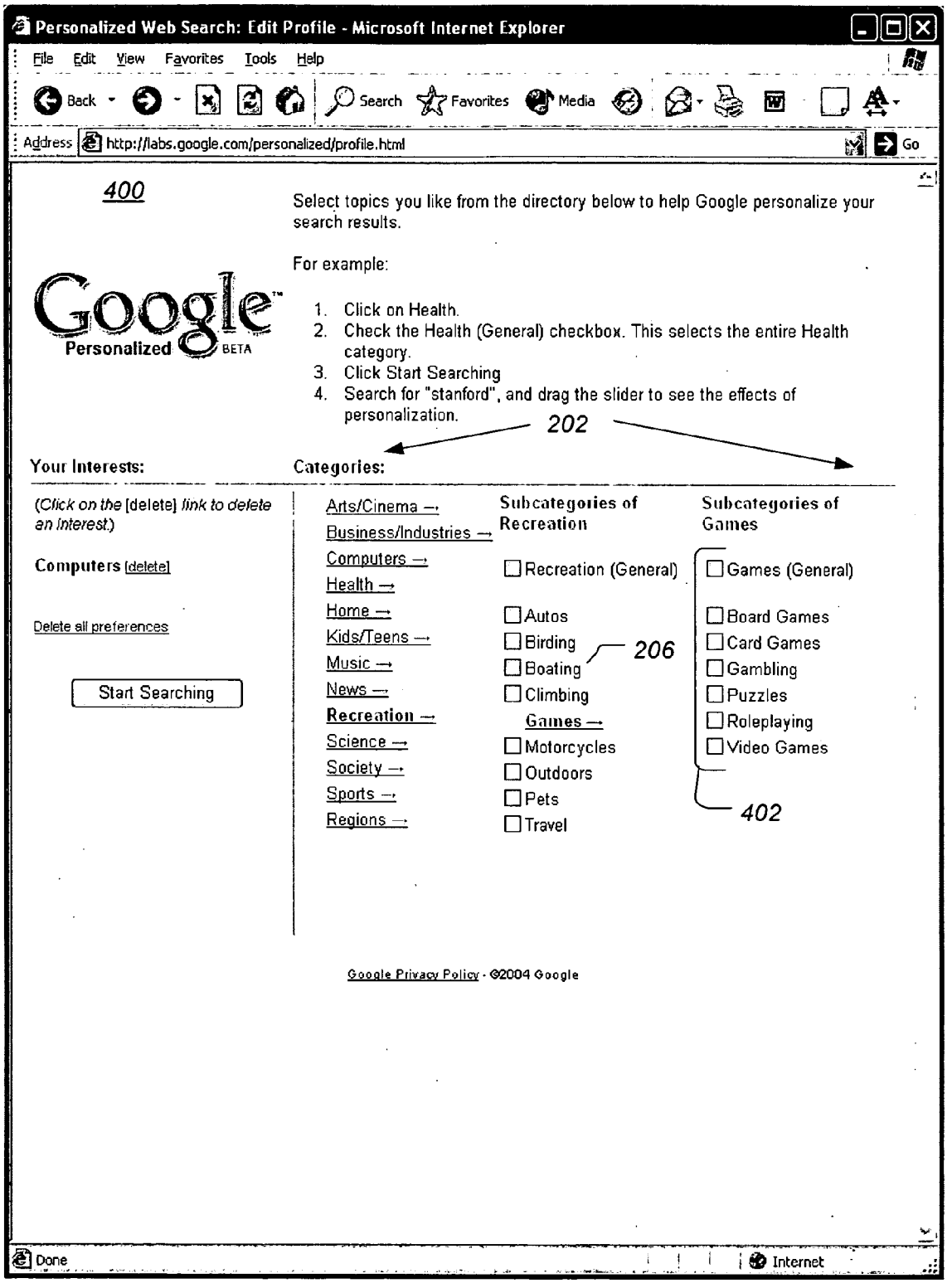

Personalized Web Search: Edit Profile - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/profile.html
Go
400
Select topics you like from the directory below to help Google personalize your search results.
For example:
1. Click on Health.
2. Check the Health (General) checkbox. This selects the entire Health category.
3. Click Start Searching
4. Search for "stanford", and drag the slider to see the effects of personalization.
Google Personalized BETA
202
Your Interests:
Categories:
(Click on the [delete] link to delete an interest.)
Computers [delete]
Delete all preferences
Start Searching
Arts/Cinema →
Business/Industries →
Computers →
Health →
Home →
Kids/Teens →
Music →
News →
Recreation →
Science →
Society →
Sports →
Regions →
Subcategories of Recreation
Recreation (General)
Autos
Birding
Boating
Climbing
Games →
Motorcycles
Outdoors
Pets
Travel
206
Subcategories of Games
Games (General)
Board Games
Card Games
Gambling
Puzzles
Roleplaying
Video Games
402
Google Privacy Policy - ©2004 Google
Done
Internet

FIG. 4

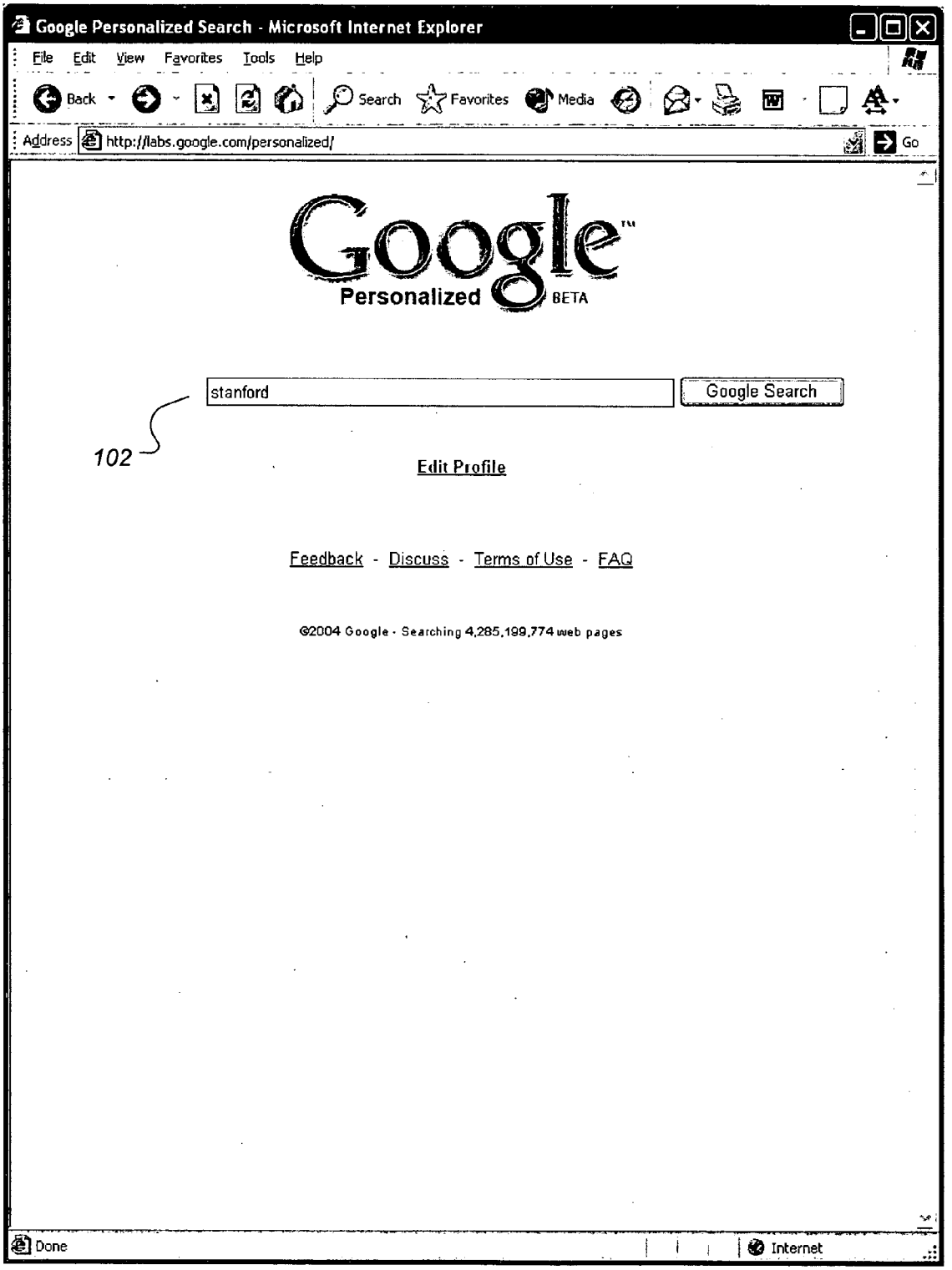

Google Personalized Search - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/
Go
Google
Personalized BETA
stanford
Google Search
102
Edit Profile
Feedback - Discuss - Terms of Use - FAQ
©2004 Google - Searching 4,285,199,774 web pages
Done
Internet

