

# **Розробка веб-сайту агрегатора новин з елементами веб-скрапінгу засобами Ruby on Rails**

**Виконавець:**

**Манастирецький Любомир Васильович**

**Керівник:**

**к.т.н., доц. Ващишак Сергій Петрович**

# Мета і завдання дослідження

Метою даного дослідження є розробка веб-сайту агрегатора новин з використанням веб-скрапінгу на основі Ruby on Rails.

Для досягнення поставленої мети вирішуються наступні завдання:

- Провести аналіз сучасних методів та інструментів веб-скрапінгу;
- Розробити архітектуру веб-сайту агрегатора новин;
- Реалізувати сервіс веб-скрапінгу для автоматичного збору новин з різних веб-ресурсів;
- Інтегрувати сервіс веб-скрапінгу в архітектуру веб-сайту;
- Провести тестування розробленого веб-сайту.

# Огляд та аналіз веб-сайтів агрегаторів новин Ruby on Rails

У мережі налічується безліч веб-сайтів, присвячених Ruby та Ruby on Rails, кожен із яких має свою унікальну привабливість: вони пропонують навчальні матеріали, огляди бібліотек і фреймворків. Проте, багато з них страждають від кількох суттєвих недоліків - застарілий дизайн, мала кількість новин, рідкісні оновлення новин, а також незручність перегляду архіву інформації, що ускладнює доступ до необхідної інформації для розробників.

## Ruby Weekly

ARCHIVES | LATEST | RSS

Want to subscribe? Enter your address here

Subscribe now »

Easy to unsubscribe at any time. Your e-mail address is safe — here's our privacy policy.

« Prev

Next »

#643 — MARCH 2, 2023

[READ ON THE WEB](#)

TOGETHER WITH  pganalyze

### Ruby Weekly

```
eval($s="t='eval($s='+d=34.chr;s=3
2.chr+$s*i=8;v=$VERSION|eval('begin;v=V
ERSION;rescue;v||RUBY_VERSION;end');f='?'*8
+'A|'+?'*20+'G7c'+?'*15+'ABCXx@-@_'OpGxCxp@-p0
xS10-G7c7q?xCPAP1q?x_x_c_x_c_x0BHCg7G7qA1_l_`GcP0xC1H
NFccq@`_lOF@`?q?x_@x_x_`GB`0`0-G7C@qCxCXP@EIG-C7p0IC7
p0IC7APIA-HNN`ccxCI@L@B`+"GpGpc@p?x_`GB`??_?F010B@
xCIP`?7c7q?HPx@-@_`G@`???7@LA`?q?x?x@i_l_0-GC`
xA-@_@GBD`').unpack('c*');w=4+v.Length*u=
15;r=10.chr;j=0;while-24+w*us-i=1+i
;x=1%w;x>0||t=t+d+'+'+r+d;k=
i/w%12>2&&x%u-3&&x%u+i
/w*11-34+'-'...'.
index(c=v[
x/u,1]
)!!c.hex +3)*991
l0; k=f [k/6 ]][k%
6]; t=t +s[
k*j =k+ j,1
]end;pr int (t+
d+` ).s pli
t.j oin [0,
609 ])# Y.E. '+r)
").split .join)#
```



