

MVVM має знайомі нам компоненти (Model, View) та новий (ViewModel) (Модель уявлення).

Висновки

Було досліджено та проаналізовано найбільш поширені види патернів. Кожен з них краще використовувати в той чи іншій ситуації. Немає єдиного вірного патерна .

При розробці клієнт-серверного додатку важливо вибрати патерн, який повністю задовільнить всі потреби розробників та самого продукту . Під час розробки клієнт-серверного додатку найчастіше використовують такі шаблони як Adapter, Observer та Singleton.

Література

1. Електронний ресурс. – Режим доступу: <http://citforum.ua> / Портал аналітичної інформації в галузі інформаційних технологій CitForum
2. Інформаційно-комунікаційні технології. Веб - сайт ООН [Електронний ресурс]. – Режим доступу: <http://www.un.org/ru/development/ict/index.shtml>
3. Електронний ресурс - Режим доступу: <http://www.pil-network.com/#uk/> Microsoft Partners in Learning

АСТИСОВА Т. І., МНОЖИНСЬКИЙ Б. Г.

РОЗРОБКА ІНФОРМАЦІЙНО-ПОШУКОВОЇ СИСТЕМИ ДЛЯ АВТОМАТИЗОВАНОГО ЗБОРУ ДАНИХ

ASTISTOVA T. I., MNOZHINSKY B. H.

DEVELOPMENT OF INFORMATION AND SEARCH SYSTEM FOR AUTOMATED DATA COLLECTION

Annotation. The relevance of the task related to the creation of search engines is determined by a number of factors:

- *in connection with the growing amount of information, new search tools are needed;*
- *it is necessary to improve the quality (relevance) of search information.*

It is known that there is a significant inconsistency between the information needs of a specific user and their expression in the form of information search requests prepared for one or another information system

The task of the work is the development of the structure of a new search system for the Internet and the development of information search algorithms with the determination of the relevance of found documents

Keywords: Sites, algorithm, search engines, relevance, searching system, DBMS, document indexing,

Вступ

Актуальність задачі, пов'язаної із створенням пошукових систем визначається низкою факторів:

- у зв'язку з зростаючим обсягом інформації потрібні нові пошукові

засоби;

- необхідно підвищити якість (релевантність) пошукової інформації.

Відомо наявність суттєвої неузгодженості між інформаційними потребами конкретного користувача та їх виразом у вигляді запитів на пошук інформації, підготовлених для тієї чи іншої інформаційної системи

Не кожен користувач здатний чітко висловити свої інформаційні потреби. Для цього необхідний певний рівень знань саме в тій галузі, деяку інформацію з якої бажає отримати користувач.

Основна частина

Проблема, з якою стикається користувач під час побудови запиту, як сформулювати запит. Ця проблема пов'язана з необхідністю опанувати мову запитів конкретної інформаційної системи.

Під час роботи з пошуковою системою користувач висловлює свої інформаційні потреби мовою запитів даної системи та сподівається отримати у результаті пошуку посилання найбільш релевантні запиту документи.

Існують два основні підходи до пошуку інформації. Перший полягає у використанні спеціальних тематичних каталогів, а другий – пошукових машин (search engine), заснованих на індексі.

Найбільш продуктивною схемою пошуку є комплексне використання цих двох базових підходів. Тому сучасні пошукові системи надають користувачам можливість використання обох підходів.

ІПС, засновані на каталогах, називатимемо класифікаційними ІПС, а ІПС, що використовують пошукові машини, - словниковими.

В основі складання тематичних каталогів лежить принцип класифікації, тобто розподіл документів (або цілих сайтів) з ієрархії тематичних рубрик (класифікатору).

Каталоги можуть містити реферати документів та короткі описи сайтів. Кількість класифікованих документів мізерно мала в порівнянні з їх загальною кількістю у Web, тому каталоги не можуть дати вичерпних відомостей з певної тематики

Здійснюючи пошук необхідно взаємодіяти з культурами розробників класифікатора та систематизаторів.