FIG. 5

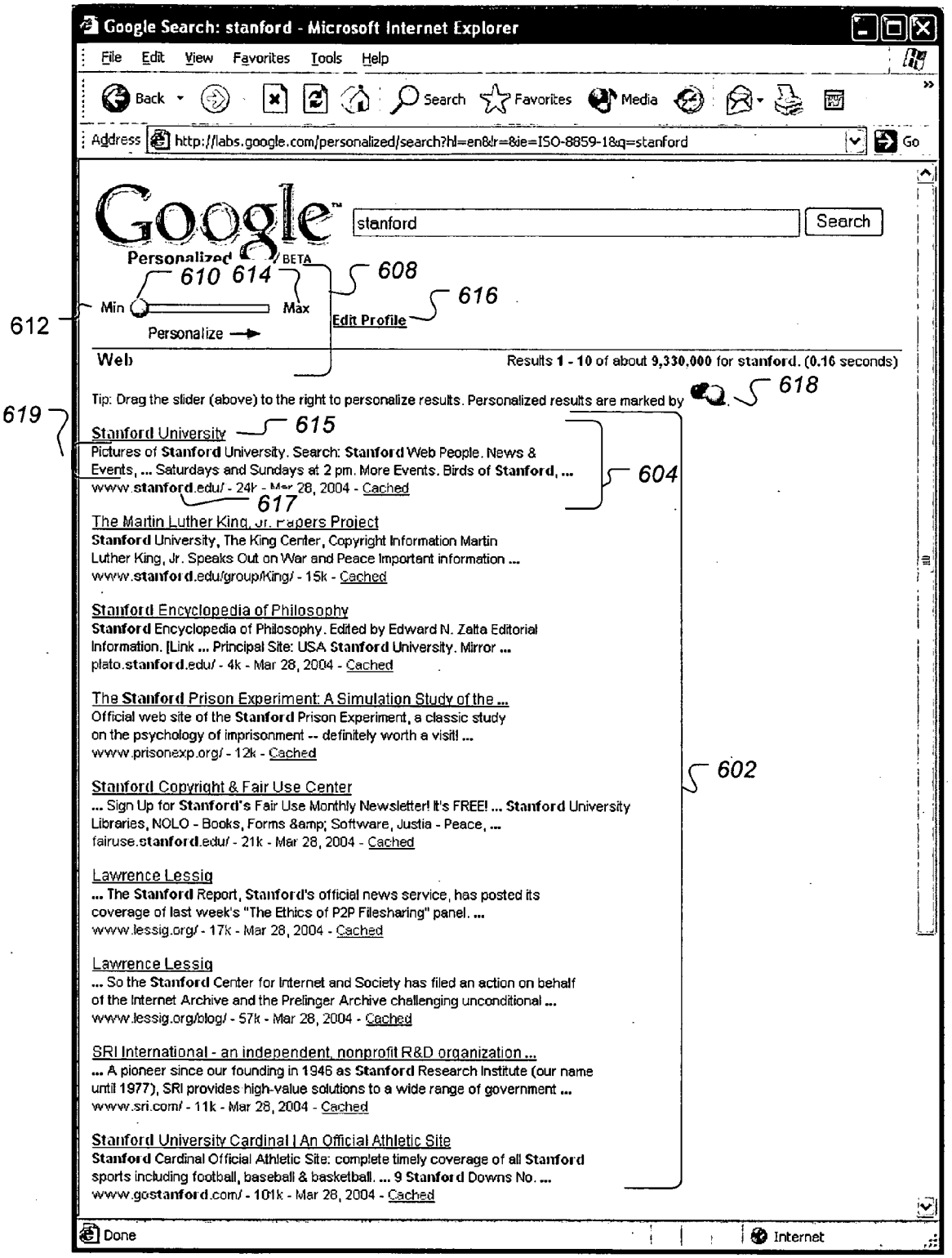

Google Search: stanford - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/search?hl=en&lr=&ie=ISO-8859-1&q=stanford
Go
Google
Personalized BETA
stanford
Search
610 614
608
616
612
Min
Max
Edit Profile
Personalize
Web
Results 1 - 10 of about 9,330,000 for stanford. (0.16 seconds)
Tip: Drag the slider (above) to the right to personalize results. Personalized results are marked by
618
619
Stanford University
615
Pictures of Stanford University. Search: Stanford Web People. News & Events, ... Saturdays and Sundays at 2 pm. More Events. Birds of Stanford, ...
www.stanford.edu/ - 24k - Mar 28, 2004 - Cached
617
604
The Martin Luther King, Jr. Papers Project
Stanford University, The King Center, Copyright Information Martin Luther King, Jr. Speaks Out on War and Peace Important information ...
www.stanford.edu/group/King/ - 15k - Cached
Stanford Encyclopedia of Philosophy
Stanford Encyclopedia of Philosophy. Edited by Edward N. Zalta Editorial Information. [Link ... Principal Site: USA Stanford University. Mirror ...
plato.stanford.edu/ - 4k - Mar 28, 2004 - Cached
The Stanford Prison Experiment: A Simulation Study of the ...
Official web site of the Stanford Prison Experiment, a classic study on the psychology of imprisonment -- definitely worth a visit! ...
www.prisonexp.org/ - 12k - Cached
602
Stanford Copyright & Fair Use Center
... Sign Up for Stanford's Fair Use Monthly Newsletter! It's FREE! ... Stanford University Libraries, NOLO - Books, Forms & Software, Justia - Peace, ...
fairuse.stanford.edu/ - 21k - Mar 28, 2004 - Cached
Lawrence Lessig
... The Stanford Report, Stanford's official news service, has posted its coverage of last week's "The Ethics of P2P Filesharing" panel. ...
www.lessig.org/ - 17k - Mar 28, 2004 - Cached
Lawrence Lessig
... So the Stanford Center for Internet and Society has filed an action on behalf of the Internet Archive and the Prelinger Archive challenging unconditional ...
www.lessig.org/blog/ - 57k - Mar 28, 2004 - Cached
SRI International - an independent, nonprofit R&D organization ...
... A pioneer since our founding in 1946 as Stanford Research Institute (our name until 1977), SRI provides high-value solutions to a wide range of government ...
www.sri.com/ - 11k - Mar 28, 2004 - Cached
Stanford University Cardinal | An Official Athletic Site
Stanford Cardinal Official Athletic Site: complete timely coverage of all Stanford sports including football, baseball & basketball. ... 9 Stanford Downs No. ...
www.gostanford.com/ - 101k - Mar 28, 2004 - Cached
Done
Internet

