Czytaj lub słuchaj artykuł online

Zeskanuj kod QR, aby otworzyć artykuł na XO.PL

Źródło artykułu:

https://xo.pl/it-na-wesolo-inet-search

INTERNET

Wyszukiwarki, czyli bibliotekarz, który próbuje skatalogować cały Internet

Pani Agnieszka postanowiła sprawdzić, czy Grzesiu i jego kolega odróżniają wyszukiwarkę od przeglądarki. Szybko odkryła, że czeka ją najdłuższa wycieczka do biblioteki w dziejach Internetu. 🔎📚🤣

— Jakiej wyszukiwarki używasz? — zapytała pani Agnieszka.

— Google'a — odpowiedział Grzesiu.

— A ja Opery — oznajmił jego kolega.

Zapadła cisza. Agnieszka odłożyła kubek.

— Opera to PRZEGLĄDARKA.

— A co za różnica?

— Taka, że właśnie pomylił pan autobus z informacją, na którym przystanku wysiąść!

Przeglądarka to pojazd, wyszukiwarka to informator

Pani Agnieszka rozłożyła na biurku mapę miasta.

— **Przeglądarka internetowa** to program, który otwiera i wyświetla strony WWW. Chrome, Firefox, Opera, Edge i Safari — to przykłady przeglądarek.

— A Google?

— Zależy, co pan ma na myśli. **Google Search** to wyszukiwarka. **Google Chrome** to przeglądarka. Jedna firma ma oba produkty.

— A Microsoft?

— **Bing** to wyszukiwarka Microsoftu, a **Edge** to jego przeglądarka.

— Czyli mogę w Operze używać Google?

— Tak. Albo Binga, DuckDuckGo czy innej dostępnej wyszukiwarki.

— A w Chrome mogę szukać Bingiem?

— Oczywiście, jeśli tak skonfiguruję przeglądarkę.

Przeglądarka jest jak samochód, którym jedziesz pod adres strony. Wyszukiwarka jest jak punkt informacji, który pomaga znaleźć ten adres, gdy go nie znasz.

— Czyli „jeżdżę Google'em” to odpowiedź na inne pytanie?

— Jeśli chodzi o Google Search, to nawet nie jest samochód!

Po co wyszukiwarka, skoro jest pasek adresu?

— To dlaczego wpisuję w Operze „przepis na pierogi” i od razu dostaję wyniki? — spytał kolega Grzesia.

— Bo przeglądarki łączą dziś pole adresu z wyszukiwaniem. Gdy wpiszesz słowa zamiast znanego adresu WWW, przeglądarka może wysłać zapytanie do ustawionej wyszukiwarki.

— Czyli Opera sama nie przeczesuje Internetu?

— Samo wpisanie zapytania w jej pasek nie oznacza, że to Opera prowadzi własną wyszukiwarkę. Korzysta z wybranej usługi wyszukiwania.

— A gdy wpiszę `https://xo.pl/`?

— Przeglądarka próbuje otworzyć wskazaną witrynę bez potrzeby szukania jej w katalogu wyszukiwarki.

— Czyli mogę znać adres i w ogóle nie pytać Google'a?

— Tak. Tak właśnie działa bezpośrednie odwiedzanie stron.

— To po co wyszukiwarka?

— Gdy nie zna pan adresu albo szuka informacji rozrzuconych po wielu witrynach.

— Czyli tak naprawdę przez lata pytałem bibliotekarza, myśląc, że rozmawiam z kierowcą?

— Dokładnie.

Bibliotekarz nie czyta całego Internetu w sekundę

Agnieszka zaprowadziła Grzesia do wielkiej biblioteki. Po podłodze jeździły małe roboty z pieczątkami.

— O! Odkurzacze! — ucieszył się Grzesiu.

— Nie. To **roboty indeksujące**, zwane crawlerami. Odwiedzają strony, odkrywają odnośniki, pobierają informacje i przekazują je do dalszego przetwarzania.

— To każde moje pytanie wysyła je w podróż przez cały świat?

— Nie. Wyszukiwarka zwykle odpowiada, korzystając z wcześniej przygotowanego **indeksu**. To ogromny katalog opisujący znalezione zasoby. Roboty stale go aktualizują.

— Czyli bibliotekarz nie czyta milionów książek dopiero wtedy, gdy poproszę o jedną?

— Właśnie. Ma kartoteki i system wyszukiwania.

— A jeśli strona powstała przed sekundą?

— Może jeszcze nie zostać odkryta ani zindeksowana.

— A jeśli ktoś skasował stronę?

— Stary wpis może przez pewien czas jeszcze widnieć w wynikach.

