SComCat: Cataloguing Open-source Scholarly Communication Technologies

Paul Walk
Director & Founder, Antleaf
Email: paul@paulwalk.net

·ANTLEAF·

www.antleaf.com

Contents

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

ANTLEAF

COAR Confederation of Open Access Repositories

NEXT GENERATION LIBRARY PUBLISHING

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

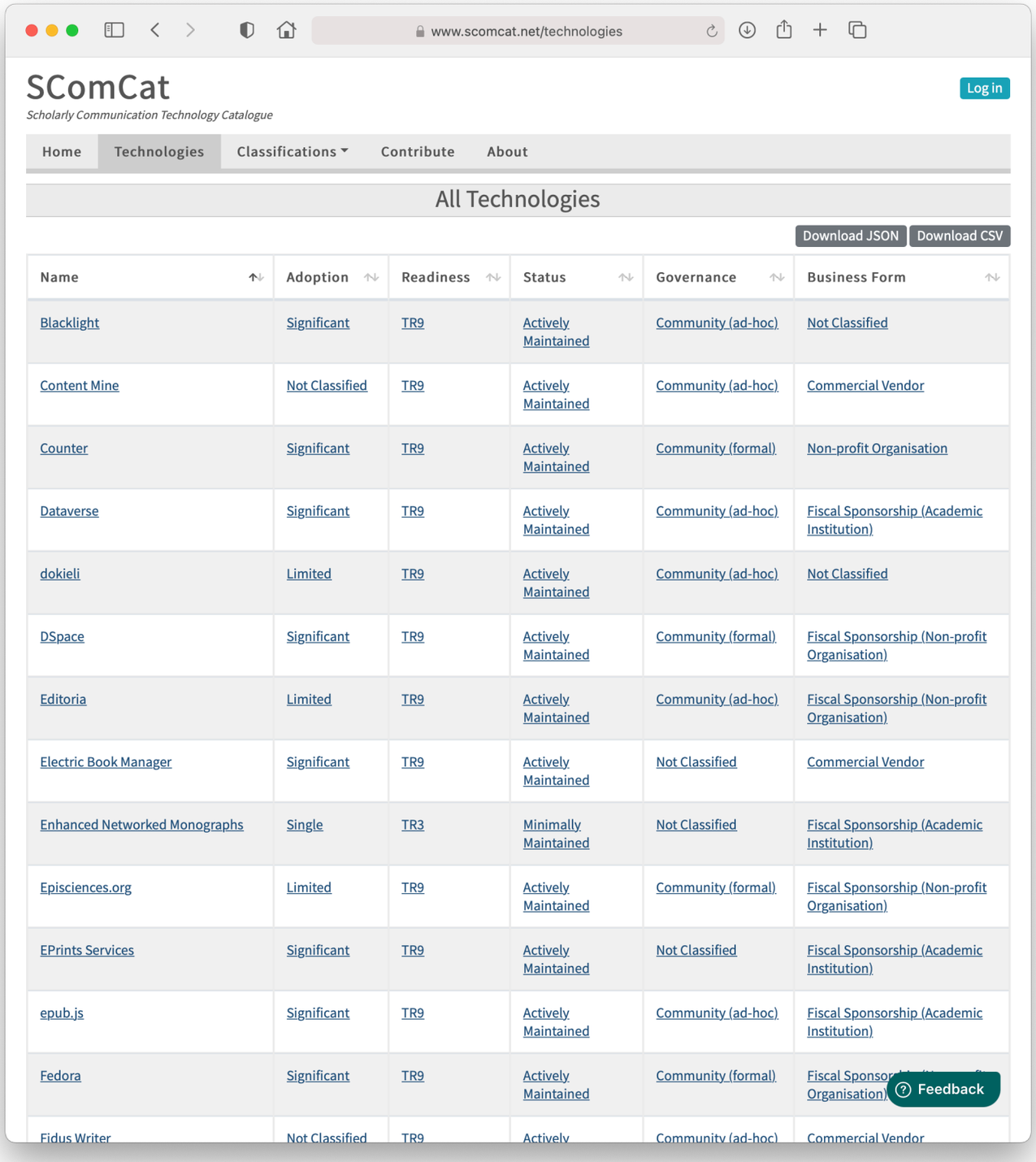
Design - overview

- Scholarly Communication Technology Catalogue
  - **"technologies" = software, standards, services**

