Engineer

RSS

старіший із двох XML-форматів, якими пишуть стрічку

Що це таке

RSS — це XML-формат для публікації вебстрічки: один елемент `channel`, що описує сайт і містить список елементів `item`, кожен із яких — запис із посиланням, датою й текстом, за адресою, яку програми для читання стрічок завантажують за розкладом. Специфікація розшифровує назву як Really Simple Syndication, багато хто каже «RSS-стрічка» про стрічку в будь-якому форматі, а RSS 1.0, «RDF Site Summary», — це окремий формат на основі RDF, а не ранніша версія того, яким користуються сьогодні.

Формат почався в Netscape, яка опублікувала RSS 0.90 у березні 1999 року і 0.91 у липні того самого року. UserLand Software випустила власну 0.91 у червні 2000 року, потім 0.92, а в серпні 2002 року — RSS 2.0. Berkman Klein Center при Гарвардській школі права опублікував версію 2.0.1, яку написав Дейв Вайнер, у липні 2003 року, і відтоді текст веде RSS Advisory Board за ліцензією Creative Commons. Дорожня карта цієї ради каже, що формат для всіх практичних цілей заморожено на версії 2.0.1. Рада відтоді публікувала переглянуті редакції документа, чинну — у березні 2009 року, але нові потреби закривають розширення у власних просторах імен XML, а не нові версії RSS.

Документ невеликий. Його корінь — елемент `rss`, атрибут `version` якого має дорівнювати `2.0`, а всередині — єдиний `channel`. Канал повинен мати `title`, `link` і `description`. Усі елементи запису необов’язкові, якщо запис має принаймні заголовок або опис. Усе інше, від `pubDate` і `guid` до `image` і `ttl` каналу, з’являється тоді, коли цього хоче видавець:

```
<?xml version="1.0" encoding="UTF-8"?>
<rss version="2.0">
  <channel>
    <title>Нотатки для прикладу</title>
```

```
    <link>https://example.com/</link>
    <description>Короткі нотатки про роботу серверів.</description>
    <item>
      <title>Переїзд бази даних</title>
      <link>https://example.com/notes/db-move/</link>
      <guid isPermaLink="true">https://example.com/notes/db-move/</guid>
      <pubDate>Sat, 17 Oct 2026 09:00:00 GMT</pubDate>
      <description>Чому переїзд зайняв один вечір.</description>
    </item>
  </channel>
</rss>
```

Застосуйте поняття до рішення

Застосуйте «Що таке RSS?» до конкретного рішення. Ми можемо пов'язати поняття із системою, яку ви створюєте або купуєте. Швидкий сайт, який легко знайти в пошуку і який ваші працівники редагують самі, — без рахунку за кожну зміну в тексті.

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. Розробила 100 вебзастосунків на HTML/HTML5 — 100 вебзастосунків побудовано

Розробила 100 вебзастосунків на HTML/HTML5 Зв'язатися з нами →

Чому це важливо

RSS нестрогий, і саме цю нестрогість видавцеві треба розуміти. Елемент `description` запису може містити звичайний текст або екранований HTML, і ніщо у файлі не каже, що саме, тож вирішує програма для читання. Елемент `guid` необов’язковий, тож без нього програма сама з’ясовує, чи запис новий. Дати пишуть за RFC 822, стандартом електронної пошти 1982 року, рядками на зразок `Sat, 17 Oct 2026 09:00:00 GMT` з англійськими назвами днів і місяців, а специфікація досі дозволяє двозначний рік. До того ж формат не має способу оголосити базову URL-адресу, тож відносному посиланню всередині запису немає надійної основи, щоб його розв’язати.

Відповідь ради — RSS Best Practices Profile, опублікований у жовтні 2007 року: рекомендації для документів, які працюють з усім широким колом програм, що їх читають. Профіль просить `guid` для кожного запису й абсолютні URL-адреси в посиланнях і описах. Він повідомляє, що єдині форми дат, з якими впоралися всі перевірені агрегатори, були в універсальному часі й закінчувалися на `GMT`, `+0000` або `-0000`. Він також описує `atom:link` з `rel="self"`, що вказує власну адресу стрічки. Стрічка, написана за профілем, залишає програмам набагато менше здогадок.

Два розширення несуть більшість того, що додає сучасна стрічка. Елемент `atom:link` запозичено з Atom. Елемент `content:encoded`, з модуля, який спершу написали для RSS 1.0, містить повний HTML запису. Профіль описує використання обох разом, повного тексту в `content:encoded` і коротшого підсумку в `description`, а його автори з'ясували, що коли запис мав обидва, шість із восьми перевірених агрегаторів показували `content:encoded` і пропускали підсумок. Елемент `enclosure` прикріплює до запису медіафайл трьома обов'язковими атрибутами: його `url`, `length` у байтах і MIME-тип у `type`.

