

---

# Информационна система за следене и регистриране на цитирания на български научни публикации

Елисавета Тричкова

*Information system for tracking and registration in citations of Bulgarian scientific publications: A lot of activities in today's dynamic world concern effective information exchange. This paper aims to present Web service information system for search and retrieval in distributed databases. The paper develops one concrete decision for using information systems of three-hierarchical client/server model. The functional part of the system is written on PHP server language, for database is used MySQL and for Web server is used Apache. These are open source products which are systematically independent.*

**Key words:** Database, Information systems, MySQL, PHP, Web service.

## ВЪВЕДЕНИЕ

Определяне на цитиранията на научни резултати сред научната общност у нас се осъществява от собствени наблюдения на авторите на научни резултати. Достъпът на наши учени до водещи световни индекси като Science Citation, Inspect и др. е ограничено по финансови причини, като се изисква поддържане на висок годишен абонамент. Такъв децентрализиран подход не позволява да се оценят и идентифицират значими научни резултати. Предлаганата разработка реализира възможност за информационно регистриране и автоматично определяне на цитиранията в български научни разработки, които адресират чуждестранни и наши учени.

Настоящата статия има за цел да представи информационна услуга в Интернет - „Български технически индекс за цитирания” - позволяваща осъществяването на селективно търсене, следене, регистриране, обработване и извличане на информация за цитирания на български научни резултати в областта на технически инженерни науки. Представено е конкретно решение, което използва трислоен модел на клиент/сървър информационна система. Набляга се на динамичното генериране и предоставяне на информация. За осъществяване на приложението са разработени няколко софтуерни модули посредством PHP сървърен език. Чрез тези модули се реализира SQL заявка (request) и SQL отговор (response), необходими за функционирането му. След получаването на SQL заявката се претърсва базата данни MySQL за да бъде извлечена желаната от потребителя информация. Информацията се съхранява във вид на отделни таблици. Данните от базата се обработват и структурират в HTML формат с цел отговаряне на SQL заявката като SQL отговор [1], [2].

## WEB УСЛУГА ЗА СЛЕДЕНЕ И РЕГИСТРИРАНЕ НА НАУЧНИ РЕЗУЛТАТИ

### 1. Три-йерархичен клиент-сървър информационен модел

Този клас информационни системи реализират модела на клиент-сървър с три йерархични нива: ниво на представяне, приложно ниво и ниво за обработка на данни [3].

Набляга се на динамичното генериране и предоставяне на информация. При отправено потребителско запитване сървърът стартира обработваща програма (ASP, CGI, Servlet, PHP), която генерира HTML страница в реално време и я изпраща на клиента. Обработващата програма може да осъществява връзка с една или няколко бази от данни. Това позволява да се реализира интерактивна комуникация между потребителя и приложението. Базата от данни изпълнява функции, процедури и други дейности, свързани с поддържането и предоставянето на актуална информация от и за потребителите на информационната система.

• **Първо ниво:** клиент, който функционира чрез общодостъпните браузери Internet Explorer, Netscape, Opera.

• **Второ ниво:** Web сървър. Това ниво има две нови особености спрямо модела клиент-сървър с две нива.

---

> Първо: операционната система на компютъра има допълнителни програмни средства за връзка и информационен обмен с базите от данни. Такива са например прилаганите драйвери ODBC (Open DataBase Connector) и JDBC (Java DataBase Connector) . Тези модули реализират прехвърлянето на данни между сървъра и базата от данни.

> Второ: Web сървърите са усложнени, така че да може да стартират за изпълнение програми (приложения). Тези програми са наречени сървърски и позволяват да се усложни алгоритмичната обработка на обмяната на информация между клиента и сървъра. Съответно тези програми статичния Web сървър не може да реализира. За целите на удобно съвместяване работата на web сървъра със сървърните програми са разработени специални алгоритмични езици.

- **Трето ниво:** реализирано е от системата за управление на бази от данни. Използвани бази от данни в информационни системи в Интернет са: Microsoft Access, Microsoft SQL Server, Oracle, Informix, Sybase. Характерна особеност е, че не съществува пълна съвместимост между различните Web сървъри, сървър езици, операционни системи и бази от данни.

Като предимства на този клас информационни системи може да се изтъкне активната комуникация между клиента и информационната система в реално време. Недостатък при този тип приложения е, че свойствата на информационната система силно се влияят от начина на проектиране и реализиране на базата от данни, особено при лошо структурирана база от данни.

