

Критерии за търсене оценяване на информация в Интернет

гл.ас.д-р Полина Михова

Директор БП "Обща и приложна биология"

Департамент Медикобиологични науки

pmihova@nbu.bg



Какви цели и задачи си поставям с тази лекция?

1. Интернет в България накратко
2. Умения и компетенции на един търсач в уеб
3. Основни "филтри", по които да преценим мигновено сайта
4. Представяне на основните характеристики на една информация в Интернет
5. Критерии за качество на съдържанието на информация
6. Добри трикове за търсене в Google
7. Списък с препратки към главните търсачки в Интернет

Някои характеристики на Интернет

- неограниченост по тематика и използвани източници
- динамичност – средният живот на уеб-страниците е 44 дни – 2 години
- актуалност
- английски език – 48% от всички уеб-сайтове
- липса на унификация по отношение на формата на файловете
- значителна степен на дублиране на документите, поради наличието на огледални сайтове

- но и липса на гаранции за качество, достоверност,
- разпръснатост в милиони компютри

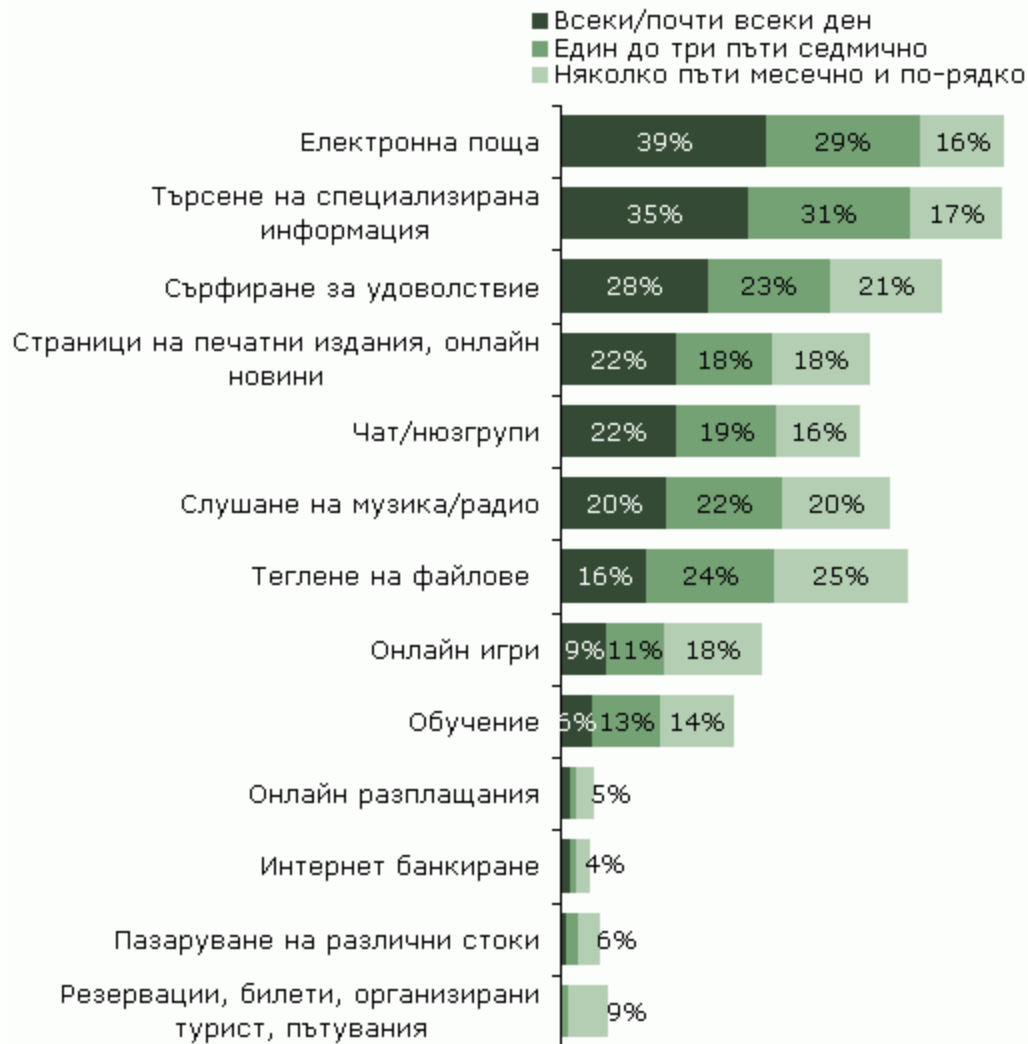
БГ и Интернет потреблението

Internet Users in the European Union

EUROPEAN UNION	Population (2010 Est.)	Internet Users, Latest Data	Penetration (% Population)	User Growth (2000-2010)	Users % Table
Austria	8,214,160	6,143,600	74.8 %	192.6 %	1.3 %
Belgium	10,423,493	8,113,200	77.8 %	305.7 %	1.7 %
Bulgaria	7,148,785	3,395,000	47.5 %	689.5 %	0.7 %
Cyprus	1,102,677	433,800	39.3 %	261.5 %	0.1 %
Czech Republic	10,201,707	6,680,800	65.5 %	568.1 %	1.4 %
Denmark	5,515,575	4,750,500	86.1 %	143.6 %	1.0 %
Estonia	1,291,170	969,700	75.1 %	164.5 %	0.2 %
Finland	5,255,695	4,480,900	85.3 %	132.5 %	0.9 %
