



# SComCat: Cataloguing Open-source Scholarly Communication Technologies

---

Paul Walk

Director & Founder, Antleaf  
Email: [paul@paulwalk.net](mailto:paul@paulwalk.net)



[www.antleaf.com](http://www.antleaf.com)

# Contents

---

- Introduction and acknowledgements
- Purpose
- Design
- Features
- Sustainability

# Introduction and acknowledgements

---

- **About SComCat**

- Developed by Antleaf for the Confederation of Open Access Repositories (COAR) as part of the Next Generation Libraries Publishing project
- Licensed under a Creative Commons Attribution 4.0 License

- **Acknowledgements**

- Funding
  - Arcadia, a charitable fund of Lisbet Rausing and Peter Baldwin.
- Data Modelling & Collection
  - Ilkay Holt, COAR
  - Sarah Lippincott, Born-Digital



# Purpose

---

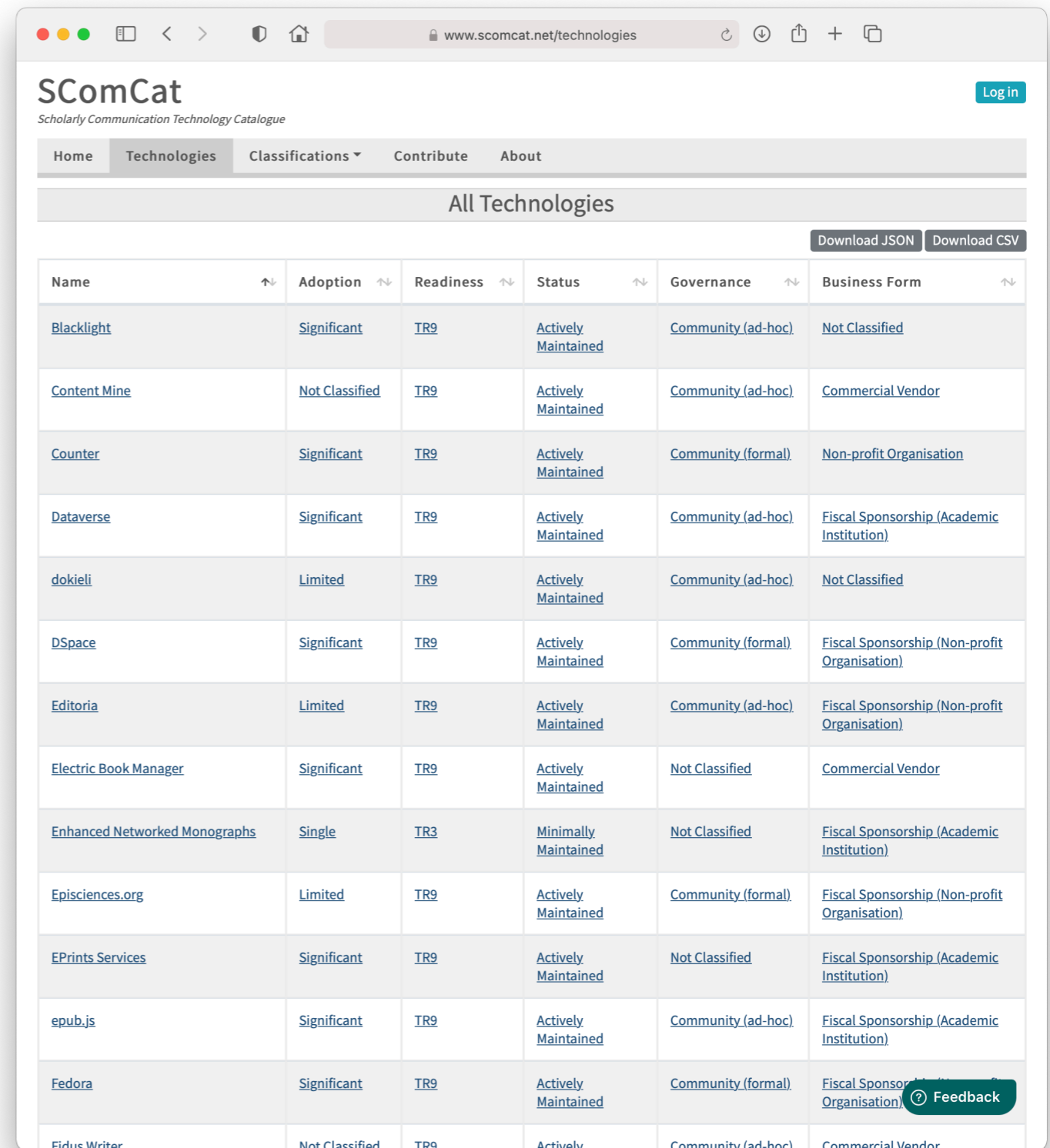
- Initially conceived as a "scan" of open-source software and related technologies in the library publishing space
- Seeded the scan with data from the *Mind The Gap* report, used with permission
- Evolved to become a sustainable, dynamic, "living" database and a **tool** to help users locate software they can use



