

System and method for searching, storing and publishing information

Andre S. Lessa
andre.lessa@lessaworld.com

Feb/01/2005

ABSTRACT

This paper describes a system and method for searching, storing and publishing information. The present invention generally comprises of a system that is designed to allow users to store and publish content as well search for information based on relevant information gathered from analyzing any previously known content created by the user or found to be relevant to the user.

Note: The content of this document has been used as the foundation for a provisional patent filled on February 1, 2005.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates generally to the field of systems for use in the Internet industry in order to better prioritize relevant content provided to the user. Moreover, it pertains specifically to such systems for the online searching and publishing industries where it has become required to understand what the user perceives as relevant prior to dumping an incredible amount of non-relevant information on the user interface.

Description of the Related Art

Currently, there are no systems capable of combining the experiences of searching, storing, and publishing information while at the same time being able to prioritize content that is relevant to the user who is accessing the system.

The search industry is growing as never. However, all companies rely on the same approach to provide results. The ranking of the search results are based on the general concept that what important for most people is going to be important for everyone.

The electronic publishing industry is also going through an expansion phase by means of the Blog industry. However, as more information is posted on the web, the more difficult it gets to expose the most relevant content to the online user. So many content sources are available today that it has become extremely hard to filter out unwanted information while being able to preserve content that matches the user preference.

One of the biggest challenges in the market today is to bring content personalization to the searching arena.

In order to provide the personalized results necessary to enhance the user experience, a platform needs to be created to not only store information deemed to be relevant to users but also use such information to identify other sources of content that might be identified to be relevant to the user.

BRIEF SUMMARY OF THE INVENTION

In view of the limitations now present in the prior art, the present invention provides a new and useful system and method for searching, storing and publishing information which is simpler in construction, more universally usable, unique and more versatile in operation than known systems of this kind.

The purpose of the present invention is to provide a new system and method for searching, storing and publishing information that has many novel features not offered by the prior art systems that result in a new system and method for searching, storing and publishing information which is not apparent, obvious, or suggested, either directly or indirectly by any of the prior art systems.

The present invention generally comprises of a system that is designed to provide a simple and straight-forward electronic user interface that allows users to store and publish content as well search for information based on relevant information gathered from analyzing any previously known content created by the user or found to be relevant to the user.

By storing content, the user submits the content for storage in a remote storage area, and the content becomes accessible to the user only. By publishing the content, the user submits the content to a remote storage area, but this time, the content is made available to the user and others to access. No matter whether the information is being stored or published, the data behind that content reflects what is relevant to the user at that point in time. The third capability uses the relevance concept as its driving factor.

In other words, the search capability offered by this system becomes tightly integrated with the storing and publishing elements as the information being searched is ranked based on what has been demonstrated to be relevant to the user based on the information that has been either stored or published. The system learns what terms, phrases, and pages are more relevant to each unique user.

The foregoing has outlined, in general, the physical aspects of the invention and is to serve as an aid to better understanding the more complete detailed description which is to follow. In reference to such, there is to be a clear understanding that the present invention is not limited to the method or detail of construction, fabrication, material, or application of use described and illustrated herein. Any other variation of fabrication, use, or application should be considered apparent as an alternative embodiment of the present invention.

A principal object of the present invention is to provide a system and method for searching, storing and publishing information that will overcome the deficiencies of the prior art devices.

An object of the present invention is to provide a system and method for searching, storing and publishing information where an interface for user interaction is provided, said interface comprising: a data entry field, a search button, a store button, and a publish button.

An object of the present invention is to provide a system and method for searching, storing and publishing information where personalized search results are generated based on data identified to be relevant to the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where search results are improved based on relevant information identified by a learning mechanism (a learning engine).

An object of the present invention is to provide a system and method for searching, storing and publishing information where a learning mechanism is used to identify information deemed to be relevant to the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where a learning mechanism is used to identify terms, phrases and documents deemed to be relevant to the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where information deemed to be relevant to the user is identified based on information that has been previously stored by the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where information deemed to be relevant to the user is identified based on information that has been previously published by the user or to the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where information deemed to be relevant to the user is identified based on the aggregated feedback obtained from examining the contents of a collection of user documents.

