Малка приказка за търсещия бот

Спомняте ли си, когато Хари Потър и неговите приятели Хармаяни и Рон попаднаха на кучето пазач в забранения коридор на Хогуортс. Много искаха да разберат какво пази, но единственото което знаеха е, че пазеното е свързано с някой си Никълъс Фламел. Естествено не знаеха кой е Никълъс Фламел и първото което направи Хармаяни е да накара и тримата да търсят информация за него в библиотеката. Спомняте си, че търсенето не беше много успешно. Обаче си мисля, че ако имаха компютри и интернет нещата щяха да са различни. Според мен днес, като търсим информация за нещо почти винаги започваме от търсене в интернет.

Има много възможности компютъра да ти помогне в търсенето, но най-често използвания начин е да попитаме Google. Всички подобни на Google помощници се наричат търсачки. Направо изглежда като магия – написваш каква информация ти е необходима и търсачката ти показва списък с много страници в които пише нещо за това. Понякога списъка е толкова дълъг, че се налага да кажеш какво по-конкретно търсиш, като напишеш не една-две думи, а повече. Можеш и да използваш знака (-) за да кажеш, че следващата дума не трябва да се среща в показаните страници. Или пък да използваш знака (\*) за да кажеш, че се интересуваш не само от това което пише преди знака, а и от всичко, което започва с него но продължава както си иска.

Първата търсачка е създадена почти веднага със стартирането на интернет и е използвала каталог на страниците. Каталога на страниците е точно като каталозите в библиотеките. В библиотеките всяка книга е описана на едно картонче, а картончетата са подредени в чекмеджета по различни начини – например по година в която е издадена книгата, или по име на автора или по заглавие на книгата. Така ако в една библиотека има милион книги, но ти търсиш точно „Хари Потър и философския камък“ библиотекарката ще ти каже да намериш чекмеджето с буква Х, после в него да потърсиш заглавието на книгата и на картончето, което ще намериш, ще пише къде се намира книгата – например коридор F, ред 12, стелаж 5. И така ти няма да разглеждаш всичките милион книги докато търсиш твоята, а бързо ще прегледаш картончетата (които са подредени, нали запомни) и ще откриеш търсената книга.

Скоро след това хората решили, че търсенето и подреждането на информацията за това какво съдържат страниците в интернет трябва да се повери също на компютри и създали търсещи роботи. Всеки търсещ в мрежата робот се наричат накратко бот. Търсещият бот постоянно обхожда мрежата –всеки ден, всеки час, всяка минута и подрежда страниците в каталози. Това се нарича индексиране, но всъщност си е едно обикновено подреждане по различни признаци. Когато всичко е подредено е по-лесно да намериш каквото търсиш. Не само при компютрите е така. Ако стаята ти е разхвърляна по-трудно си намираш играчките. И ако майка ти е като мен значи често чуваш правилата – дрехите веднага да отидат в чекмеджето с дрехите, а играчките в коша с играчки. И в чекмеджето с дрехи трябва да има ред – чорапите отделно от блузите☺. И блузите подредени на купчинки – една купчинка за тези с къс ръкав и една за тези с дълъг. Това са все признаци за подреждане, а търсещия бот спазва много такива правила. Обаче резултатът е страхотен – само казваш от какво се интересуваш и Google ти показва къде да го намериш.

Ботът разглежда страницата и решава кои от думите в нея са важни и ги записва на „картончето“ на тази страница, след това отива на следващата страница. Тъй като преместването от страница в страница прилича на движението на паяк, който си плете мрежата, понякога търсещия бот се нарича паяк. Търсещия бот се грижи за това да прецени дали тази страница ще е важна за теб или не. Това се нарича рейтинг на страницата. Колкото по-висок рейтинг има една страница, толкова по-напред в списъка ще ти я покаже търсачката. За да не се налага да преглеждаш целия списък.

Изобщо търсещия бот върши страхотна работа. Но за да се научи на това е трябвало някой да измисли правилата по които търсещия бот да решава, коя следваща страница да посети и как да и сложи рейтинга. Търсещия бот, както и всички останали действия, които автоматите и програмите предприемат, всъщност са им наредени от хората. Всички те правят това, което хората са им казали да правят. Никак не е лесно да се научи един търсещ бот да взема решения. Колкото по-добре хората които са го създали са му казали какво да прави, толкова по-добре ботът ще си върши работата. Ако един бот е отличен търсач, то това означава, че хората които са го написали са били отлични програмисти. За това, ако искаш да напишеш собствен бот, който да търси нещо в мрежата по-добре започни да си подреждаш стаята, чекмеджетата и слагай нещата в ред, който ти помага да ги намираш по-лесно ☺.