

Петко Димов

# SEO - ОПТИМИЗАЦИЯ НА САЙТОВЕ ЗА ТЪРСЕЩИ МАШИНИ

Как да управляваме информацията в социалните медии на Българското Интернет пространство?



София 2019



**Петко Димов**

# **SEO - Оптимизация на сайтове за търсещи машини**

**Как да управляваме информацията  
в социалните медии на Българското  
Интернет пространство?**



**София, 2019**

## Съдържание

1. Рецензии.....	7
2. Съкращения.....	10
3. Въведение и основни определения. Какво е SEO?.....	12
4. Принцип на работа на търсещите машини.....	17
5. Видове алгоритми за класиране в търсачките.....	22
5.1. Page Rank .....	22
5.2. Trust Rank.....	29
5.3. Други основни актуализации на алгоритъма.....	32
5.3.1. Google Panda.....	33
5.3.2. Penguin.....	33
5.3.3. Колибри.....	33
5.3.4. RankBrain.....	34
6. Оператори за търсене, символни команди и URL модификатори в Google.....	38
6.1. Символни команди в Google.....	38
6.2. Основни оператори за търсене.....	39
6.3. URL модификатори.....	41
6.4. Потребителски шаблони за търсене.....	42
7. SEO концепция и план за оптимизация на нов сайт.....	43
8. Избор на домейн и хостинг на сайт.....	48
8.1. Избор на домейн.....	48
8.2. Избор на хостинг.....	53
9. Избор на ключови думи.....	54
10. Мета тагове и код на сайта.....	60
10.1. Изисквания за мета тагове.....	60

## SEO

10.2. Особенности на кода, които влияят на вътрешните фактори ..62	62
10.3. Структурирани данни и микроформат Schema.org.....64	64
11. URL адреси.....66	66
12. SEO копирайтинг и съдържание на сайта.....69	69
12.1. SEO копирайтинг.....69	69
12.2. Видове уеб съдържание.....75	75
12.2.1. Уеб статия.....76	76
12.2.2. Ревюта, оферти и съобщения за пресата.....78	78
12.2.3. Постинг.....79	79
12.2.4. E-Mail.....80	80
12.3. Елементи и форматиране на Уеб текст.....81	81
12.3.1. Слоган и лого.....83	83
12.3.2. Заглави.....83	83
12.3.3. Подзаглавия.....86	86
12.3.4. Основен текст.....87	87
12.4. Стил, дизайн и принципи на рекламния текст.....96	96
12.5. Предпазване от кражба на уеб текст.....104	104
13. SEO архитектура и навигация на сайта.....105	105
14. Измерване и подобряване скоростта на отваряне на страниците в сайта .....112	112
15. Оптимизация на изображения и видео.....116	116
15.1. Употреба на снимки.....116	116
15.2. Употреба на видео.....118	118
15.3. Flash технология.....118	118
16. Link Building – техники за набиране на връзки.....122	122
17. Източници на уеб трафик.....126	126

18. Сайтове за социални мрежи.....	131
18.1. Facebook.....	139
18.2. Instagram.....	144
18.3. Twitter.....	147
18.4. Други социални мрежи.....	148
18.5. Организиране, планиране и автоматизация на съдържанието в социалните медии.....	151
18.6. Проследяване и анализ на информацията в социалните медии.....	156
19. Употреба на сателитни сайтове.....	161
20. Локално насочване на уебсайт.....	163
21. SEO за WordPress.....	165
22. Сигурност на сайт.....	168
23. Черно и сиво SEO.....	172
24. Измерване и анализ на ефективността на сайт.....	174
25. Управление на онлайн репутацията в сферата на отбраната...184	
25.1.Операции за постигане на информационно превъзходство..184	
25.2. Информационни оръжия.....	193
25.3. Интернет тролове и техники на тролене.....	200
Заключение.....	203
Литература.....	204
Приложение 1.....	215