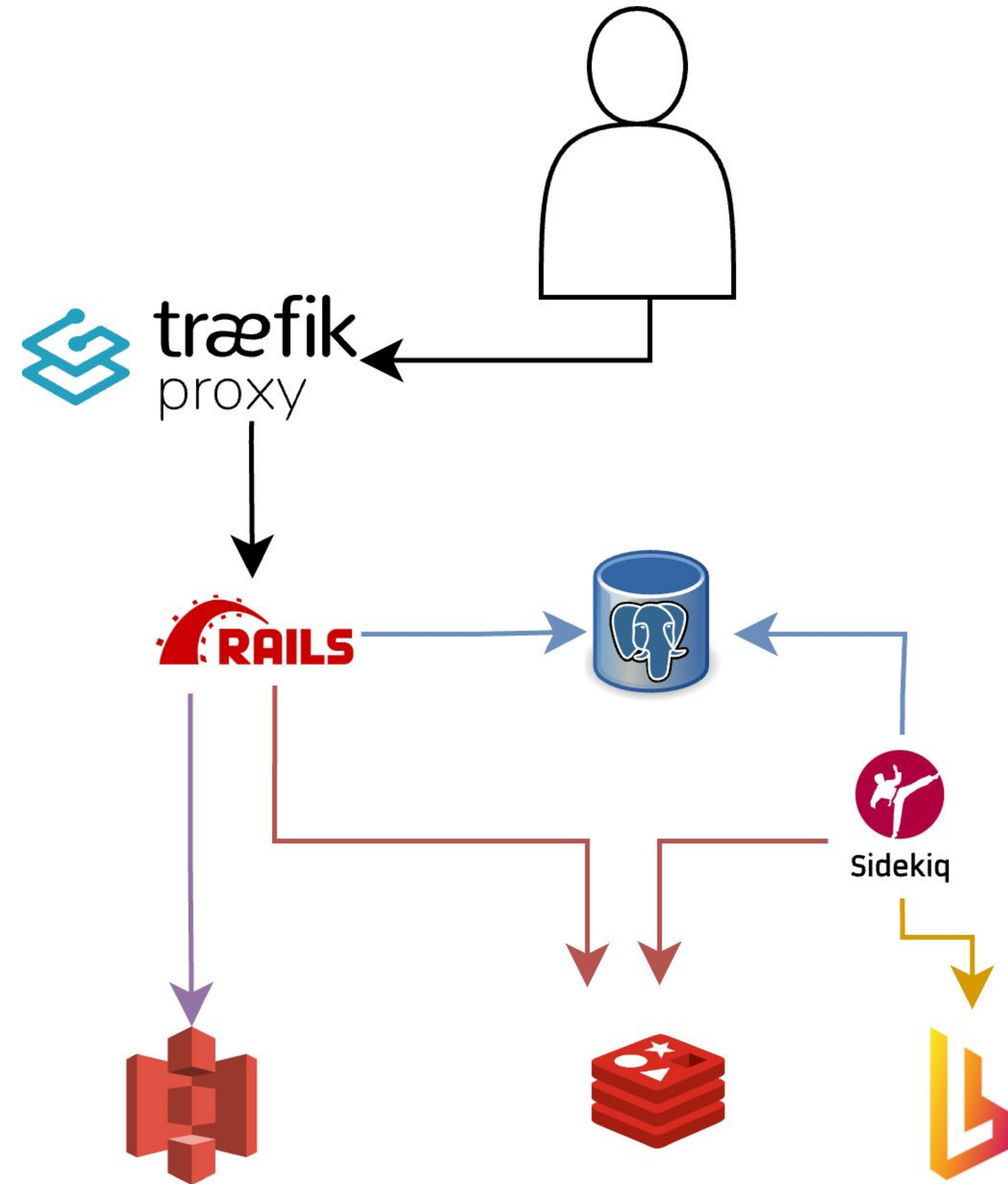
[Ruby Turns 30 Years Old; Celebrate with a Quine!](#) — Thirty years ago [in a chat room far, far away](#), the name 'Ruby' came into being for Matz's then-nascent language. This milestone has been celebrated by many including [Matt Sears](#), [DHH](#), [Charles Nutter](#), in a (Japanese) [audio message by Matz himself](#), and [many others](#), but I had to feature this amazing effort to create a quine, of sorts, that works on "all currently known CRubys" (right back to Ruby 0.49!). It's worth hitting "Translate to English" on this page to learn a few things about Ruby's history.