FIG. 6

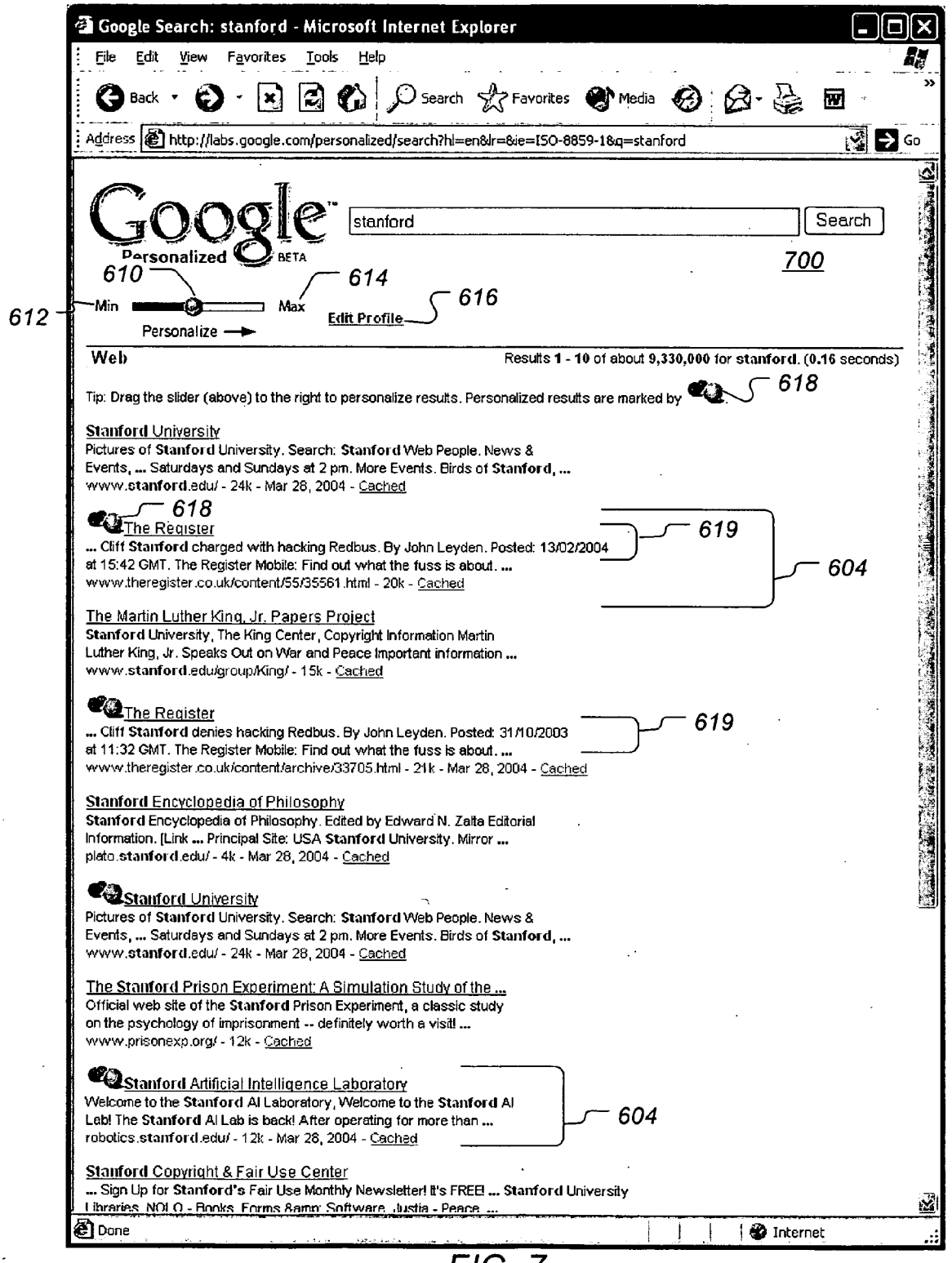

Google Search: stanford - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/search?hl=en&lr=&ie=ISO-8859-1&q=stanford
Go
Google Personalized BETA
stanford
Search
700
610
612
Min
Max
614
Personalize
Edit Profile
616
Web
Results 1 - 10 of about 9,330,000 for stanford. (0.16 seconds)
Tip: Drag the slider (above) to the right to personalize results. Personalized results are marked by
618
Stanford University
Pictures of Stanford University. Search: Stanford Web People. News & Events, ... Saturdays and Sundays at 2 pm. More Events. Birds of Stanford, ...
www.stanford.edu/ - 24k - Mar 28, 2004 - Cached
618
The Register
... Cliff Stanford charged with hacking Redbus. By John Leyden. Posted: 13/02/2004 at 15:42 GMT. The Register Mobile: Find out what the fuss is about. ...
www.theregister.co.uk/content/55/35561.html - 20k - Cached
619
604
The Martin Luther King, Jr. Papers Project
Stanford University, The King Center, Copyright Information Martin Luther King, Jr. Speaks Out on War and Peace Important information ...
www.stanford.edu/group/King/ - 15k - Cached
The Register
... Cliff Stanford denies hacking Redbus. By John Leyden. Posted: 31/10/2003 at 11:32 GMT. The Register Mobile: Find out what the fuss is about. ...
www.theregister.co.uk/content/archive/33705.html - 21k - Mar 28, 2004 - Cached
619
Stanford Encyclopedia of Philosophy
Stanford Encyclopedia of Philosophy. Edited by Edward N. Zalta Editorial Information. [Link ... Principal Site: USA Stanford University. Mirror ...
plato.stanford.edu/ - 4k - Mar 28, 2004 - Cached
Stanford University
Pictures of Stanford University. Search: Stanford Web People. News & Events, ... Saturdays and Sundays at 2 pm. More Events. Birds of Stanford, ...
www.stanford.edu/ - 24k - Mar 28, 2004 - Cached
The Stanford Prison Experiment: A Simulation Study of the ...
Official web site of the Stanford Prison Experiment, a classic study on the psychology of imprisonment -- definitely worth a visit! ...
www.prisonexp.org/ - 12k - Cached
Stanford Artificial Intelligence Laboratory
Welcome to the Stanford AI Laboratory, Welcome to the Stanford AI Lab! The Stanford AI Lab is back! After operating for more than ...
robotics.stanford.edu/ - 12k - Mar 28, 2004 - Cached
604
Stanford Copyright & Fair Use Center
... Sign Up for Stanford's Fair Use Monthly Newsletter! It's FREE! ... Stanford University Libraries. NOLO - Books, Forms & Software. Justia - Peace ...
Done
Internet