France	64,768,389	44,625,300	68.9 %	425.0 %	9.4 %
Germany	82,282,988	65,123,800	79.1 %	171.3 %	13.7 %
Greece	10,749,943	4,970,700	46.2 %	397.1 %	1.0 %
Hungary	9,992,339	6,176,400	61.8 %	763.8 %	1.3 %
Ireland	4,622,917	3,042,600	65.8 %	288.1 %	0.6 %
Italy	58,090,681	30,026,400	51.7 %	127.5 %	6.3 %
Latvia	2,217,969	1,503,400	67.8 %	902.3 %	0.3 %
Lithuania	3,545,319	2,103,471	59.3 %	834.9 %	0.4 %
Luxembourg	497,538	424,500	85.3 %	324.5 %	0.1 %
Malta	406,771	240,600	59.1 %	501.5 %	0.1 %
Netherlands	16,783,092	14,872,200	88.6 %	281.3 %	3.1 %
Poland	38,463,689	22,450,600	58.4 %	701.8 %	4.7 %
Portugal	10,735,765	5,168,800	48.1 %	106.8 %	1.1 %
Romania	21,959,278	7,786,700	35.5 %	873.3 %	1.6 %
Slovakia	5,470,306	4,063,600	74.3 %	525.2 %	0.9 %
Slovenia	2,005,692	2,003,136	64.8 %	332.8 %	0.3 %
Spain	46,505,963	29,093,984	62.6 %	440.0 %	6.1 %
Sweden	9,071,055	8,007,000	88.3 %	107.5 %	1.0 %

Защо се ползва Интернет в България?

Направления, по които се използва Интернет: Януари 2009, 18+



* Сред потребителите на Интернет

Умения на Интернет изследователя

- търпение
- упоритост
- добра памет
- организационни умения
- творчество
- решителност
- познаване на спецификата и ограниченията на информационните ресурси в мрежата

Cybrarian

- сравнително малък информационен шум
- подбрани са най-добрите ресурси
- дава се аналитично описание
- обхванатите теми са развити в голяма дълбочина
- сравнително лесно издирване на справочни ресурси
- няма опасност от получаване на прекалено много, често безполезни като обем, резултати

Качество на информацията

- За каква аудитория
- Кой ще я ползва
- Каква е дълбочината

Какви видове информация и носители можем да намерим в интернет пространството?