- **Plugin** (plug-in) – подпрограма, която се добавя към главна програма, за да увеличи нейните възможности;
- **RSS** – Really Simple Syndication, формат позволяващ издърпване на информация от сайт, който предлага тази функция;
- **SEO** – Search Engine Optimization, Оптимизация за търсещи машини;
- **SEO Copywriting** е дейност извършвана от експерти, състояща в написване на текстовете в сайта, така че да отговарят на условията за класиране в начало от търсачките при търсене.
- **SMM** – Social Media Management, Оптимизация за социални мрежи;
- **SERPs** – Search Engine Result Pages, Страниците с резултатите при търсене в Интернет търсачка;
- **SPAM** – нежелано електронно съобщение, изпратено до множество потребители с рекламна или друга цел, често негативно възприемано от получателя.
- **SSL** – Secure Sockets Layer, криптографски протокол за връзка клиент-сървър;
- **Tag** (Tag или етикет) – Обикновено за SEO значение има фразата мета тагове, което всъщност представляват сбор от етикети, важни за търсачките и четими само от тях.
- **TF** – Trust Flow, метрика за измерване на доверието;
- **TR** – TrustRank, Доверителен ранг;
- **Traffic** (трафик на сайта) – брой уникални посетители на даден сайт, подава се дневно, седмично и месечно, т.е. разбит за определен период от време;
- **URL** – Uniform Resource Locator, стандартизиран уникален адрес сочещ към онлайн ресурс.
- **W3C** – World Wide Web Consortium, консорциум за международни интернет стандарти

### 3. Въведение и основни определения. Какво е SEO?

Девизът на Военна академия „Г. С. Раковски“ гласи: **знанието е сила**, а глобалната интернет мрежа е най-голямото хранилище на знание в днешния свят! След изобретението на Гутенберг през 1440 г. печатните книги са съхранявали това знание, но с появата на цифровата ера базите данни на уебсайтовете са се превърнали в основният източник на човешкото знание. Нещо повече, те са и инструмент за комуникация, което ги превръща в оръжие за масово поразяване на съвременното информационно общество. Този факт може да доведе, както до някои неблагоприятни последици за националната сигурност, така и до редица възможности за постигане на целите на държавата<sup>1</sup>. Това налага задълбоченото изучаване на интернет пространството и постоянно подобряване на уебсайтовете в него.

**Интернет** е глобална система от свързани компютърни мрежи, която служи за отдалечен достъп до голямо разнообразие от информационни ресурси и услуги. Тази мрежа се ражда през 1969 г. в резултат на програма на отбраната на САЩ, наречена APRANET (Advanced Research Projects Agency Network), чрез свързването на компютърните мрежи на няколко американски университета с цел създаване на устойчива и оцеляваща комуникационна система в случай на война. Първоначално тя се използва за изпращане на електронна поща, създаване на пощенските списъци, нюз групи, бюлетини въз основа на комуникационните протоколи TCP/IP. През 1984 г. е представена системата за имената на домейните (domain name addressing system), а през 1988 г. стават възможни разговорите в реално време, с помощта на развитието на протоколите Internet Relay Chat.

През 1990 г. APRANET престава да съществува и се ражда световната **уеб мрежа** – **World Wide Web**, която по идея на Тим Бърнърс-Лий представлява взаимно свързани хипертекстови документи, до които може да се достигне чрез компютърната мрежа с помощта на програми наречени уеб браузъри. Основното звено в

---

<sup>1</sup> Доктрина на въоръжените сили на Република България, София, 2017

тази мрежа са уебсайтовете, които се дефинират като интернет базирани приложения, състоящи се от множество уеб страници с общ URL адрес, който често се състои само от името на домейна или IP адреса и пътя до основната директория. Към настоящия момент Google е индексирала около 1,3 милиарда уебсайта във видимото интернет пространство<sup>2</sup>.

В първоначалният вариант на уеб мрежата т.н. **Уеб 1.0**, ресурсите в нея са достъпни с помощта на търсачки (като Google, Bing, Yandex и т.н.). В последствие се увеличава частта от уеб приложенията, в които взимат участие всички заинтересовани участници и потребителите започват да изразяват себе си чрез създаденото от тях съдържание. По този начин се ражда т.нар. уеб 2.0.

**Уеб 2.0** е наименование създадено от Тим О'Рейли през 2004 г. Съдържанието в този сегмент на интернет се създава най-вече от самите потребители, затова те се наричат „новите медии“ и в тях делът на видео и фото съдържанието нараства с ускорени темпове<sup>3</sup>. Днес почти всички уебсайтове са обвързани с останалите социални платформи и публикуваното в тях съдържание автоматично се разнася в социалните мрежи.