## 2. Решение и софтуерна реализация

Решението ще задоволява напълно функционалните изисквания на информационната системата, а именно:

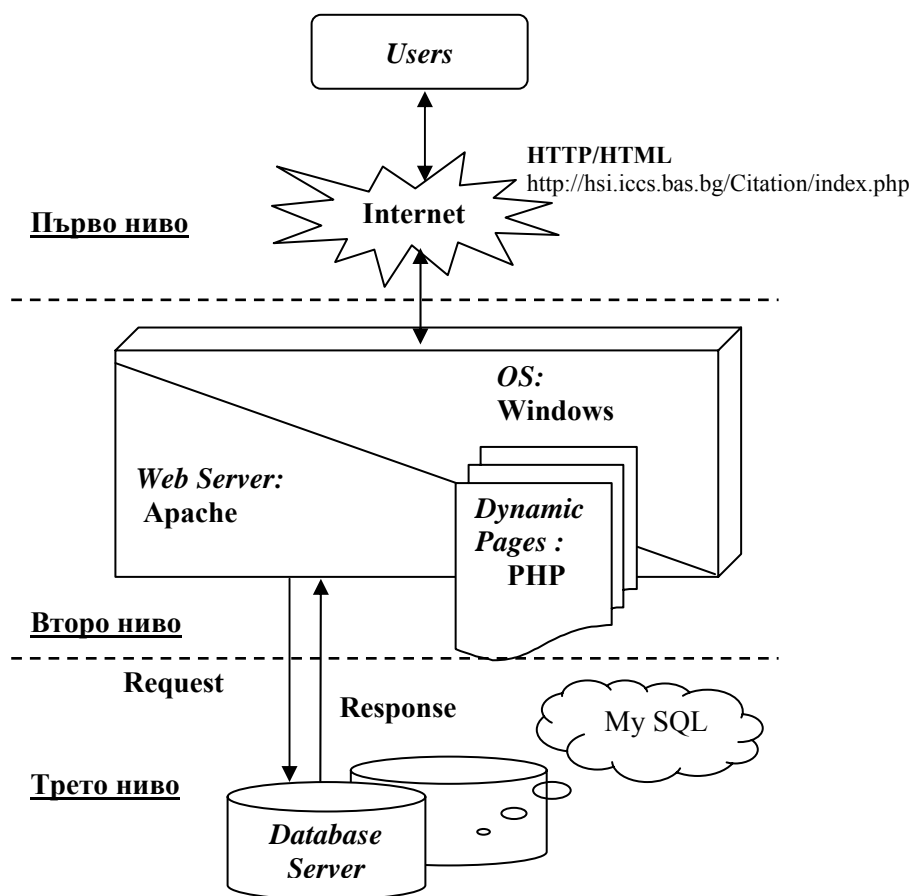
- Работа с разпределени бази данни - да допуска възможност за работа с бази данни разположени на сървър.
- Логическо групиране на данните - да позволява да се обработват всички запитвания на логическата група база данни.
- Разграничение на достъпа - системата да бъде способна да предоставя различни нива на достъп за ползвателите на информация.
- Проверка и контрол - системата е длъжна да може да събира статистика по запитванията на ползвателите .
- Откритост - системата да позволява да се разширява лесно и да бъде базирана на открити стандарти и протоколи.
- Простота в обслужването - за потребителите системата да предоставя прости интерфейси за достъп към информацията.
- Съдържанието на базата да бъде в подходяща електронна форма.
- Да дава on-line достъп на потребителите и информация за извършени обновления в приложението.
- Да дава информация при възникнали въпроси при търсенето.

По-конкретно приложението е указано с архитектура от фиг.1 и има три логически обособени йерархични нива.

За реализацията на първото функционално ниво – клиента (на фигурата е обозначен като Users) – се използва обикновен браузер (например Internet Explorer или Netscape), който се свързва с главната сървърска част на приложението. В тази част от информационната система комуникацията се извършва по HTTP протокол.

Второто ниво е изградено под Windows операционна система и се реализира чрез Web сървъра, който обслужва информационната система. Набляга се на динамичното генериране и предоставяне на информация. При отправено потребителско запитване сървърът стартира обработваща програма (ASP, CGI, Servlet, PHP), която генерира HTML страница в реално

време. Избран е Web сървърът Apache, а като обработваща програма се използва PHP интерпретатор за програмните модули от страна на сървъра.



Фиг.1 Архитектура на приложението

Третото ниво е нивото на базата данни. То е реализирано чрез MySQL database. MySQL сървърът е разработен да управлява и поддържа големи бази данни много по-бързо от съществуващите за това решения и в последно време се използва много успешно в среди с високи изисквания относно производителността. В конкретната база данни се съдържат две отделни таблици. В тези две таблици съответно се съдържа подробна информация за цитирания на научни резултати на български автори в областта на технически инженерни науки.

