CrowdLit: Crowd-powered Literature Search

Zhenyao Cai, Anthony Phonethibsavads, Shayan Doroudi

University of California, Irvine zhenyaoc@uci.edu, tonyphonet@gmail.com, doroudis@uci.edu

Abstract

As scientific literature is rapidly expanding, researchers are faced with the difficulty of finding all the relevant literature to contextualize and influence their own research. There is a proliferation of tools using artificial intelligence to help aid researchers in literature search, but we believe these tools may still miss literature (e.g., ones coming from other disciplines). To address this situation, we introduce Crowd Literature Search, a pipeline that involves crowdworkers or laypeople in the literature search process. We developed CrowdLit, a platform for enabling Crowd Literature Search. We describe various features in CrowdLit and how they can potentially facilitate literature search, scientific communication, and involving the general public in scientific research.

Introduction

The rapid growth of scientific publications makes it nearly impossible for researchers to read all existing literature (Bornmann, Haunschild, and Mutz 2021). This poses the risk of overlooking key studies, leading to potential gaps in knowledge, or what Swanson (1986) has termed "undiscovered public knowledge." While automated systems have been developed to recommend relevant papers, they often miss cross-disciplinary studies and rely too heavily on a researcher's judgment, which can inadvertently introduce biases (Kang et al. 2023; Choe et al. 2021). To address this, we introduce the Crowd Literature Search (CLS) pipeline to leverage the collective intelligence of both researchers and the general public to discover a broader spectrum of literature. By involving laypeople (e.g., crowdworkers, citizen scientists, or undergraduate students), it offers a multidimensional perspective on literature discovery, while also enabling public participation in scientific processes. We believe CLS can have learning benefits for both researchers and laypeople; researchers can benefit by improving their scientific communication skills while laypeople can get a low-stakes entry point into research, learn literature search skills, and learn about research that aligns with their interests. We developed CrowdLit, a platform designed to streamline and enhance the CLS experience. A feasibility study of CLS involving professional researchers and crowdworkers showed that while researchers found the CLS approach useful, especially for early-stage projects, nonresearchers appreciated the chance to delve into research.

We believe a demonstration of CrowdLit at the HCOMP conference will provide benefit to researchers studying human computation and collective intelligence, as well as valuable feedback to us in furthering our platform.

Crowd Literature Search and CrowdLit

The Crowd Literature Search (CLS) pipeline is designed to integrate crowdworkers¹ into the literature search phase of the research process. The CLS pipeline is characterized by five stages, forming an iterative cycle that fosters continuous collaboration between researchers and crowdworkers.

To implement the CLS Pipeline, we developed CrowdLit, a web-based platform where researchers can create and deploy literature search tasks for their research projects to be completed by crowdworkers. CrowdLit is implemented with NodeJS, MongoDB, React, and features that use various APIs including OpenAI's GPT, CrossRef, Semantic Scholar, Hugging Face. Below we describe the five stages of CLS and how they were implemented in the CrowdLit platform:

Stage One: Researchers Initiate Tasks Researchers begin by uploading project information. This includes the overarching research question they're investigating, key terms essential for the literature search, a detailed description of their projects, clarifications on specific jargon and terminologies, and tips on effective literature searching.

In addition to allowing researchers to upload project information, CrowdLit tries to help researchers improve their science communication using large language models. Specifically, we provide a series of options for researchers to improve their project description including "Evaluate your description and get some feedback from AI," "Add more details in your current description," "Restructure the description," and "Detect [jargon] from description." Each of these tools calls the OpenAI GPT API using a carefully-crafted system message to give feedback to researchers or automate parts of the research communication process. For example, automatically detecting jargon would pre-populate a

¹In what follows we use the term "crowdworker" to refer to any person contributing to the literature search tasks. While we recruited crowdworkers from Prolific in our feasibility study, we note that these individuals do not need to be paid crowdworkers; they could be citizen scientists or students.