— Bibliotekarz ma opóźnienie?

— Tak. I proszę nie zgłaszać na niego reklamacji do działu zegarków.

Robot przy drzwiach i kartka robots.txt

Jeden z robotów stanął przed drzwiami z napisem „Nie wchodzić”.

— A to co? — spytał Grzesiu.

— Właściciele witryn mogą publikować wskazówki dla robotów, między innymi w pliku `robots.txt`. Mogą też używać innych mechanizmów ograniczających indeksowanie.

— Czyli kartka na drzwiach zamyka wszystko na klucz?

— Nie. **robots.txt nie jest zabezpieczeniem dostępu do poufnych danych.** To zestaw reguł dla współpracujących robotów, a nie mur chroniący pliki przed każdym odwiedzającym.

— To gdzie trzymać tajne dokumenty?

— Za rzeczywistą kontrolą dostępu, a nie tylko za uprzejmą prośbą do crawlera.

— A wszystkie strony trafiają do wyszukiwarki?

— Nie. Część jest nieodkryta, zablokowana, niedostępna po zalogowaniu albo po prostu pominięta.

— Czyli brak wyniku nie dowodzi, że czegoś nie ma?

— Dokładnie. Podobnie jak brak pana skarpetek w szufladzie nie dowodzi, że nigdy ich nie kupił.

— Proszę nie wracać do skarpetek.

Ranking, czyli kto stoi pierwszy na półce

Grzesiu poprosił bibliotekarza o „najlepszą pizzę w mieście”. Na ladę spadło tysiąc kartek.

— Którą mam przeczytać?!

— Wyszukiwarka stara się uporządkować wyniki według przewidywanej przydatności. Bierze pod uwagę wiele sygnałów, zależnie od systemu: treść, dopasowanie do pytania, jakość i inne czynniki.

— Czy pierwsza strona to prawda objawiona?

— Absolutnie nie. **Wysoka pozycja nie gwarantuje prawdziwości ani bezstronności informacji.**

— A czemu sąsiad ma inne wyniki?

— Znaczenie mogą mieć język, region, ustawienia, historia lub inne elementy kontekstu. Różne wyszukiwarki mają też różne indeksy i mechanizmy ustalania kolejności.

— Czyli bibliotekarz nie podaje całego Internetu w jednej obiektywnej kolejności?

— Nie.

— A jeśli będę pytał coraz głośniej?

— To tylko zirytuje ludzi przy innych stolikach.

Reklama w bibliotece?

Na pierwszej stronie wyników Grzesiu zobaczył kilka pozycji oznaczonych jako sponsorowane.

— O! Bibliotekarz poleca właśnie te?

— Część wyników to **reklamy**, wyświetlane zgodnie z zasadami danej usługi. Powinny być odróżnione od zwykłych rezultatów wyszukiwania.

— Czyli ktoś zapłacił za miejsce w części reklamowej?

— Tak. To nie jest to samo co organiczny ranking.

— A SEO?

— **SEO**, czyli optymalizacja pod wyszukiwarki, polega między innymi na tworzeniu przydatnej, dostępnej technicznie treści, którą roboty mogą zrozumieć. Nie oznacza automatycznego kupienia sobie pierwszego miejsca.

— Jeśli napiszę na stronie słowo „pizza” tysiąc razy, to zostanę królem pizzy?

— Raczej zostanie pan królem irytującego tekstu.

— A jeśli ktoś reklamuje lekarstwo, to znaczy, że działa?

— Nie. Wyniki i reklamy warto oceniać krytycznie, zwłaszcza gdy chodzi o zdrowie, finanse lub bezpieczeństwo.

— Czyli wyszukiwarka znajduje informacje, ale nie bierze za mnie odpowiedzialności za myślenie?

— Wreszcie dotarliśmy do sedna.

Google, Bing, Yahoo i reszta biblioteki

Przy wejściu wisiało kilka szyldów.

— **Google Search**, **Microsoft Bing**, **Yahoo Search**, **DuckDuckGo**, **Brave Search**, **Ecosia**… — czytał Grzesiu.

— Aż tyle bibliotek?

— To różne usługi wyszukiwania. Mogą inaczej pokazywać wyniki, inaczej traktować prywatność i opierać się na własnych indeksach lub częściowo na rezultatach partnerów.

— Czyli Yahoo nadal istnieje?

— Tak, jako usługa wyszukiwania. Nie wszystko, co pamięta pan z dawnych lat Internetu, zniknęło.

— DuckDuckGo to przeglądarka czy wyszukiwarka?

— Marka oferuje wyszukiwarkę, ale także własną przeglądarkę. Nazwa firmy nie zawsze oznacza jeden rodzaj programu.