- **Суетно публикуване** – това може да представлява специфичен документ, който съдържа информация от висока важност, но не е преминал през двоен “сляп” процес на редактиране или не е разпространен в специализираната област, за която е.
- **Сива литература** – памфлети, предпечатни издания, технически доклади
- Професионални асоциации
- Резултати от настоящи екстерименти
- **Рекламни и PR материали**
- **Мултимедия**

Първични ресурси и източници

- Артефакти (монети, растения, фосили, мебели, средства, облекло, всичко от времето на провеждане на експеримента);
- Аудио записи;
- Дневници;
- Комуникация по е-майл;
- Интервю;
- Журналини статии, преминали редакция;
- Писма;
- Статии във вестници;
- Оригинални документи (акт за раждане, завещание, акт за граждански брак);
- Патенти;
- Снимки;
- Просидинги на срещи и конференции;
- Записи на организации, правителствени агенции;
- Речи;
- Проучвания – маркетингови, потребителски;
- Видео записи;
- Произведения на изкуството, архитектурата, литературата, музиката;
- Уеб-сайтове.

Вторични източници

- Библиографии;
- Биографични документи;
- Коментари, критики;
- Речници и енциклопедии;
- Истории;
- Монографии;
- Учебници (считани също и за третични);
- Уеб-сайтове (считани също и за първични).

Третични източници

- Алманаси;
- Хронологии;
- Директории;
- Книги с описание на факти;
- Пътеводители;
- Индекси, абстракти, библиографии, които локализират първични и вторични източници;
- Ръководства;
- Учебници.

Първични филтри

- Домейн или сайт
- Период от време
- Език
- Заглавия
- Съдържащ се конкретен текст
- Формат
- Наличие на графични илюстрации
- Наличие на хипервръзки
- Адрес на он-лайн документи
- Големина

Основни характеристики на информацията в Интернет

- достоверна
- защитена от неоторизиран достъп
- достъпна във всеки сегашен или бъдещ момент по регламент
- да подлежи на контрол и сравнение
- да се регистрира и отчита при стандартни (стандартизирани) всеобщо приети правила (промените да се регистрират)
- да има начини за нейния параметричен текущ контрол (за видове грешки)
- Задължително е да се цитира авторството, времето и мястото на достъп

Критерии за качество на съдържанието на информация

- авторитет и репутация на авторите/източника – квалификация, CV на автора; официален орган, институция, академична личност
- актуалност – видове дати – на създаване на документа; на въвеждането му в интернет; последна актуализация, период на валидност на данните; политика за осигуряване актуалността на инфо; валидност на хипервръзки от сайта
- точност – има ли грешки, от какъв характер са; условия, при които важи инфо; чужди източници – достъп до тях; при научно изследване – методи и обработка на данните; мотивация на източника на данните; проверимост на инфо; възможност за проверка;
- информационна стойност – уникалност, първична инфо; факти, посока на хипервръзките, информационна стойност, анотации
- завършеност – грешки – видове и количество; тронави конструкции, да няма препратки към уеб-сайтове under constructions; библиографско описание; анотации; реферат; валидни/невалидни връзки;

-
- ❑ обективност – пристрастност, фирма-реклама;
 - ❑ безпристрастност на фирмата-държател на сървъра;
 - ❑ ясна разграниченост на рекламата – на една или на повече фирми;
 - ❑ хипервръзки с различни от авторските виждания и възгледи
 - ❑ удобство на предвижване – навигация, бутони, бързо и лесно възпроизводство на отделни фрагменти на екрана; електронна книга; еднофайлова версия;
 - ❑ възможност за копиране, разпечатване и съхранение, възможности за търсене в обемисти файлове по дадане параметър, ключова дума или др.
 - ❑ подпомагане на потребителя – Help, FAQ, e-mail, tel.N, информиране по е-пощата за актуализации
 - ❑ обхват и пълнота – дълбочина, очевидни пропуски; огарничения (език, регион, период време). За справки – списък на съдържание, план на сайта, налични хипервръзки, библиография,