FIG. 7

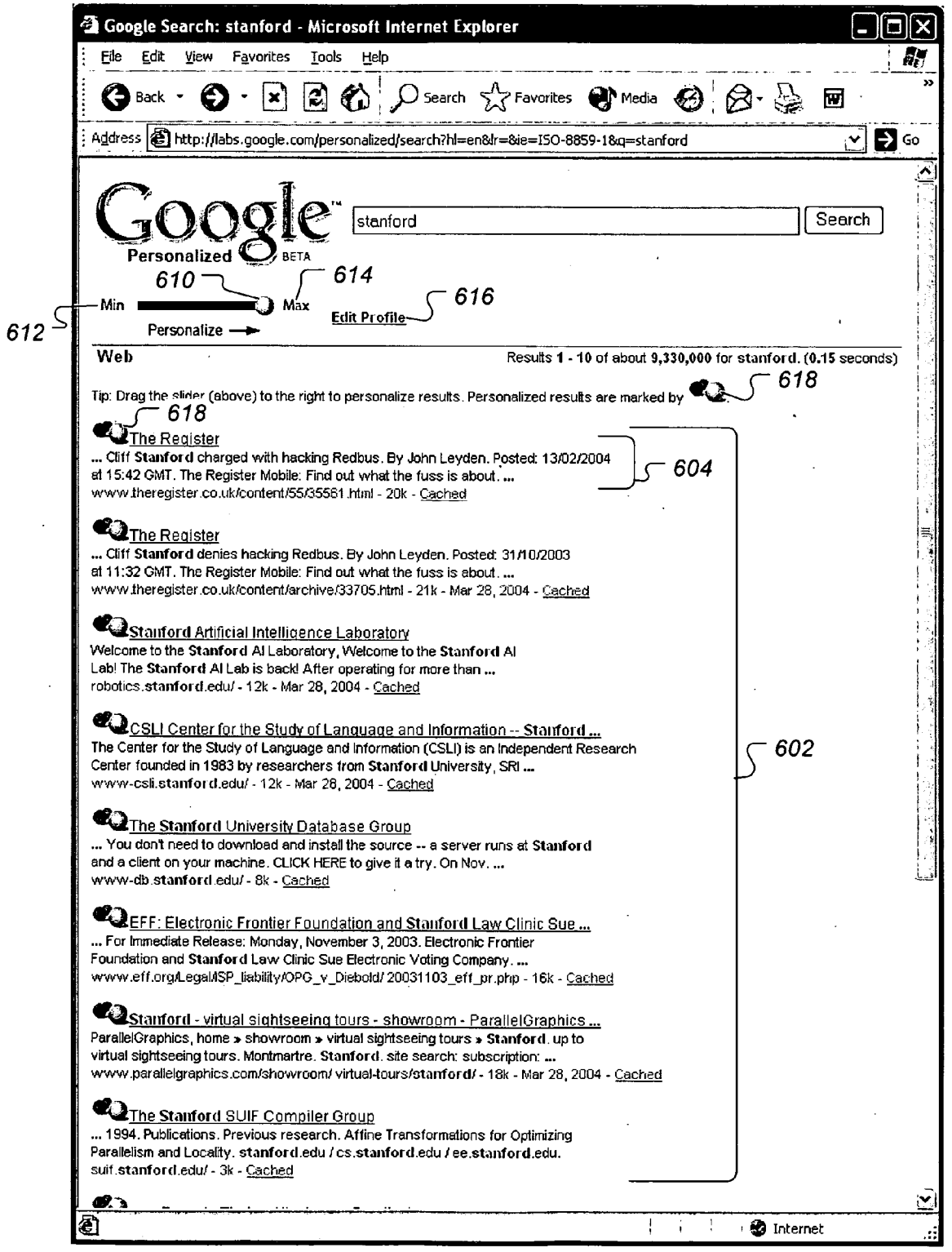

Google Search: stanford - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/search?hl=en&lr=&ie=ISO-8859-1&q=stanford
Go
Google Personalized BETA
stanford
Search
610
614
612
Min
Max
Personalize
Edit Profile
616
Web
Results 1 - 10 of about 9,330,000 for stanford. (0.15 seconds)
Tip: Drag the slider (above) to the right to personalize results. Personalized results are marked by
618
618
The Register
... Cliff Stanford charged with hacking Redbus. By John Leyden. Posted: 13/02/2004 at 15:42 GMT. The Register Mobile: Find out what the fuss is about. ...
www.theregister.co.uk/content/55/35561.html - 20k - Cached
604
The Register
... Cliff Stanford denies hacking Redbus. By John Leyden. Posted: 31/10/2003 at 11:32 GMT. The Register Mobile: Find out what the fuss is about. ...
www.theregister.co.uk/content/archive/33705.html - 21k - Mar 28, 2004 - Cached
Stanford Artificial Intelligence Laboratory
Welcome to the Stanford AI Laboratory, Welcome to the Stanford AI Lab! The Stanford AI Lab is back! After operating for more than ...
robotics.stanford.edu/ - 12k - Mar 28, 2004 - Cached
CSLI Center for the Study of Language and Information -- Stanford ...
The Center for the Study of Language and Information (CSLI) is an independent Research Center founded in 1983 by researchers from Stanford University, SRI ...
www-csli.stanford.edu/ - 12k - Mar 28, 2004 - Cached
602
The Stanford University Database Group
... You don't need to download and install the source -- a server runs at Stanford and a client on your machine. CLICK HERE to give it a try. On Nov. ...
www-db.stanford.edu/ - 8k - Cached
EFF: Electronic Frontier Foundation and Stanford Law Clinic Sue ...
... For Immediate Release: Monday, November 3, 2003. Electronic Frontier Foundation and Stanford Law Clinic Sue Electronic Voting Company. ...
www.eff.org/Legal/ISP_liability/OPG_v_Diebold/20031103_eff_pr.php - 16k - Cached
Stanford - virtual sightseeing tours - showroom - ParallelGraphics ...
ParallelGraphics, home » showroom » virtual sightseeing tours » Stanford. up to virtual sightseeing tours. Montmartre. Stanford. site search: subscription: ...
www.parallelgraphics.com/showroom/virtual-tours/stanford/ - 18k - Mar 28, 2004 - Cached
The Stanford SUIF Compiler Group
... 1994. Publications. Previous research. Affine Transformations for Optimizing Parallelism and Locality. stanford.edu / cs.stanford.edu / ee.stanford.edu.
suif.stanford.edu/ - 3k - Cached
Internet

FIG. 8

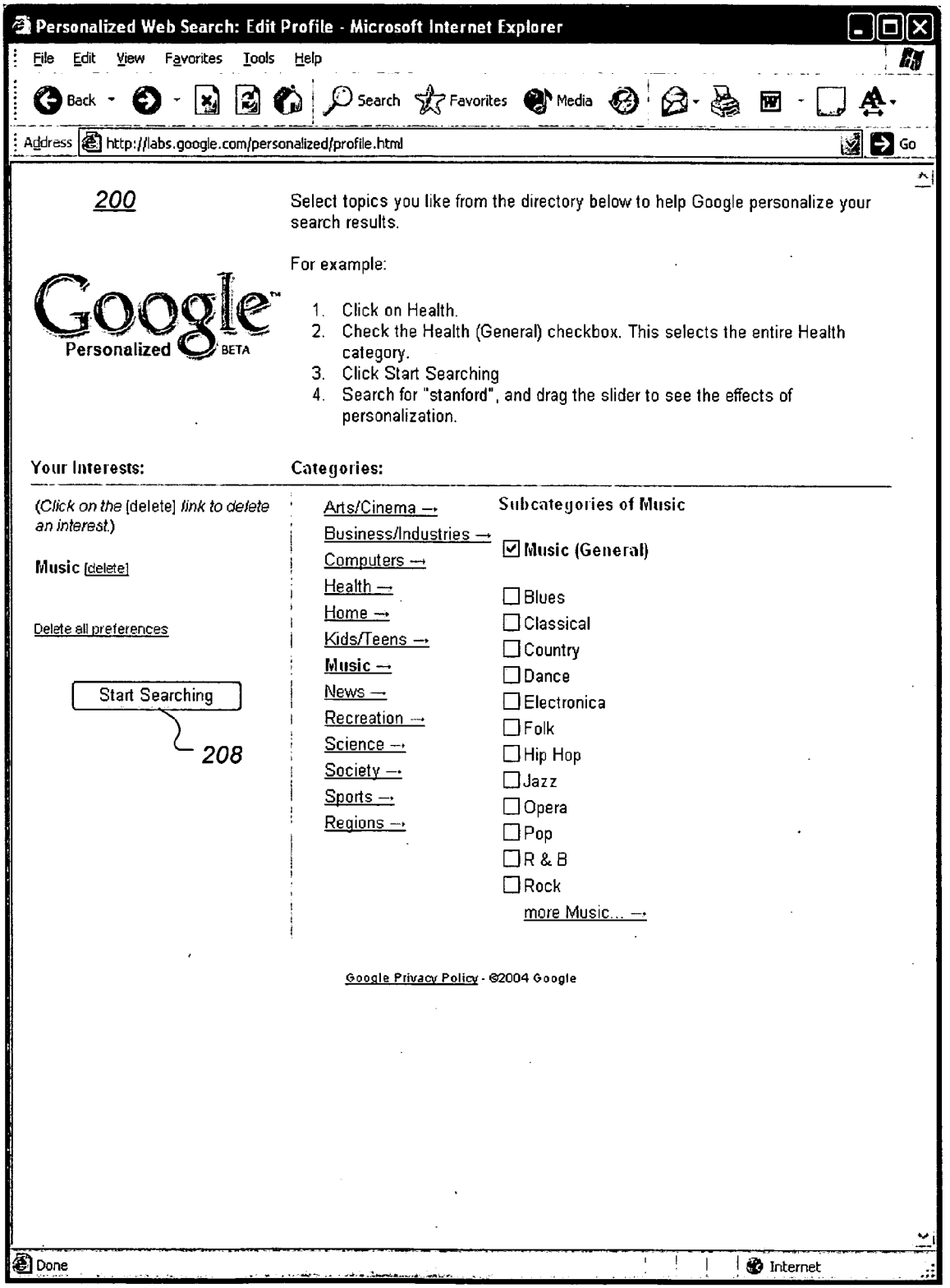

Personalized Web Search: Edit Profile - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/profile.html
Go
200
Select topics you like from the directory below to help Google personalize your search results.
For example:
1. Click on Health.
2. Check the Health (General) checkbox. This selects the entire Health category.
3. Click Start Searching
4. Search for "stanford", and drag the slider to see the effects of personalization.
Google Personalized BETA
Your Interests:
(Click on the [delete] link to delete an interest.)
Music [delete]
Delete all preferences
Start Searching
208
Categories:
Arts/Cinema →
Business/Industries →
Computers →
Health →
Home →
Kids/Teens →
Music →
News →
Recreation →
Science →
Society →
Sports →
Regions →
Subcategories of Music
Music (General)
Blues
Classical
Country
Dance
Electronica
Folk
Hip Hop
Jazz
Opera
Pop
R & B
Rock
more Music... →
Google Privacy Policy - ©2004 Google
Done
Internet