<https://mindthegap.pubpub.org>

# Design - overview

- Scholarly Communication Technology Catalogue
  - **"technologies" = software, standards, services**
- Categories for classification and **faceted search**
- **Dependency mapping** - i.e. this technology is depended on by some other technology
- Web-based relational database, built & deployed with common components
  - Rails, SQLite, Docker

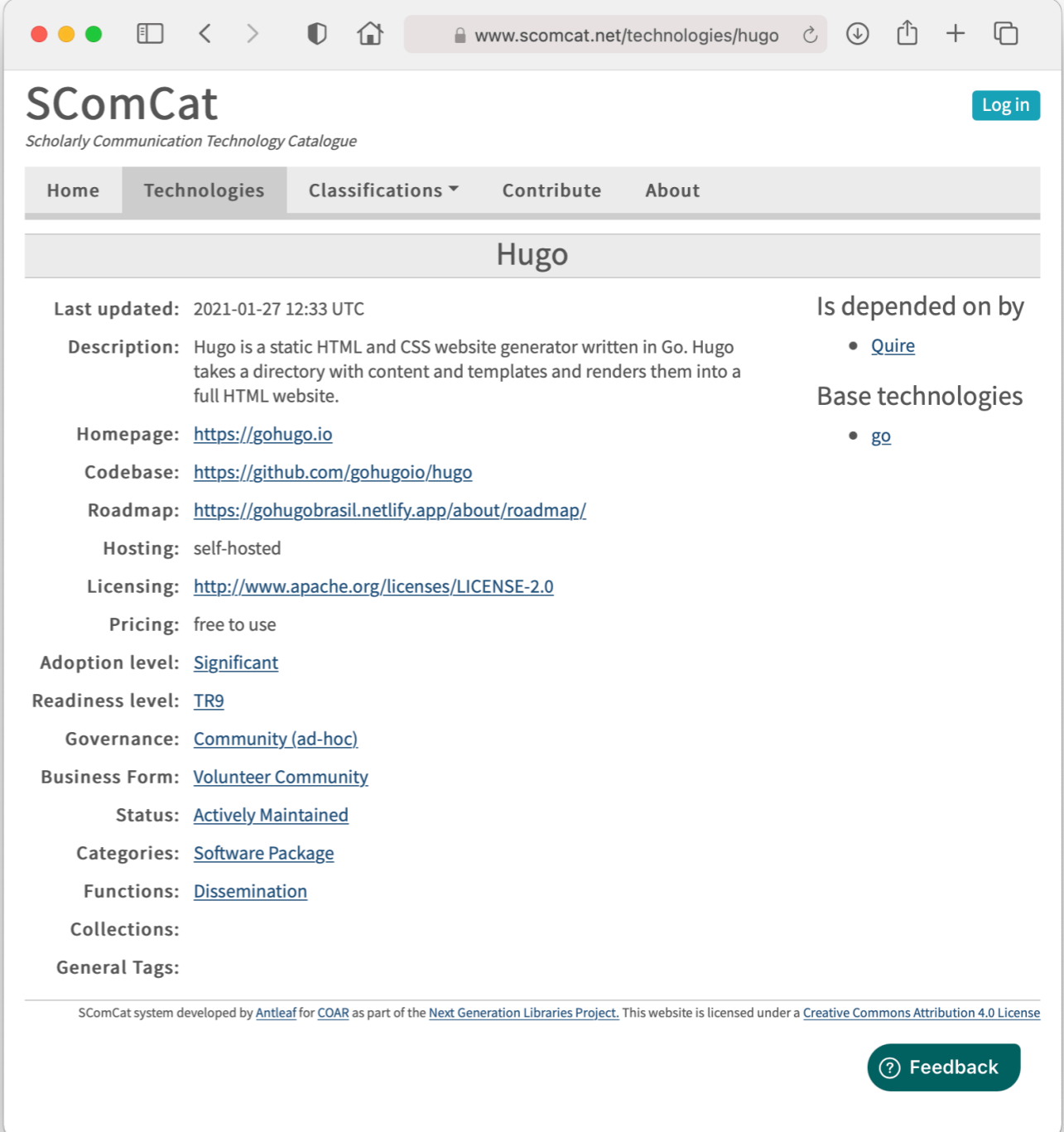


The screenshot shows the SComCat website interface. The browser address bar displays 'www.scomcat.net/technologies'. The page title is 'SComCat' with the subtitle 'Scholarly Communication Technology Catalogue'. A navigation menu includes 'Home', 'Technologies', 'Classifications', 'Contribute', and 'About'. A 'Log in' button is visible in the top right. Below the navigation, there is a section titled 'All Technologies' with 'Download JSON' and 'Download CSV' buttons. The main content is a table with columns: Name, Adoption, Readiness, Status, Governance, and Business Form. The table lists various technologies with their respective attributes.

Name	Adoption	Readiness	Status	Governance	Business Form
<a href="#">Blacklight</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Not Classified</a>
<a href="#">Content Mine</a>	<a href="#">Not Classified</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Commercial Vendor</a>
<a href="#">Counter</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (formal)</a>	<a href="#">Non-profit Organisation</a>
<a href="#">Dataverse</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Fiscal Sponsorship (Academic Institution)</a>
<a href="#">dokieli</a>	<a href="#">Limited</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Not Classified</a>
<a href="#">DSpace</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (formal)</a>	<a href="#">Fiscal Sponsorship (Non-profit Organisation)</a>
