

Vyhledávače

Otázky a odpovědi

Co je to internetový vyhledávač? Jak funguje?

Internetový vyhledávač je služba, která nám umožňuje pracovat s internetem při vyhledávání informací tak, jak to dnes běžně známe. Dříve si lidé museli pamatovat přesné adresy webových stránek, protože nebyl žádný nástroj, který by jim umožnil jednotlivé webové stránky postupně vyhledávat nebo procházet. Webových stránek však bylo stále více a tento způsob prohlížení internetu se velmi rychle stal nepoužitelným. Internetové vyhledávače byly tedy vytvořeny především proto, aby bylo **vyhledávání na internetu efektivnější a pohodlnější**. Jako uživatelé zadáváme do vyhledávače **klíčová slova** a jsou nám vráceny výsledky, které systém vyhledávače vyhodnotil jako nejvhodnější. Výsledkem vyhledávání však nemusí být jen webové stránky, může se jednat i o **videa, obrázky** nebo **další informace**.

Co jsou algoritmy?

Algoritmus je návod, postup nebo posloupnost jednotlivých kroků, které umožní vyřešit zadanou úlohu, algoritmy mají nějaký výsledek. Nejčastěji se toto slovo vyskytuje ve spojitosti s programováním, ale už z principu můžeme jako algoritmus chápat celou řadu všedních věcí, které kolem sebe vidíme v běžném životě. Dobrým příkladem může být vaření dle receptu, které nás vede krok po kroku až k požadovanému výsledku. Základním stavebním kamenem algoritmů jsou **instrukce a příkazy, které mají být naplněny v přesném pořadí**.

Proč jsou algoritmy tajné?

Zveřejnění vyhledávacích algoritmů by znamenalo **odhalení obchodního tajemství** a současně je na internetu mnoho spamových, nebezpečných a závadných webů, které se na první místa ve výsledcích vyhledávání pokoušejí dostat **pomocí nekalých praktik**. Zveřejnění algoritmů by výrazně zjednodušilo hledání způsobů, jak obcházet různá pravidla, a výsledky vyhledávání by zobrazovaly weby, o které nestojíme a nezajímají nás.

Jaké vyhledávače jsou v ČR nejrozšířenější?

Drtivou část (až 96 %) vyhledávání na českém internetu si mezi sebou dělí **Google** a **Seznam**. Ve zbylých procentech jsou zastoupeny ostatní vyhledávače jako Yahoo a Bing. Procentní podíl Seznamu a Googlu je přibližně v poměru 25:75. Měřit to je však složité a proměnlivé, záleží například, z jakých zařízení se uživatelé připojují a o jaký přístup k vyhledávání jde (viz níže).

V čem se jednotlivé vyhledávače mezi sebou liší?

Existují dva základní přístupy vyhledávání informací na internetu a klasifikace webových stránek. Prvním z nich jsou **vyhledávače**, druhým jsou **předmětové katalogy**. Dnes již toto striktní dělení není možné, řešení bývají smíšená, ale stále převažují charakteristiky jednoho nebo druhého přístupu. Vyhledávače fungují následovně: tzv. **indexovací roboti** prochází celý web, a pokud naleznou novou stránku nebo obsah, indexují ji (uloží) do své databáze (tzv. indexu). Takovou stránku automaticky označí klíčovými slovy, přiřadí jí určité charakteristiky a tím později umožní tuto stránku vyhledat. Pokud hledáme patřičná **klíčová slova**, vyhledávač nám nabízí právě tyto vhodné weby. Slabinou tohoto přístupu je, že eviduje i různé pochybné stránky, protože automatické mechanismy je nedokážou správně vyhodnotit. Tak funguje například Google. Naopak **předmětové katalogy** jsou vytvářeny převážně manuálně. Obsah internetu je v tomto případě zpracováván **lidmi**, nikoliv roboty. Výsledky předmětových katalogů jsou zpravidla mnohem užší než u vyhledávačů. Neobsahují tolik výsledků, výsledky jsou však lépe strukturované a řazené a jako uživatelé postupujeme různými kategoriemi. To oceníme nejvíce v momentě, kdy víme, co přesně hledáme. Propracovaný předmětový katalog má například Seznam, který však ovládá i funkci vyhledávání. To jsou všechny ty záložky jako Firmy.cz atp. Google nic takového nemá.

Jak jsou výsledky vyhledávání řazené? Vidíme všichni stejné výsledky?

Výsledky vyhledávání jsou řazené **algoritmy** a každý vidíme **jiné výsledky na stejná klíčová slova / dotazy**. Stojí za tím vyhledávací algoritmus, který se snaží výsledky vyhledávání **přizpůsobit „na míru“ tomu, kdo se ptá**. Má to hned dva důvody. Na straně jedné je díky nim internet pro uživatele **přívětivější**, protože snáze najdeme, co potřebujeme. Na straně druhé jde o **zisk**. Velká část příjmů internetových společností plyne z reklamy, která je, stejně jako výsledky vyhledávání, personalizovaná. Jinými slovy „nabízená na míru“. To spolu souvisí. Čím lépe se vyhledávači podaří nabídnout nám relevantní obsah, tím více času na internetu strávíme. Internetové společnosti tím také získají více prostoru pro to „podstrčit“ nám reklamu, která nás pravděpodobně zaujme.

Jak mohou provozovatelé stránek ovlivňovat, aby se jejich web zobrazil mezi prvními?