НОН кодът в медицината е публикуван на 29 езика

1. Предоставени възможности за откриване на авторския екип
2. Ясно дефиниране правата за собственост върху предложената информация
3. Цитиране източниците, на които се позовава сайтът
4. Публикуване ролята и правата на спонсорите или финансово заинтересованите лица, които имат отношение към информацията
5. PR-активност и значимост (роля) на рекламните продукти – да се разграничат безусловно от професионалната информация
6. Цели – точно описание и дефиниция по адресанти
7. Леснота на достъп до информацията
8. Дата на създаване на нов домейн информация
9. Дата на последното обновяване на конкретни данни и информация
10. Дата за създаване на нов домейн информация или дата на последно обновяване на конкретна информация
11. Акредитация на автора като професионалист или на институцията, която представлява
12. Цитирани акредитационни документи на авторите
13. Професионални сдружения, в които участват авторите
14. Лесна навигация в частите на сайта и към хипервръзките му

-
15. Достъпно указание с обяснения на вътрешната схема за работа
 16. Осигурени линкове към други специализирани сайтове – дори и популяризиращи алтернативни позиции и информация
 17. Стратегия за представяне на цитираните източници
 18. Балансирани доказателства – статистика, базирани на научни доказателства изводи, най-добри практики
 19. Подходящ (научен, наукообразен) стил на писане и терминологичен език
 20. Осигурен механизъм за обратна връзка потребител/издател
 21. Факс-номер
 22. Е-мейл адрес
 23. Публикувани права на издателя
 24. Авторски права върху информацията
 25. Дали и как е представен официално процесът за избор на дадена статия за публикуване -рецензии, редакторски екип (ако има!), други критерии
 26. Ясна и прецизна йерархия на доказателствения материал, на който са базирани заключенията.

Някои препоръки от практиката

- alexa.com – поставете проверявания от вас линк в търсещото поле на посочения адрес
- Ще видите:
 - Оценка на травика
 - Анализи
 - Статистика на сайта включително история на сайта, и други линкове към тази страница
 - Информация за контакти или собственици на домейна
 - Архивът на уеб страниците - "Wayback Machine"

-
- Погледнете детайлно url и домейна
 - Огледайте **"About us," "Philosophy," "Background," "Biography," "Who am I,"**
 - Автори и дата на сайта
 - Как се представя информацията – авторска ли е, цитирана ли е, преразказана ли е?
 - Връзки към други източници – организация, работещи ли са, добри ли са препратките?

Кой се свързва към тази страница?

- Има ли много препратки?
- Към какви сайтове се свързва?
- Какво пък е тяхното съдържание?
- При онлайн продажба се слагат знаци на световно признати марки SSL сертификати, като VeriSign
- Колко бързо може да свали дадена информация?
- Бърз ли е достъпът до конкретния URL?

По каква причина страницата е качена в уеб пространството?

- Информативна, дава факти, дава данни ?
- Обяснява, убеждава?
- Продава?
- Споделя?
- Открива?
- Адекватно ли е съдържанието на страницата на това в библиотеката?

Още някои препоръки

□ **ДОСТЪПНОСТ**

- Достъпен ли е сайтът с ASCII browser като Lynx?
- В стандартен html формат ли е, или е необходимо да си използват допълнителни приложения?
- Използва ли стандартни мултимедийни приложения?
- Има ли страници, които отнема твърде много време за зареждане?
- Достъпен ли е сайтът винаги, или често е претоварен?
- Стабилен ли е, или URL се променя?
- Безплатни ли са всички страници или има страници, които изискват такса?

□ **ДИЗАЙН**

- Добре оформени ли са отделните страници, или трябва да скролвате завинаги?
- Появяват ли се важни инструкции преди да кликнете на определени линкове?
- Работят ли всички части на сайта?
- Интуитивно ли използвате сайта, или някои негови части не са ви ясни?
- Можете ли с яснота да кажете къде се намирате и бързо да преминавате от една локация към друга?
- Дружелюбен ли е изгледът му?
- Концептуално и съдържателно повече ли е от една книга?
- Може ли потребителят да взаимодейства спокойно със сайта?
- Обезопасени ли са взаимоотношенията – личната информация?

