The Journal of Open Source Software

CART: A Tool for Making Paper Relevancy Screening Easier

Nathan Reitinger [1]

1 University of Maryland, USA

DOI: 10.21105/joss.06494

Software

- Review
- Repository
- Archive

Editor: Nikoleta Glynatsi

Reviewers:

- @atanikan
- @jhculb

Submitted: 20 February 2024
Published: 16 July 2024

Summary

CART (Culling Abstracts for Relevancy in Teams) provides a way for a team of individuals to complete the first step common to most systematization of knowledge (SoK) projects[1]: culling, in teams, a large set of research papers (i.e., titles and abstracts) based on per-project relevancy criteria. Existing software solutions either do not fit this need in a usable way—leading to error-prone voting—or do not offer the necessary functionality to work on the problem effectively.

CART provides both functionality and usability: *Functionally*, CART accommodates teams of any size (i.e., accommodating race conditions), review requirements of any number (e.g., mandating that each paper is reviewed by n team members), and dynamic participation rates (e.g., team members may participate in as little or as much reviewing as they like). *Usability-wise*, CART provides a gamified user interface, making the process of reviewing papers easier, and therefore more accurate and efficient.

Statement of Need

A systematization of knowledge or systematic review paper is a popular type of paper found across multiple disciplines: medicine, computer science, sociology, psychology, economics, political science, marketing, and genetics. This type of paper most commonly involves a two-stage process. In the first stage, researchers gather a large set of papers—easily over 10,000—and then manually filter those papers (based on paper titles and abstracts) according to predefined relevancy criteria (Franz et al., 2021; Reitinger & Mazurek, 2023; Singhal et al., 2023; Thomas et al., 2021). For example, imagine a researcher who wanted to learn about legally-sufficient data sanitization tools (Bellovin et al., 2019). A paper on the prevalence of a particular web tracking technique like canvas fingerprinting (Reitinger & Mazurek, 2021) would be marked as *irrelevant*, but a paper on statutory differential privacy requirements (Reitinger & Deshpande, 2023) would be marked as *relevant*. After this first step is complete—often resulting in a set of around 100 papers—most researchers then move on to analyzing papers in-depth, the second stage, using qualitative coding techniques well-fit by existing software.

Surprisingly, software to support this first step of paper culling is lacking. Existing tools are either too generalist (i.e., resulting in ineffective paper reviewing) or too fine-tuned (i.e., resulting in tradeoffs impacting functionality). Generalist tools like Microsoft Excel or Google Sheets could allow teams of researchers to check the relevancy of snippets of text, but these tools can be clunky (i.e., given a scale of more than 10,000 papers in consideration) and not usable (i.e., reviewing hundreds of rows in a spreadsheet is error-prone and fatiguing). Other tools, like Atlas.ti (Smit & Scherman, 2021) or maxQDA (Franzosi et al., 2013), could allow teams of researchers to “code” for relevancy, but do not naturally accommodate dynamic

[1] These projects are also called systematic reviews.

participation (e.g., team members might not be able to review the same amount of papers) or variable review rates (e.g., sometimes you want all team members to review all papers, and sometimes you want at least two, but you do not set who those team members are). Enter CART.

CART

CART provides a way for teams to accomplish the goal of culling a large set of research papers. Team members may contribute as little or as much as they want, the experience of reviewing papers is gamified, and CART prioritizes protection against data loss, error, and data corruption. Most importantly, CART is fully functional without requiring server space (although this is certainly possible with an appropriately provisioned environment) and easily tunable for a project's specific needs.

CART utilizes Flask, ngrok, and a web interface to allow one team member to create a server that other team members can access. As shown in Figure 1, CART has two primary functions.