Заморожений формат — це й перевага. Формат, що не змінюється з 2003 року, — стала ціль, і стрічку, написану за ним сьогодні, читають програми, написані за тим самим текстом багато років тому. Atom закриває більшість цих прогалин самою специфікацією, тож аргумент за те, щоб публікувати ще й RSS, стосується програм для читання, а не строгості: деякі беруть лише RSS, а RSS-адресу, роздану роки тому, досі опитують.

Як це виглядає на практиці

Цей сайт публікує RSS 2.0 за адресою `/uk/index.xml` поруч зі стрічкою Atom і JSON Feed. Сторінка подяк пояснює, чому XML-стрічок дві: усі три містять повний текст, а XML-стрічок дві, а не одна, бо деякі програми для читання стрічок підтримують лише один із двох форматів, і неможливо передбачити, який саме. Та сама сторінка веде RSS-стрічку до валідатора стрічок W3C поруч зі стрічкою Atom, щоб кожен міг сам запустити перевірку, не вірячи сайту на слово.

Сторінка про читання сайту з термінала розглядає стрічки як дані. Вона каже, що портфоліо публікується у трьох форматах, кожен із повним текстом кожного запису, а не анонсом, і завантажує RSS-стрічку самим лише `curl`:

```
curl -s https://engineer.company/uk/index.xml
```

Де виникають проблеми

Зображення каналу зникає, а валідатор заперечує. Програма для читання показує назву стрічки без зображення, а сервіс W3C попереджає про формат зображення. Елемент `image` в RSS — найсуворіший елемент для зображень серед трьох форматів стрічок: лише GIF, JPEG або PNG, щонайбільше 144

пікселі завширшки й 400 заввишки. Картка для соцмереж у WebP для нього не годиться. RSS-шаблон цього сайту натомість вказує в `image` свою PNG-іконку на 144 пікселі, а коментар поруч фіксує, що картка WebP викликала попередження `ImageUrlFormat`, яке помітили, коли повторно перевірили стрічку після зміни.

Записи стоять у неправильному порядку, і ніхто про це не повідомляє. Програма показує нові записи під старими або з неправильною датою. Дати в RSS — це рядки, тож дата в ISO 8601, шаблон дати з одруківкою чи назва дня іншою мовою однаково дають правильно сформований XML, і помічає це лише розбирач дат у програмі. У Hugo метод шаблону `Format` для дати пише англійські назви, а функція `time.Format` їх локалізує, і `pubDate` з українськими назвами днів і місяців уже не є датою RFC 822. Цей сайт форматує кожну `pubDate` методом за шаблоном дати, що лежить в одному частковому шаблоні разом із шаблоном дати для Atom, а коментар там каже, що одруківка в будь-якому з них — не помилка збирання, а стрічка, яка розбирається й сортується неправильно в чужій програмі.

Іконки в записі нічого не малюють, а валідатор перелічує помилки. Повний текст у `content:encoded` залишає свою сторінку, тож усе, що від неї залежить, зникає. SVG-іконка, яка посилається на спрайт, визначений деінде на сторінці, у програмі для читання лишається порожньою, а сервіс W3C позначає сам елемент `svg`. Тому стрічки цього сайту отримують заголовки без фірмової шестерні-роздільника; коментар у тому шаблоні фіксує, що валідатор видав дванадцять повідомлень `Non-html tag: svg` і шість скарг на атрибут `role`.

Посилання всередині запису нікуди не ведуть. Підписник переходить за посиланням у записі й потрапляє на сторінку, якої немає, або взагалі нікуди. Посилання відносно кореня на зразок `/uk/notes/` працює на вебсторінці, бо браузер знає адресу сторінки, але документ RSS не може оголосити базову URL-адресу. Профіль каже, що опис не повинен містити відносних URL-адрес і що агрегатор може спробувати розв’язати таку адресу відносно `link` каналу; програма, яка цього не робить, залишає посилання зламаним. Пишіть кожне посилання в HTML стрічки як абсолютну URL-адресу.

Дві категорії з однаковим словом виглядають як одна. Запис, позначений у двох словниках, показує термін двічі, або програма зливає два терміни, що означають різне. Елемент `category` в RSS має необов’язковий атрибут

`domain`, який називає словник, звідки походить термін. RSS-шаблон цього сайту пише його на кожній категорії, бо, як каже коментар, без нього терміни, спільні для двох словників, читалися б як дублікати.

Що почитати далі

- RSS 2.0 Specification — текст формату від RSS Advisory Board, елемент за елементом, з дорожньою картою, яка оголошує його замороженим.
- RSS Best Practices Profile — рекомендації ради й те, що вона з'ясувала, перевіряючи справжні агрегатори.
- RSS History — кожна версія від 0.90 Netscape до чинного тексту, з датами й видавцями.
- RDF Site Summary (RSS) 1.0 — окремий формат на основі RDF, що має ту саму назву.
- RFC 822 — стандарт електронної пошти 1982 року, чий синтаксис дати й часу RSS використовує досі.



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 RSS?» 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/rss/