BioSoDa Bioinformatics Software Database

ABSTRACT

When trying to solve bioinformatics problems, an early and critical decision is whether to use existing software or to develop new software. Identification of suitable existing software is complicated by the lack of a centralized repository or registry of bioinformatics software and consequently many projects waste resources by replicating existing software. The proposed project aims to address this problem by developing a curated bioinformatics software database (BioSoDa) targeted to the needs of the bioinformatics community.

The primary existing sources for information about available bioinformatics software are journal articles and website. Searching either of these sources is problematic for a number of reasons:

- The searcher is relying on text searches which require that the searcher predict appropriate words or terms that will be used in the article or webpage
- Software packages often have multiple functions so it might be easy to predict search terms based on the primary function but the minor functions might be discussed insufficient detail to enable those functions to be picked up via a text search
- Not all relevant journals are available online
- Many journals are only available on a subscription basis so they are only searchable via search engines based at the journal publishers’ website and not via common search engines such as Google
- Even if a journal is available online and has somehow been indexed by a search engine, the content is usually only available via subscription in which case it is still inaccessible to many researchers.

A curated registry would address many of these problems.

Prior to this internship, initial design of the user interface and database schema has been completed. The goal for this internship will be to complete Release Version 1 of this project.