- Categories for classification and **faceted search**

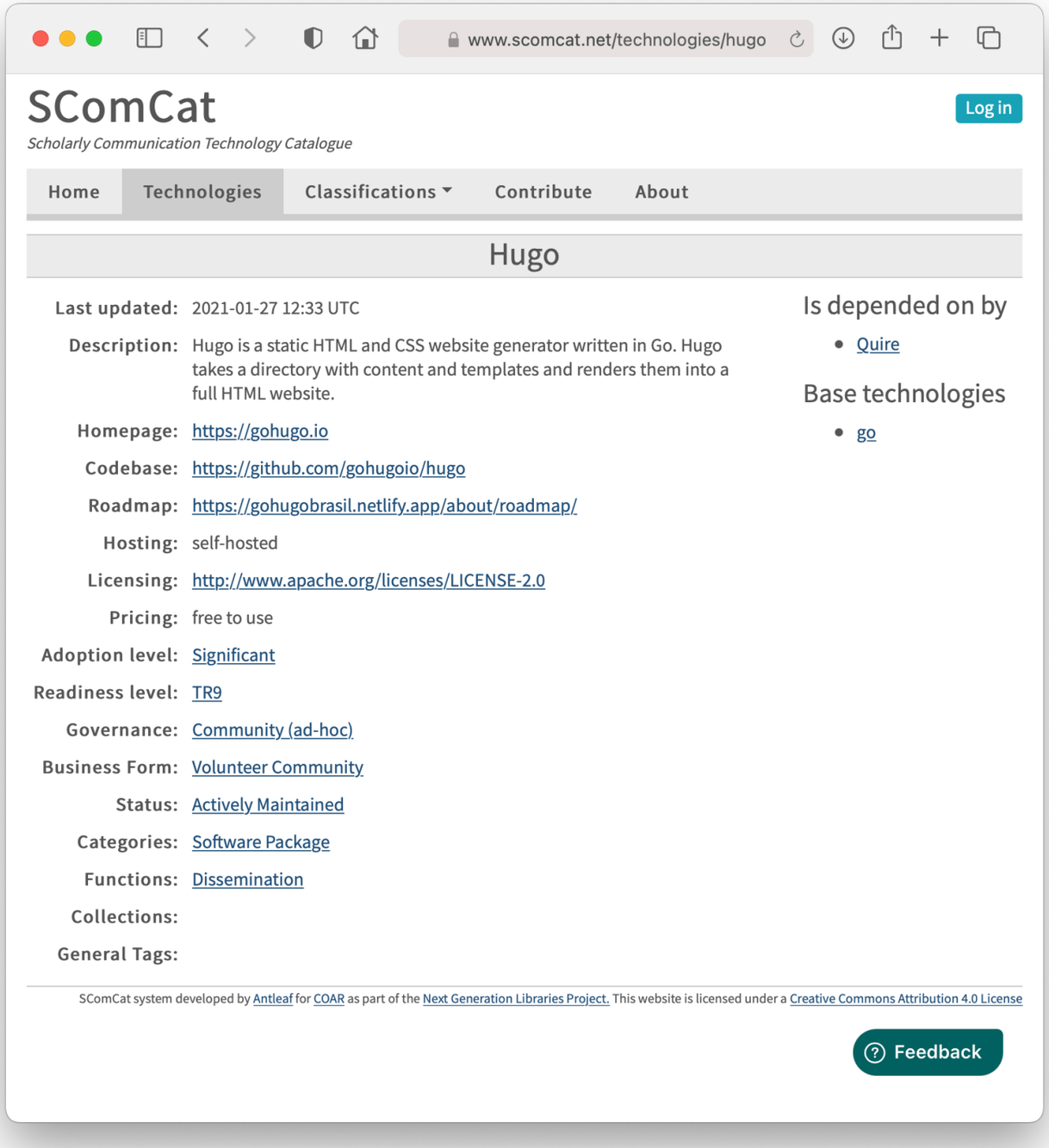
- **Dependency mapping -** i.e. this technology is depended on by some other technology

- Web-based relational database, built & deployed with common components
  - Rails, SQLite, Docker