Existují **placené** (SEM – search engine marketing) i **neplacené** (SEO – search engine optimization) **metody**, kterými lze dosáhnout lepších výsledků ve vyhledávání. **Placené metody** jednoduše umožní posunout určité weby výše, protože někdo zaplatil za jejich inzerci. Jako uživatelé můžeme poznat, zda nějaký web za zobrazení platí, příspěvky bývají zobrazené jako „ad“, „reklama“ či podobně. Formy placení však mohou být různé – existují i modely, kdy svůj web poskytneme jako reklamní prostor a za to se ocitneme výše ve výsledcích. Pokud chceme ovlivnit výsledky vyhledávání a dostat se na lepší místa, můžeme to do značné míry ovlivnit i sami – pomocí **neplacené metody** (SEO). Je důležité dobře naplnit web informacemi, které nevidí uživatel, ale vyhledávač ano. Jedná se například o klíčová slova, která jsou s obsahem našeho webu spojená. Vyhledávač díky těmto informacím lépe ví, co od našeho webu vlastně může očekávat. Lépe si s naším webem rozumí, a tím pádem ho upřednostní. Pokud bychom však svůj web naplnili informacemi a klíčovými slovy, která jsou často vyhledávaná, ale s naším webem nemají nic společného, vyhledávač to časem odhalí a za tento podvod náš web naopak velmi upozadí.

Jaká osobní data o nás vyhledávače sbírají?

Problematika sběru dat se netýká jen vyhledávačů, ale i dalších služeb a aplikací, neboť jsou všechny různými způsoby propojené a umí spolu komunikovat. Jako příklad můžeme uvést sledování videa: když se jako přihlášený uživatelé na YouTube budeme dívat na video na svém notebooku a zastavíme ho, můžeme později pokračovat třeba z mobilního telefonu přímo tam, kde jsme přestali. **Vyhledávače jako takové mohou například sledovat naši aktivitu, čas strávený na stránce nebo polohu.** Informace z e-mailu či kalendáře se mohou propsat až do našeho vyhledávání, pokud máme služby propojené. O tom, jak a jaká data se sbírají, rozhoduje řada věcí. Důležitým faktorem například je, zda **jsme přihlášení svým uživatelským účtem** do služeb. Google nabízí přehled našich přihlášených aktivit včetně časového záznamu: na jaká videa jsme se dívali, čemu jsme dali lajk nebo dislajk, co jsme komentovali, jaké stránky jsme navštívili, kde jsme byli a tak podobně. Vše je k vidění zde: <https://myactivity.google.com/myactivity> Často můžeme také vidět, že nás webová stránka informuje o tom, že pracuje s cookies. Cookies jsou malé a užitečné soubory, které umí například samy vyplnit kontaktní či objednávací formulář, který jsme už jednou vyplnili, aniž bychom museli vyplňovat znovu. To však není záležitost vyhledávání ani vyhledávačů, ale webových prohlížečů.

Lze sběr osobních dat vyhledávači nějak minimalizovat?

Pokud chceme eliminovat množství dat, která jsou o nás sbírána a zaznamenávána, můžeme například využít funkce **anonymního prohlížení**. Takové prohlížení nám ani zdaleka nezaručí stoprocentní anonymitu směrem do internetu, ale významně **zredukuje množství dat, která o nás sbírá prohlížeč nebo služby**. Problematika sběru dat se však netýká jen vyhledávačů. Opět nezapomínejme na to, že služby či aplikace jsou dnes propojené různými způsoby a umí spolu komunikovat. Je tedy důležité **dbát i na to, jaké aplikace si stahujeme, jaká oprávnění od nás chtějí a co jim dovolujeme**. Google také ve svém pokročilém nastavení umožňuje omezit či zakázat sběr určitých dat. Musíme se pak ale smířit s tím, že výsledky vyhledávání nebudou tolik přesné, jak jsme většinou zvyklí. Existují také vyhledávače jako například www.duckduckgo.com, které se profilují tím, že žádná data o uživatelích nesbírají.

Jak zpřesnit výsledky vyhledávání?

Čím přesněji je položený dotaz, tím relevantnější odpověď lze vyhledávačem získat. Jednoslovná hesla nemusí vždy stačit, a tak je potřeba v dotazu specifikovat mnohem více podmínek a pravidel, které zužují portfolio odpovědí a filtrují pryč ty méně relevantní. Tyto složitější dotazy se ve vyhledávačích vytvářejí pomocí **správné dotazovací syntaxe** (skladby) a **operátorů**. Takto můžeme zpřesnit dotaz **přidáním dalších klíčových slov** (operátor AND), **výběrem** některých z nich (operátor OR), nebo například **vyločením stránek** dle určitých pravidel. Zároveň lze v dotazu omezit odpovědi jen na určitý typ souboru, umístění textu, umístění v konkrétní URL adrese, zda se klíčové slovo vyskytuje v metadatech nebo zda se vyskytuje nějaký prvek v určitém rozsahu. Kromě syntaxe a operátorů se lze i přímo v nastavení vyhledávání omezit i na konkrétní jazyk či časové období.

Odpovědi vypracoval: David Kudrna, Národní úřad pro kybernetickou a informační bezpečnost

Odkazy:

Příručka informační gramotnosti

<https://gramotnost.info/>

Kurz informační gramotnosti pro studenty středních škol

<https://kisk.phil.muni.cz/onlife>

Doporučená literatura:

ČERNÝ M., CHYTKOVÁ D., MAZÁČOVÁ P., ŠIMKOVÁ G.: Informační vzdělávání pro učitele. Brno: Flow, 2015.

ČERNÝ M.: Digitální informační kurátorství jako univerzální edukační přístup. Brno: Masarykova univerzita, 2017. Dostupné: <https://munispace.muni.cz/book?id=962>