Finding and Learning From Ethical Breaches By Scientists

Harold "Skip" Garner Jr. and members of his laboratory

Medical Informatics, Translation, Training and Ethics (MITTE) Office, Virginia Tech, Blacksburg, Va.

Progress in science depends on productive, reproducible and especially accurate research. However, the research, its reporting and funding is not always of the highest ethical caliber. Why? Because scientists are human, too. They are susceptible to pressures like any other profession, the pressure to win grants and produce manuscripts which are then reported on their resume, and play an important role in their professional advancement. How can this behavior best be monitored, so that we can devise ways to encourage all scientists to adhere to the highest ethical standards? One approach, our approach is to employ a type of text analytics, text similarity searching, to detect and analyze professional publication plagiarism and duplication; and it can also be employed to detect and understand other types of unethical behavior, including grant and contract fraud. During this presentation I will show how the technology detects potentially unethical text (publications and grant submissions that have inappropriate high similarity). I will then discuss how this behavior is quantified as well as a sociological analysis of the impact of this behavior on all the stakeholders (authors, editors, granting agencies, and we, the rest of the scientific community and the public at large). I will also show how other types of unethical behaviors are often coupled, including falsification of data, fabrication, and improper authorship. The most important take away message will be that our electronic society makes it possible to continuously check the literature and other scientific writing (grant applications) for ethical issues, so it is important to master correct scientific behavior, for at some time problems will be discovered.

Keywords: plagiarism detection, ethical violations, text analytics