Czytaj lub słuchaj artykuł online

Zeskanuj kod QR, aby otworzyć artykuł na XO.PL

Źródło artykułu:

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

INTERNET

Cache, czyli dlaczego przeglądarka pokazuje wczorajszą gazetę

Grzesiu wrócił do hotelu „Pod Ciasteczkiem” i poprosił recepcjonistę o dzisiejszą gazetę. Dostał wczorajszą. Na pierwszej stronie wciąż widniała informacja o uroczystym otwarciu hotelu, choć od tamtej pory zdążył już odpaść szyld i dwa razy zmienił się kierownik. Recepcjonista twierdził, że wszystko jest w najlepszym porządku. W końcu gazeta była całkiem czysta. 📄🤣

Gazeta z wczoraj jest jak nowa

— Dzień dobry, poproszę aktualny rozkład śniadań — powiedział Grzesiu.

— Już podaję — recepcjonista sięgnął pod ladę.

— Ale tu jest napisane, że śniadanie trwa do dziesiątej! Dzisiaj przesunęli na jedenastą.

— Ten egzemplarz wygląda znakomicie.

— BO JEST WCZORAJSZY!

— Za to nie musiałem biec do drukarni.

Tak właśnie działa **pamięć podręczna przeglądarki**, czyli cache. Przeglądarka może przechować pobrane wcześniej zasoby strony — na przykład obrazy, arkusze CSS, skrypty JavaScript, a czasem także dokumenty HTML — żeby przy następnej wizycie nie pobierać wszystkiego od początku.

— Czyli pan trzyma kopie gazet pod ladą?

— Tak. Szybciej się je wydaje, oszczędzamy transfer i nie trzeba za każdym razem fatygować drukarza.

— A kiedy gazeta się zmienia?

— Wtedy, panie Grzesiu, zaczyna się cała zabawa.

Strona nowa, przycisk stary

Hotelowy informatyk poprawił stronę z rezerwacjami. Zmienił kolor przycisku i naprawił błąd w formularzu. Na serwerze wszystko było gotowe.

— No to działa? — spytał kierownik.

— U mnie działa.

— A u Grzesia?

Grzesiu odświeżył stronę i zobaczył stary przycisk.

— JA TU MAM NADAL POPRZEDNIĄ WERSJĘ!

— Proszę nie krzyczeć, może przeglądarka korzysta z zachowanej kopii pliku CSS albo JavaScript — odpowiedział informatyk.

Strona internetowa to często zestaw wielu osobnych plików. Nawet jeśli dokument HTML jest aktualny, jeden ze starszych zasobów może sprawić, że witryna wygląda lub zachowuje się inaczej niż zamierzał autor. Dlatego znane administratorom zdanie „u mnie działa” nie zawsze kończy dyskusję.

— To mam zepsuty komputer?

— Niekoniecznie. Czasem komputer po prostu zbyt gorliwie pomaga.

— To niech przestanie!

Kto decyduje, jak długo trzymać gazetę?

Informatyk przywiesił na gazetach karteczki.

— Ta ulotka może leżeć godzinę. Ten obrazek dłużej. A tego dokumentu nie przechowujemy.

— Skąd recepcjonista ma to wiedzieć? — spytał kierownik.

— Z instrukcji dostarczonej przez serwer.

W odpowiedziach HTTP serwer może używać nagłówka `Cache-Control`. Przykładowo `max-age=3600` pozwala uznać odpowiedź za świeżą przez określony czas, zwykle 3600 sekund. `no-cache` nie znaczy „nie zapisuj”: pozwala przechować kopię, ale przed ponownym użyciem wymaga jej weryfikacji. Z kolei `no-store` mówi, by nie przechowywać odpowiedzi w pamięci podręcznej.

— Czyli „no-cache” nie oznacza braku cache?

— Zgadza się.

— Kto to nazwał?!

— Proszę nie pytać. Tak działa informatyka.

Ostateczne zachowanie zależy również od przeglądarki, kontekstu żądania i innych dyrektyw. To wskazówki protokołu, nie kalendarz przyklejony do każdej gazety na zawsze.

Proszę sprawdzić, czy to nadal to samo wydanie

Następnego dnia recepcjonista znalazł pod ladą gazetę bez wyraźnej informacji, czy nadal jest aktualna.

— Wyrzucić? — spytał Grzesiu.

— Najpierw zadzwonię do wydawcy.

— Znowu będzie pan pobierał całą gazetę?

— Nie, zapytam, czy się zmieniła.

