Online Q&A in STEM Education: Curating the Wisdom of the Crowd

Project

With the constantly changing landscape of online information, specifically generated by non-experts, there is a dire need to guide and instruct students about evaluating and using such information. This need is even more prominent in STEM (Science, Technology, Engineering, and Math) education where assessment of sources for factual information, blending various scientific data and methods while evaluating the source and applicability of each, and using them in addressing real life problems is paramount. The proposed project will integrate crowdsourced information (user-generated in online environments) and the wisdom of librarians and other information professionals to enhance the experience of STEM learners.

The Association of College and Research Libraries (ACRL) *Information Literacy Standards* Task Force (ACRL, 2012) recommended that the next iteration of the standards must acknowledge the role of students as content creators and curators, and acknowledge the value of information in diverse formats. Online Q&A has become an increasingly popular method for students to seek, create, and curate information (Shah et al., 2009; Gazan, 2011). While there is a clear need to help STEM learners assess the quality of their sources, it is also possible that students may be discovering valuable content from online Q&A sites that is appropriate to their learning objectives, yet not acknowledged by instructors or Virtual Reference (VR) librarians due to its nontraditional nature. The sheer number of students' online Q&A citations (Turnitin, 2012) makes both quality assessment of Q&A content and a more nuanced understanding of the role of online Q&A services in STEM education worth investigating. At the same time, incorporating crowdsourced information into education must be done with caution. Library practitioners and scholars are in a unique position: both are accustomed to evaluating socially constructed information, but also understand how specific communities define quality information in diverse ways. The purpose of this project, similarly, sits at a unique intersection of LIS, education, and social sciences. It investigates the social phenomenon of people constructing and seeking information through online Q&A; uses theories and expertise from LIS to evaluate and curate this information; and applies the resulting knowledge to STEM education by augmenting students' learning experiences with curated wisdom of the crowd.

The PIs propose to leverage and investigate students' demonstrated use of social Q&A (SQA) and community-based Q&A (CQA) services in learning environments, with the goal of identifying areas of potential crossover that would result in better library and educational services. Specifically, the proposed project will: (1) investigate the use of online Q&A services and develop a theoretical framework that contextualizes them within the broader domain of information behavior; (2) develop methods for evaluating the relevance, quality, and impact of content created and shared in Q&A environments; and (3) test the outcomes of the first two stages in STEM education environments via a formative design-evaluation cycle.

Key Personnel

PI: Chirag Shah, Assistant Professor, Rutgers University; PI: Rich Gazan, Associate Professor, University of Hawaii; Internal Advisory Board: Nick Belkin, Marie Radford, Dan O'Connor, Michael Lesk from Rutgers University; External Advisory Board: Lynn Connaway, OCLC, Stephen Freeland, University of Maryland Baltimore County, Wayde Oshiro, Leeward Community College, Joel Downs, CBS Interactive; Industry Partner: Brainly.com – an online Q&A service for students and educators.

Goals and Outcomes

The relative pros and cons of expert and peer-based Q&A services open up new opportunities to create hybrid solutions. The key to creating such solutions for STEM education is to: (1) gain knowledge about assessing user-generated content in online Q&A (**Goal-1**); and (2) use that knowledge to augment/support STEM students' online information seeking (**Goal-2**). The proposed project will meet goal #1 with the help of

experts (librarians), and #2 by developing a new tool for testing and deployment in libraries and STEM classrooms. Specifically**,** the project will address the following research questions (RQs).

RQ1: How can online Q&A activities be explained and investigated as information seeking behavior?

RQ2: What are appropriate quality metrics for online Q&A content, specifically related to STEM?

RQ3: How would the addition of an online Q&A quality assessment process help meet the rigorous requirements of STEM education?

The proposed work to address these RQs will produce the following **outcomes**.

1. A new theoretical framework to understand and explain online Q&A behavior;
2. An extension to existing models of information seeking and behavior;
3. A methodology to assess content quality in online environments; and
4. A method and a tool to incorporate curated Q&A content into STEM education.

Impact

The proposed project will add substantially to our understanding of people's, especially that of STEM students', information-seeking practices in online environments through asking and answering questions. This new knowledge will be grounded by empirical data about assessing content quality from diverse sources and embedded in an important information-seeking domain: STEM education.

Through an integrated research and education plan, this research program will contribute directly and significantly to support for online content assessment in higher education. Because the PIs will be testing and evaluating the proposed tool specifically in classroom and library instruction environments, that tool, released as open-source software as part of the project's dissemination plan, will be ready to be used and further developed by others in higher education.

Finally, the project will provide a valuable contribution to our national goal of encouraging and better supporting STEM education, and making curated information more accessible and applicable to larger audiences. **This also addresses IMLS's strategic goal #1 of providing engaging and empowering learning experiences.** The proposed project will help connect people with professionals, search experts with domain experts, and information seekers with information facilitators in an effective and meaningful way: promoting and supporting crowdsourced contributions, assessed and curated by experts.

Work Plan

The work will be carried out by the means of three user studies: (1) ***Study-1 to address RQ-1 (year-1)****:* search and browsing log data collection and interviews with 40 students of online Q&A services to investigate their motivations and expectations for using such services as well as to obtain an understanding of how they assess and utilize user-generated content; (2) ***Study-2 to address RQ-2 (year-2)****:* collection of objective assessments with the help of the crowd (300 novice assessors using crowdsourcing service Mechanical Turk) and subjective assessments with the help of experts (10 librarians) to build statistical models and tools to evaluate social media and crowdsourced content; and (3) ***Study-3 to address RQ-3 (year-3)****:* a lab study with 40 STEM students and a classroom study with 300 STEM students to evaluate the efficacy of developed tools and services for automatic content assessment that augment students' learning processes.

Budget

The proposed project will span three years, from 10/1/2016 until 9/30/2019. The total estimate of costs related to personnel, participant support, and project management is $654,139, of which $204,837 **(more than 31%) will be costl shared** by the participating institutions. The direct cost ($449,302) to IMLS will include tuition reimbursement ($77,784) and fellowship stipend ($168,195) for two PhD students; faculty/ staff support ($79,086); human subject payments ($4,400); conference travels ($23,386); consultancy fees ($3,000); fringe benefits ($21,304), and indirect expenses ($72,147).