KUJ-MA-ME



# Структура веб-сайту rails.news

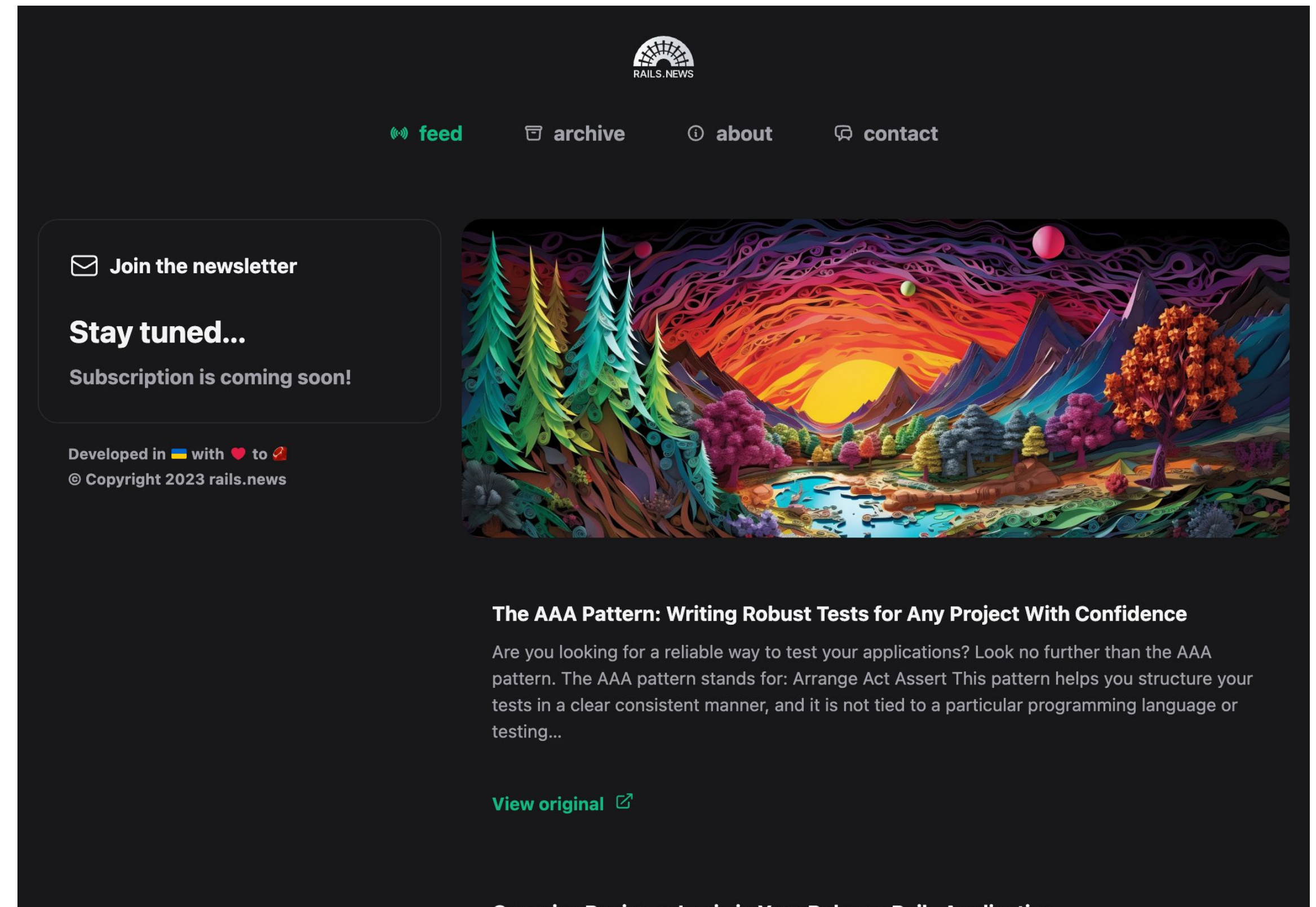
- Веб-сайт (моноліт Ruby on Rails) - основний інтерфейс для користувачів
- БД PostgreSQL- використовується для зберігання даних користувачів та новин
- БД Redis - використовується для зберігання інформації, необхідної для роботи Sidekiq, а також для кешування даних
- Сервіс Browserless - використовується для реалізації скрапінгу веб-сторінок
- Воркер Sidekiq - використовується для розпаралелювання та виконання важких задач у фоновому режимі
- Сховище S3: використовується для зберігання та постачання статичного контенту, такого як зображення, відео, PDF-файли та інше
- Reverse-proxy Traefik - високопродуктивний та гнучкий reverse-proxy сервер



# Результати виконання

Головна сторінка веб-сайту "rails.news" структурно складається з головного шаблону та містить дві основні секції:

- Секція "newsletter", де користувачі мають можливість підписатися на отримання новин від веб-сайту. Проте, наразі, функція підписки ще не доступна;
- Секція новин, в якій публікуються статті та новини, що стосуються Ruby on Rails. Кожна стаття включає заголовок, короткий опис та посилання на оригінальне джерело.