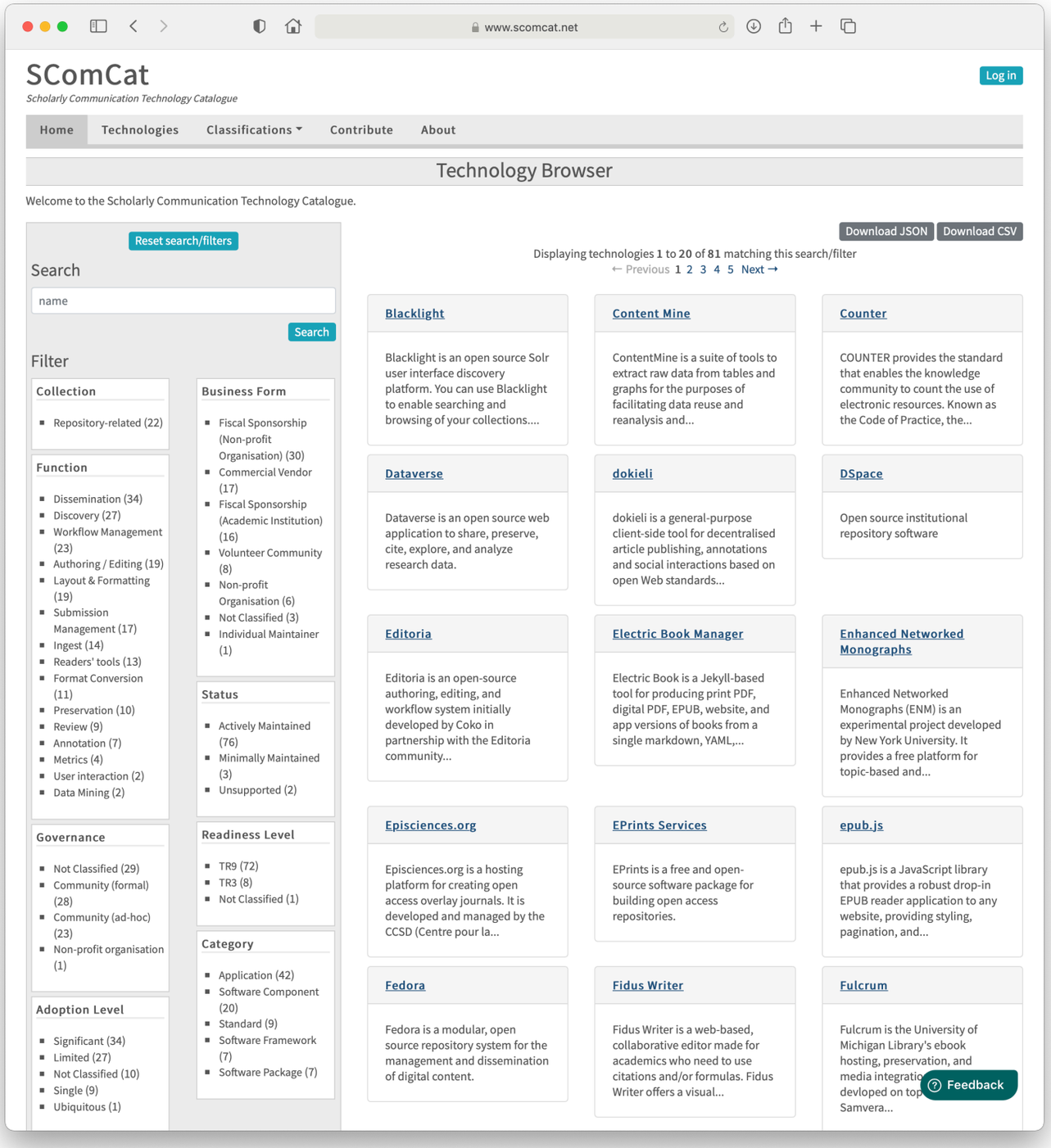
Design - metadata & classifications

- 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?



