

Как да търсим информация в мрежата

Тамара Лазарова*

Национален център по заразни и паразитни болести

„Интернет е като библиотека без каталог, където всички книги скачат и се местят всяка нощ.“

В последните десет години Интернет предлага достъп до огромни количества информация от всякакъв характер. Само няколко натискания на клавиатурата ни делят от милиарди текстове и илюстрации по различни научни теми. С възхода на Интернет и лавинообразното увеличаване на информационните ресурси, всички се сблъскваме с проблема как селективно и резултатно да използваме тези ресурси. Много често определят безплодното търсене в мрежата като търсене на игла в купа сено.

Проблемът е, че липсва централен каталог на документите в Интернет.

Обработката и предоставянето на информация в мрежата се извършва по два основни начина:

– Първият е с преобладаващото участие на изкуствения интелект

– Вторият е с преобладаващото участие на човешкия интелект.

В зависимост от това, двата основни вида инструменти за обработка и предоставяне на информация в мрежата са:

– **Търсещи машини (търсачки) – компю-**

търни програми, които извършват машинна обработка и предоставяне на документи

– **Тематични указатели, портали, виртуални библиотеки – продукт на целенасочен подбор от страна на библиотекари и тесни специалисти в дадена област**

Съществуват два вида търсачки: „индивидуални“ и „метатърсачки“.

Всички индивидуални търсещи машини разполагат с по-голяма или по-малка база данни, съдържаща копия от уеб документи, които техните „роботи“ са обходили. Някои от най-популярните са: Alta Vista, Yahoo, Google.

За разлика от тях, метатърсачките не използват роботи за събиране на информация и не разполагат със своя база данни. Те изпращат запитване до няколко от индивидуалните търсачки в интернет и получените резултати обобщават в общ списък. Такива са известните Vivisimo, Mamma, Excite.

Тематичните указатели са колекции от хипервръзки към Интернет ресурси, изготвени от информационни специалисти, които ги подбират и организират в предметни (тематич-

*E-mail: tamara.lazarova@gmail.com

ни) категории. Те използват определени критерии за селекция и търсят в базите данни по определена тема или по специфичен отрасъл. По този начин те увеличават възможността за намиране на релевантна информация в дадена област, защото включват в каталога информация, която не се улавя от общите машини. Някои от популярните медицински указатели са: Medscape (<http://www.medscape.com/home>), Netdoctor (<http://www.netdoctor.co.uk>), Doctor's Guide (<http://www.docguide.com>). Вече съществуват и български медицински портали като Арсмедика (<http://www.arsmedica.bg>), Медицина (<http://medicina.start.bg>), Медицина и здраве (<http://www.medicine.primasoft.bg>).

Виртуалните библиотеки предлагат селектирани, отлични източници на информация, препоръчани от експерти в съответната област. Различават се от обикновените указатели по това, че съдържат много по-малко ресурси, поради по-високата степен на селективност. Едни от най-известните и често използвани медицински виртуални библиотеки са: PubMed Central (<http://www.pubmedcentral.nih.gov>), SciCentral (<http://www.scicentral.com>) и MedlinePlus (<http://medlineplus.gov>).

Все повече библиотеки дигитализират и предоставят в Интернет определени заглавия или отделни сбирки от техните колекции. В тях сами, без посредничеството на библиотекар, можем да извършваме справки: какви заглавия са публикувани по интересуващата ни тема, да търсим по годините на публикация и според езика на публикацията. Търсене в електронните каталози вече можем да правим и в страниците на някои български библиотеки: Националната библиотека „Св.Св. Кирил и Методий“ (<http://www.nationallibrary.bg>), Централната библиотека на БАН (<http://www.cl.bas.bg>), Библиотека на СУ „Св. Климент Охридски“ (<http://ns.libsu.uni-sofia.bg/index.html>), Централна медицинска библиотека (http://www.medun.acad.bg/cmb_htm/cmb1_home_bg.htm).