	e	Researcher3	Normal : BIUS: HEEE & BIA						
Toward a Learning Edit Project	Evaluato Result SAVE PUBLISH		We explore how crowdravines can be trained to tackle complex crowetourcing tasks. We are particularly interested in training novice workers to partom well on adving tacks in initiations when the space of strategies is long and workers needs to discover and by different strategies to be successful. In all to operative, a perform a comprision of the different training intraleys. For comparison we be such challenges, we allow that providing expert examples is an elective time of training usupassing other times of training in nearly all measures of interest. Therewer, but thinking relies croses to domain appression, which may be operative tacking. Therefore, and a successful training in creatly all measures of interest. Therewer, but thinking relies croses to domain appression, which may be operative to discing. Therefore, a factor of the discing.						
Basic Information *	Description*		experiment we study the feasibility of training workers in the absence of domain experite. We show that having workers validate the work of their poer workers can be own more offective than having them review expert examples if we only present solutions filtered by a threshold length. The results suggest that considering examples are then the state of						
Research Question *	Normal \Rightarrow B $I \ \sqcup \Rightarrow$ 39 $\models \equiv \equiv \equiv \Rightarrow $ $\Rightarrow \blacksquare \ I_x$		This approach of having workers validate the work of their overs and Allering solutions by length can be more effective than using excert, examples, indicating that leveraging the cohective intelligence of over workers can be harnessed for training in an automated manner.)						
Keywords * Description and Jargons *	We explore how convelocates can be trained to lack's complex convelocating tasks. We are particularly interested in training novice workers to particularly interesting the training tasks of the tra	ce-by-	Evaluate your description and get some feedback from AI						
Related Articles Researcher Tip	of interest. However, such training mixes on access to domain expertise, which may be opportive or laking. Therefore, in a generative study the teachility of training workers in the absence of domain expertise. We show that having workers in the absence of domain expertise. We show that having workers that having workers that having show approximately for each present autocons the EXPLAN EXAMPLE AN largery. The results suggest that coversourced solutions of per versions may be harmosed in an automated training papeline.		It printent: The motivation and main findings of your research are well articulated. However, the impact of your research could be further emphasized. For instance, you could discuss how this research could revolutions: the way convolve/kms are trained, leading to more efficient and effective convolvement/or. Additionally, the methods used in your research						
Research Team	Evaluate your description and get some feedback from AI		are not learned, reading on those trademices according to consider the second second second your research are not clearly outlined. It would be helpful to provide a brief description of the five different training strategies you compared and how you conducted your experiments.						
UNPUBLISH PROJECT	Add more details in your current description								
	Restructure the description.		Your writing is clear and concise. However, using analogies or real-world examples could help make the concepts more relatable to the reader. For example, you could compare the process of training convolutions'to tolexing a cleas of students, where different teaching methods yield different results. Also, the relevance of your research to the real-world could be highlighted more. You could discuss how your findings could be applied in various industries that rely or conviducing.						
	Jargon Explanation DETECT FROM DESCRIPTION	cture: enucture of your writing is clear and easy to follow. However, dividing the text into sections with headings and esedings could improve readability. For instance, you could have separate sections for the motivation, main gs, impact, and methods of your research.							
	· ·		Normal ÷ B I U © 99 (E E E • % E T _k						
Identify jargon and p Provide jargons and clear explanations, which could help people understand the topic domain-specific exp			When is the study? We solve the revealed of the study of the study is complex considerating basis. We are particularly interested in training model workers to gatefine well on solveg tasks in situations where the space of strategines is large and workers need to discover and by different strateging to be accessful. Why is the important? In a lot of approximit, the particular of the different banking strateging is for complex web strates that banking when the the lot of approximit, the particular of the different banking strateging is for complex web strates challenges, we show that have the strates that the strates of the strateging strateging is the strate strateging is the strates of the strates that when the strates that the strateging strateging is the strateging is the strate strateging is the strate strateging is the strateging strates are strateging as the strates from the strateging is the strate strateging is the strate strateging is the strateging strates are strateging as the strates from the strateging is the strate strateging is the strateging strateging is the strateging strateging as a strateging experiment of the strateging is the strateging strateging is an associated strateging in associated strateging is the strateging strateging is the strateging strateging is the strateging strateging is the strateging strateging strateging strateging is the strateging strateging strateging strateging strateging strateging strateging is the strateging						
	Related Article		Findings or potential impact Wo show that having workers valiable the work of their peer workers can be even more effective than having them review expert examples if we only present solutions filtered by a threshold length. The results suggest that crowdsourced solutions of peer workers may be harnessed in an automated training peptine.						
	Enter DOI (e.g., 10.1145/2858036.2858268) or related work FROM DOI		What is the study method? In a first septement, we portrom a comparison of five different training strategies. In a second experiment, we study the feasibility of training workers in the absence of domain expertise.						
			Jargon Explanation DETECT FROM DESCRIPTION						
	Augor A	crowdworkers individuals who perform tasks or jobs through an online platform or crowdsource							
	Researcher Tip		crowdsourcing tasks tasks or projects that are outsourced to a large group of people, typically through						
		espert esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individuals with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrations provided by individual with specialized invanted esamples or demonstrating provided by individu							
	Provide tips to novices for finding better quality articles. e.g. any specific databases or venues for them to explore? Is		automated training pip a system or process that automatically trains and develops workers or individuals						
	there a specific year range to consider? Can website articles be included as well? etc.								

