

Contact Information

Publications@ipcg.com

ipCapital Group, Inc.
400 Cornerstone Drive, Suite 325
Williston, VT 05495
United States of America
(802) 872-3200

TITLE

Retrieving Highly Relevant Data using a Natural Language Search

ABSTRACT

An OS search filter can learn the commonly executed preferences of an OS level user. When a user executes a database search or internet search, the user may exhibit tendencies to open certain file types or use certain search parameters. The operating system can learn these tendencies and automatically apply filters to retrieve results that conform to these parameters.

1. BACKGROUND

Problem or Opportunity

In using a database or web browser, users want to locate stored information that is highly relevant to their query. Current search methodologies rely heavily on the use of Boolean operators which can be unwieldy or confusing for a user. An alternative methodology is called "natural language searching". "Distinctly different from the current, non-linguistic search processing models in use, natural language search uses ordinary, familiar phrases and sentences, not 'computerese.' This approach typically leads to more accurate and effective search results."¹ However, one of the major difficulties with natural language searching is ambiguity in written and spoken language that is normally resolved via context, grammar, tone or other clues.² Due to these limitations, the results retrieved via current natural language searching systems are suboptimal. In response to this, a better natural language search system is needed.

Background Publications

Described below are a number of systems and methods that attempt to provide a solution to the problem described above. However, these systems and methods are unable to fully address the problem of retrieving highly specific search results using natural language searches.

¹ <http://www.powerset.com/news/parc.pdf>

² http://www.parc.com/research/projects/natural_language/

PARC Research is creating linguistic algorithms to understand the search terms context and rewrite the search based on its interpretation of natural language terms.³ It remains to be seen whether this approach can effectively increase the accuracy of natural language searching.

Google appears to be focusing on the order and importance of words in the natural language search, rather than improving the natural language interface.⁴ This is a very different approach and involves changing the web browser.

Both of these solutions involve overhauling the search engine rather than the user OS.

US Patent Number 7403938 details an invention which breaks natural language queries into tokens, and processes them based on patterns in those tokens. This invention is extensible to retrieving information from an operating system. However, it does not use the operating system to keep historical records of a user's queries and search results to increase the accuracy of natural language searches.

2. SUMMARY OF INVENTION

Invention Summary

This invention uses a new technique to increase the relevance of retrieved results from natural language searching. This invention involves the Linux OS acting as a search filter that can be turned on or off. The OS keeps a record of a user's previous searches. This record contains information about the user's searching preferences such as: document types; information source type (i.e. blog, .com site, major news outlet); date limitation etc. After a statistically significant number of searches, wherein the OS can accurately determine the user's preferences, the OS compiles this information and serves as a filtering mechanism wherein it interacts with the browser software to compile and display only the search results that match the user's previous search habits.

In an alternative embodiment, the OS identifies the file types that can be opened by the User Machine. This is done by monitoring the document reading applications installed on the machine. When a search is executed, the OS filters out any documents that are not supported by the User Machine.

Unique Concepts

This invention has the unique ability for smart filtering to occur at the OS level to reflect user preferences or capabilities.

3. DESCRIPTION OF THE INVENTION

³ http://www.parc.com/research/projects/natural_language/

⁴ <http://www.searchenginejournal.com/natural-language-search-not-a-priority-for-google/6130/>

Figure 1 depicts the system for Retrieving Highly Relevant Data using a Natural Language Search.

The User Device consists of any personal computing device, such as a desktop computer, laptop, or mobile device.

The I/O Devices are input and output devices through which the user and User Machine may communicate. I/O Devices may include a keyboard, mouse, monitor, or removable disk drive.

The OS is an operating system, such as Linux, residing on the User Machine. The OS manages interactions with hardware, applications, and the user.

The Search Program is a program designed to identify and retrieve information from a data source such as a web browser.

The Adaptive Filter is a learning module that is contained within the OS that stores information about user searches. When the Adaptive Filter is turned on, it can filter results retrieved by the Search Program to incorporate user preferences.

The User Preferences Registry is a module within the Adaptive Filter where the OS stores the User's search preferences.

The On/Off Switch is controlled by the User and can turn the Adaptive Filter on or off.

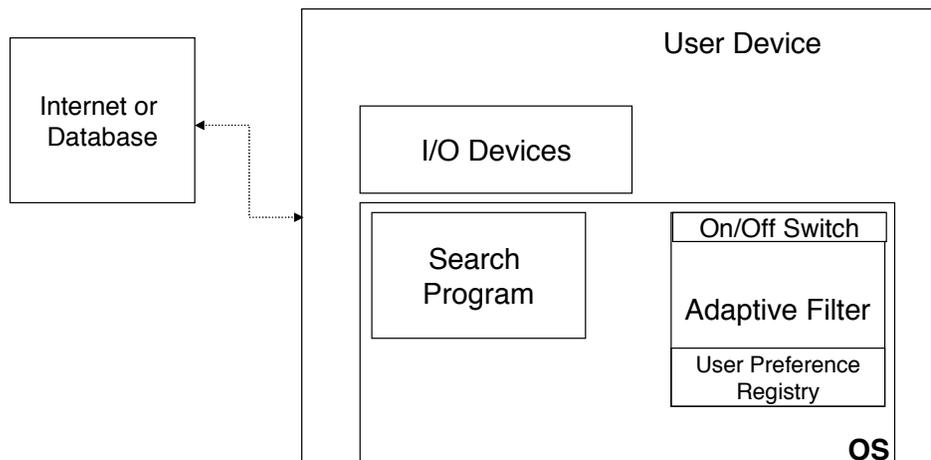


Figure 1. System for Retrieving Highly Relevant Data using a Natural Language Search

Figure 2 provides a flowchart depicting the operation of the Adaptive Filter for Highly Relevant Natural Language Search Results.