□ **СЪДЪРЖАНИЕ**

- Информативно ли е заглавието? Ясни ли са целта, обхвата и рамките на сайта?
- Прецизирани ли са заглавията на публикациите, или се използва жаргонна терминология?
- Организиран ли е към нуждите на потребителите или отразява вътрешна йерархия?
- Съдържанието отговаря ли на целите?
- Графично и мултимедийно добре ли е организиран сайтът – изскачат ли постоянно рекламни банери, информационни стикери и др?
- Цитирани ли са източниците – достъпни ли са и отговарят ли на съдържанието?

Пример за “Отговорност (Authority)”

Добър

- <http://www.ci.mankato.mn.us/>

□ Не толкова добър

- <http://descy.50megs.com/mankato/mankato.html>

Пример за “Accuracy”

□ Добър

- <http://www.genderandaids.org/>

□ Не толкова добър

- <http://147.129.226.1/library/research/AIDSFACTS.htm>

Пример за “Обективност”

□ Добър

- http://www.hgfound.org/res_disorders.html

□ Не толкова добър

- Отклоняващ се

<http://www.martinlutherking.org/>

Пример за “Навременност”

□ Добър

- http://www.amriglobal.com/products_and_services/products_desc.cfm?prodID=1

□ Лош

- <http://www.cmpharm.ucsf.edu/cohen/>

Пример за “Покритие и обхватност”

- Добър и не толкова добър?
 - <http://www.skysports.com/>
 - <http://www.sportsrumble.com/>

Пример за “Дизайн”

□ Добър

- www.samford.edu

□ Не толкова добър

- <http://www.naturalsolutions1.com>

Как да търсим пълноценно в Google

1. Кавички.
2. Тире, т.е. минус, поставено пред някоя дума, изолира всички сайтове, които я съдържат.
3. Събиране. Много търсачки игнорират думи или знаци, които се срещат често, защото това прави търсенето тромаво и неprecizно. Ако сложите "+" пред дадена важна дума обаче, тя със сигурност ще присъства.
4. Думата "или", т.е. OR (задължително с главни букви) на английски, пести време. Ако потърсите "почивка Анталия OR Кушадасъ", резултатите от двете дестинации няма да се смесят.
5. Дата. Ако търсите в сайт, където информацията често се повтаря с едни и същи термини и действащи лица (например спортна медия), добре е да ограничите търсенето до определена дата. Някои сайтове имат полета за търсене около дадена дата, но други нямат. При тях добавяйте като ключови думи периода, който ви интересува, например "Милан победи 3:0 Януари 1999"
6. Задавайте въпроси.
7. Смесвайте, тоест да комбинирате команди. Няколко плюса или минуса и две три кавички в едно търсене прецизират резултатите.

-
8. Конкретика. В случай, че търсите статия в голям сайт, но удряте на камък, ограничете Google-резултатите само от него или от неговия домейн с командата "site:". Така например "прием site:www.uni-sofia.bg" намира директно страницата за прием на нови студенти на СУ Климент Охридски.
 9. Търсите даден линк? Просто добавете "inanchor:". Командата е полезна при търсене на мейл-адреси, които често се изписват като връзка под името на човека.
 10. Търсене в заглавието. С командата "intitle:" търсите само в заглавията на страниците.
 11. "intext:" е обратното на "intitle:". С тази команда прескачате заглавията и търсите само в текста на статиите.
 12. С командата "link:" може да видите кои страници са сложили линкове към вашата страница или някоя друга, която ви интересува.

-
13. Калкулатор. Google е точен калкулатор от онзи тип, които имат функции с логаритми, синуси и всички елементи, които се моля да не ме срещат често във всекидневието. Напишете уравнение $"12+3456*7-890="$ и вижте резултата. При мен излезе 23 314.
 14. Търсач на пратки. Напишете номера на пратката, FedEx или UPS , и Google ще ви покаже информация къде се намира в момента.
 15. Конвертор на мерни единици. "meters in 70 yard"
 16. Брокер. Не че често се налага, но Google може да следи и стоковата борса. Напишете AAPL или MSFT и мигновено ще получите борсовите индекси на Apple и Microsoft.
 17. Карта. Напишете името на улицата, която търсите, града, плюс ключовата дума "map" в секцията за търсене на картинки (<http://images.google.com>) и ще получите карта на местността.
 18. Учител по правопис. Ако напишете някоя дума неправилно, той ще ви я покаже, както трябва. Но не се доверявайте прекалено много. Google показва правописа на думи, които могат да бъдат намерени в речниците.
 19. Речник. Като напишете "define:" и след това непознатата думи или акроним [примерно CPU или RAM], търсачката ще претърси целия Интернет и ще извади пред вас няколко обяснения за това, което сте въвели.

