Engineer

hreflang

атрибут, що пов'язує сторінку з її перекладами

Що це таке

Атрибут `hreflang` на посиланні повідомляє, якою мовою написано сторінку на іншому його кінці. HTML дозволяє ставити його на елементи `a`, `area` і `link`, але відомим його зробило одне застосування: набір елементів `link rel="alternate"` у `head` сторінки, по одному на кожну мовну версію, яким сайт повідомляє пошуковим системам, що кілька адрес містять той самий вміст різними мовами.

```
<link rel="alternate" hreflang="en" href="https://example.com/guide/" />
<link rel="alternate" hreflang="da" href="https://example.com/da/guide/" />
<link rel="alternate" hreflang="x-default" href="https://example.com/guide/" />
```

Стандарт HTML називає цей атрибут «суто рекомендаційним». Його значення мусить бути коректним мовним тегом за BCP 47, а браузер має визначати мову завантаженої сторінки з самої сторінки, а не з посилання, яке до неї привело. У поєднанні з `rel="alternate"` атрибут `hreflang`, що відрізняється від мови самого документа, позначає ціль як переклад.

Пошукові системи перетворили цю рекомендаційну примітку на протокол. Google описує три рівноцінні способи передати набір: елементи `link` у `head`, заголовок HTTP `Link` для PDF та інших файлів без `head` і записи в XML-карті сайту. Кожна версія перелічує саму себе й усі інші версії, з повними абсолютними адресами. Пара враховується лише тоді, коли обидві сторінки називають одна одну, тож чужий сайт не може оголосити себе перекладом ваших сторінок. Коди складаються з мови за ISO 639-1 і, за потреби, регіону за ISO 3166-1, а сам регіон без мови недійсний. Зарезервоване значення `x-default` вказує сторінку для читачів, чия мова не збігається з жодною з інших, і, як зазначає Google, задумане воно для сторінок вибору мови.

Із ним легко сплутати двох сусідів. Атрибут `lang` каже, якою мовою текст перед читачем; `hreflang` каже, якою мовою сторінка, на яку веде посилання. Google зазначає, що не використовує жодного з них, щоб визначити мову сторінки, тож `hreflang` не зробить сторінку данською; він повідомляє пошуковій системі, що данська сторінка, яку та вже розпізнала, належить до однієї групи з англійською. А канонічна URL зливає дублікати в одну проіндексовану адресу, тоді як `hreflang` позначає мовні альтернативи й допомагає пошуковій системі обрати, яку версію показати.

Застосуйте поняття до рішення

Застосуйте «Що таке hreflang?» до конкретного рішення. Ми можемо пов'язати поняття із системою, яку ви створюєте або купуєте. Програма, що природно читається на кожному ринку, де ви продаєте, — з правильними датами, валютами й сортуванням імен.

Переглянути наші послуги Зв'язатися з нами →

Перевірте припущення за терміном

Використайте пояснення та пов'язані роботи, щоб сформулювати запитання для проєкту. Ми можемо перевірити припущення разом до вибору інструмента. Створила англо-данську систему i18n — 12,027 повідомлень на локаль — англійською та данською

Створила англо-данську систему i18n Зв'язатися з нами →

Чому це важливо

`hreflang` фіксує одне рішення: чи утворюють мовні версії сторінки одну групу, чи їх оцінюють як окремі сторінки. Окремо кожна версія змагається сама за себе, і пошукова система може показати данському читачеві англійську сторінку або вивести обидві версії в тих самих результатах. Оголошені групою, версії сприймаються як один вміст, і пошукова система може запропонувати ту, що відповідає мовним налаштуванням читача.

Група має сенс лише тоді, коли версії справді містять той самий вміст. Сайтові однією мовою `hreflang` не потрібен узагалі, а два мовні розділи, що різняться по суті, теж не група. Регіональні варіанти з невеликими відмінностями, як-от сторінки `en-GB` і `en-US` з різними цінами, утворюють групу так само.

Ціна — облік, і він зростає з квадратом кількості мов. Кожна версія кожної сторінки несе повний список, тож коли один переклад додають, переносять чи знімають, змінюється `head` усіх сторінок-сестер, а список, правильний на більшості сторінок і хибний на кількох, ламається тихо: пари, що ведуть в обидва боки, використовуються, решта відкидається, і жодної видимої помилки немає. Тому `hreflang` погано пасує ручному редагуванню

й природно — генераторові, який виводить набір кожної сторінки з одного списку. Від `hreflang` залежить, яку з кількох версій побачить користувач пошуку.

Як це виглядає на практиці

Цей сайт збирає свої мовні версії з одного дерева вмісту, тож кожна сторінка має набір відповідних перекладів, які треба оголосити.

Один шаблон вирішує, які переклади оголошувати, і його рішення використовують у кількох місцях: у наборі `hreflang` у `head`, у тегах `og:locale:alternate`, якими користуються попередні перегляди посилань, у списку альтернатив у карті сайту, у `inLanguage` у структурованих даних сайту і, після ще одного віднімання, у мовній смузі, яку бачить читач. Список містить і саму сторінку, а отже дає те посилання на себе, якого вимагають правила. Коди називають лише мови, як-от `en`, `da` і `uk`, тож кожну версію пропонують усім читачам цієї мови, де б вони не були, а англійський запис пишеться вдруге як `x-default`. І `head`, і карта сайту несуть набір, згенерований з того самого списку, тож розійтися вони не можуть.