<a href="#">Editoria</a>	<a href="#">Limited</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Fiscal Sponsorship (Non-profit Organisation)</a>
<a href="#">Electric Book Manager</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Not Classified</a>	<a href="#">Commercial Vendor</a>
<a href="#">Enhanced Networked Monographs</a>	<a href="#">Single</a>	<a href="#">TR3</a>	<a href="#">Minimally Maintained</a>	<a href="#">Not Classified</a>	<a href="#">Fiscal Sponsorship (Academic Institution)</a>
<a href="#">Episciences.org</a>	<a href="#">Limited</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (formal)</a>	<a href="#">Fiscal Sponsorship (Non-profit Organisation)</a>
<a href="#">EPrints Services</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Not Classified</a>	<a href="#">Fiscal Sponsorship (Academic Institution)</a>
<a href="#">epub.js</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Fiscal Sponsorship (Academic Institution)</a>
<a href="#">Fedora</a>	<a href="#">Significant</a>	<a href="#">TR9</a>	<a href="#">Actively Maintained</a>	<a href="#">Community (formal)</a>	<a href="#">Fiscal Sponsorship (Non-profit Organisation)</a>
<a href="#">Fidus Writer</a>	<a href="#">Not Classified</a>	<a href="#">TR9</a>	<a href="#">Actively</a>	<a href="#">Community (ad-hoc)</a>	<a href="#">Commercial Vendor</a>

# Design - metadata & classifications

- Basic metadata
  - *Homepage URL, license, codebase URL...*
- Classification metadata:
  - Readiness Level
    - Is this ready to use?
  - Adoption Level
    - How widely is this used?
  - Governance
    - How is this governed?
  - Business Form
    - How is this supplied?
  - Status
    - Is this actively supported/developed?
  - Category
    - What type of software/system is this?
  - Function
    - What function(s) does this provide?