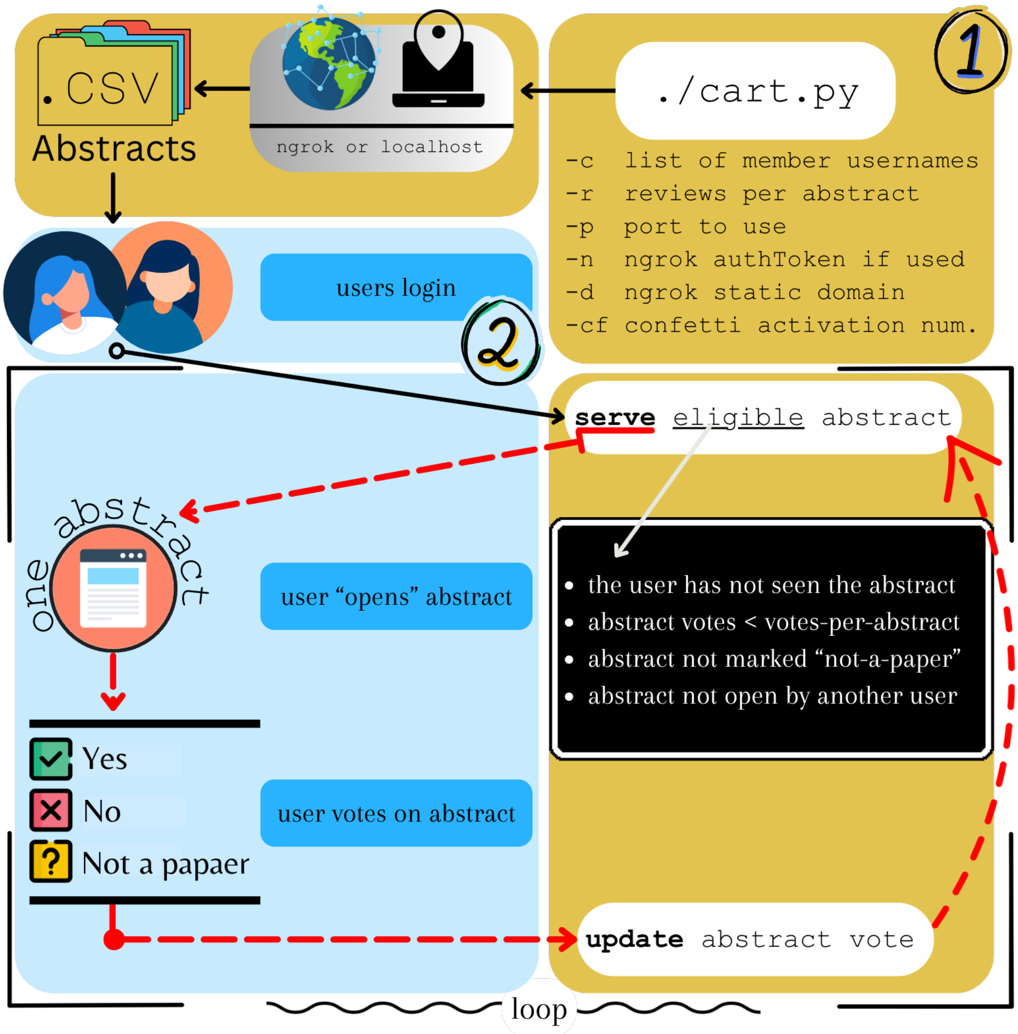

Figure 1: Architecture of CART.

First, CART runs a Flask application responsible for serving and updating papers (i.e., represented as an abstract, title, URL, and a few metadata fields). A team member will only be served

The Journal of Open Source Software

an abstract if: (1) the paper is not one that the team member has seen before: (2) the paper has not yet met the required amount of votes per paper; (3) the paper has not been flagged as "not" a research paper (i.e., a proceedings introduction, extended abstract, or similar document); and (4) the paper is not currently being viewed by another user. After a vote is cast, CART will update a log noting who voted, what their vote was, and how many votes have been cast for the paper. This process of serving and updating papers will loop until there are no more eligible papers available.

Figure 2: Web interface tabs: home, progress, history, information, and account.

Second, CART creates a gamified user experience for reviewing papers. The web interface makes voting on a paper simple (i.e., button click or keyboard press), allows users to set personal goals, and rewards (with confetti) users on every 50th (changeable) vote submitted. The web interface also includes tabs (Figure 2) for **voting** on papers (home), viewing the team's **progress** (bar chart, total vote counts), viewing personal voting **history** (last 50 edits, permitting a change-of-vote), logging information **about** the project (collaborative editing with backups, where users may write guidelines for voting), and viewing **account** details (logging in and out). Moreover, the web app is optimized to provide information in a timely fashion (e.g., storing 'open' papers in logs rather than reviewing the full corpus of possible papers to find open papers) to enable speedy reviewing.

References

Bellovin, S. M., Dutta, P. K., & Reitinger, N. (2019). Privacy and synthetic datasets. *Stan. Tech. L. Rev.*, *22*, 1. https://doi.org/10.31228/osf.io/bfqh3

Franz, A., Zimmermann, V., Albrecht, G., Hartwig, K., Reuter, C., Benlian, A., & Vogt, J. (2021). SoK: Still plenty of phish in the sea—a taxonomy of user-oriented phishing interventions and avenues for future research. *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 339–358. https://doi.org/10.26083/tuprints-00020675

Franzosi, R., Doyle, S., McClelland, L. E., Putnam Rankin, C., & Vicari, S. (2013). Quantitative narrative analysis software options compared: PC-ACE and CAQDAS (ATLAS.ti, MAXqda, and NVivo). *Quality & Quantity*, *47*, 3219–3247. https://doi.org/10.1007/s11135-012-9714-3

Reitinger, N., & Deshpande, A. (2023). Epsilon-differential privacy, and a two-step test for quantifying reidentification risk. *Jurimetrics: The Journal of Law, Science & Technology*, *63*(3), 263–317. https://www.americanbar.org/content/dam/aba/publications/Jurimetrics/spring-2023/epsilon-differential-privacy-and-a-two-step-test-for-quantifying-reidentification-risk.pdf

Reitinger, N., & Mazurek, M. L. (2021). ML-CB: Machine learning canvas block. *Proc. Priv. Enhancing Technol.*, *2021*, 453–473. https://doi.org/10.2478/popets-2021-0056

Reitinger, N., & Mazurek, M. L. (2023). SoK: Considerations in measuring compliance with privacy regulations. *7th Workshop on Technology and Consumer Protection (ConPro 2023)*. https://conpro23.ieee-security.org/papers/reitinger-conpro23.pdf

Singhal, M., Ling, C., Paudel, P., Thota, P., Kumarswamy, N., Stringhini, G., & Nilizadeh, S. (2023). SoK: Content moderation in social media, from guidelines to enforcement, and research to practice. *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&p)*, 868–895. https://doi.org/10.1109/eurosp57164.2023.00056

Smit, B., & Scherman, V. (2021). Computer-assisted qualitative data analysis software for scoping reviews: A case of ATLAS.ti. *International Journal of Qualitative Methods*, *20*, 1–3. https://doi.org/10.1177/16094069211019140

Thomas, K., Akhawe, D., Bailey, M., Boneh, D., Bursztein, E., Consolvo, S., Dell, N., Durumeric, Z., Kelley, P. G., Kumar, D., & others. (2021). Sok: Hate, harassment, and the changing landscape of online abuse. *2021 IEEE Symposium on Security and Privacy (SP)*, 247–267. https://doi.org/10.1109/sp40001.2021.00028