Поділ між машинами й читачами навмисний. На попередньому етапі реалізації українську версію оголошували пошуковим роботам і вносили до карт сайту, але мовний перемикач у верхній панелі її не пропонував. Цю відмінність контролював один прапорець конфігурації: пошукові роботи отримували взаємні посилання на українську версію, а змінювався лише мовний перемикач для читача. Другий прапорець, для мови, яку вже збирають, але яка ще не готова, прибирає її з обох списків і позначає її сторінки `noindex`. Сторінка такою мовою теж не оголошує сестер, бо `hreflang` від неоголошеної сторінки до оголошеної — це пара, яку пошукова система бачить лише з одного кінця.

Копії підкоряються тій самій логіці. Сторінка на одному з тематичних субдоменів, чия канонічна URL вказує на основний сайт, узагалі не виводить `hreflang`: набір читають на тій сторінці, яку пошукова система прийняла як канонічну, а копія, що називала б переклади свого хоста, суперечила б власній канонічній URL. Перемикач мов на ній усе одно працює, бо читач, якому потрібна данська, однаково має її отримати.

Перевірка зібраних сторінок вимагає, щоб кожна сторінка з альтернативами містила `x-default` і власну мову, а кожна альтернатива була зібраною сторінкою оголошеною мовою й вела назад. Правило про посилання на себе потрібне не для галочки. Генератор має два списки з майже однаковими назвами: `.Translations`, що не містить поточної сторінки, і `.AllTranslations`, що містить, і помилка в одне слово прибирає власну мову кожної сторінки і з `head`, і з мовної смуги водночас, тож вони й далі узгоджені між собою. На навмисно зламаній збірці ця одна вимога виявилася єдиною в перевірці, що зловила 528 із 792 сторінок.

Де виникають проблеми

Користувачам пошуку показують версію не тією мовою. Данського читача ведуть на англійську сторінку, або обидві версії стоять поруч. Пари ведуть в один бік: данська сторінка називає англійську, але англійська досі перелічує лише ті мови, що були раніше, а пошукові системи ігнорують пару без зворотного посилання. Будуйте набір кожної сторінки з одного списку, що містить і саму сторінку, і перевіряйте взаємність на зібраних сторінках.

Анотації цілої мови нічого не дають. Розмітка на місці й є коректним HTML, але версії трактують як не пов'язані між собою. Коди означають щось інше, ніж задумано: `ua` чи `dk` там, де мови мають коди `uk` і `da`, `be` в значенні Бельгії, хоча це код білоруської мови, або `en-UK` замість `en-GB`, бо `UK` — зарезервований код регіону, який Google ігнорує. Беріть мову з ISO 639-1, додавайте регіон з ISO 3166-1 лише там, де вміст справді різниться між країнами, і пам'ятайте, що коректний тег BCP 47 Google може й не підтримувати.

Переклади випадають з індексу. У пошуку з'являється лише мова оригіналу, а переклади трактують як його дублікати. Кожен переклад називає своєю канонічною URL оригінал, а це просить злити переклад з оригіналом, тоді як його `hreflang` просить його показувати. Робіть кожну мовну версію власною канонічною URL; Google радить, щоб канонічна сторінка була тією самою мовою.

Сторінку, що просила лишитися поза пошуком, однаково оголошують. Чорнеткова мова з позначкою `noindex` з'являється серед альтернатив опублікованих сторінок або сама перелічує їх у своєму `head`. Оголошення й директиву задано в різних місцях, тож сайт дає пошуковому роботові дві

відповіді про ту саму адресу. Вирішуйте для кожної мови окремо, оголошена вона в обидва боки чи в жоден, тримайте це рішення в одному параметрі, який читають усі списки, і звіряйте з ним зібрані сторінки.

Що почитати далі

- HTML Standard: the hreflang attribute — що атрибут означає на посиланні і чому браузер не має йому довіряти.
- HTML Standard: link type “alternate” — як `hreflang` і `type` описують переклад або інший формат того самого документа.
- RFC 5646: Tags for Identifying Languages — синтаксис BCP 47, якому має відповідати значення атрибута.
- Google Search Central: localized versions of your pages — три способи передати набір, правило взаємності та `x-default` .

Словник

Обирайте за фактичними умовами

Опишіть робочий процес, наявні інтерфейси та користувачів, яким потрібна зміна. Ми можемо визначити перевірюваний результат і спосіб передачі, який дасть вашій команді змогу підтримувати його.

Дивитися, скільки це коштує Зв'язатися з нами →

Зробіть термінологію практичною

Запитайте, як «Що таке hreflang?» стосується вашої організації. Конкретного прикладу з поточної системи достатньо для початку розмови.

Переглянути наші послуги

welcome@engineer.company +45 50 10 61 84 Копенгаген, Данія

Візьміть цю сторінку із собою

Візитівка

Дбайливо зроблено з любов'ю — D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Авторське право © 2025 – дотепер · Engineer ApS · Усі права захищено.

Надруковано з https://engineer.company/uk/glossary/hreflang/