Най-лесният начин за намиране на информация в интернет е по известен адрес. Ако не знаем адреса на помощ идват търсачките. Търсенето започва с ключова дума или списък от

думи, без да е необходимо да се знае интернет адреса, на който се намира желаната информация. Търсещите средства указват адресите на страниците, в които се намират думата или думите, които търсим. Такова търсене е търсене по буквален текст и не винаги удовлетворява изискванията ни. Например търсене на тема Наполеон може да ви даде информация за Наполеон Бонапарт, груги Наполеоновци, но също така може да откриете страници на музеи и хотели на името на Наполеон, съществителни коне и напитки с това име и др. За да се избегнат тези неудобства при търсене на информация е достатъчно да се знаят няколко основни правила:

- Колкото по-конкретно е формулирано търсенето, толкова по-точни ще бъдат резултатите

- Зададените думи в единствено число се тълкуват от търсачките двояко – както в единствено, така и в множествено число.

- Повечето търсачки тълкуват малките букви и като малки, и като големи. Препоръчва се да се въвеждат само малки букви.

- Повечето търсачки игнорират някои малки думи (предлози, съюзи, определителни членове) като a, an, as, at, of, on, the, to. Те не се отчитат дори в рамките на една неделима фраза като „Crime and punishment“ („Престъпление и наказание“).

За да постигнете по-добри и по-точни резултати, когато използвате търсещите машини е полезно да използвате езика на изразите за търсене (логически оператори). Част от тях са описани в таблица 1.

В резултат от съществуващата конкуренция и желанието за предоставяне на изчерпателна информация, различните средства за търсене все повече се опитват да улеснят потребителите си като непрекъснато подобряват интерфейса на страниците си и правят използването на логическите оператори много по-лесно и ефективно. Добър пример в това отношение е световно известната търсачка GOOGLE (www.google.com). Тя е много популярна и в България, тъй като в началото на 2002 г. предостави търсене на редица езици,

Таблица 1. Логически оператори за търсене в Интернет.

	Какво прави:	Как се изписва:	Резултат:
(+) плюс	търси страници, в които има и двете гуми	nosocomial + infection	страници, които съдържат и двете гуми
(-) минус	изключва от резултатите дадена гума или израз	infection-nosocomial	страници, които съдържат инфекции, които не са носокомиални
OR или-или	търси страници, които съдържат едната или другата гума	infections nosocomial OR hospital	страници, които съдържат и nosocomial и hospital infections
(„ ”) кавички	търси точната фраза	„nosocomial infections“	намира страници, които съдържат точната фраза „nosocomial infections“
(~) сходни гуми	търси гумата и всичките ѝ синоними	~ nosocomial infections	намира страници, които съдържат гумите nosocomial infections, както и infection in hospitals, urinary tract infections, surgical infections и гр.
(.....) много-точие	търси в определен обхват	epidemiology 2000...2007	намира страници, които съдържат гумата epidemiology през периода 2000-2007 година
(*) апостроф	замества дадена гума	“Vladimir * Putin”	намира страници, които съдържат гума между Vladimir и Putin (Vladimir Vladimirovich Putin)
define:	Търси определение на гумата (термина) или фразата.	define:nosocomial	Търсачката ще претърси Интернет и ще извади пред вас няколко дефиниции на търсения термин
filetype:	Търси ключовите гуми само в определен вид файлове (pdf, ppt, xls и гр.)	nosocomial filetype: ppt	Ще се покажат само страници, съдържащи Power Point презентации, съдържащи гумата nosocomial
site:	Ограничава търсенето само в определен сайт.	nosocomial site: www.bulnoso.com	Ще покаже страници, съдържащи гумата nosocomial само от уеб сайта на Булнозо