**Уеб 3.0** се пови през последните години като "мрежа от мобилни приложения", в което потребителите просто установяват връзка между притурката си и ресурса (услуга, портал и т.н.) чрез интернет. Понастоящем се говори за „**Интернет на нещата**“, който се състои от вградени информационни блокове, свързани чрез интернет, с контролни центрове на различни обекти от физическия свят. Така, например, към световната мрежа са свързани всевъзможни домакински уреди, производствени линии, системи за управление на водите, за отопление, електропренасяне и така нататък.

Голяма част от уебсайтовете са в т.н. „**невидим интернет**“, този не се индексира от търсачките. Според наличната информация около 90% от цялото научно, техническо, технологично, финансово, икономическо и обществено съдържание се намира именно тук.

---

<sup>2</sup> <http://www.internetlivestats.com/> [прегледан 30.01.2018]

<sup>3</sup> O'Reilly, T. What is Web 2.0. Design Patterns and Business Models for the Next Generation of Software, p. 30, 2005

Невидимия интернет има и т.н. **тъмна мрежа**, която се използва от различни видове престъпници, чрез софтуер наречен Тор. Понастоящем той се използва главно за незаконни дейности, престъпления в киберпространство, трафик на наркотици, оръжия и за осъществяване на целенасочени действия за подкопаване на държавния суверенитет.

Специален сегмент е така наречената **мрежа от пари**. Световната тенденция е намаляването на оборота на плащанията в брой и прехода към електронни пари във всичките им форми и системи от рода на „Биткойн“.

По този начин цифровата среда на интернет пространството има сложно картографиране, където отделни сегменти се развиват сами, независимо от общите тенденции.

Защитата на информацията в интернет пространството се характеризира с три основни параметъра:

- *Конфиденциалност* – това са номерата на банкови сметки, лични данни, пароли, лична и служебна кореспонденция и всичко, което би трябвало да се вижда само от този, за които е предназначено.

- *Достоверност* – тя може да обхваща, както същността и целостта на данните, така и техния произход. Примери в тази посока има достатъчно много – фалшифициране на избори, умишлено манипулирани статистически данни, а във военната област – дезинформация на противника.

- *Достъпност* – информацията е ценна именно с това, че може да бъде използвана. Затова основен проблем за всяка информационна система е да се намери златната среда между нейната защита и използваемост. Именно това противоречие създава уязвимостите в информационните системи.

Към месец декември 2017 г. в Интернет има над 3,8 милиарда потребители, които всяка секунда правят по 2,3 милиарда търсения. По данни на Националния статистически институт за 2017 г. цели 67,3% от българските домакинства имат широколентов достъп до

интернет, като 98 % от тях използват търсачката Google.BG за търсене на информация в световното интернет пространството<sup>4</sup>.

Днес, в съвременното информационно общество не е достатъчно просто да имате уебсайт, а е жизнено необходимо той да излиза на първите места в резултатите на търсачките, защото огромната част от техните потребители отварят само сайтове, показани на първа страница в резултатите от търсене.

Дейността по популяризиране и подобряване условията за тяхното класиране се обозначава със съкращението - SEO.

Акронимът **SEO** идва от английското съкращение „**Search Engine Optimization**“, което на Български език се превежда „**оптимизация на уебсайтове за търсещи машини**“.

Първата употреба на този термин датира от средата на 90-те години с появата на първите мета търсачки в интернет пространството. След това масово навлиза в онлайн бизнес услугите, с което SEO се превръща във важна част от маркетинговия микс.

Така в информационните технологии се появи ново направление, а в науката за **Интернет маркетинг** се създадоха разделите: *оптимизация за търсачки, оптимизацията за социални медии, управление на онлайн репутацията и платената реклама на клик, обозначавани със следните съкращения SEO, SMO, ORM и PPC.*

SharePoint експерта Mastykarz определя SEO като общото название на редица дейности, които се фокусират върху повишаване на класирането на даден уебсайт в резултатите от търсене<sup>5</sup>. По отношение на областта на приложение, той разделя SEO на две категории: свързани с технологиите и свързани със съдържанието.