An object of the present invention is to provide a system and method for searching, storing and publishing information where information deemed to be relevant to the user is identified based on the analysis of user-generated information, such as blog entries, emails, instant messages, general documents, and voice transcripts.

An object of the present invention is to provide a system and method for searching, storing and publishing information where information deemed to be relevant to the user is identified based on the analysis of the content of certain documents or sections within such documents as defined by the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved by extracting, following and parsing the content behind document links provided by the user.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved based on information classified by the user as private.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved based on information classified by the user as public.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved based on information obtained from monitoring the user actions that relate to the user interface.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved based on information obtained from examining collections of documents.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data is improved based on user feedback obtained from monitoring the user behavior.

An object of the present invention is to provide a system and method for searching, storing and publishing information where the identification process of the relevant user data may or may not be improved based on the aggregated relevancy of terms, phrases, and documents originated from observing what is relevant to a group of users.

Another object of the present invention is to provide a system and method for searching, storing and publishing information that is more universally functional in today's market than the prior art devices.

It is intended that any other advantages and objects of the present invention that become apparent or obvious from the detailed description or illustrations contained herein are within the scope of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The present invention comprises of a computer system capable of combining general searching, storing and publishing functionalities based on what it understands to be relevant to the general user.

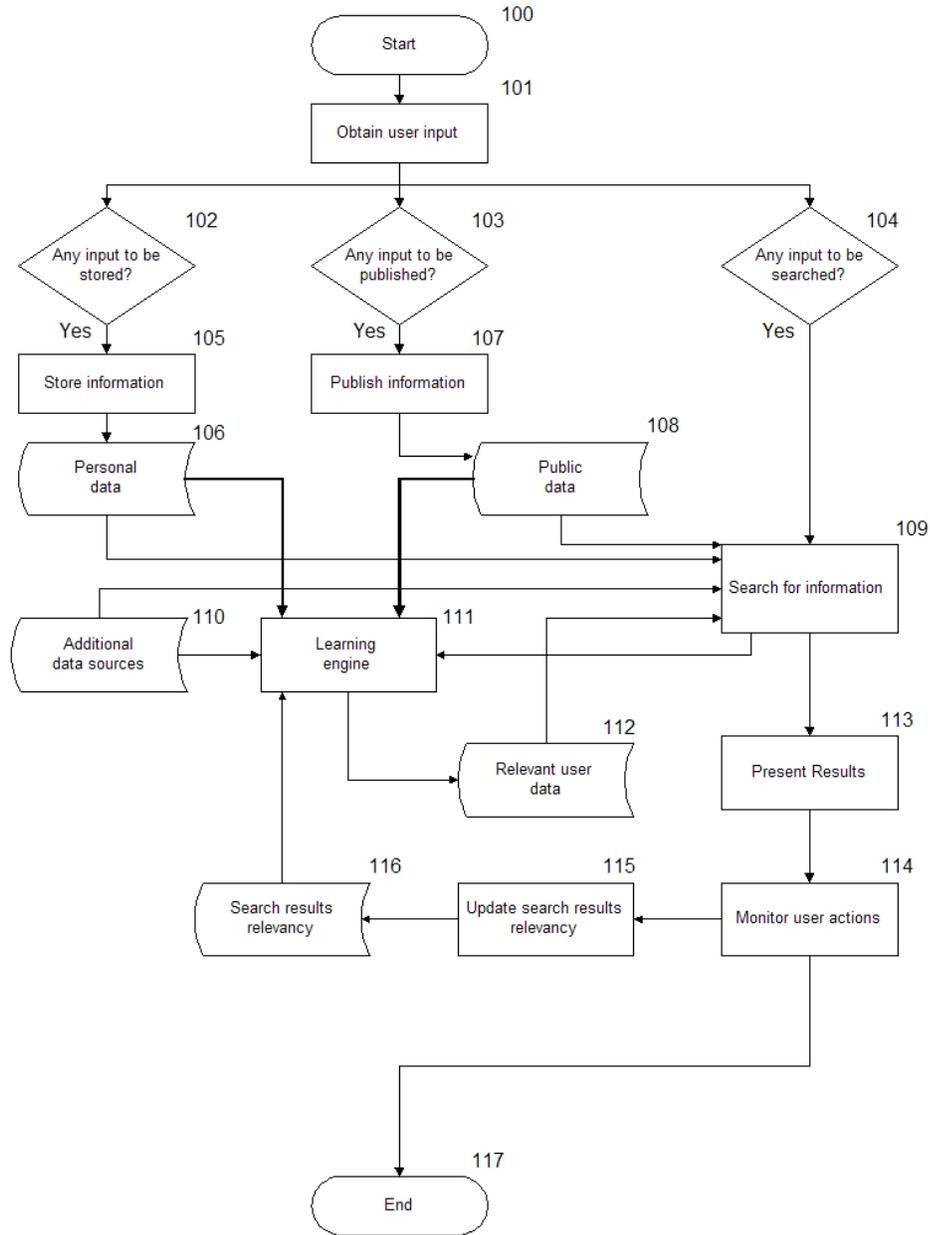


Fig. 01

FIG.1. illustration of the overall process flow of the invention

Fig. 1 illustrates the overall process, which is described next. In one embodiment, once the user connects to a remote system **100**, the application provides an interface that allows the user to type in certain information that is considered to be relevant for the user at that moment. The user interface for obtaining the user data **101** is shared among the searching, storing, and publishing functionalities. The data entry field is the same one for all three options. Along with the proper data entry field, additional interface objects, such as buttons, are provided to allow the user to select the action that needs to be associated with the information been provided.