### 3. Изпълнение на запитванията и извеждане на резултатите

Информационната система ще позволява на потребителите да извършват селективно търсене по определени критерии, да осъществяват регистриране, обработване и извличане на информация за цитирания на български научни резултати в областта на технически инженерни науки. Това приложение ще използва MySQL база данни за съхранение на информацията за цитиранията и PHP скриптове за генериране на HTML страници за представяне на съдържанието.

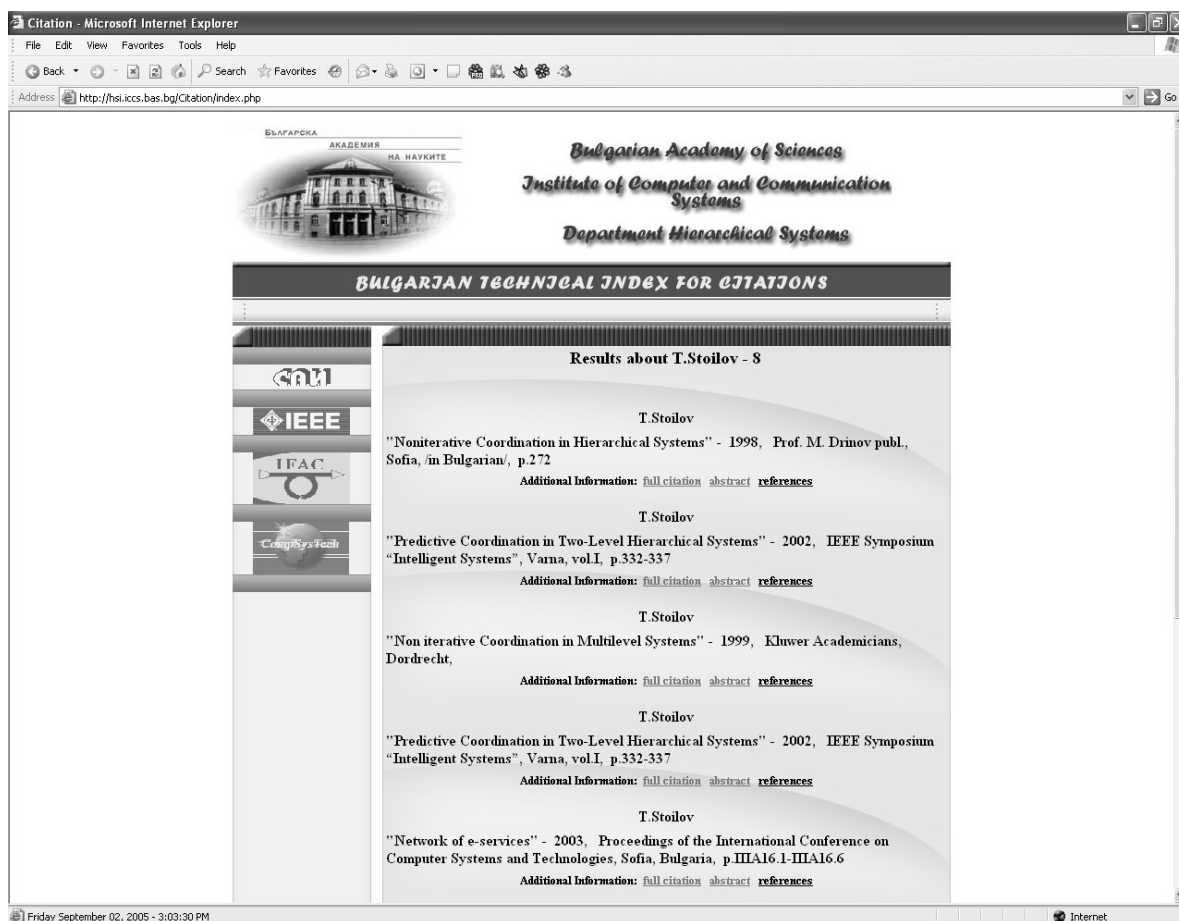
Записите в индекса за цитирания се съхраняват в таблица „citation” в базата данни. Тази таблица е създадена посредством phpMyAdmin. Това е скрипт за управление на база данни. Таблицата съдържа петнадесет колони и всеки ред в таблицата е уникално определен от полето rowid.

Главната страница на приложението съдържа формуляр, чрез който се осъществява търсене, обработване и извличане на информация от базата данни.

Търсенето се осъществява по три основни критерия: по име на автор; по заглавие на публикация; по година на публикуване; Всеки потребител има възможност да осъществи търсенето попълвайки текстовото поле съответстващо на един от трите критерии, да комбинира два от тези критерии или да използва всичките три критерия едновременно.