FIG. 9

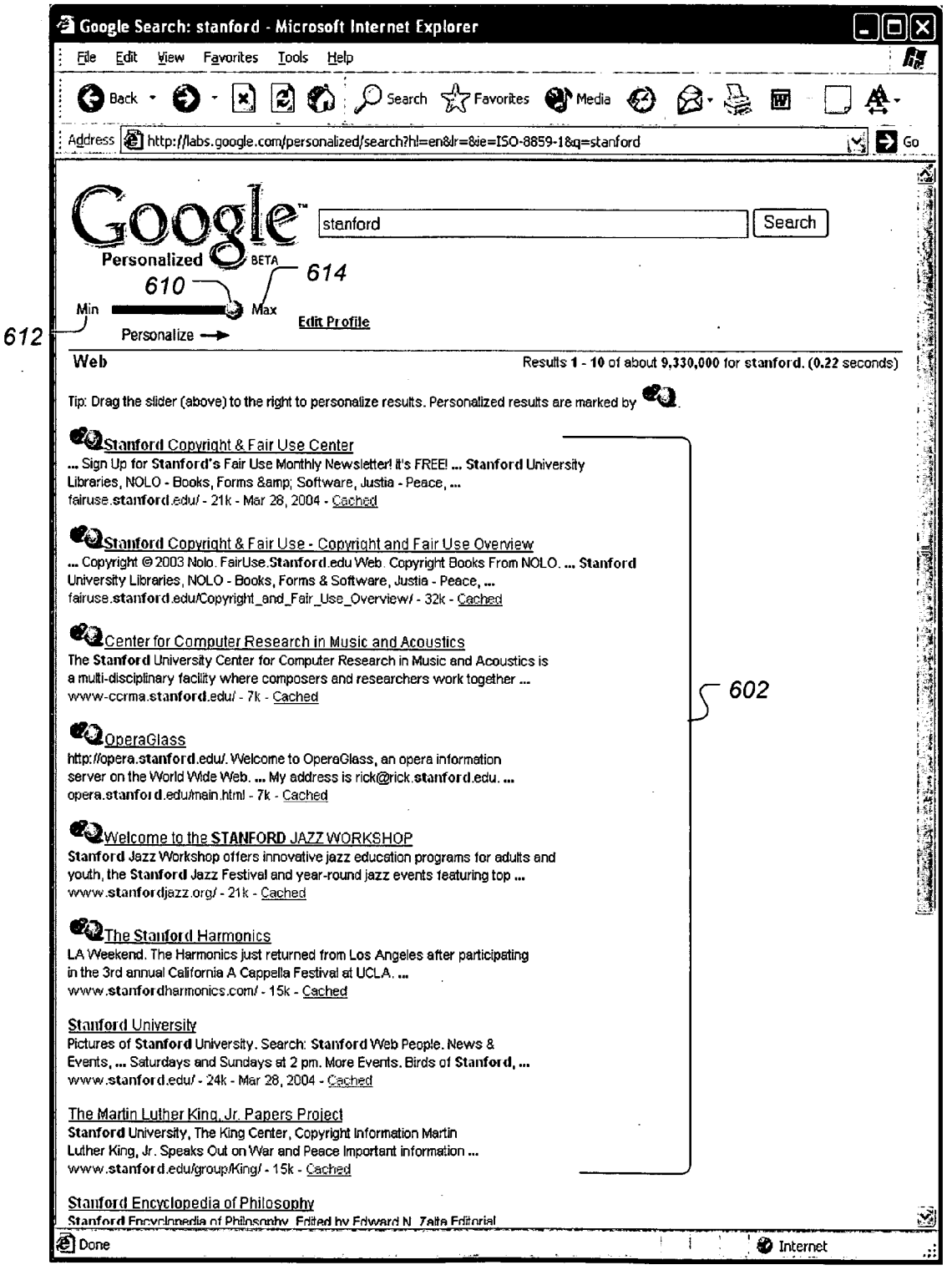

Google Search: stanford - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://labs.google.com/personalized/search?hl=en&lr=&ie=ISO-8859-1&q=stanford
Go
Google Personalized BETA
stanford
Search
610
614
Min
Max
612
Personalize
Edit Profile
Web
Results 1 - 10 of about 9,330,000 for stanford. (0.22 seconds)
Tip: Drag the slider (above) to the right to personalize results. Personalized results are marked by
Stanford Copyright & Fair Use Center
... Sign Up for Stanford's Fair Use Monthly Newsletter! It's FREE! ... Stanford University Libraries, NOLO - Books, Forms & Software, Justia - Peace, ...
fairuse.stanford.edu/ - 21k - Mar 28, 2004 - Cached
Stanford Copyright & Fair Use - Copyright and Fair Use Overview
... Copyright © 2003 Nolo. FairUse.Stanford.edu Web. Copyright Books From NOLO. ... Stanford University Libraries, NOLO - Books, Forms & Software, Justia - Peace, ...
fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/ - 32k - Cached
Center for Computer Research in Music and Acoustics
The Stanford University Center for Computer Research in Music and Acoustics is a multi-disciplinary facility where composers and researchers work together ...
www-ccrma.stanford.edu/ - 7k - Cached
602
OperaGlass
http://opera.stanford.edu/. Welcome to OperaGlass, an opera information server on the World Wide Web. ... My address is rick@rick.stanford.edu. ...
opera.stanford.edu/main.html - 7k - Cached
Welcome to the STANFORD JAZZ WORKSHOP
Stanford Jazz Workshop offers innovative jazz education programs for adults and youth, the Stanford Jazz Festival and year-round jazz events featuring top ...
www.stanfordjazz.org/ - 21k - Cached
The Stanford Harmonics
LA Weekend. The Harmonics just returned from Los Angeles after participating in the 3rd annual California A Cappella Festival at UCLA. ...
www.stanfordharmonics.com/ - 15k - Cached
Stanford University
Pictures of Stanford University. Search: Stanford Web People. News & Events, ... Saturdays and Sundays at 2 pm. More Events. Birds of Stanford, ...
www.stanford.edu/ - 24k - Mar 28, 2004 - Cached
The Martin Luther King, Jr. Papers Project
Stanford University, The King Center, Copyright Information Martin Luther King, Jr. Speaks Out on War and Peace Important information ...
www.stanford.edu/group/King/ - 15k - Cached
Stanford Encyclopedia of Philosophy
Stanford Encyclopedia of Philosophy. Edited by Edward N. Zalta Editorial
Done
Internet

FIG. 10

VARIABLE PERSONALIZATION OF SEARCH RESULTS IN A SEARCH ENGINE

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. §119(e) of Application Ser. No. 60/557,783 filed on Mar. 29, 2004, which is incorporated by reference herein. This application also incorporates by reference U.S. patent application Ser. Nos. 10/646,331 (filed on Aug. 22, 2003) and 10/641,489 (filed on Aug. 14, 2003); all of the above applications are commonly owned along with this application.

FIELD OF INVENTION

[0002] This invention relates in general to online search engines, and in particular to personalization of search results.

BACKGROUND OF INVENTION

[0003] The current state of the art in online search engines is highly advanced in its ability to retrieve documents (e.g., web pages or portions thereof, images, etc.) that are responsive to the terms of a query. Search engines today can quickly retrieve specific documents that match the terms of the query. However, current search engines often return documents that while accurately correspond to the specific terms of the query, do not in any way reflect the user's underlying interests. Thus, two different users, one who is very interested in sports, and another who is interested in politics, will obtain exactly the same results to a given query, say "drug testing in baseball," even though the first user may be more interested in learning about which teams have implemented drug testing, while the latter user is more interested in learning about legislation related to drug testing in sports. Thus, there is a need for providing a mechanism and methodology for personalizing search results in accordance with the interests of the users.

[0004] Further, while a user may have particular set of interests that may be useful in processing their search query, a user may not want to always have such interests influence the query results. Thus, it would be desirable to a provide a mechanism and a methodology by which the user can variably adjust the degree to which his interests influence the results of a given search query.

SUMMARY OF THE INVENTION

[0005] The present invention provides allows a user to variably adjust the degree to which his or her personal interests influence the presentation of results of a given search query. The present invention has multiple embodiments, including a system for providing a search engine, a methodology for processing a search query, and a user interface for the user to variably adjust the degree of influence.

[0006] For a user's perspective, the present invention provides a mechanism for personalizing search results. The user defines a set of interests, for example as a set of topics or categories of interest to the user, as an interests profile. The user then submits a search query to a search engine. The search engine retrieves documents that are relevant to the query and presents these as search results. At first (though not necessarily) the search results are ranked without regard to the user's interests. This enables the user to see the relevance of the documents without influence by his personal interests. The user can variably control ranking of the documents in the search result set. In a preferred embodiment, this is done by direct manipulation of graphical control element, a personalization control. This graphical element can have various graphical representations, and in one embodiment it is represented as a slider. Adjusting the control element adjusts the degree to which the results are ranked according to the user's interests, as expressed in his interest profile. In response to changes in the position of the control element, the search results are reranked, thereby allowing the user to see how his interests affect the ordering of the results.