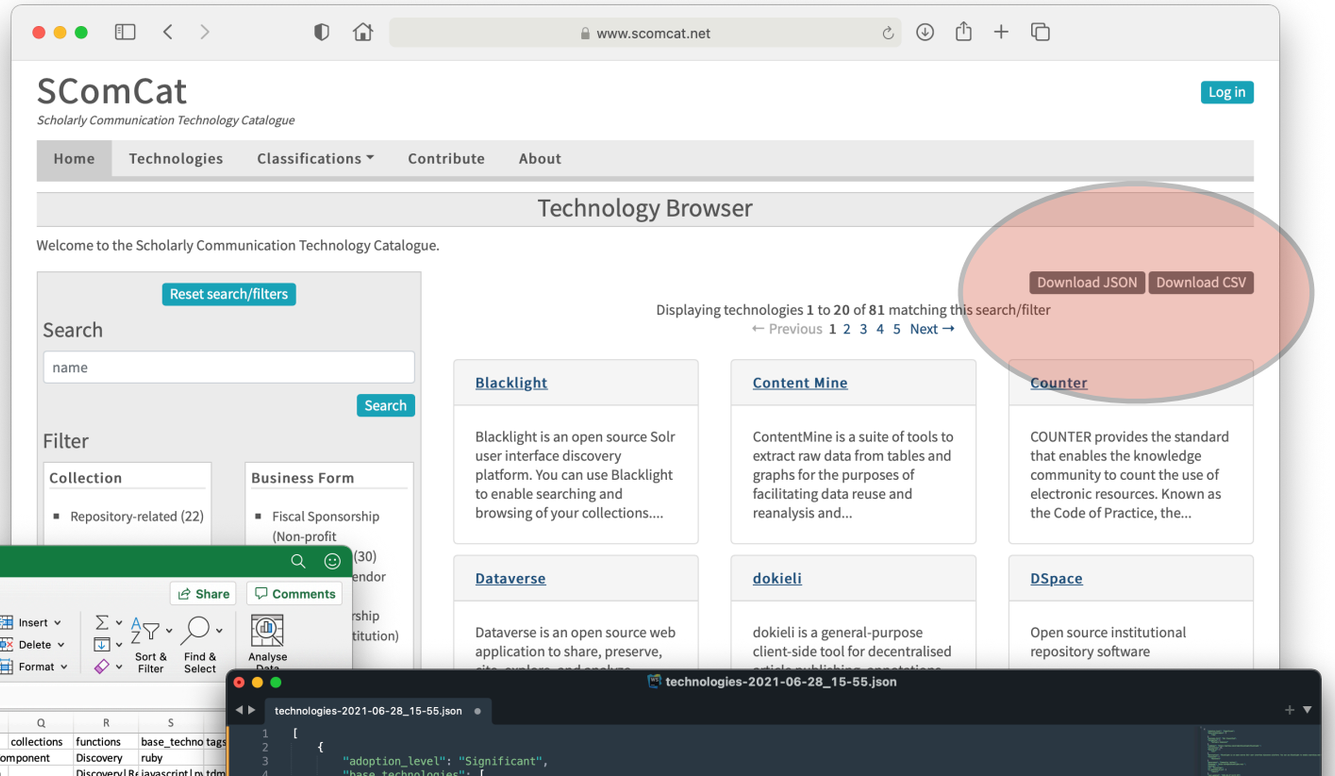
Features - faceted search

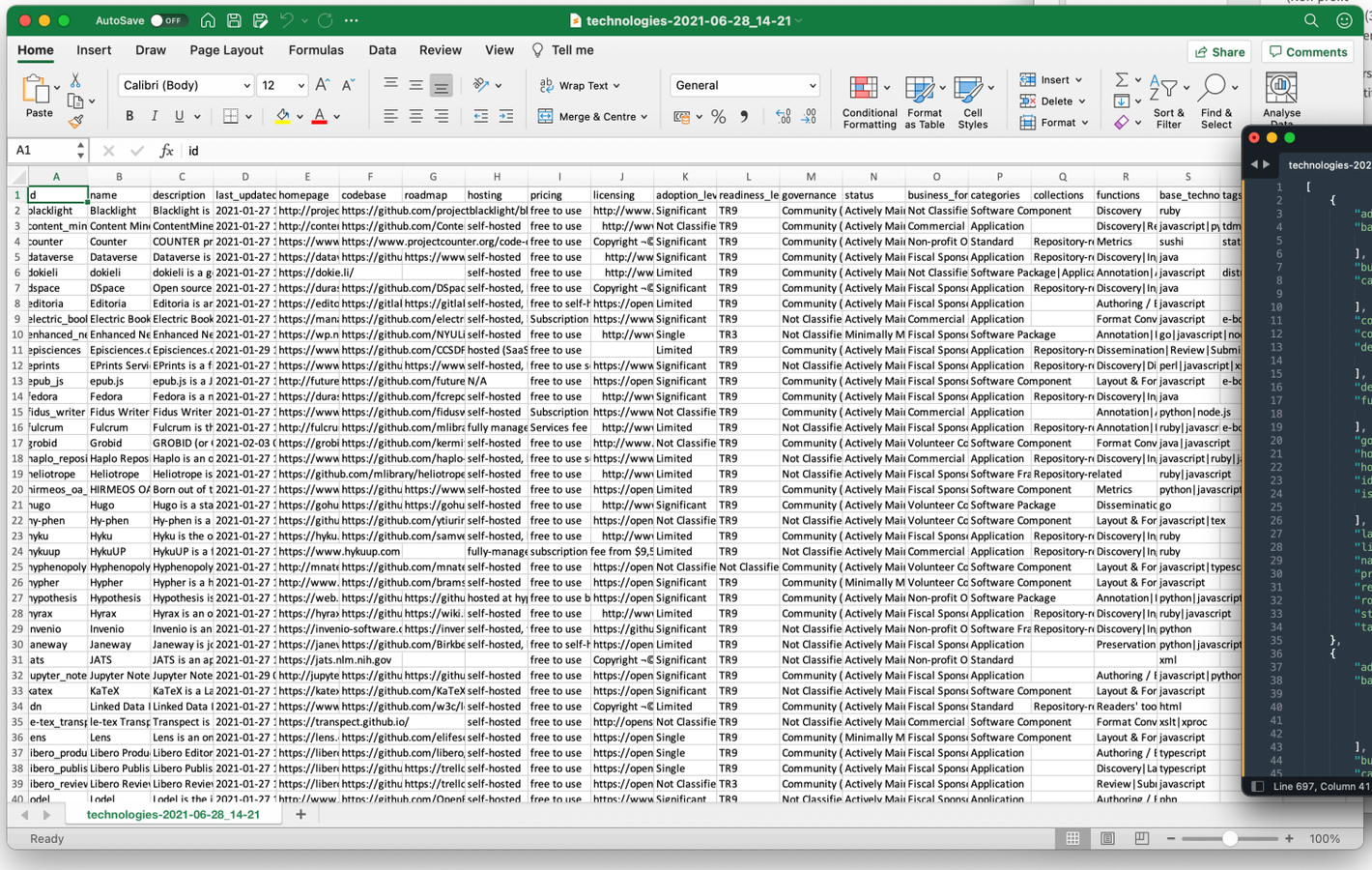
- Adoption Level
  - *ubiquitous, limited...*
- Readiness Level
  - *TR3, TR9...*
- Governance
  - *vendor, community (formal)...*
- Business Form
  - *commercial vendor, volunteer community...*
- Status
  - *actively maintained, unsupported...*
- Category
  - *software package, standard, framework...*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