Культурні проблеми, пов'язані з використанням класифікаційних ІПС, призвели до створення ІПС словникового типу.

З погляду «недосвідченого» користувача словникова ІПС – це засіб пошуку документів за ключовими словами. Сценарій пошуку простий: користувач за допомогою мови запитів висловлює те, що він хоче знайти, і буквально за кілька секунд отримує список посилань на документи, що задовольняють його запити.

З погляду внутрішньої організації, словникова ПС складається з двох частин, зазвичай, працюючих паралельно.

Перша частина (індексуєчий агент (agent), павук (spider), робот (robot)) відповідальна за індексування Web-документів, а друга (пошукова машина) здійснює пошук документів за індексом відповідно до запитів користувачів

Основна ідея словникової ПС - створити словник (індекс) із слів, які у документах Інтернет, у якому кожному слову відповідатиме список документів, які містять. Якщо пошук слів у такому словнику виконується швидко, можна відмовитися від дорогих послуг розробників класифікаторів і систематизаторів.

Індексуєчий агент методично переглядає документи Web, переходячи від одного документа до іншого за допомогою гіпертекстових посилань усередині документів. До кожного зустрінутого документа виробляється виділення індексної інформації та збереження їх у базі даних, званої індексом.

З філософської точки зору кожен документ можна уявити, що складається з двох частин: «важливої» і «неважливої».

$$\{\text{Документ}\} = \{\text{Важлива частина}\} \cup \{\text{Неважлива частина}\} \\ \{\text{Неважлива} \cap \{\text{Важлива частина}\} \text{ частина}\} = \emptyset$$

В ідеалі до індексу має потрапляти лише «важлива» частина документа, а «неважлива» частина губитися. Однак алгоритми індексування, що застосовуються різними ПС, відрізняються.

Для зберігання індексу часто використовують системи управління базами даних (СУБД), де пошук по індексу здійснюється з допомогою вбудованих засобів СУБД.

Інформаційний потік або масив L представляють у вигляді матриці розмірності $N \times D$, де в якості рядків виступають пошукові образи N документів.

$$L = \left\{ \forall i = \overline{1, N}; j = \overline{1, D} : b_{ij} = \begin{cases} 0, & t_j \notin l_i \\ 1, & t_j \in l_i \end{cases} \right\}$$

де: t_i – термін, l_i – документ.

При такому розгляді можна сформулювати процедуру звернення до інформаційної системи так:

$$L \times q = r,$$

де: q – вектор запиту, r – вектор відгуку системи на запит.

Це традиційне визначення процедури пошуку документів ПС було введено для вирішення проблеми автоматичного індексування документів, але виявилось надзвичайно корисним і для опису пошуку.

Щоб стандартизувати роботу пошукових служб, було розроблено стандарт з обов'язковою підтримкою будь-якою пошуковою системою єдиних команд запитів, що локалізують вузол за його доменним ім'ям, а документ - за URL.

Найпопулярніша у світі пошукова машина AltaVista (<http://www.altavista.com/>) містить 11 мільярдів слів, витягнутих із 30 мільйонів Web-сторінок.

Виновки

Єдиної оптимальної схеми пошуку в Інтернеті не існує. Залежно від специфіки необхідної інформації, для її пошуку слід використовувати відповідні пошукові служби. В принципі, звичайно, можна завжди користуватися якоюсь однією пошуковою системою, але чим грамотніше підібрані пошукові служби і складено запит на пошук інформації, тим якіснішими будуть результати пошуку.

Література

1. Кожичкін Євген Сергійович. Пошукова система, заснована на нейронній мережі. [Електронний ресурс]. — Режим доступу: <http://miem.edu.ua/>
2. Капустін В.А. Основи пошуку інформації в Інтернеті / Методичний посібник [Електронний ресурс]. – Режим доступу: <http://www.edc.samara.ua/gr/basics.htm>
3. Талантов Михайл Професійний пошук в Інтернеті: повнота, достовірність, швидкість. [Електронний ресурс]. – Режим доступу: http://nur.yamal.ua/internet/search/prof_search01.shtml