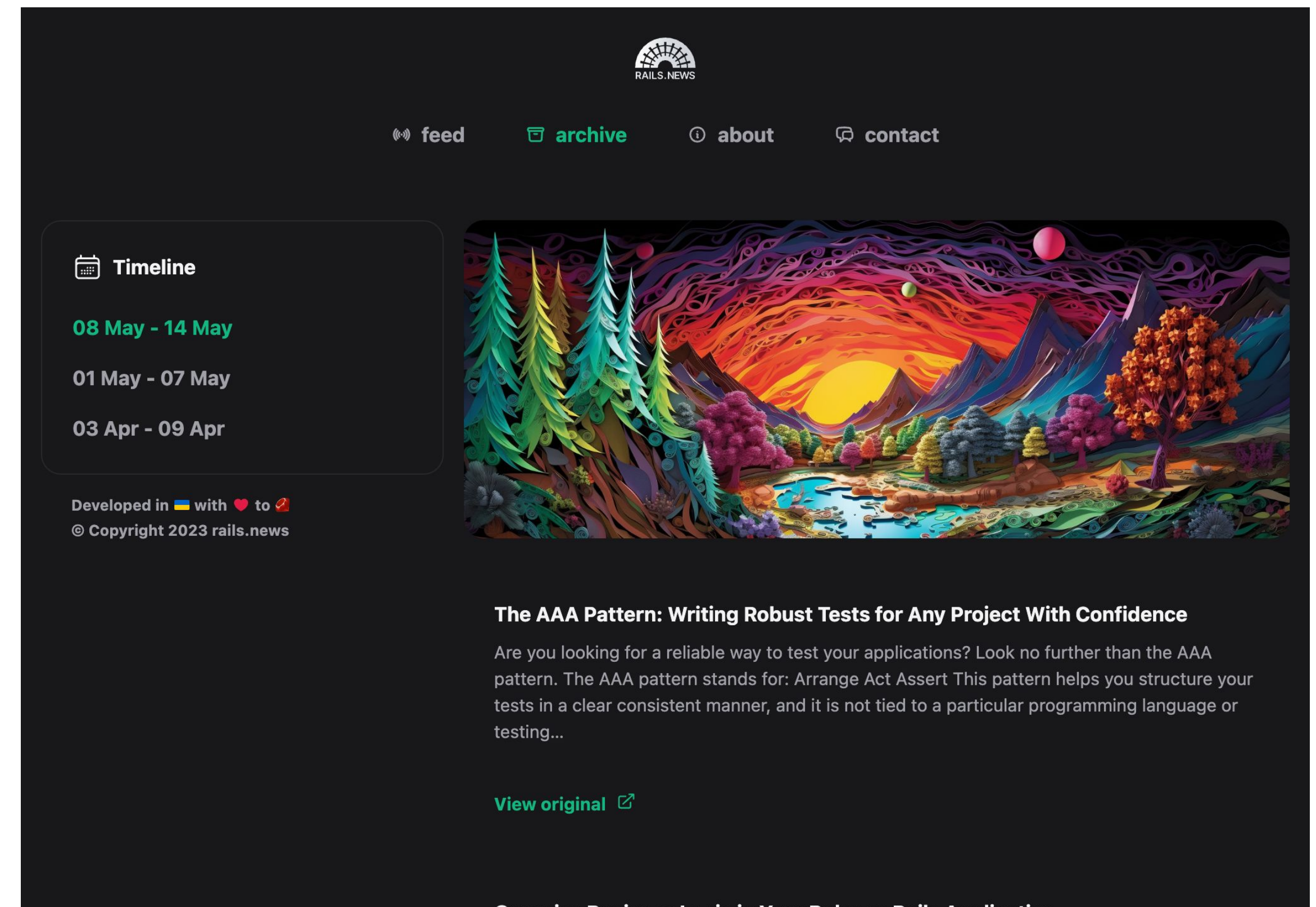




# Результати виконання

Сторінка “Архів” веб-сайту "rails.news" структурно складається з головного шаблону та містить дві основні секції:

