SPARK Your Systematic Review: Automating Literature Collection, Filtering, and Topic Discovery with Python

Cameron Frederick Atkinson

[1] School of Social Sciences, University of Tasmania, Tasmania 7005, Australia
[2] Natural Hazards Research Australia, PO Box 116, Carlton South, VIC 3053, Australia

AI-pocalypse Now: Automating the Systematic Literature Review with SPARK (Systematic Processing and Automated Review Kit) – Stage One:

SPARK automates reviews by integrating Python scripts for data gathering, organisation, and filtering. It automates literature searches, removes duplicates, applies keyword screening, and employs LDA topic modelling to identify key themes, streamlining research efficiency and accuracy.

New Method to Conduct a Systematic Literature Review

SPARK automates systematic literature review (SLR) workflows in four stages:

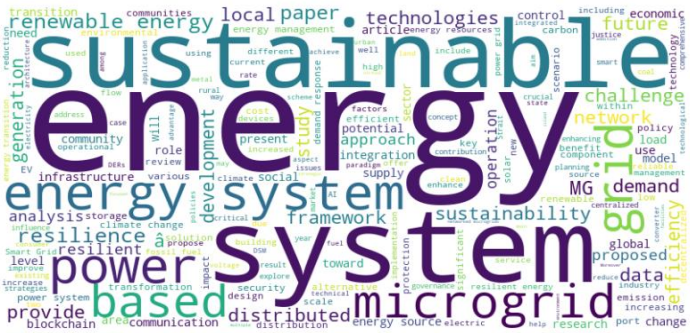
1. Gathering articles via API queries from Web of Science and Scopus, saving results to structured CSVs.
2. Organising & filtering data by removing duplicates through DOI matching and fuzzy title comparison, ensuring a clean dataset.
3. Keyword screening applies user-defined terms to refine relevant articles.
4. LDA topic modelling iterates over included abstracts to identify hidden themes and keywords, helping structure data extraction templates.

SPARK ensures reproducibility, scalability, and transparency, making automated SLRs more efficient, structured, and research-driven.

SPARK for Energy Resilience and Sustainability

SPARK is applied to a SLR on energy resilience and sustainability, automating data collection, filtering, and topic discovery.

Using Web of Science and Scopus, it retrieves, organises, and analyses research on energy governance, resilience and sustainability, ensuring efficient, transparent, and scalable evidence synthesis for policy and academic insights.



AI Enhances Research, it does not Replaces Researchers

SPARK enhances efficiency and transparency in SLRs by automating time-consuming tasks like literature retrieval, filtering, and topic discovery.

Unlike Generative AI, SPARK's hard-coded tools eliminate the risk of hallucinations, maintaining research accuracy and integrity. The modular structure allows researchers to adapt search and screening criteria for domain-specific needs.

While API access and programming knowledge may be barriers, SPARK's structured approach bridges these gaps by providing reproducible, adaptable scripts.

By integrating human oversight with automation, SPARK accelerates evidence synthesis, topic mapping, and data extraction, demonstrating the future of AI-assisted research automation in SLRs.

Further information

For additional information scan the QR code or contact:
Cameron Frederick Atkinson, PhD Candidate, University of Tasmania
Cameron.Atkinson@utas.edu.au

Australian Government

UNIVERSITY of TASMANIA