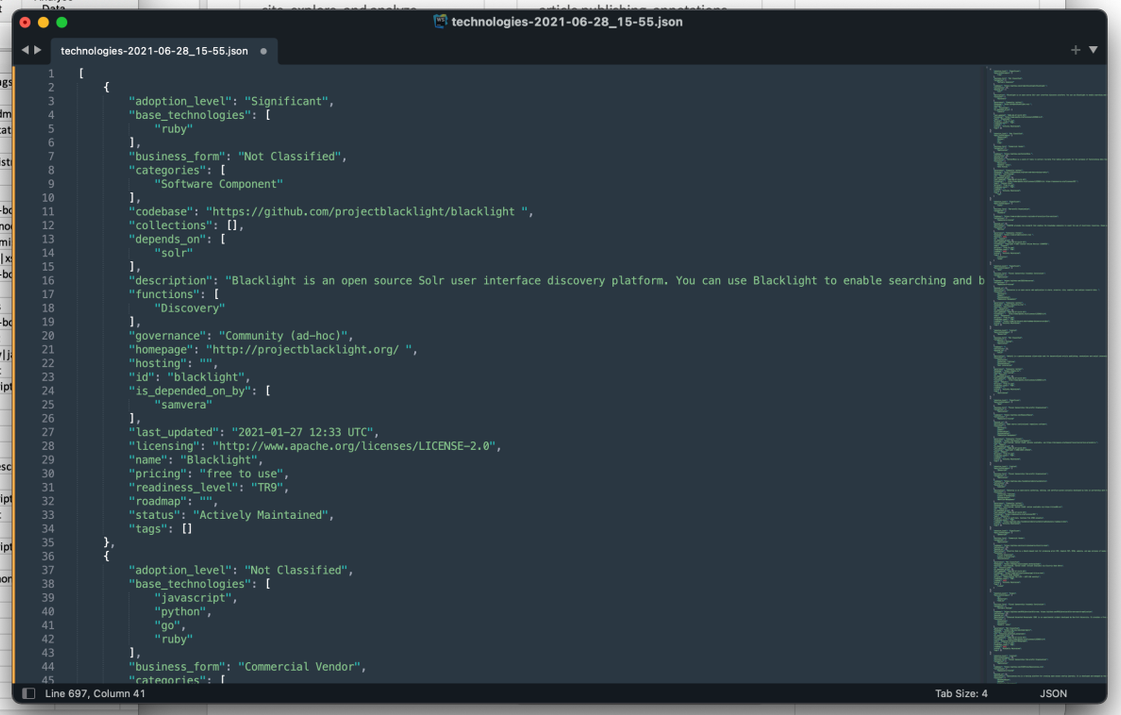


Features - structured data exports

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

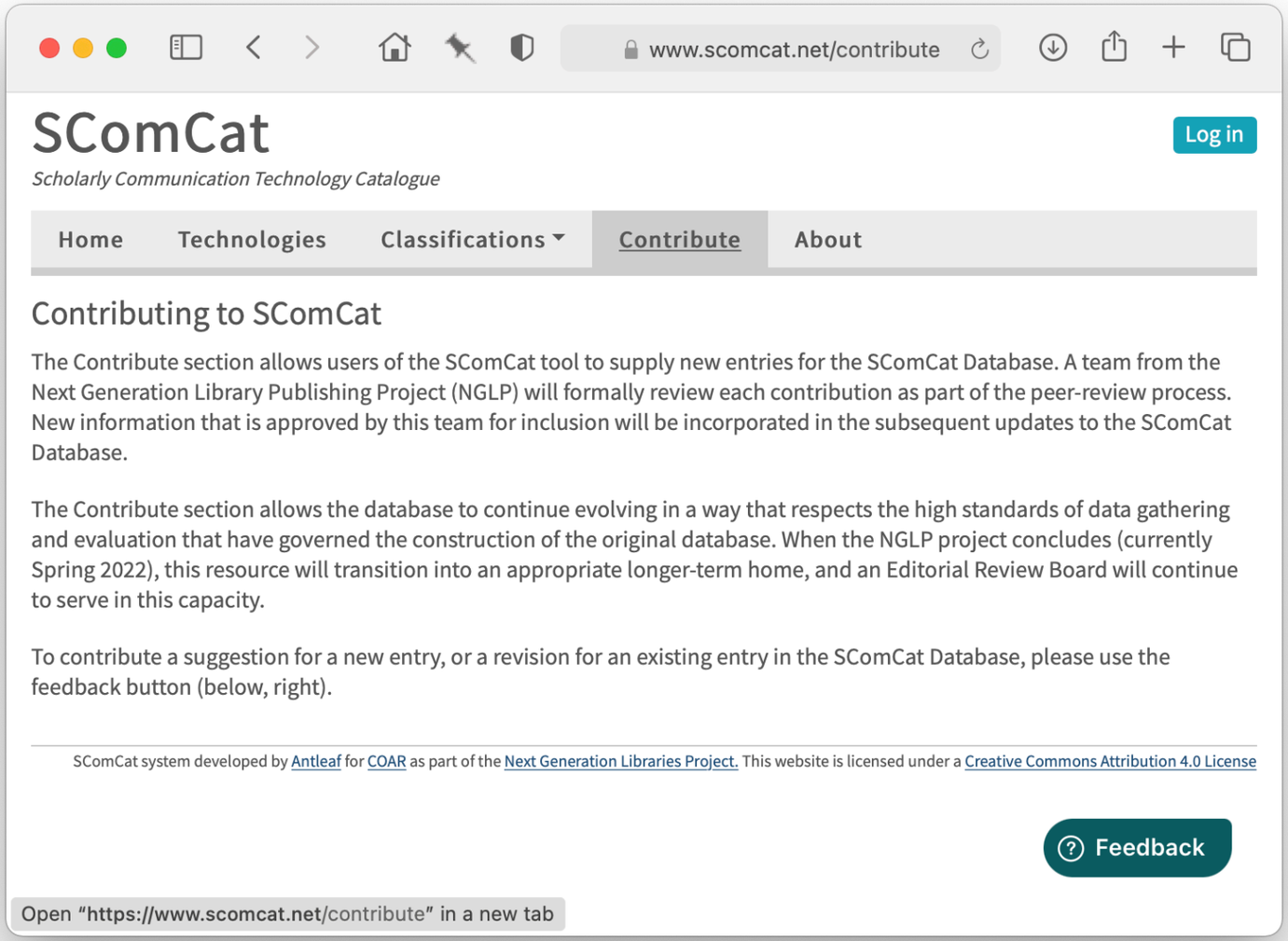