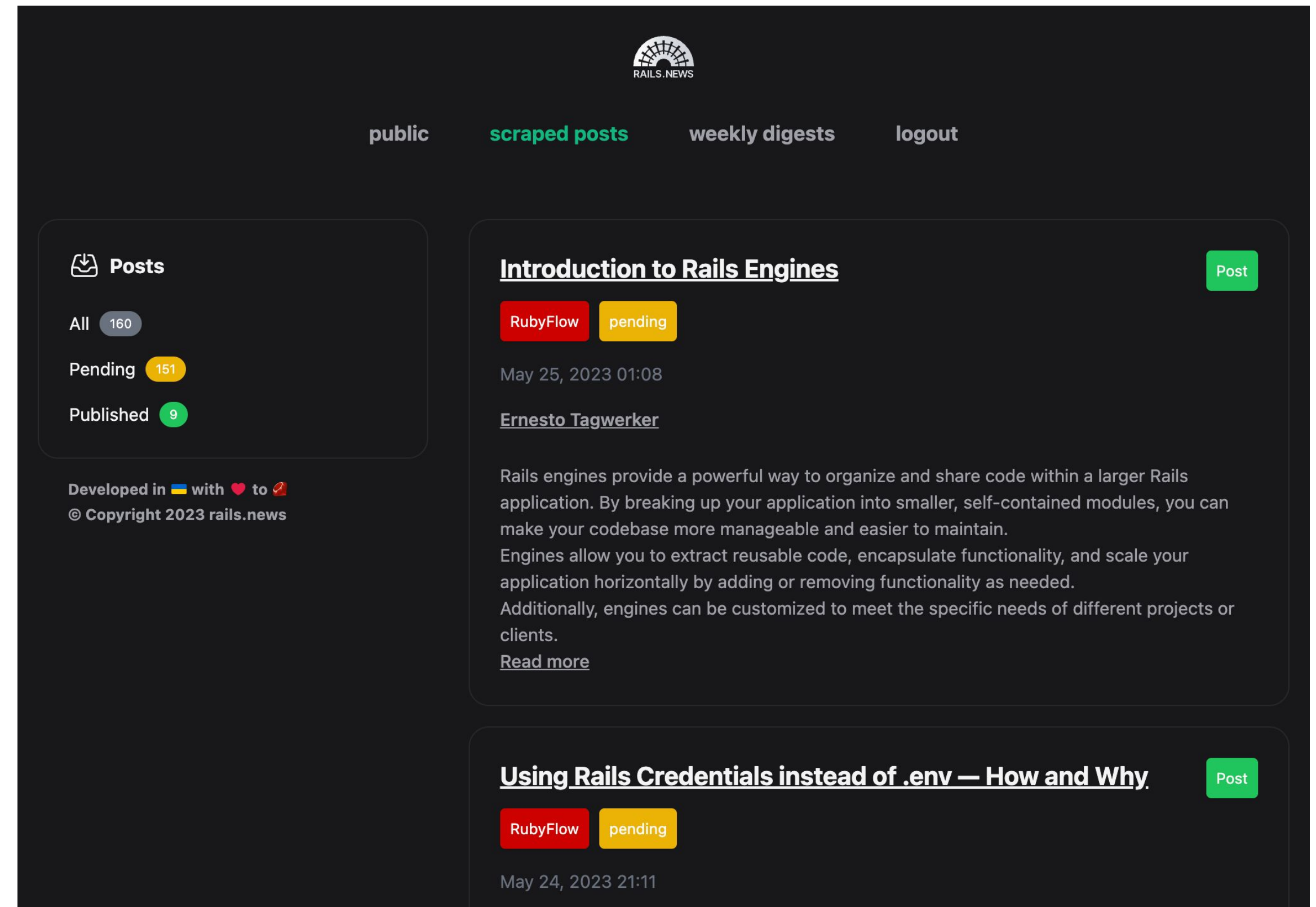
- Секція "Timeline" яка відображає дати публікації новин та статей, представлених у форматі "день місяця - день місяця”;
- Секція статей, яка містить список статей та новин, що були опубліковані за вказаних дат. Кожна стаття включає заголовок, короткий опис та посилання на оригінальне джерело.