Докато Андерсън твърди, че **SEO е технически, аналитичен и творчески процес** за подобряване на видимостта на даден уебсайт в търсачките. Основната му функция е да стимулира

---

<sup>4</sup> Достъп на домакинствата в интернет <http://www.nsi.bg/bg/content/2808> /достъп-на-домакинствата-до-интернет [пегледан 30.01.2018]

<sup>5</sup> Search Engine Optimization in SharePoint 2013 <https://blog.mastykarz.nl/search-engine-optimization-sharepoint-2013/> [пегледан 30.01.2018]

повече посещения на уебсайта, които се превръщат в интеракции (взаимодействие или продажби)<sup>6</sup>.

В крайна сметка може да се обобщи, че **от техническа гледна точка SEO е непрекъснат процес, комплекс от дейности за развитие и подобряване видимостта на уебсайтовете и тяхното класиране в органичните (алгоритмични) резултати след провеждане на търсене в търсачката.**

От **научна гледна точка SEO** е свързано с изучаване на това как работят търсещите машини, как тяхната работа влияе на развитието на web базираните информационни системи и управлението на информацията, която търсят хората в Интернет пространството.

Според насочеността на техническите дейности за развитие на уебсайтовете SEO се разделя на два основни вида:

**On-page SEO – вътрешна оптимизация.** Това са дейностите ориентирани към конфигуриране на структурните елементи вътре в уебсайта спрямо техническите изисквания на търсещите машини<sup>7</sup> с цел подобряване на вътрешните фактори, влияещи на общото му класиране.

**Off-page SEO – външна оптимизация.** Тя включва комплекс от дейности и мероприятия, които се прилагат извън страниците на уебсайта с цел увеличаване на неговия ранг или входящ трафик и в крайна сметка да за повишаване видимостта и популяризиране на сайта.

#### **Вътрешните фактори** включват:

- различни параметри свързани с кода (валидност по W3C, намаляване тежестта на HTML и CSS файловете, включване на мета тагове),
- структурата на сайта и наличие на sitemap, SEO приятелски URL адреси,

---

<sup>6</sup> Anderson, S. How to write SEO – Friendly Web Copy for Google. <http://www.hobo-web.co.uk/how-to-write-seo-friendly-website-content-for-google/> [Accessed 11 March 2017.]

<sup>7</sup> Technopedia explains on-page optimization. <https://www.techopedia.com/definition/1548/on-page-optimization> [прегледан 17.01.2017]

- параметри на текста (дължина, плътност, уникалност, наситеност), изображенията и т.н.

**Външните фактори** се определят като изграждане на онлайн авторитет и репутация на съответния сайт в Интернет пространството. Най-често това се постига с техника наречена link building за изграждане на хипервръзки от други сайтове с релевантна тематика. За набирането на такива връзки се използват портали, търсачки, социални мрежи, директории за статии и всички други възможни места с отчитане на социалните сигнали и поведенческите фактори на потребителите.

В крайна сметка SEO е иновативна технология, която има определени общи принципи за повишаване ефективността на уебсайтовете и е постоянно променящ се процес, който трябва да се поддържа непрекъснато през целия живот на сайта.

Тази технология е необходима за да се направят уебсайтовете по-лесни за ползване от хората и по-разбираеми за роботите на търсачката. Това помага на търсещите машини да разберат за съдържанието на уебсайта, да го класифицират и да го предложат по полезен начин на потребителите.

## 4. Принцип на работа на търсещите машини

Търсещите машини се намират в центъра на информационните операции в интернет, като всеки ден насочват милиарди интернет потребители към уебсайтове, които се появяват в горната част на страницата с резултатите от търсене.

Според повечето автори търсачките са **"приложения, които използват ключови думи и фрази за намиране на информация в Интернет"**<sup>8</sup>. Тези приложения включват няколко основни компоненти<sup>9</sup>:

---

<sup>8</sup> Tilley, S. & Rosenblatt, H. Systems Analysis and Design. Eleventh edition. Boston, MA: Cengage Learning

<sup>9</sup> <https://sup-port.google.com/uebmasters/answer/1061943?hl=en> [прегледан 17.01.2017]

- интерфейс за търсене;
- уеб паяци (ботове, работи);
- бази данни;
- алгоритми за търсене и извличане на информация.

