

## Correlation between Funding and Publications: A Pre-Step towards Future Research Prediction

**Authors :** Ning Kang, Marius Doornenbal

**Abstract :** Funding is a very important - if not crucial - resource for research projects. Usually, funding organizations will publish a description of the funded research to describe the scope of the funding award. Logically, we would expect research outcomes to align with this funding award. For that reason, we might be able to predict future research topics based on present funding award data. That said, it remains to be shown if and how future research topics can be predicted by using the funding information. In this paper, we extract funding project information and their generated paper abstracts from the Gateway to Research database as a group, and use the papers from the same domains and publication years in the Scopus database as a baseline comparison group. We annotate both the project awards and the papers resulting from the funded projects with linguistic features (noun phrases), and then calculate tf-idf and cosine similarity between these two set of features. We show that the cosine similarity between the project-generated papers group is bigger than the project-baseline group, and also that these two groups of similarities are significantly different. Based on this result, we conclude that the funding information actually correlates with the content of future research output for the funded project on the topical level. How funding really changes the course of science or of scientific careers remains an elusive question.

**Keywords :** natural language processing, noun phrase, tf-idf, cosine similarity

**Conference Title :** ICCSIB 2016 : International Conference on Cybermetrics, Scientometrics, Informetrics and Bibliometrics

**Conference Location :** Barcelona, Spain

**Conference Dates :** October 03-04, 2016