[0007] In one embodiment, the ranking of search results operates to increase the relevance score of documents that are associated with the directory of topics. Each topic in the directory is associated with a group of documents (e.g., websites) which are deemed relevant to the topic. For each document (again, website), a boost value is provided which indicates the degree of relevance of the site to the topic. All pages within a given site get that site's boost value. The boost value may be understood as a scaling factor that represents the relevance of the site to the topic. During the query processing, the set of documents is returned, and scored without regard to their inclusion in the topic directory. In addition, a separate score for each document is determined in which the document's underlying relevance score is adjusted by its associated boost value if any, for example by multiplying the relevance score by the boost value, thereby scaling the relevance of the document to the query. Generally, only those documents which belong to sites listed in the directory will have a boosted relevance score, and other documents in the search results will not. The reranking of the documents in accordance with their boosted relevance scores is then a function of the position of the personalization control, either from no influence (where the boost value has no effect on the relevance score) to complete influence, wherein boost value for each document is applied 100% to the relevance score, and any number of intermediate levels of application. The user can thus immediately rerank the documents in the search results by adjusting the position of the control.

[0008] A system embodiment includes a topic directory for providing topics (categories) to a user, a search engine and index for searching indexed websites in response to a user's query, and a personalization component for determining for the search results at least one (and preferably a set) of rankings of the search results in accordance with the user's interests. A user interface embodiment includes a topical directory, with various topics, and graphical controls (e.g., checkboxes) in which the user can select topics of interests, and a graphical control that allows the user to adjust the search result ranking in accordance with his interest profile.

[0009] The present invention is next described with respect to various figures and diagrams and technical information. The figures depict various embodiments of the present invention for purposes of illustration only. One skilled in the art will readily recognize from the following discussion that alternative embodiments of the illustrated

and described structures, methods, and functions may be employed without departing from the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] **FIG. 1** is an entry web page for a search engine.

[0011] **FIG. 2** is web page for editing a user's personal profile of topics.

[0012] **FIG. 3** is another web page of topics, and subtopics.

[0013] **FIG. 4** is another web page of subtopics.

[0014] **FIG. 5** shows the entry web page for the search engine with the query "stanford".

[0015] **FIG. 6** shows the initial results of the search on query "stanford", along with a control **608** for controlling a degree of influence of the user's interest profile in ranking the search results.

[0016] **FIG. 7** shows the ranking of the search results of **FIG. 6** after the user has adjusted the control **608** to an intermediate position.

[0017] **FIG. 8** shows the ranking of the search results of **FIG. 6** after the user has adjusted the control **608** to a maximum position.

[0018] **FIG. 9** shows a web page of topics after deletion of a "Computer" topic and selection of a "Music" topic by the user.

[0019] **FIG. 10** shows the ranking of the search results of **FIG. 6** following the change in topic interests in **FIG. 9**.

DETAILED DESCRIPTION

[0020] Referring now to **FIG. 1** there is shown the entry (or landing page) **100** of a search engine. The search engine includes a query entry field **102** in which a user enters the terms of a search query. A search button **106** submits the query to the search engine for processing. The first time the user accesses the search engine, he will not have defined a user profile for personalizing the search results, and a message is displayed to that effect. The user can access the Edit Profile link **104** in order to establish their personal interests.

[0021] Clicking on the Edit Profile link **104** accesses the page **200** illustrated in **FIG. 2**. This page shows the top level topics **204** of a topical directory **202**. The topics in the directory can be, for example, the topics defined in the Open Directory Project, found at http://dmoz.org/ or can be some other topical directory. (In this disclosure the term "topic" and "category" are interchangeable, as are "subtopic" and "subcategory".) The top level topic links may have a set of subtopics, as indicated by the arrows **206**. At any time, the user can begin a search by clicking on the Start Searching button **208**. Clicking on a particular top level topic takes the user to a page associated with that topic's subtopics.

[0022] **FIG. 3** illustrates the result of this step, where the user has clicked on the topic "Computers" on page **200**. As shown in **FIG. 3**, the subtopics **302** of the topic "Computers" have been displayed in the right hand column. Next to each subtopic **302** is a checkbox **314**. The user can select the checkbox **314** to indicate his interest in a specific subtopic. Alternatively, the user can select the checkbox **316** for the entire subtopic itself, "Computers (General)". This indicates that the user is interested in all of the subtopics of the topic "Computers". The set of topics selected by the user may be called a user profile, a personalization profile, or an interest profile. On the left side of **FIG. 3**, after the user has selected one or more topics of interest, these topics are listed in the interests field **306**. Next to each topic name is a delete link **308** that allows the user to delete individual topics. The user can also delete all topics by selecting the "Delete all preferences" link **310**.

[0023] The subtopics themselves can have any number of additional levels of subtopics. For example, **FIG. 4** illustrates that the subtopic "Games" has further subtopics **402**, "Board Games,""Card Games," etc. These were presented to the user in response to the user clicking on the subtopic arrow **206** next to the "Games" topic label. Again, the user can select either the entire subcategory by selecting checkbox **404**, or any of the individual subtopic checkboxes **406**. In one embodiment, in addition to the top level topics **204**, there are two subtopic levels. This provides sufficient breadth and granularity of topical coverage, without imposing unnecessary computational complexity.

[0024] Referring again to **FIG. 3**, the user has indicated his interest in computers generally. Clicking on the Start Searching button **208** brings up search page **100** again, which no longer shows the message requesting the user to establish a personalization profile. At this point, the user enters a search query in the search field, here the query "stanford", as shown in **FIG. 5**. The search engine performs a search of its index and provides the search results. **FIG. 6** illustrates the initial search results page **600**. Here the search results **602** comprise a set of documents **604**, each document indicated by a URL **617**, along with a link **615** to the page, and a snippet **619** of the page text to show a portion of the context to the document to the user. Near the top of the page is the personalization control **608**. In the context of this application, "documents" are understood to be any type of media that can be indexed and retrieved by a search engine, including web pages, images, multimedia files, text documents, PDFs, and so forth. No limitation as to the scope of the invention is implied by the use of the generic term "documents."

[0025] The personalization control **608** allows the user to adjust the ranking of the individual documents/pages of the search results with respect to his interests, preferably as expressed in the personalization profile. The personalization control may be explained as controlling a degree of influence or contribution of the user's interest profile to the ranking of the search results. **FIG. 6** illustrates the control **608** in an embodiment where it is a slider type control that includes a control handle **610** that is moved along a linear dimension between a minimum value **612** to a maximum value **614**. This type of control **608** is generally very intuitive and easy for users to understand how it operates. Alternative embodiments include dial type controls, similarly providing a variable adjustment between a minimum and maximum value, and radio buttons each with an associated predetermined adjustment amount (e.g., five radio buttons, associated respectively with 0%, 20%, 40%, 80%, and 100%). Those of skill in the art of user interface design will appreciate that other types of controls may be implemented to achieve essentially the same effect.

[0026] In the example of **FIG. 6**, the personalization control **608** is set all the way to its minimum value **612**. The search results accordingly are ranked on their relevance to the query and query independent information such as PageRank, as disclosed in U.S. Pat. No. 6,285,999, "Method for node ranking in a linked database", which is incorporated by reference herein in its entirety and for all of its teachings. Notice that the search results are generally related to Stanford University.

[0027] Referring now to **FIG. 7**, there is shown the ranking of the search results **602** after the user has adjusted the slider **610** to an intermediate position about halfway between the minimum and maximum values. Now, the individual documents **604** are reranked, and those that are more related to the user's interest (previously identified as "Computers (General)" are more highly ranked. In order to for the user to distinguish the documents that are more related to his interests, those documents are indicated by a personalization icon **618** next to the link text **615**. Notice that with the intermediate position of the slider **610**, about every other document is a personalization result, though this is not required. Notice further that the snippet **619** text has text that relates to computers (e.g., "hacking") even though that term was not included in the original search query.

[0028] Referring now to **FIG. 8**, there is shown the ranking of the search results **602** after the user has adjusted the slider **610** to its maximum position **614**. First, in this example all of the documents **604** in the search results **602** reflect the user's personalization profile, and are annotated with the personalization icon **618**. Second, the documents in the search results relate to computers and Stanford University, thereby reflecting the user's interests in computers, as it relates to the query "stanford". For example, the first two results are related to "hacking" at Stanford; the next three documents relate specifically to various computer science projects at Stanford. Thus the third document is for Stanford's Artificial Intelligence Laboratory, the fourth document is for Stanford's Center for the Study of Language and Information, and the fifth document is for Stanford's Database Group. The remaining documents are likewise related to the "Computers" interest of the user. Notice further, that what was previously the first (e.g., most relevant) result, the link to Stanford University's home page, when there was no personalization applied (**FIG. 6**) and when only intermediate personalization was applied (**FIG. 7**), no longer even appears in the top eight results.