Непосредствено под трите текстови полета с основните критерии са разположени пет под критерия обозначени с тип на въвеждане - checkbox. Срещу всеки checkbox съответства името на конференция на която е бил цитиран съответния доклад. За да улесни своето търсене потребителят може да маркира checkbox с желания от него под критерии.

След натискане на бутона „Search”, като преди това са избрани и зададени желаните критерии, се отваря страница съдържаща резултатът от търсенето, фиг.2.



Фиг.2 Страница от приложението показваща резултатите от направено търсене

Извличат се записи от базата данни сочещи колко на брой пъти търсеният автор е цитиран и съответно са изведени и показани резултати за всяка една от цитираните публикации на същия автор. Тези резултати съдържат информация за заглавието на публикацията, година на публикуване, място или издателство на публикуване.

Под всеки изведен резултат са посочени три линка с допълнителна информация за цитирания автор. При избор на линк „references” на потребителя ще бъде показана визитна картичка за цитирания автор. При избор на линк „abstract” на потребителя ще бъде показано кратко резюме на публикацията на цитирания автор. При избор на линк „full\_citation” от базата данни се извлича подробна информация както за цитирания автор и неговата публикация така и за цитирания автор и неговата публикация. Под всеки един запис е прикачена и самата статия или доклад под формата на „pdf” формат. Потребителят има възможност да отвори и разгледа публикацията, както и да я съхрани и използва.

---

От главната страница на приложението чрез натискане на линк „add new item” е предоставена възможност за въвеждане на нови записи. Дава се потребителско име и парола, които се попълват в страница съдържаща идентификационна форма.

Формата за попълване на нови записи е разделена на две части. Първата част е предназначена за попълване на информация за статия или доклад на цитирания автор, а втората част е предназначена за попълване на информация за статия или доклад на цитиращия автор.

След натискане на бутона „add item”, ако формата е попълнена коректно, се извежда съобщение за успешно въведен запис. В противен случай се извежда съобщение за грешка.

Понастоящем в системата се следят годишните издания от конференциите “Автоматика и информатика”, Компютърни системи и технологии CompSysTech, Системи за автоматизиране на инженерния труд SAER, международните конференции организирани от Съюза по Автоматика като IEEE Conferences on Intelligent Systems 2002, 2004, IFAC Symposium DECOM 2004 и др.

### **ЗАКЛЮЧЕНИЕ**

Представена е информационна услуга в Интернет - „Български технически индекс за цитирания” - предназначена за селективно търсене, следене, регистриране, обработване и извличане на информация за цитирания на български научни резултати в областта на технически инженерни науки. Проектирана е Web базирана информационна система, която използва трислоен модел на клиент/сървър системно-алгоритмична архитектура. Акцентирано бе на динамичното генериране и предоставяне на информация. За осъществяване на приложението бяха разработени няколко софтуерни модули посредством PHP (акроним от Hypertext Preprocessor). Чрез тези модули бе реализирана SQL заявка (request) и SQL отговор (response), необходими за функционирането му. Също така бе представена софтуерна реализация на приложението и предложена архитектура за работа с разпределени бази данни, базираща се конкретно на продуктите с отворен код PHP, MySQL, Apache. Описаният модел, в различна степен, подобрява и улеснява начина за предоставяне на информация на потребителя и цели бързо и динамично предаване на информацията, бързина в съхраняването и търсенето на данните, сигурност в операциите. Запълването на системата с актуални данни и информация се осъществява от упълномощени представители от Организационните комитети на Научните форуми, издаващи материалите на конференциите.

Работата е отчасти подкрепена от Европейската комисия, проект №FP6-027178.

### **ЛИТЕРАТУРА**

[1] Trichkov K. Distributed searching in global network. Proceedings of IFAC DECOM-TT 2004, Bansko, Bulgaria 3-5 October, ICCS-BAS, p.235-238

[2] Ivanova E., Web Service Architectures for Distributed Search in Databases, Proceedings of the II International IEEE Conference “Intelligent Systems’04”, Bulgaria, Varna, June, 22-24, 2004, ISBN 0-7803-8278-1, p.572 - 575.

[3] Ivanova Z.,K.Stoilova, T.Stoilov. Systems and Algorithmic Models of Information Services in Internet. Journal of Automatics and Informatics, №3, 2003, pp.13-18 (in Bulgarian).

### **За контакти:**

н.с. III ст. Елисавета Тричкова, Институт по Компютърни и Комуникационни Системи- БАН, e-mail: elisaveta@hsi.iccs.bas.bg