# $25,000 BANNER Challenge

The millions of biomedical publications that exist are a valuable, but cumbersome resource. Identifying those documents that are relevant to particular genes, proteins, diseases, drugs, or organisms is currently a costly, human intensive process. The Scripps Research Institute, together with the NASA Tournament Lab at Harvard, is launching a contest to develop new algorithms to aid in the automated named entity recognition (NER) of biomedical objects.

The goal of this contest is to develop an algorithm to generate keywords based on the natural language text of publication abstracts. A set of manually annotated abstracts has been created to aid in algorithm development and scoring.

Full problem details can be found at XXX-URL-XXX.

To participate, entrants are required to register for the contest between XXXXX and XXXX.

# Contest Format

A total of $25,000 prizes will be distributed in the form of room prizes to the winners in virtual competition rooms. All registered participants will be randomly divided into competition rooms of about XXX participants. There are three types of rooms:

• The race. The winners are the first two coders to submit a solution with score greater than XXX. No prizes will be awarded if no submissions reaches a score of at least XXX.

• The tournament. The winners are the two coders with the highest score at the end of the submission period.

• The tournament with a reserve. Only submissions that score at least XXX are eligible for prizes. The winners are the two coders with the highest score at the end of the submission period, given that the score is at least XXX. No prizes will be awarded if no submissions reaches a score of at least XXX.

This is an experimental contest format and thus will be an unrated event.

As part of the experimental contest format, participating in this contest also involves participating in a research project. Your participation is voluntary. If you choose to participate, you will need to complete several surveys including a registration survey and a final survey. You may discontinue your participation at any time. Please see below for more information about how your survey responses will be used.

The data collected, including survey responses, analysis of code submissions, and communications on the public forum of the challenge, will be used for research purposes. The substance of your survey responses will not affect your eligibility for winning a prize in this or future contests.

Only the immediate project team at TopCoder and researchers at Harvard University will see your individual data. Data will only be shared in an anonymous form in which individuals cannot be identified.

If you have any questions about the use of your information, please contact: Andrea Blasco (Ablasco@fas.harvard.edu). Harvard University has a Standing Committee on the Use of Human Subjects in Research (CUHS) to which complaints or problems concerning any research project may, and should, be reported if they arise.