[0029] In the examples of **FIGS. 6-8** the user can dynamically adjust the ranking of the documents in real time, without the search engine having to recompute the search results. This further aids the user's ability to explore the search results.

[0030] As shown on **FIGS. 6-8**, an Edit Profile link **616** allows the user to edit the interests in his user profile at any time. **FIG. 9** illustrates where the user has returned to the topic directory page **200** as illustrated in **FIG. 2**, and deleted his existing topic interest in Computers, and created a different interest in "Music". The user then again selects the Start Searching button **208**. The search engine provides an updated set of search results, which documents will be the same as before, since the query term has not changed. However, as of the last search, the user has positioned the control **610** for full personalization, and hence the search engine applies this setting and ranks the documents according the existing (and new) personalization profile, this time for "Music."**FIG. 10** illustrates the dramatic difference this makes in the search results. Here, the top six results are all related to music and arts, and none of the previous results relating to computers make the top of the search result list. This shows how the user's personalization profile can significantly alter the ranking of a given set of documents.

[0031] In one embodiment of the invention, the personalization of search results is provided by the following methodology. Those of skill in the art of search engine design can readily understanding the basic structure of a search engine and its operation in which this methodology can be implemented.

[0032] 1) When the user selects topics (categories), they are added to a cookie containing a list of his interests IN (where IN is short for "interests"). The interests can be expressed as either text tokens (e.g. "computers"), or as indices into a directory, where each topic/subtopic has an index number. When the user accesses the site, the system retrieves this interest list from the cookie. Alternatively, the site can provide a user registration page, in which the user registers with a user name and password, and the user's interests are stored at the search engine site in a database, and retrieved in response to the user login.

[0033] 2) Each time the user issues a search query Q, the list IN is sent to a frontend web server.

[0034] 3) Then the frontend web server issues two requests to a backend search engine (Q, null) and (Q, IN), which returns to the frontend web server two results sets, R_base (normal) and R_in (personalized), respectively.

[0035] 4) The base query (Q, null) is the same as searching for Q on the search engine without any personalization, so R_base is the result set that the search engine would normally return in response to the query. As noted above, the search results accordingly are ranked on their relevance to the query and any other factors, such as PageRank, as disclosed in U.S. Pat. No. 6,285,999, "Method for node ranking in a linked database", which is incorporated by reference herein in its entirety and for all of its teachings.

[0036] 5) The personalized query (Q, IN) contains the user's interest list. The search engine follows a similar processing control flow as a normal query, but it activates a personalization component during the scoring function, which "boosts" documents that are relevant to the set of interests IN.

[0037] In one embodiment of the invention, these boosts are implemented using a set of maps (one map for each topic/subtopic in the directory), of the form:

[0038] {Site-->boost}

[0039] So for example, the category "Health" has a map of the form

[0040] {Site-->boost}

[0041] which contains entries such as:

[0042] nih.gov—>5.8

[0043] cdc.gov—>7.9

[0044] med.stanford.edu—>3.5

[0045] where nih.gov, cdc.gov, and med.Stanford.edu are various sites that have been determined (either manually or automatically) to be related to the topic "Health". In other words, for each topic in the directory, there is a set of sites that have been determined to be relevant to the topic, and for each of these sites, a boost value is defined.

[0046] The boost for the sites listed in the topical directory is generally determined as follows:

[0047] a) A "site graph" is generated where nodes of the graph are sites (basically, pages on the same host) and edges between nodes are weighted based on the number of pages from one site that link to pages on another. This same type of graph can be used to compute all topic boost maps.

[0048] b) For each topic in the directory, say "Health", a number of sites are selected as "start sites" **S0** whose home page is listed in the Open Directory. For example, for a university like Stanford, start sites may be selected as any site ending in .stanford.edu.

[0049] c) A computation is run in two passes:

[0050] i) first identify a set of sites **S1** that are linked-to heavily by those sites in **S0**, with each site in **S1** assigned a weight according to how heavily it's linked-to by sites in **S0**.

[0051] ii) then identify those sites **S2** that are linked-to heavily by those sites in **S1**, weighted as with **S1**.

[0052] d) The sites in **S2** are boosted with their assigned weights.

[0053] When a personalized query comes in with an interest set (e.g., {Health, Music}), then the personalization component consults the respective maps, and then for query results (i.e. documents) that are in these maps, the personalization components boosts the IR score for them, so that they would get ranked higher. The IR score is the information retrieval score that the search engine provides for each document, and can be a function of document relevance and other factors, such as PageRank.

[0054] For example if http://www.med.stanford.edu/research/ were one of the results for query Q, and IN included "Health," then that result would get a multiplicative boost of 3.5 to its IR score, because that page is on the site med.stanford.edu. For multiple interests, the values in the corresponding maps are multiplied together for a combined effect. Further details of one implementation for determining the boost for sites in the directory boost map are described in U.S. patent application Ser. No. 10/646,331 incorporated by reference herein.

[0055] 6) The frontend web server receives R_base and R_in. R_base represents the results **602** the user sees when the personalization control **608** is set at 0 (or its minimum position **612**). R_in represents the results **602** the user sees when the slider is at its maximum position **614**, e.g. 10. In one embodiment, the control has 11 positions, but any number of positions may be used. The rankings that are shown to the user between 0 and 10 are interpolated; this means using R_base and R_in, the personalization component constructs intermediate rankings (R_1 . . . R_9) by moving each individual result along a virtual path from its position when the control **608** is at its minimum position to when the control **608** is at its maximum position.

[0056] 7) How the slider effect is provided to the user:

[0057] The frontend web server appends the results in R_in that are not in R_base to R_base to get a list of all unique results (either in R_base or R_in) called R_all. The frontend web server then assigns consecutive identification numbers to the results in R_all that serve as identifiers. The frontend then sends to the user's browser two items:

[0058] a) the mapping {id->result} for each of the results in R_all, containing for each document in the results its identification number; and

[0059] b) N vectors that denote the rankings the user should see, in terms of result IDs, where N is the number of positions of the slider, and each vector represents the ordering of the results based on the degree of personalization. In this example, there would be 11 vectors corresponding to positions **0** through **10** of the control **608**. In one embodiment Vector 0 is [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10], reflecting the un-personalized ranking of the results, assigned with the local IDs starting with R_0. Vector 10 is the ordering in the fully personalized results R_in; for instance, it might be [15, 17, 5, 0, 1, 2, 3, 4, 8, 9, 10], where documents 15 and 17, for example, are documents in the relevant map of the topic directory whose IR score was boosted according to the boost values in the map, and the value associated with the control **608**'s position. Transmitting the N ranking vectors to the client allows the client browser to quickly recomputed the rankings, rather than imposing that calculation on the server.

[0060] 8) Finally, Javascript code in the result page causes the user's browser to reorder the search results using the appropriate one of the 11 positional ranking vectors in response to the user changing the position of the control **608**. Further details of one embodiment of steps 6-8, above, are described in U.S. application Ser. No. 10/641,489, incorporated by reference herein.

[0061] The present invention has been described in particular detail with respect to one possible embodiment. Those of skill in the art will appreciate that the invention may be practiced in other embodiments. First, the particular naming of the components, capitalization of terms, the attributes, data structures, or any other programming or structural aspect is not mandatory or significant, and the mechanisms that implement the invention or its features may have different names, formats, or protocols. Further, the system may be implemented via a combination of hardware and software, as described, or entirely in hardware elements. Also, the particular division of functionality between the various system components described herein is merely exemplary, and not mandatory; functions performed by a single system component may instead be performed by multiple components, and functions performed by multiple components may instead be performed by a single component.

[0062] Some portions of above description present the features of the present invention in terms of algorithms and

symbolic representations of operations on information. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. These operations, while described functionally or logically, are understood to be implemented by computer programs. Furthermore, it has also proven convenient at times, to refer to these arrangements of operations as modules or by functional names, without loss of generality.

[0063] Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "calculating" or "determining" or "identifying" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system memories or registers or other such information storage, transmission or display devices.

[0064] Certain aspects of the present invention include process steps and instructions described herein in the form of an algorithm. It should be noted that the process steps and instructions of the present invention could be embodied in software, firmware or hardware, and when embodied in software, could be downloaded to reside on and be operated from different platforms used by real time network operating systems.

