Proposal Status | MAIN

Organization: University of Alabama Tuscaloosa

Review #1

Proposal Number:	1629871
NSF Program:	COMPUTING RES INFRASTRUCTURE
Principal Investigator:	Carver, Jeffrey C
Proposal Title:	CINew: Advanced Systematic Literature Review Authoring Infrastructure for Software Engineering
Rating:	Multiple Rating: (Good/Fair)

REVIEW:

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to intellectual merit.

The proposed infrastructure, SAInT, provides support for researchers performing systematic literature reviews (SLRs) in software engineering. This includes mechanisms to help researchers define SLR protocols, search multiple databases for relevant articles, identify duplicates, select articles for inclusion in the study, extract data from the articles, and support collaboration, among others.

Strengths and Weaknesses:
+ SLRs are useful and are becoming increasingly popular in software engineering. As the investigators point out, performing a careful literature review is time consuming, so automation can be helpful. With support of a CRI-P grant, the investigators have consulted with many researchers who have written SLRs and have identified many of the challenging aspects of such work.
- The proposed SAInT architecture has nine components. While some of these, such as facilities to search multiple databases, are fairly straight-forward, others, especially quality assessment, automated text mining, synthesis, and qualitative/quantitative report generation, seem extremely challenging. The proposal does not give a clear sense of how these will be addressed. Are there some particular natural language processing techniques that will be used? How mature are they and what needs to be done to train them for this domain?
Another problem with the plan (though not as serious as the above issue) is the assumption that all relevant literature will be available for search; unfortunately some software engineering literature is protected by copyrights that could preclude inclusion in the repository.
+/- If successful, this infrastructure would be useful, both to researchers preparing SLRs and to researchers, students, and software engineers who'd like to read SLRs. However, I do not think the work will be transformative.
- The proposal lays out a careful plan, but, as noted above, it seems overly ambitious.
+/- The team is well qualified in terms of understanding the challenges of SLRs. If they really plan to do things like text mining and summarization, they should bring in some researchers with Natural Language Processing expertise.
+/- The university seems to have good general purpose computing infrastructure, including large servers. It is not clear whether this will be enough to host the project or whether additional servers or cloud resources are needed.

In the context of the five review elements, please
evaluate the strengths and weaknesses of the proposal with respect to broader impacts.

Systematic Literature Reviews have proved valuable in other fields, notably medicine, and are becoming more popular in software engineering. If successful, this project can make it easier for researchers to produce SLRs and thereby make it easier for researchers, students, and practitioners to keep up with the most important developments and trends in the literature.
It is possible that the infrastructure would be useful for researchers producing/consuming SLRs in other fields, but the proposal does not clearly articulate which aspects of SAInT will be specific to software engineering SLRs and which can be generalized.

Please evaluate the strengths and
weaknesses of the proposal with respect to any additional solicitation-specific review criteria, if
applicable

+/- If development of the infrastructure is successful, it will facilitate literature reviews, which, in turn, can help target research to the right problems; however the infrastructure will not directly support new research
+/- The infrastructure is targeted to software engineering literature reviews, which is well within the core of CISE; however, although the proposal focuses on SE SLRs, it is not clear whether the fundamental issues that SAInT

addresses are specific to software engineering or computer science. Generality could be a strength, but it probably would make the core natural language processing tasks even harder.
+ The investigators have done a very good job of soliciting input from the SE SLR community on the key challenges that make producing SLRs hard. They also have a good plan for involving the community in development of the tool through two workshops and early release of parts of the system to beta testers.
+ The proposal includes letters of support from researchers at 10 universities and a statement of support signed by representatives of 9 other organizations.
+ The proposal discusses several related SLR tools and shows how SAInT expands on their capabilities and/or integrates them.
+/- Although there is a detailed management plan, as discussed above, I think it underestimates the technical difficulty of the work. The proposal does not discuss the issue of how to sustain the infrastructure after the NSF funding is used up.

Summary Statement

The proposed infrastructure for facilitating systematic literature reviews could, if successful, make it easier to conduct such reviews and thereby help provide useful resources to researchers, students, and software engineers. However, some of the hardest aspects of reviewing literature involve text analysis, understanding, and summarization tasks that are very difficult.

Back to Proposal Status Detail

National Science Foundation
4201 Wilson Boulevard, Arlington, Virginia 22230, USA
Tel: 703-292-5111, FIRS: 800-877-8339 | TDD: 703-292-5090