Crowdnotation: A Crowdsourcing Annotation Tool for Genomics Studies

Diogo Matos da Silva[1], Helder Takashi Imoto Nakaya[1]

1 USP

Abstract

When authors deposit a microarray study into GEO database, they are free to describe the experiments using their own words. This author-based annotation method follows no defined ontology or classification standards, creating enormously difficulty for others when querying studies based on specific key words. Only human manual curation can properly annotate studies, as several computer programs have already attempted and failed at this task. We are developing a web-based tool named Crowdnotation, where students can remotely annotate the studies. In return, any student participating in the annotation will be included as an author in our publications. We believe that this alone serves as a strong incentive for a significant number of students. More importantly, to make the annotation process more attractive and enjoyable, Crowdnotation will have several gamification features, such as scores, rankings, friends and badges. This will promote engagement and ultimately improves performance and efficiency. Annotation accuracy will be ensured through consistency among peers. Students will learn about the different types of experimental designs and will gain significant knowledge in the field of modern molecular biology. We expect this worldwide web-based community strategy will create a massive collaborative network of students working towards a common goal. In addition, students from remote areas of developing countries will gain a valuable opportunity to be part of a scientific collaboration and subsequent publication. In the future, this tool may be easily adapted to answer various scientific questions and for that, we will have the contact information of several thousand future scientists. Finally, all annotation data will be organized into a structured and open-accessed database, and the codes for Crowdnotation will be freely available for future developers.