If the user opts for storing information **102**, the data provided through the user interface is parsed and consequently saved **105** into a data repository of personal data **106**, data itself classified as private to the user.

If the user opts for publishing information **103**, the data provided through the user interface is parsed and consequently saved **107** into a data repository of public data **108** to be made available to all system users, including the original user that posted the information.

If the user opts for searching for specific information **104**, the data provided through the user interface (the query) is parsed **109** and used to feed a learning engine **111** with the terms and/or phrases provided by the user.

The learning engine **111** is used to identify what type of data is relevant to the user and in order to achieve its goal it needs to constantly monitor several distinct sources of relevant information.

- (a) The query itself – Obtained via the user interface **101**.
- (b) Personal Data **106** –Terms and Phrases stored by the user for personal usage.
- (c) Public Data **108** – Terms and Phrases stored by the user for personal and public usage.
- (d) Additional Data Sources **110** - Terms and Phrases of general information that are used to identify relationships and synergies between terms and phrases from the user universe.
- (e) Search Results Relevancy **116** – Terms and Phrases found on selected search results that have been identified as being relevant or not relevant to the user.

By comparing and ranking the different weights of terms and phrases found within all these different sources of data, the learning engine **111** becomes capable of generating a logical representation that shows what type of data is relevant to the end-user **112**.

The end goal of the search process **109** is to be able to look at a general collection of documents originated from one or more data sources **110** and be able to identify the results that better match the type of content the user sees as relevant. Once the result set is identified and properly ordered, it gets presented to the user by means of an electronic user interface **113**.

After presenting **113** the initial list of results to the user, the system starts monitoring the actions **114** performed by the user in order to better understand what results better match what the user is looking for. By identifying the content

behind the material that a user has selected to visit and/or by monitoring the type of feedback created by certain results presented to the user, the system is capable of updating the relevancy level **115** of the terms and phrases contained in some of the results listed in the result set. The terms and phrases identified as being relevant to the user are put aside **116** so they can be used as an additional data feed by the learning engine **111**.

