

Rogue Scholar has an API

Martin Fenner 

Published October 9, 2023

Citation

Fenner, M. (2023). Rogue Scholar has an API. *Front Matter*. <https://doi.org/10.53731/ar11b-5ea39>

Keywords

Feature, Rogue Scholar



Copyright

Copyright © Martin Fenner 2023. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Front Matter

The Rogue Scholar science blog archive has launched a dedicated API today, publicly available at <https://api.rogue-scholar.org> and complementing the [website](#).

Rogue Scholar had an API before but with two important limitations.

- **Serverless.** The API at <https://rogue-scholar.org/api> uses [serverless](#) technology, which isn't a good fit for long-running resource-intensive processes.
- **GitHub Actions.** GitHub Actions are used for DOI registrations. They can be triggered at specific times, but [not more than once every 5 min](#).

The new API overcomes these limitations once it is fully implemented by the end of the year. The version released today implements HTTP **get** requests, supports anonymous users, and provides the same information that is also available via the Rogue Scholar [website](#). The API is implemented as a Python [Quart application](#) (an async Python web micro-framework heavily inspired by Flask), hosted on the [fly.io](#) platform, and available as Open Source software via [GitHub](#), [PyPi](#), and [Zenodo](#). More work is needed to allow users to run the API locally, as the API requires data from the database (Postgres) and search index (Typesense), which both also use Open Source software but need authentication for access. The simplest way to get started with the Rogue Scholar API is to use the [OpenAPI endpoint](#) with the [Swagger UI](#):

Rogue Scholar API 0.6.2 OAS 3.1
/openapi.json

default ^

GET /blogs Get all blogs. v

GET /blogs/{slug} Get blog by slug. ^

Parameters Try it out

Name	Description
slug <small>required</small>	
string (path)	slug

Responses

Code	Description	Links
200	Media type application/json	No links

[Rogue Scholar API](#)
[Swagger UI](#)

In the coming weeks, I will work on improving the Rogue Scholar API in the following important areas:

Integration of DOI registration

This is currently done via GitHub Actions and a [Ruby gem](#) automatically converts the blog post metadata into Crossref XML needed for DOI registration. This works fine but doesn't easily scale to 100s or more DOI registrations or updates per day, and is more difficult to integrate with other workflows compared to an internal API. The goal is to switch to the DOI registrations via a background task triggered by the API and using the Python [metadata conversion library](#) that I wrote at the beginning of the year.

Conversion of blog posts to ePub or PDF

Converting the science blog posts archived in Rogue Scholar into ePub, PDF, or other formats using the [Pandoc](#) universal document converter would enable several interesting use cases, for example storing blog posts locally with a reference manager or generating collections by blog, author, or topic.

Metadata conversion

The API released today continues the Rogue Scholar integration with DOI content negotiation to convert blog post metadata into different formats such as BibTeX or formatted citations. We could offer [additional metadata conversions](#) not currently implemented by DOI content negotiation such as [Schema.org](#) JSON-LD.

Data Science using science blogs

Finally, the new API enables data scientists to explore science blogs in more detail. With close to 10,000 science blog posts from 60 different blogs going as far back as 2005 and available as full-text with an open license ([CC-BY](#)), many interesting questions can be explored. I will start with a Jupyter notebook that provides a more detailed analysis than the [Rogue Scholar stats page](#), taking as inspiration the work of the [Journal of Open Source Software](#). I am particularly interested in the more than [750 blog posts that include references](#) in their metadata, as to the best of my knowledge that kind of bibliometric analysis has never been done.

References

Fenner, M. (2023). *front-matter/rogue-scholar-api: Initial public release (v0.6,2)* [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.8433679>

Smith, A. M. (2023). *JOSS publishes 2000th paper*. Journal of Open Source Software Blog. <https://doi.org/10.59349/zh4g1-q7e26>