Figure 1: CrowdLit offers AI tools to help researchers make their project descriptions more accessible to the public.

set of domain-specific or advanced words along with definitions that would be shown to crowdworkers; researchers can choose to remove or modify any of these terms and definitions. The project creation interface is shown in Figure 1.

Stage Two: Crowdworker Sense-making At this stage, crowdworkers immerse themselves in the diverse research projects provided by the researchers. They can choose to work on the projects they are interested in. In CrowdLit, we present crowdworkers with the descriptions and terminologies uploaded by researchers (possibly with the help of AI). For jargon, they can hover over the terms to see the definitions.

Stage Three: Active Search Phase The crowdworkers then start their literature search. Using the guidance and key terms provided, they begin their search for relevant literature. They can also interact with other crowdworkers to learn from their searching strategies.

In CrowdLit, users can either manually upload papers or upload just the DOI and the platform will automatically fetch the article information using the CrossRef API. We also have brief tutorials that users can click on to get some guidance about how to conduct literature search (e.g., how to use Google Scholar or conduct keyword-based searches).

AI for g	group Task <u>Collected Literature</u>		
Upvotes 🛧	Comments Title	Year 🛧	Venue
16 0	Contextualizing AI Education for K-12 Students to Enhance Their Learning of AI	2021	Springer Link
心。	Using chatbot as a goal-setting facilitator in an online course: students' performance and teachar's nerresting (use) Comments on literature search strategies	2022	International Conference on Teaching, Assessment, and Learning for Engineering
心 0	from other crowdworkers. Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Metivation (Link)	2022	Journal of educational computing research
1 <u>6</u> 0	Piloting Antificial Intelligence (Al) to Facilitate Online Discussion in Large Online Classes: A Case Study (Link)	2022	Cases on Innovative and Successful Uses of Digital Resources for Online Learning
ić i C	Chatbots and Flipped Learning: Enhancing Student Engagement and Learning Outcomes port and Collaboration Link		Research Gate
а	ticles they find helpful.		Rows per page: 10 ∞ 1–10 of 30 < >

Figure 2: Review and upvote on literature retrieved by other crowdworkers.

Finally, users can also click on the Collected Literature tab at the top of the page. This will take them to another page, where they can review and upvote on literature retrieved by other crowdworkers (see Figure 2).

Stage Four: Feedback Loop and Refinement Researchers take an active role again during this stage. They evaluate the articles sourced by the crowdworkers, assessing them for relevance and quality. Constructive feedback can

						e Researcher3
AI for group	Edit Project Eval	uate Result				EXPORT
	Popularity \uparrow	Relevance	Title	Year 🛧	Citations	Venue
EVALUATED	3	0.64 Se	Using chatbot as a goal-setting facilitator in an online course: students' performance and teacher's perception (Link) mantic relevance score	2022		https://ieeexplore.ieee.org
evaluater	3	0.58 ber of til	Moderator Chatbot for Deliberative Discussion (Link) mes retrieved by crowdworkers	2021	Citat	Proc. ACM Hum. Comput. Interact.
EVALUATED	2	0.52	Conversation Technology With Micro-Learning: The Impact of Chattoch-Based Learning on Students' Learning Motivation and Performance Link	2020	48	Journal of Educational Computing Research
EVALUATED		0.37	Contextualizing AI Education for K-12 Students to Enhance Their Learning of AI Literacy Through Culturally Responsive Anonnacheo (194)	2021	Rows p	Springer Link wr.page: 10 = 1–10 of 25 < 1

Figure 3: Researchers evaluate and give feedback to crowd-workers.

then be provided to the crowdworkers about their selections and any identified gaps. This stage also sees researchers refining and adjusting their project information, in case the direction of their research shifts or becomes clearer.