During the time while the system is monitoring the user actions **114**, the system constantly tries to improve the result set made available to the user. Whenever the user decides to cease improving his search, the system understands the user reached its goal and the process is considered to be finalized **117**.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the user will use an Internet-enabled computer device to access the system. By visiting a specific Uniform Resource Identifier, such as an Internet URL, the user will have access to a user interface where information can typed in and the proper action (Search, Store, or Publish) can be initiated.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the user will use an intranet network to access the system. By visiting a specific Uniform Resource Identifier, such as a URL, the user will have access to a user interface where information can typed in and the proper action (Search, Store, or Publish) can be initiated.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the user will use a mobile device to access the system. By using the browsing capabilities of such device, the user will select a specific Uniform Resource Identifier, such as a URL, and then have access to a user interface where information can typed in and the proper action (Search, Store, or Publish) can be initiated.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, when presented with a list of identified relevant documents, the user indirect feedback is used to change the weight of certain terms, concepts, or documents in order to refine the search process. For instance, when a list of ten results is presented, if the user skips the top two selecting the 3rd option, indirectly the user said that the first two documents do not seem able to contain the requested information. Behind the scenes, the system here described is able to identify the terms that belong to the skipped document or documents and reduce their overall relevancy within the context of the results when refining the search.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, when presented with a list of identified relevant documents, the user direct feedback is used to change the weight of certain terms, concepts, or documents in order to refine the search process. For instance, a document may be presented along with an interface aid to allow the user to say whether the document relates to the searched topic or not. Depending on the user feedback, a positive or negative weight might be applied to the document terms when refining the search.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the user might use a collection of documents temporarily or permanently stored in the system in order to initialize the term weights for a specific set of search requests.

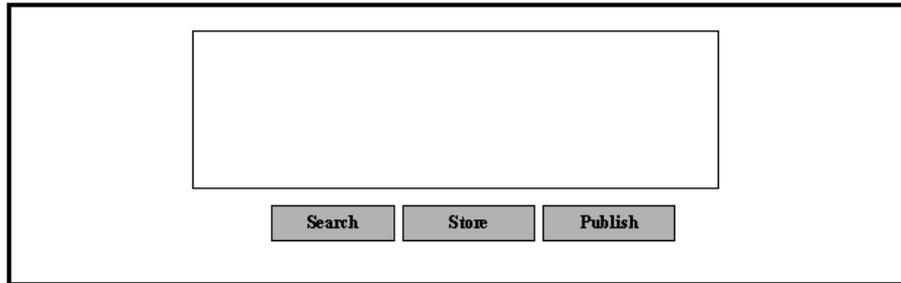


Fig. 02

FIG.2. illustrates a suggested user interface for the system

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the user might use a user interface like the one illustrated in Figure 2 where in addition to the standard Search button, a Store and a Publish buttons have been introduced. Also, a larger data entry field has been introduced to allow users to type in multi-line text. This represents what the user sees in the browser.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the system can extract web links found in user documents, follow them and parse the content of the remote page in order to identify the relevant content behind that page that was found to be relevant to the user. Such relevance information can then be added to the relevant user data repository maintained by the system.

In an alternate embodiment of operation of the present invention, in addition to the description of the core embodiment, the system can use a section of a document as a reference for what is relevant to the user. For instance, if the user is visiting a certain document and only one paragraph is found to be important to the user, the user can highlight and copy such paragraph alone to the system along with a reference to where it came from - in other words, a sophisticated way to use the Store/Publish capability. The information is then saved into one of the user data repositories along with a link that allows the user or users to visit the original document if necessary.

It will also be understood that, in addition to accept an input manually typed in by the user, the system can be used to automatically obtain the input directly from other sources. In other words, not only can the input be provided by a real person, but it can also be automatically obtained from an electronic mean, such as a database or a document.

It is further intended that any other embodiments of the present invention that result from any changes in application or method of use or operation which are not specified within the detailed written description or illustrations contained herein yet are considered apparent or obvious to one skilled in the art are within the scope of the present invention. It is also understood that the embodiments here described can be combined in order to create new embodiments.