In step 1, the User activates a Search Program to retrieve information from the Internet or a Database.

In step 2, if the Adaptive Filter has been activated by the On/Off Switch, proceed to step 3. If the Adaptive Filter is not active, proceed to step 4.

In step 3, the Adaptive Filter filters Search Program results to display based on stored user preferences.

In step 4, the User chooses files to open or web sites to view based on the results retrieved by the Search Program.

In step 5, the Adaptive Filter records the commonalities of which search result the User views and saves the User's preferences within the User Preference Registry.

In step 6, the Adaptive Filter uses the User's choices stored within the User Preference Registry to establish parameters with which to filter future search results.

In step 7, if the User desires to make another search return to step 1. If not, the method ends.

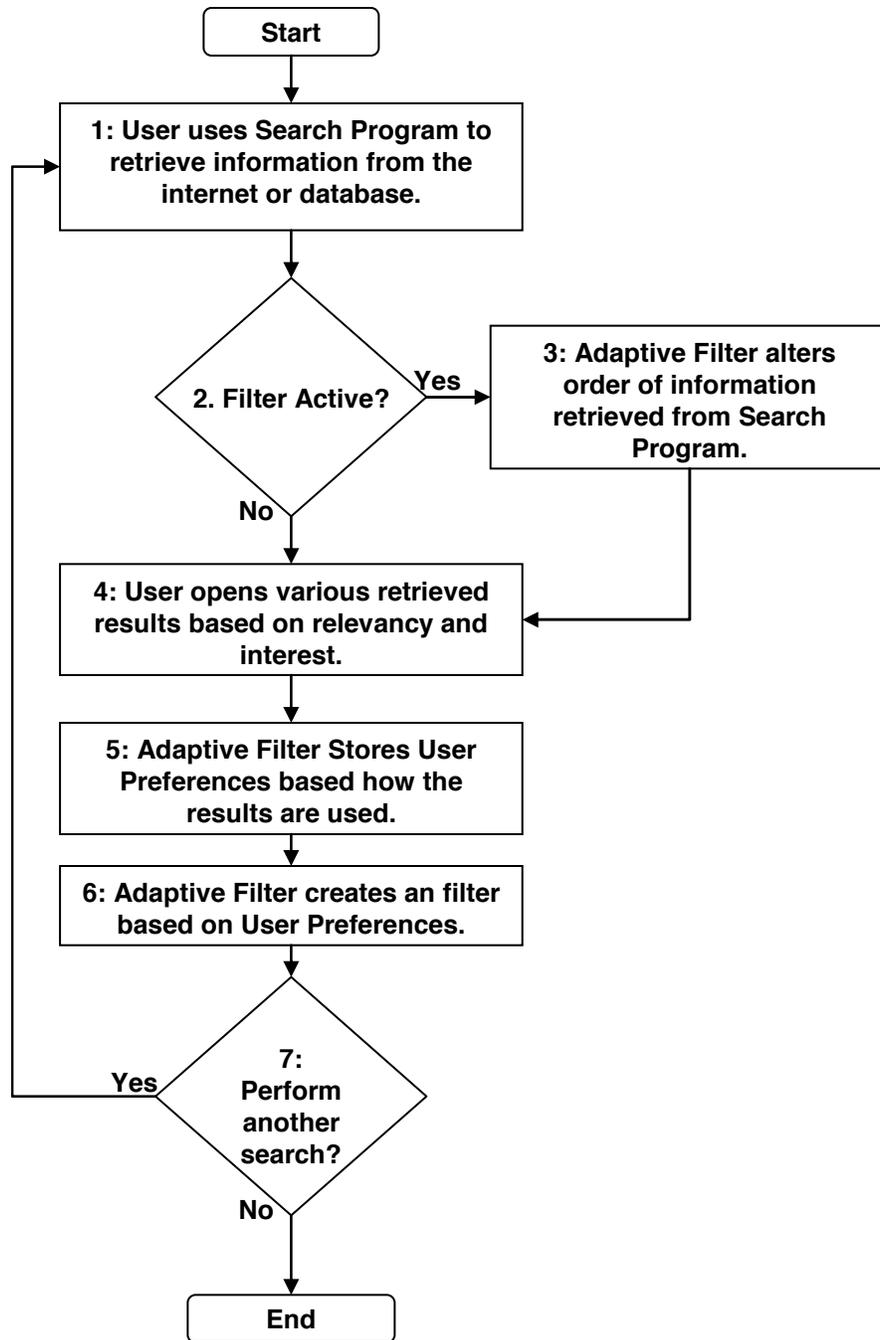


Figure 2. Method for Retrieving Highly Relevant Data using a Natural Language Search