In CrowdLit, researchers access literature gathered by crowdworkers. We use the Semantic Scholar API to retrieve paper details such as abstracts, summaries, venues, and citation counts. Through the Hugging Face API, we used the 'paraphrase-mpnet-base-v2' model to calculate the semantic relevance between the researchers' research questions and the paper titles and abstracts. This process could help researchers swiftly filter out irrelevant papers; see Figure 3.

Stage Five: Search Strategy Enhancement With feedback in hand, crowdworkers return to their literature search. They refine and optimize their approach, ensuring it's better aligned with the researcher's objectives. This iterative process promotes continuous improvement and could result in a more tailored set of literature for the researchers.

Feasibility Study

We have conducted a study to evaluate the feasibility of CLS pipline as well as our platform. Below we discuss the study design and preliminary results.

Study Design

We recruited researchers from a mailing list of faculty members at a research university, a Slack channel of researchers interested in literature search, and others in the authors' personal networks. We recruited crowdworkers on Prolific. During our study, each researcher posted 1-2 of their research projects that need literature collection on our platform, after which we tasked crowdworkers with the literature search. Over a 5-day period, both groups engaged with the web system. Crowdworkers were instructed to discover more articles, engage with interactive tutorials on literature search, and review feedback from both researchers and their peers, while researchers provided feedback, evaluated the retreived literautre and refined their project descriptions. We interviewed researchers immediately after they created their project submissions and at the end of the study. We also distributed questionnaires to crowdworkers, capturing their impressions both immediately after their initial system interaction and after their 5-day experience.

Study Results

We collected data from 10 researchers and 55 crowdworkers. Among the crowdworkers, 36 completed the first questionnaire while 19 completed both questionnaires and did work over the 5-day period. Through thematic analysis of the interviews and open-ended questions, we found that:

- 8 out of 10 researchers thought the platform was useful and expressed a desire to use it in the future. They thought it would be useful for gaining a broader understanding of a field through diverse literature, explore existing studies for early-stage projects, and help crowdworkers enhance their research and searching skills.
- 50 out of 55 crowdworkers felt they gained from the experience, including improving their skills in searching and evaluating literature and understanding progress of state-of-the-art research. Three participants noted gaining insights into the nature of research.
- In terms of collaboration between researchers and crowd-workers, all 19 crowdworkers who completed the second questionnaire felt they were at least slightly involved in the research, with 14 of them feeling they were at least moderately involved. Most crowdworkers found researcher feedback helpful and many changed their methods as a result of it. Conversely, researchers felt a more limited sense of collaboration. One advocated for extended interaction, while two wished for direct conversations with crowdworkers to better understand their viewpoints, backgrounds, and interests.

Conclusion and Future Work

In this work, we introduced CrowdLit, a web-based platform that implements the CLS pipeline which is designed to engage crowdworkers in collaborative literature discovery alongside researchers. As a demo, we plan to have researchers from the HCOMP community try CrowdLit and give feedback on its various features. They can either act as researchers and try our project creation interface (including tools to help them with scientific communication) or act as citizen scientists who try searching for literature on other researchers' projects. In future work, we aim to test CrowdLit in educational settings with high school and undergraduate students to assess its potential for improving search skills and fostering STEM and research interests.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. (2033868). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Bornmann, L.; Haunschild, R.; and Mutz, R. 2021. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*, 8(1): 1–15.

Choe, K.; Jung, S.; Park, S.; Hong, H.; and Seo, J. 2021. Papers101: Supporting the discovery process in the literature review workflow for novice researchers. In 2021 IEEE 14th Pacific Visualization Symposium (PacificVis), 176–180. IEEE.

Kang, H. B.; Soliman, N.; Latzke, M.; Chang, J. C.; and Bragg, J. 2023. ComLittee: Literature Discovery with Personal Elected Author Committees. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–20.

Swanson, D. R. 1986. Undiscovered public knowledge. *The Library Quarterly*, 56(2): 103–118.