Serwer może oznaczyć zasób identyfikatorem `ETag` albo datą modyfikacji `Last-Modified`. Przy ponownym sprawdzaniu przeglądarka może wysłać warunkowe żądanie, np. `If-None-Match` lub `If-Modified-Since`. Jeśli kopia pozostaje aktualna, serwer może odpowiedzieć `304 Not Modified`, bez ponownego przesyłania pełnej treści.

— I co powiedział wydawca?

— „Nie zmieniono”.

— To po co dzwoniliśmy?

— Żeby wiedzieć, że nie trzeba jechać po kolejną paczkę.

— Czyli telefon był tańszy od ciężarówki.

— Wreszcie pan rozumie.

Odśwież, odśwież mocniej, a potem nie panikuj

Grzesiu nacisnął przycisk odświeżania.

— Nadal stare.

Nacisnął ponownie.

— NADAL STARE!

— Dwukrotne klikanie nie zawsze robi większe wrażenie na serwerze — westchnął informatyk.

Zwykłe odświeżenie strony często powoduje sprawdzenie aktualności zasobów, ale szczegóły zależą od przeglądarki. Istnieje też **twarde odświeżenie**, które zwykle wymusza pobranie zasobów z pominięciem części zwykłych kopii podręcznych. Skróty klawiaturowe i zachowanie różnią się między przeglądarkami i systemami.

— A przycisk „Wyczyść dane przeglądania”?

— To cięższa artyleria. Można usunąć zapisane pliki z pamięci podręcznej, ale proszę wybrać właściwą kategorię.

— A jak zaznaczę wszystko?

— Można przy okazji usunąć także cookies i wylogować się ze stron.

— TO JA CHCIAŁEM TYLKO NOWY PRZYCISK, NIE NOWE ŻYCIE!

Pamięć podręczna to nie to samo co **ciasteczka**, historia odwiedzin czy zapisane hasła. Przed czyszczeniem warto sprawdzić, co dokładnie zostanie usunięte.

Informatyk też ma swoje sztuczki

Kierownik hotelu zażądał rozwiązania, które nie wymaga tłumaczenia każdemu gościowi, jak mocno naciskać odświeżanie.

— Możemy odpowiednio ustawić zasady cache — wyjaśnił informatyk. — A plikom, które zmieniamy, nadawać nowe adresy albo nazwy z numerem wersji.

— Jak „gazeta-2026-11-08.css”?

— Mniej więcej. Często używa się też skrótu zawartości pliku.

Gdy adres zasobu się zmieni, przeglądarka traktuje go jak nowy plik. Dzięki temu statyczne pliki z wersją w nazwie mogą być długo przechowywane, a mimo to po aktualizacji strona odwołuje się do świeżej wersji. To częsta metoda unikania problemów ze starym CSS lub JavaScriptem.

— A jeśli nie zmienimy adresu?

— Wtedy trzeba zadbać o sensowne reguły weryfikacji i ważności.

— Proszę to zapisać w procedurach.

— Już zapisane.

— Gdzie?

— W dokumencie, którego pan wczoraj nie odświeżył.

Raport: po co nam ten schowek?

Recepcjonista w końcu przyniósł aktualną gazetę.

— No! Dało się? — triumfował Grzesiu.

— Dało. Tylko proszę pamiętać, że gdybyśmy każdemu codziennie drukowali wszystkie te same obrazki, hotelowe łącze miałoby dość jeszcze przed śniadaniem.

Cache przeglądarki przyspiesza otwieranie stron i ogranicza zbędne pobieranie danych. Problem pojawia się wtedy, kiedy stara kopia zostaje użyta tam, gdzie powinna pojawić się nowa. Dobre ustawienia serwera i rozsądne zarządzanie wersjami plików sprawiają, że przeglądarka oszczędza czas bez podawania wczorajszych wiadomości jako dzisiejszych.

— A cookies?

— To numerki do szatni, nie gazety.

— A historia?

— Spis tego, co pan czytał.

— Czyli mam trzy różne rzeczy?

— Tak.

— Informatyka powinna rozdawać więcej szufladek.

Dyrekcja hotelu „Pod Ciasteczkiem” informuje: recepcja wdrożyła nowe zasady przechowywania gazet. Stare egzemplarze będą sprawdzane przed ponownym wydaniem, a nowe otrzymają daty i oznaczenia. Grzesiu z dumą otworzył aktualną stronę z rozkładem śniadań. Następnie zapisał ją na pulpicie jako zrzut ekranu i przez trzy kolejne dni narzekał, że nic się nie aktualizuje. 🗞️💻🤣❤️