— A Brave?

— Podobnie: Brave Browser jest przeglądarką, Brave Search — wyszukiwarką.

— A Ecosia?

— To wyszukiwarka, która może korzystać z różnych dostawców wyników. Nie każda usługa musi samodzielnie indeksować cały Internet.

— Czyli najpierw sprawdzam nazwę PRODUKTU, a nie tylko producenta?

— Dokładnie. Google i Chrome to nie synonimy, Bing i Edge też nie.

— Szkoda. Lubiłem mieć wszystko w jednej szufladzie.

Prywatność, czyli kto widzi moje pytania?

Grzesiu podszedł do bibliotekarza i wyszeptał: — Dlaczego kot patrzy na mnie jak kierownik?

— Proszę pamiętać, że pytania wysyłane do wyszukiwarki trafiają do jej usługi. Zakres zapisywanych informacji i personalizacji zależy od dostawcy oraz ustawień.

— A jak włączę tryb incognito w przeglądarce?

— **Tryb prywatny nie oznacza anonimowości w Internecie.** Zazwyczaj ogranicza lokalne zachowywanie historii sesji, ale nie czyni zapytań niewidocznymi dla wyszukiwarki, odwiedzonych witryn czy sieci, z której pan korzysta.

— Czyli nie mogę ukryć się przed światem przyciskiem z kapeluszem?

— Nie w ten sposób.

— A można zmienić domyślną wyszukiwarkę?

— W większości popularnych przeglądarek tak. Warto sprawdzić ustawienia i politykę prywatności wybranej usługi.

— A kot naprawdę mną zarządza?

— Tego wyszukiwarka nie rozstrzygnie. Proszę zapytać kota.

AI i odpowiedzi, które nie zawsze mają rację

Bibliotekarz wystawił nową ladę z tabliczką „Odpowiedź od razu”.

— Teraz nawet nie muszę otwierać stron! — ucieszył się Grzesiu.

— Niektóre wyszukiwarki pokazują skrócone odpowiedzi, fragmenty witryn lub podsumowania generowane przez AI — wyjaśniła Agnieszka.

— To już nie muszę sprawdzać źródeł?

— Właśnie przeciwnie. **Podsumowanie może pominąć kontekst albo zawierać błąd.** Szczególnie ważne informacje sprawdzaj w wiarygodnych źródłach.

— A wyszukiwarka i asystent AI to to samo?

— Nie. Narzędzia mogą się łączyć, ale wyszukiwanie dokumentów i generowanie odpowiedzi to różne zadania. Jedno może korzystać z drugiego.

— Czyli bibliotekarz czasem sam opowiada historię o książkach?

— Tak, ale powinien pan upewnić się, czy rzeczywiście przeczytał właściwe książki.

— Zaczynam podejrzewać, że w tej bibliotece najważniejsza jest umiejętność czytania.

— Piękny wniosek.

Egzamin Grzesia: Opera kontra Google

Agnieszka zamknęła katalog i podała Grzesiowi kartkę.

— Egzamin. Jaka jest różnica między przeglądarką a wyszukiwarką?

— Przeglądarka otwiera strony internetowe. Wyszukiwarka pomaga mi je znaleźć.

— Przykłady?

— Przeglądarki: Opera, Firefox, Chrome, Edge, Safari. Wyszukiwarki: Google Search, Bing, Yahoo Search, DuckDuckGo.

— Czy w przeglądarce Firefox można używać wyszukiwarki Google?

— Tak!

— A w Chrome — Binga?

— Też!

— A jak znam adres strony?

— Mogę go wpisać bezpośrednio, nie pytając wyszukiwarki o drogę.

— Bardzo dobrze.

— Czyli Opera jest samochodem, Google Search drogowskazem, a Internet drogami?

— W naszym uproszczeniu — właśnie tak.

— To idę powiedzieć koledze, że ma nie wyszukiwać Operą.

— Może pan *korzystać z wyszukiwarki w Operze*. I niech pan nie zaczyna kolejnej wojny o słowa.

— Za późno. Już wysłałem mu prezentację na czterdzieści slajdów.

Dyrekcja Biblioteki Całego Internetu informuje: roboty wróciły z objazdu, indeks nadal jest niekompletny, a dział reklam nie jest działem prawdy objawionej. Grzesiu zdał egzamin z różnicy między przeglądarką a wyszukiwarką. Niestety pięć minut później zapytał, jakiego Internetu używa pani Agnieszka: „Chromowego czy Googlowego?”. Bibliotekarz zamknął bibliotekę na przerwę techniczną. 🔎📚🤣❤️