The screenshot shows a web browser window displaying the SComCat website. The URL in the address bar is [www.scomcat.net/technologies/hugo](http://www.scomcat.net/technologies/hugo). The page title is "SComCat" with the subtitle "Scholarly Communication Technology Catalogue". A "Log in" button is visible in the top right corner. The navigation menu includes "Home", "Technologies", "Classifications", "Contribute", and "About". The main content area is titled "Hugo" and contains the following information:

- Last updated:** 2021-01-27 12:33 UTC
- Description:** Hugo is a static HTML and CSS website generator written in Go. Hugo takes a directory with content and templates and renders them into a full HTML website.
- Homepage:** <https://gohugo.io>
- Codebase:** <https://github.com/gohugoio/hugo>
- Roadmap:** <https://gohugobrasil.netlify.app/about/roadmap/>
- Hosting:** self-hosted
- Licensing:** <http://www.apache.org/licenses/LICENSE-2.0>
- Pricing:** free to use
- Adoption level:** [Significant](#)
- Readiness level:** [TR9](#)
- Governance:** [Community \(ad-hoc\)](#)
- Business Form:** [Volunteer Community](#)
- Status:** [Actively Maintained](#)
- Categories:** [Software Package](#)
- Functions:** [Dissemination](#)
- Collections:**
- General Tags:**

On the right side of the page, there is a section "Is depended on by" with a link to [Quire](#), and a section "Base technologies" with a link to [go](#). At the bottom of the page, there is a footer with the text: "SComCat system developed by [Antleaf](#) for [COAR](#) as part of the [Next Generation Libraries Project](#). This website is licensed under a [Creative Commons Attribution 4.0 License](#)". A "Feedback" button is located in the bottom right corner.

# Features - faceted search

- Adoption Level
  - *ubiquitous, limited...*
- Readiness Level
  - *TR3, TR9...*
- Governance
  - *vendor, community (formal)...*
- Business Form
  - *commercial vendor, volunteer community...*
- Status
  - *actively maintained, unsupported...*
- Category
  - *software package, standard, framework...*
- Function
  - *authoring, discovery, preservation...*

The screenshot displays the SComCat (Scholarly Communication Technology Catalogue) website. The page features a navigation menu with 'Home', 'Technologies', 'Classifications', 'Contribute', and 'About'. A 'Technology Browser' section is active, showing a search bar and a 'Filter' sidebar. The sidebar includes faceted search options for Collection, Function, Governance, Adoption Level, Business Form, Status, Readiness Level, and Category. The main content area displays a grid of technology cards, each with a title and a brief description. The cards include: Blacklight, Content Mine, Counter, Dataverse, dokieli, DSpace, Editoria, Electric Book Manager, Enhanced Networked Monographs, Episciences.org, EPrints Services, epub.js, Fedora, Fidus Writer, and Fulcrum. The page also includes a 'Log in' button, a 'Reset search/filters' button, and pagination controls.

