CiteAs: Improving the Visibility of Research Software for Better Scientific Collaboration

In this talk, we are presenting CiteAs, a search tool for retrieving citation information about research software and engaging potential collaborators around them. In today's computing-intensive research fields, the practice of developing software is inseparable from the conduct of research. Models and methodologies are often implemented in software, but their scientific value has yet to be well recognized. Thus, many scientific software contributors are seeking to implement software citation, so that the scientific contribution of software work can leave its traces in the record of scientific progress.

However, software citation is more than crediting scientific software work. It reveals more information about the research process and gives pointers to other scholars to identify and access the software. In this way, the reproducibility of research is improved. Other researchers thus can get a deeper understanding of the procedures and methodologies involved in the research software. The access and shared understanding offer an entry point for further engagement with others' software work, to wit, a pathway to scientific collaboration around the software.

Among all the recommendations for constructing an appropriate software citation, including those determined by journal and conference policies and advocacy groups, we think it is important to link the software authors' citation request to the end-users' need for identifying and accessing the software. The reasons are twofold: first, this ensures the credit is given to the right people; second, this way builds a link between the software authors and the potential contributors to the software product. While this path is often absent, making it visible can get more potential team members on board for building upon the software.

Thus, we have developed CiteAs, a specialized search engine matching software authors' citation requests to end-users' information needs for the software. Taken any input about the targeted software, be it a digital identifier, or simply the name of the software, CiteAs will spider the web and return a recommended citation and the traces it finds about the software being searched: the software metadata, its project webpage, or repository, etc. By demonstrating this provenance for the citation information, CiteAs points end-users to the access to the software, also prompts software authors to share more available information about the software. As a two-sided system, the design of CiteAs aims at engaging both software authors and end-users towards more actionable transparency of the software work, and thus more possibility for assembling a distributed team building upon it. Currently, CiteAs is in its "soft launch"; and we are continuously improving it and adding more intelligent features for engaging more the potential collaborators around the scientific software products.