# Результати виконання

Сторінка зібраних постів в адмін панелі веб-сайту "rails.news" структурно складається з шаблону специфічного для адмін панелі, та містить дві основні секції:

- Секція "Posts" відображає список можливих статусів публікації та відіграє роль фільтрів;
- Секція статей містить список статей та новин, що були зібрані за допомогою Sidekiq. Кожна стаття включає заголовок, короткий опис, посилання на оригінальне джерело, дату коли оригінал було опубліковано, та посилання на форму для публікації новини на сайті.







- feed
- archive
- about
- contact

## Log in

Email

Password

Remember me

Log in



- public
- scraped posts
- weekly digests
- logout

Title \*

Introduction to Rails Engines

External url

https://rubyflow.com/p/bo65ur-introduction-to-rails-engines

```

---
id: 49281
uid: bo65ur
url: https://rubyflow.com/p/bo65ur-introduction-to-rails-engines
user:
  url: https://github.com/etagwerker
  name: Ernesto Tagwerker
  title: Introduction to Rails Engines
content: |-
  <p>Rails engines provide a powerful way to organize and share code within a larger Rails application. By breaking u
  <p>Engines allow you to extract reusable code, encapsulate functionality, and scale your application horizontally b
  <p>Additionally, engines can be customized to meet the specific needs of different projects or clients.</p>
  <p><a href="https://www.fastruby.io/blog/introduction-to-rails-engines.html?utm_source=rubyflow&utm_medium=post
timestamp: '2023-05-25T01:08:15.000Z'
  
```

- https://www.fastruby.io/blog/introduction-to-rails-engines.html?utm\_source=rubyflow&utm\_medium=post&utm\_car

```

--- !ruby/hash:ActiveSupport::HashWithIndifferentAccess
author: Kudakwashe Siziva
date: '2023-05-23T18:58:06.000Z'
description: An intro to Rails Engines are an architectural pattern that can be used
to modularize or pack a Rails application. An article about this powerful feature
in Rails.
image: https://fastruby.io/blog/assets/images/header/introduction-to-rails-engines.jpg
logo:
publisher: Introduction to Rails Engines
title: Introduction to Rails Engines - FastRuby.io | Rails Upgrade Service
url: https://fastruby.io/blog/introduction-to-rails-engines.html
  
```

Content



Rails engines provide a powerful way to organize and share code within a larger Rails application. By breaking up your application into smaller, self-contained modules, you can make your codebase more manageable and easier to maintain.

Engines allow you to extract reusable code, encapsulate functionality, and scale your application horizontally by adding or removing functionality as needed.

Additionally, engines can be customized to meet the specific needs of different projects or clients.

[Read more](#)

Create Post



# Висновки

У рамках дослідження було детально вивчено предметну область проекту, яка охоплює аналіз новин та інформації зі світу Ruby і Ruby on Rails. Дослідження існуючих аналогів показало їхні недоліки, такі як обмеженість архіву та застарілий дизайн сайтів. Особистим внеском була розробка веб-сайту агрегатора новин з використанням Ruby on Rails, що включає в себе елементи веб-скрапінгу для збору новин з різних джерел.

При виборі інструментів для розробки сайту було враховано їхню ефективність, надійність та доступність, що дозволило створити стабільний та продуктивний веб-ресурс. Використання MRSK для розгортання додатку в продакшені дозволило забезпечити його гнучкість, автономність та ефективність.

Щодо адміністративного функціоналу, було створено панель адміністратора для керування ресурсами сайту та захищено адміністративну частину сайту від несанкціонованого доступу.

В результаті, було створено повноцінний веб-сайт, що виконує всі поставлені завдання: збирає, організовує та відображає новини для спільноти Ruby і Ruby on Rails.