# Features - structured data exports

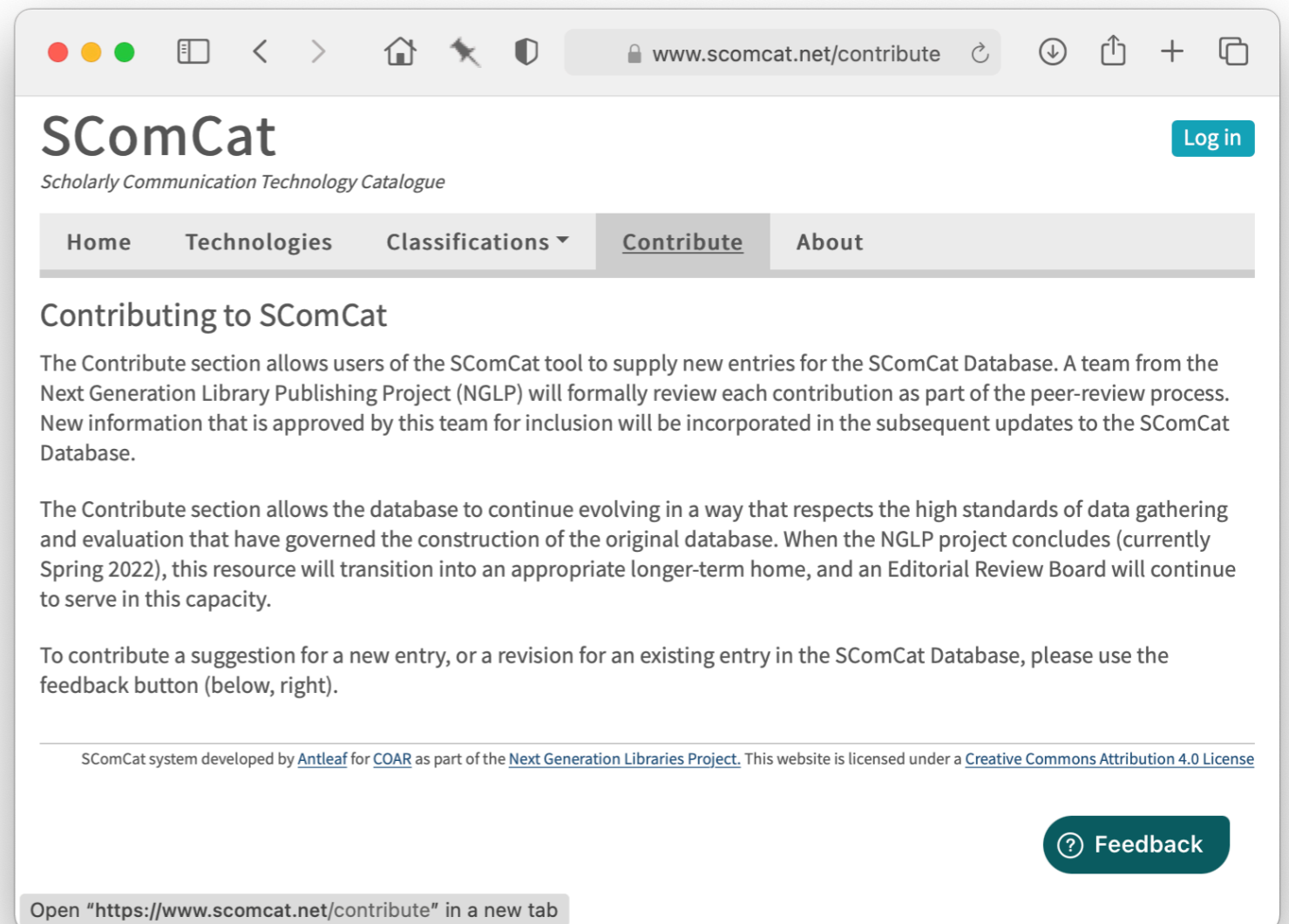
- The data is available for **human** and **machine** consumption
- Export results of faceted search (or entire catalogue) into:
  - CSV (good for MS Excel)
  - JSON

The screenshot displays the SComCat website interface. At the top, there's a navigation bar with 'Home', 'Technologies', 'Classifications', 'Contribute', and 'About'. Below this is a 'Technology Browser' section with a search bar and a 'Reset search/filters' button. The search results are displayed in a grid of technology cards, including 'Blacklight', 'Content Mine', 'Counter', 'Dataverse', 'dokieli', and 'DSpace'. A red circle highlights the 'Download JSON' and 'Download CSV' buttons. In the foreground, there are two windows: a Microsoft Excel spreadsheet showing a table of technology data with columns like 'id', 'name', 'description', 'last\_update', 'homepage', 'codebase', 'roadmap', 'hosting', 'pricing', 'licensing', 'adoption\_level', 'readiness\_level', 'governance', 'status', 'business\_for', 'categories', 'collections', 'functions', and 'base\_techno\_tags'; and a JSON viewer showing the corresponding JSON structure for the selected technology.



# Sustainability - updates & how to contribute

- NGLP project has established a help-desk to receive suggested additions/updates to the catalogue
- Users can use the "feedback" function to comment in general or on particular technologies
- Suggested additions/updates are reviewed by an NGLP editorial group