**Интерфейсът за търсене** е страницата, където се пишат заявките за търсене. Той е връзката между потребителите и базите с данни. В Google интерфейсът е индивидуален, така че потребителите могат да избират няколко опции, като новини, видеоклипове, изображения, публикации и т.н.

**Ботовете** се дефинират от Google като "общ термин за всяка автоматизирана програма (робот или паяк), използвана за автоматично откриване и сканиране на уебсайтове, като следват връзки от една уеб страница към друга<sup>10</sup>." Уеб ботовете постоянно "обхождат" мрежата за нова информация, като по график проследяват поставените връзки в първоначално създаден списък от сайтове и изтеглят новите страници във формат HTML за последваща обработка и класифициране.

**Базите данни** съхраняват информацията, извлечена от уеб ботовете (роботите, паяците) и я класифицират в подходящ за възпроизвеждане вид. Всеки път, когато правим заявки в търсачката, ние търсим информация в базите данни.

**Алгоритмите за търсене и класиране** се определят от Google като "компютърни процеси и модели, които правят търсенето значително по-лесно и дават релевантен отговор на запитванията на потребителите<sup>11</sup>". Всяка търсачка използва различни алгоритми, така че при една и съща заявка за търсене ще се видят различни варианти на резултата.

Ботовете, базите данни, интерфейсът за търсене и алгоритмите са водещи технологии за търсачката, която показва крайния резултат от тяхната работа в **интерфейса за търсене**.

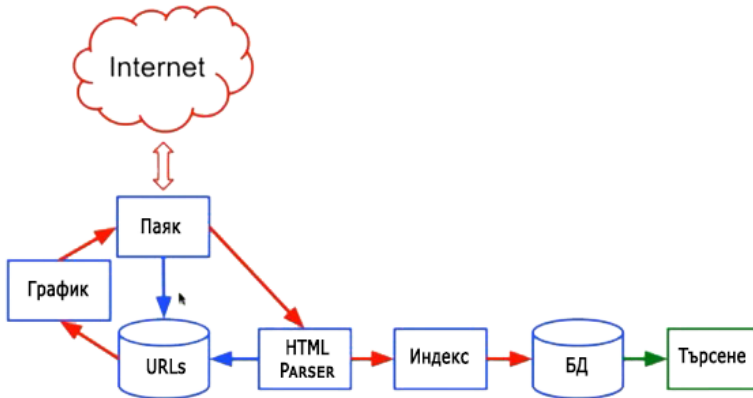
При създаването на една чисто нова уеб страница в Интернет търсещите машини не знаят за нейното съществуване и съответно тя е невидима за хората, които не влизат в нея. Тя не се появява в

---

<sup>10</sup> Google Search Console. <https://search.google.com/search-console/>

<sup>11</sup> Google Inside <https://www.google.com/in-sidesearch/howsearchworks/algorithms.html>

результатите от търсене, докато Google не я **индексира**. Това значи ботовете (обхождащите паяци) на търсачката да я намерят, обработят и да я съхранят в базата данни на Google, където се класифицира и подрежда по тематика и ключови думи.



Фиг.1 Архитектура на уеб робот

В своята работа уеб ботовете "обхождат" мрежата за нова информация движейки се по връзките между страниците. При посещение на индексирана страница те проследяват всички хипервръзки в нея и ако открият нова, неиндексирана страница сканират кода на страницата, заглавията, уеб съдържанието, търсят етикети, описания и инструкции за да научат повече за нейната тема, кой стои зад това, какво прави и какво го интересува, след което записват новите данни с определена точност в базите данни (индексират ги).

Това не е малко работа, като се има в предвид, че всяка минута се изграждат между 300 – 500 нови уебсайта, всеки от които има средно 7 страници.

По данни на Internetlivestats.com от септември 2017 г. Google обработва почти 3,8 милиона търсения всяка минута или 5,5 милиарда на денонощие, а в техния индекс се съхранява информация за над 30 трилиона уникални уеб страници и над 100