Включително и на български. Ако при повикването на Google се зареди английският ѝ вариант, от началната страница се избира “Language tools” (инструменти за търсене на даден език). Отваря се страница с падащо меню – списък на езиците. Избираме български и въвеждаме интересуващите ни понятия. В отговор на нашето запитване Google ще ни представи списък от уеб сайтове, като първи в списъка

ще се появят най-често посещаваните сайтове и тези, към които са направени най-много препратки. Така напр. при задаване на английски “Library Bulgaria” ще се зареди библиотеката на БАН, тъй като нейният сайт е най-често посещаваният и редица чуждестранни институции са направили хипервръзка (link) към този сайт. Функцията “разширено търсене” ни дава възможност за прецизиране на търсенето по език,

дата на публикуване, формат на файла както и за търсене в определени полета от страницата – адреса на страницата (URL), заглавието и т.н. (фиг.1).

Екипът на Google си е поставил грандиозната задача да организира световната информация и да я направи универсално достъпна и полезна. Тази своя мисия те осъществяват чрез различни интересни проекти. Един от тях е Google Scholar (Грамомен Google) <http://scholar.google.com>, който стартира през ноември 2004 г. в отговор на критиките към компанията, че търсачката не съответства на интересите на учени и изследователи. Сега Scholar съдържа над 2 милиона записи на книги, статии, резюмета, организирани по области на знанието. Специалистите по информационни продукти прогнозираат, че след време Google Scholar ще конкурира престижните и платени бази данни. За разлика от стандартната търсачка, Google Scholar прави справки само в надеждни академични материали, които са с проверен произход и са били публикувани в интернет пространството от университети и научни институти. След като се зададат ключовите думи, се получава препратка към милиони академични материали, в това число и офлайн източници – книги, печатни издания, резюмета и технически новини във всички области на знанието. Резултатите се сортират по индекс, по дата, по източник и по автор.

Наскоро Google обяви и друг проект – сканиране на милиони книги, които представляват фондовете на 5 от най-големите библиотеки в света. По този проект за създаване на най-голямата интернет библиотека, с Google работят университети като Оксфорд, Станфорд и Харвард, както и Обществената библиотека на Ню Йорк. Повече информация ще намерите на адрес: <http://www.google.bg/books>. Книгите с изтекло авторско право ще се предлагат в пълен текст. За пример, само от колекцията на Оксфорд ще бъдат сканирани 1,5 млн. печатни издания от 19 век. За книги под авторско право ще има извадки и цитати с връзки към книжарници и библиотеки. Това е херкулесов проект, който ще отнеме около 10 год. и може да превърне Google в най-големия собственик на електронни издания, снабдявайки изследователи и студенти с безпрецедентно средство за информация.

Ползването на Интернет като информационен ресурс и световна библиотека не изисква специални компютърни знания, а умения за изграждане на ефективен, точен и релевантен процес на информационно търсене. В това отношение, познаването на някои правила и избирането на ефективна стратегия за провеждане на информационно търсене могат да улеснят както начинаещия, така и професионалния потребител.

Google Разширено търсене		Съвети за търсене Всичко за Google
<input type="text"/>		10 резултати <input type="button" value="Google търсене"/>
Намери резултати	<input type="checkbox"/> с всичките думи	<input type="text"/>
	<input type="checkbox"/> с точната фраза	<input type="text"/>
	<input type="checkbox"/> с поне една от думите	<input type="text"/>
	<input type="checkbox"/> без думите	<input type="text"/>
Език	Показвай страници, написани на	<input type="text" value="всякакъв език"/>
Формат на файла	<input type="checkbox"/> Само резултати с формат на файла	<input type="text" value="всякакъв формат"/>
	Показвай страници, обновени през	<input type="text" value="без значение"/>
Срещания	Търсене в	<input type="text" value="навсякъде в страницата"/>
Домейн	<input type="checkbox"/> Само резултати от сайта или домейна	<input type="text"/>
	например: google.com, .org Повече информация	

Фиг. 1. Разширено търсене в Google.