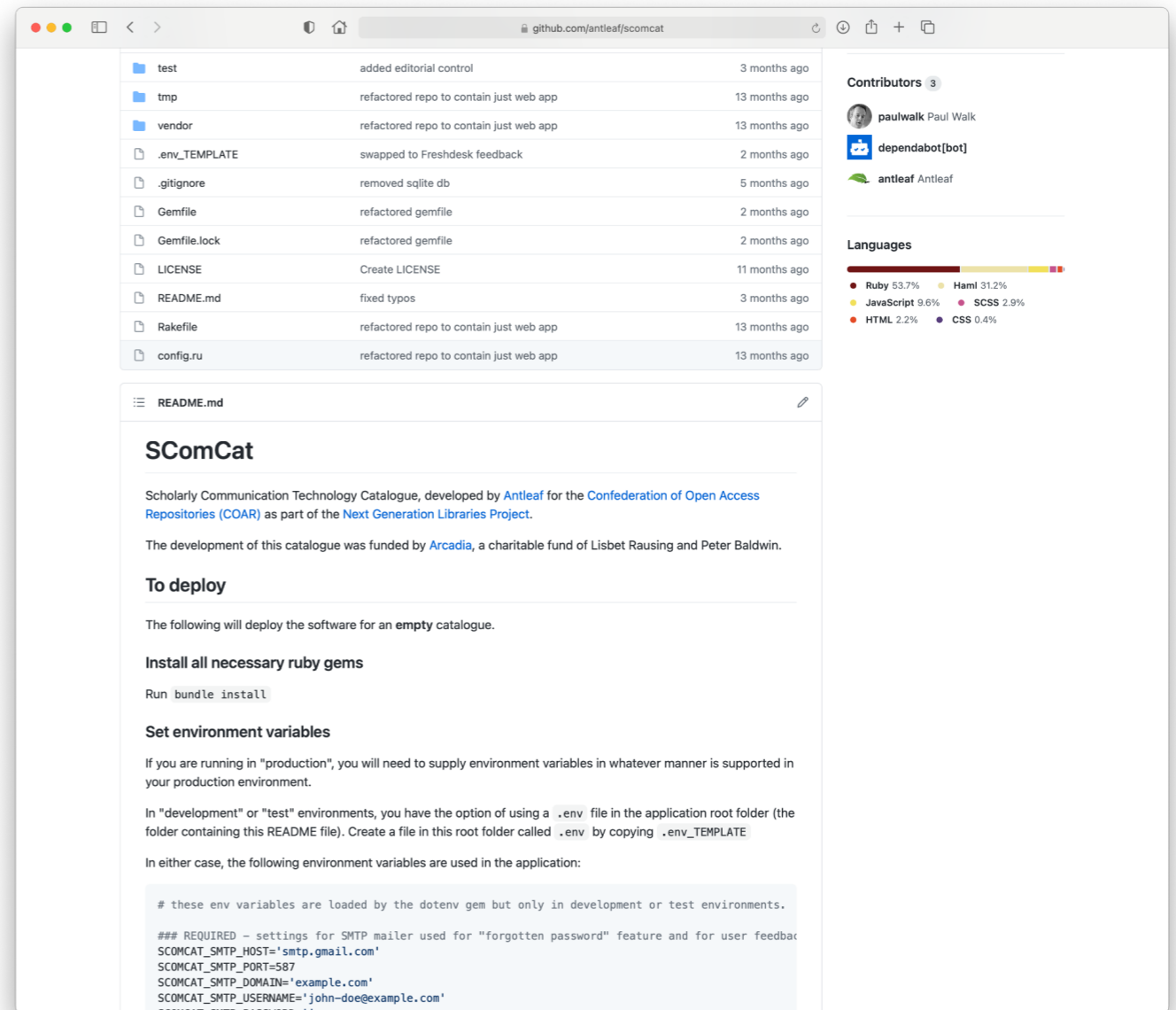


The screenshot shows a web browser window displaying the 'SComCat' website. The URL in the address bar is 'www.scomcat.net/contribute'. The page title is 'SComCat' with the subtitle 'Scholarly Communication Technology Catalogue'. A navigation menu includes 'Home', 'Technologies', 'Classifications', 'Contribute', and 'About'. The 'Contribute' section is active, featuring a 'Log in' button in the top right. The main content area is titled 'Contributing to SComCat' and contains two paragraphs of text explaining the contribution process and the project's future. A 'Feedback' button is located at the bottom right of the page. A footer note states: 'SComCat system developed by Antleaf for COAR as part of the Next Generation Libraries Project. This website is licensed under a Creative Commons Attribution 4.0 License'. A browser notification at the bottom of the window reads: 'Open "https://www.scomcat.net/contribute" in a new tab'.



# Sustainability - code and running service

- Data is licensed as CC-BY
- Software is open-source (MIT license)
- All source code on Github
- Configured to be deployed in standard containers (Docker & Kubernetes enabled)
- Maintained by COAR for NGLP until the end of the project. Actively seeking a new home for SComCat thereafter



Thank you for listening!

<https://www.scomcat.net>

[paul@paulwalk.net](mailto:paul@paulwalk.net)