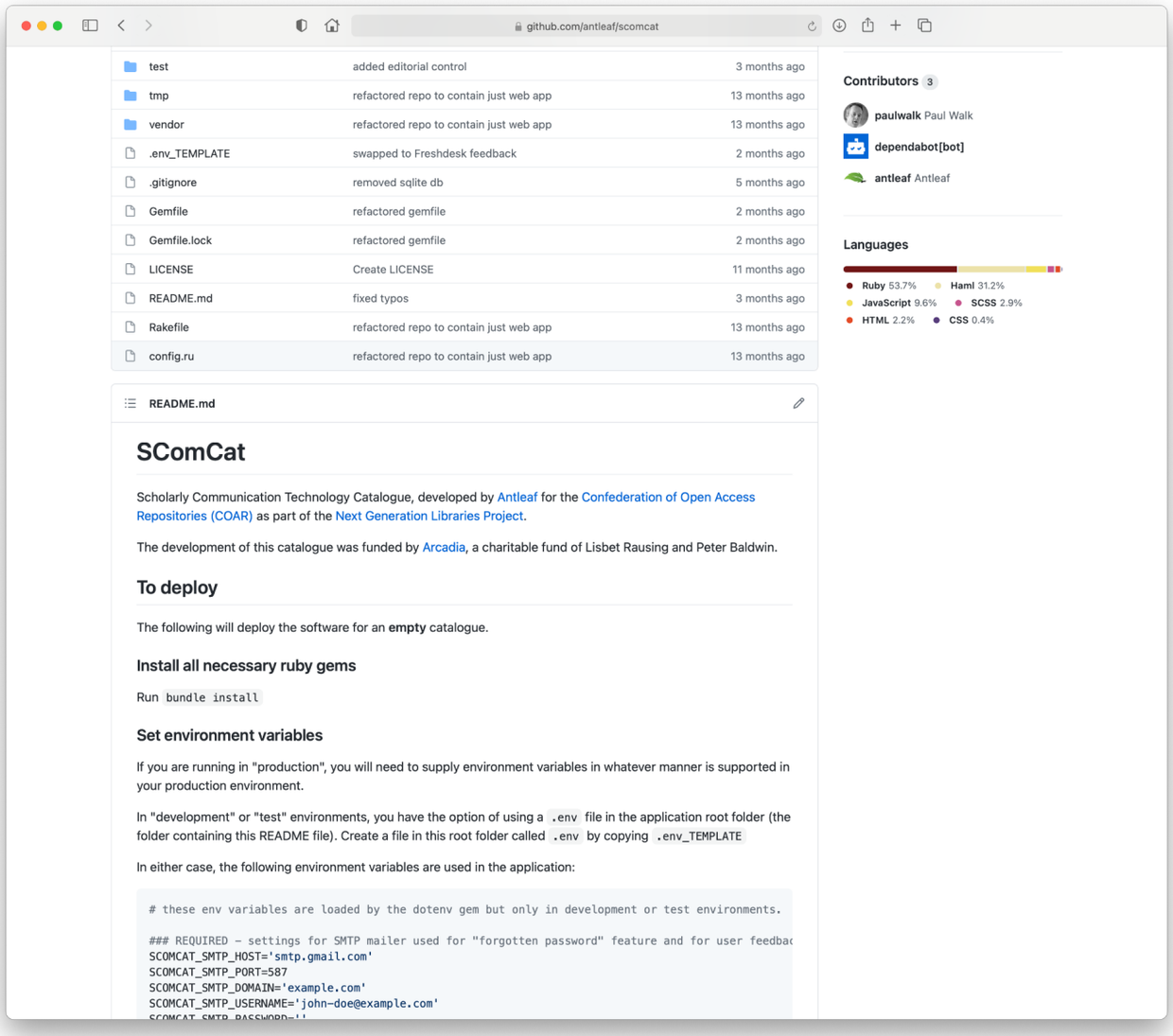
Sustainability - updates & how to contribute

- 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



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