-
20. Информационен бюлетин. Казва се Google News Alerts. Въвеждате вашия мейл адрес, после ключови думи, после Google всеки ден ще сканира новинарските заглавия за вашите ключови думи и ще ви изпраща бюлетин. Услугата е на <http://www.google.com/newsalerts>
 21. Какво ново. С Google Web Alerts научавате и за всички нови уебстраници, посветени на дадена тема.
 22. Новинарска агенция. <http://news.google.com> индексира постоянно над 5400 новинарски сайта за събития от целия свят. Идеален начин да се ориентирате новините
 23. Лаборатория с иновации. На <http://labs.google.com> ще видите работните версии на бъдещите проекти, върху които работят софтуерните инженери от Google. Дават дори да си поиграете с тях.

Списък на главните търсачки в Интернет

- [Google](#) - безспорен лидер в Интернет пространството за момента, чиито резултати при търсене зависят от значението на съдържанието и популярността на интернет сайта. Всички страници складираны в каталога на Google могат да се видят във вида, в който са индексирани.
- [Yahoo!](#) показва резултати при търсене от своята собствена търсачка, анонсирана през Февруари 2004. В момента Yahoo! притежава платената търсачка Overture и доставчика на търсеци услуги Inktomi. Чрез Overture Yahoo! е собственик също и на търсачките AltaVista и Alltheweb, и двете сред най-популярните машини за търсене в мрежата.
- [MSN Search](#) е web търсачка от Microsoft, чийто резултати се захранват от директорията LookSmart модифицирани от собствен редакторски екип. Същевременно Microsoft работи по създаване на собствена търсачка и MSN ползва собствени работи сканиране и индексирание на мрежата.
- [Ask Jeeves](#) е търсачка редактирана от хора, която позволява на потребителите да задават въпроси без да познават техниката на търсене в мрежата. При търсене Ask Jeeves включва и резултатите от търсачките Teoma и Direct Hit, които притежава..



Списък на главните търсачки в Интернет

- ❑ [AOL International](#) позволява търсене в интернет и своята директория. Резултатите при търсене за интернет сайтове идват от Open Directory а Inktomi осигурява тези за web страниците.
- ❑ [AltaVista](#) е сред най-големите търсачки в интернет. Предлага превод на резултатите и търсене на мултимедия.
- ❑ [AllTheWeb](#) търсачка целяща да индексира всичко в Web. Предлага търсене на картинки, видео, MP3 and FTP.
- ❑ [Netscape Search](#) предлага търсене в интернет като резултатите идват от Open Directory и Google.
- ❑ [LookSmart](#) е web директория съставяна от редактори, която се използва от търсачки като MSN Search, Excite и други партньори.
- ❑ [Inktomi](#) е търсачка чийто каталог е наличен само чрез нейните партньори, независимо че техните резултати малко се различават.
- ❑ [Teoma](#) е crawler-базирана търсачка, която предоставя адекватни резултати при търсене както и някои допълнителни опции.
- ❑ [WiseNut](#) е търсачка собственост на LookSmart, представя адекватни резултати чрез нейния огромен каталог от web сайтове.
- ❑ [Lycos](#): е интернет търсачка с модел на директория, някои от резултатите идват от AllTheWeb и Open Directory project.

Използвани ресурси

- <http://alpharesearch.bg/>
- <http://exworthy.tripod.com/teachreswebeval.htm>
- Въпросник за оценка на уебсайтове
- Библиография, свързана с оценяването на информацията в уеб пространството

Благодаря ви за вниманието!

pmihova@nbu.bg