[0065] The present invention also relates to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general-purpose computer selectively activated or reconfigured by a computer program stored on a computer readable medium that can be accessed by the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus. Those of skill in the art of integrated circuit design and video codecs appreciate that the invention can be readily fabricated in various types of integrated circuits based on the above functional and structural descriptions, including application specific integrated circuits (ASIC). In addition, the present invention may be incorporated into various types of video coding devices.

[0066] The algorithms and operations presented herein are not inherently related to any particular computer or other apparatus. Various general-purpose systems may also be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will be apparent to those of skill in the art, along with equivalent variations. In addition, the present invention is not described with reference to any particular programming language. It is appreciated that a variety of programming languages may be used to implement the teachings of the present invention as described herein, and any references to specific languages are provided for disclosure of enablement and best mode of the present invention.

[0067] Finally, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter. Accordingly, the disclosure of the present invention is intended to be illustrative, but not limiting, of the scope of the invention.

We claim:

1. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents in response to an interest profile of the user and a degree of influence of the interest profile to the ranking; and

presenting the ranked documents to the user.

2. The method of claim 1, wherein presenting the ranked documents further comprises:

displaying a personalization icon next to each document associated with an interest that is included in the interest profile of the user.

3. The method of claim 1, further comprising:

receiving from the user an input of a degree of influence which the user's interest profile is to have on the ranked documents;

determining a second ranking of the set of documents in response to the degree of influence; and

presenting the documents with the second ranking to the user.

4. The method of claim 3, wherein determining the second ranking is performed without retrieving a second set of documents relevant to the query.

5. The method of claim 3, wherein the second ranking is determined prior to receiving the user input of the degree of influence.

6. The method of claim 3, wherein the second ranking is determined concurrently with determining the first ranking.

7. The method of claim 3, wherein the user input of the degree of control is a change of position in a graphical representation of a control.

8. The method of claim 7, wherein the control is a slider.

9. The method of claim 7, wherein the control is a dial.

10. The method of claim 7, wherein the control is radio button.

11. The method of claim 1, wherein determining a first ranking further comprises:

retrieving the interest profile of the user from a cookie.

12. The method of claim 1, wherein determining a first ranking further comprises:

retrieving the interest profile of the user from a database communicatively coupled to a search engine.

13. The method of claim 1, wherein the interest profile of the user comprises a set of text tokens, each token corresponding to a topic of interest.

14. The method of claim 1, wherein the interest profile of the user comprises a set of topic index numbers, each index number corresponding to a topic of interest.

15. The method of claim 1, wherein determining a first ranking comprises:

scaling the information retrieval score of at least one document in the set by a boost value associated with the

document in response to the document being associated with an interested included in the interest profile of the user.

16. The method of claim 1, wherein each document has an information retrieval score, and determining a first ranking comprises:

for each document associated with a topic included in the interest profile of the user, determining a boost value for the document, and scaling an information retrieval score of the document by the boost value and the degree of influence; and

ranking the documents in the set according to their information retrieval scores.

17. The method of claim 1, wherein determining a first ranking comprises:

for at least one document in the set of documents:

determining whether a website containing the document is associated with a topic included in the interest profile of the user; and

scaling an information retrieval score of the document by the boost value associated with the website containing the document.

18. The method of claim 1, wherein determining a first ranking comprises:

for at least one document in the set of documents:

determining whether the document is included in a first website linked to by a second website, the second website associated with a topic included in the interest profile of the user; and

scaling an information retrieval score of the document by a weight value associated with the first website.

19. The method of claim 18, wherein the weight value is based on a number of links from the second website to the first website.

20. The method of claim 1, wherein determining a first ranking comprises:

for a document associated with a plurality of topics included in the interest profile of the user:

determining a website containing the document;

determining for each topic associated with the website, a boost value for a website; and

determining a final boost value for the document as a function of the boost values; and

scaling the information retrieval score of the document by the final boost.

21. The method of claim 1, wherein determining a first ranking comprises:

determining for at least one document, a boost value from a topic directory comprising a plurality of topics, each topic associated with at least one website, each website having a boost value, wherein the interest profile of the user contains at least one topic from the topic directory; and

scaling an information retrieval score of the least one document by its boost value.

22. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents independently of an interest profile of the user;

presenting the ranked documents to the user;

receiving from the user an input providing a degree of influence which the user's interest profile is to have on the ranked documents; and

presenting the documents to the user with a second ranking according to the degree of influence.

23. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents according to a minimum degree of influence of an interest profile of the user;

determining a second ranking of the set of documents according to a maximum degree of influence the interest profile of the user;

presenting the first ranking of the documents to the user;

receiving a user input providing a selection of a maximum degree of influence of the user's interest profile;

presenting the second ranking of the documents to the user.

24. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents according to a maximum degree of influence the interest profile of the user;

determining a second ranking of the set of documents according to a minimum degree of influence of an interest profile of the user;

presenting the first ranking of the documents to the user;

receiving a user input providing a selection of a minimum degree of influence of the user's interest profile;

presenting the second ranking of the documents to the user.

25. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents according to a minimum degree of influence of an interest profile of the user;

determining a second ranking of the set of documents according to a maximum degree of influence the interest profile of the user;

presenting the first ranking of the documents to the user;

receiving a user input providing a degree of influence of the user's interest profile;

determining a third ranking of the documents by interpolating between the first ranking and the second ranking according to the received degree of influence relative to the minimum and maximum degrees of influence.

26. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

determining a first ranking of the set of documents in response to an interest profile of the user and a degree of influence of the interest profile to the ranking, wherein the interest profile includes at least one topic of interest to the user;

presenting the ranked documents to the user;

receiving from the user a modification to interest profile, the change comprising at least one of a) at least one topic added to the interest profile, and b) at least one topic deleted from the interest profile;

determining a second ranking of the set of documents in response to the modified interest profile; and

presenting the documents with the second ranking to the user.

27. A computer implemented method of personalizing results of a search query to a user, comprising:

at a server:

retrieving a set of documents relevant to a query of a user;

determining a first ranking of the set of documents according to a minimum degree of influence of an interest profile of the user;

determining a second ranking of the set of documents according to a maximum degree of influence the interest profile of the user;

determining a plurality of ranking vectors, including a first ranking vector describing the first ranking of the set of documents, a second ranking vector describing the second ranking of the set of documents, and at least one intermediate ranking vector describing a third ranking of the set of documents interpolated between the first and second rankings;

at a client:

presenting the first ranking of the documents to the user according to the first ranking vector;

receiving a user input providing a degree of influence of the user's interest profile;

selecting one of the plurality of ranking vectors corresponding to the received user input of the degree of influence; and

presenting the set of documents according to the selected ranking vector.

28. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user;

ranking the set of documents according a set of topics selected by the user as representative of the user's interests and to a degree of influence of the selected topics; and

presenting the ranked documents to the user.

29. A computer implemented method of personalizing results of a search query to a user, comprising:

retrieving a set of documents relevant to the search query of the user, each document having an information retrieval score;

boosting the information retrieval score of at least one document in the set according to a boost value assigned to a website related to the document, where the website is associated with at least one topic in a set of topics selected by the user as representative of the user's interests;

ranking the set of documents according to their information retrieval scores, including both documents with boosted information retrieval scores and unboosted information retrieval scores; and

presenting the ranked documents to the user.

30. A user interface for a search engine, comprising:

a topical directory including a plurality of topics, including topics selectable by a user to indicate an interest in the topic for defining an interest profile of the user; and

a search results page, that lists search results comprising a plurality of documents relevant to a search query of a user, and including a graphical control element that enables the user to adjust a ranking of the documents in accordance with the interest profile of the user.

31. The user interface of claim 30, wherein the topical directory further comprises:

each topic associated with a plurality of websites, each website having an associated boost value for boosting an information retrieval score of a document related to that website when the document is relevant to the search query.

32. The user interface of claim 30, wherein the graphical control is a slider.

33. The user interface of claim 30, wherein the graphical control is a dial.

34. The user interface of claim 30, wherein the graphical control is radio button.

35. The user interface of claim 30, wherein the graphical control has a plurality of positions, each position associated with a degree of influence of the user's interest profile.

36. A search engine system, comprising:

a topical directory including a plurality of topics, including topics selectable by a user to indicate an interest in the topic for defining an interest profile of the user;

a search engine for retrieving search results comprising documents in response to a query provided by a user and ranking the documents according to relevance to the query; and

a personalization component that determines at least one ranking of the documents according to the user's interest profile.

* * * * *