

Unmasking Deception: The Fight Against Fake Research Papers

Fake research publications can be deadly, as shown by the many who believed Andrew Wakefield's fraudulent claim that a measles vaccine causes autism, which led to vaccine refusal, measles outbreaks, and numerous deaths worldwide. In the 25 years following the retraction of Wakefield's work, paper mills have produced over 10,000 fraudulent journal articles, causing a spike in retractions, and eroding trust in research. Our project aims to support the scientific community's quality assurance efforts in combatting paper mills, which are profit-driven and potentially illegal organizations that generate fake manuscripts resembling genuine research. In the first phase, we will employ the Delphi technique by inviting 30 experts, including 10 representatives each from the following fields (1) journal editors, (2) publishing managers, and (3) academics researching peer review processes or scientific publishing. They will be asked to share their expert opinions on good practices for detecting and preventing paper mills' fake manuscripts in the peer review process, and the challenges that make it difficult to differentiate fake papers from genuine ones. This collaborative effort aims to design a framework that enhances the current peer review process in detecting and preventing paper mill infiltration (RO1). In the second phase, we will build on routine activity theory to create and test a theoretical model of key identifiers of fake papers (examining attributes include authors, journal, reviewer and editor, and paper content) and model its diffusion pattern (RO2). We will utilise all retracted paper mill articles available in the Retraction Watch Database from inception to present as "fake papers" and match with "genuine papers" published in Q1 journals across various fields to test our model (n=20,450), analysing these data with computational linguistics, advanced statistics, and three machine learning models. Our pilot test (n=300) confirmed significant differences between genuine and fake papers among some attributes. Armed with an evidence-based framework, theory and vast data, we will develop an AI screening tool to detect fake research to aid researchers, editors, reviewers, publishers, and general public. Removing fraudulent studies will enable researchers to base their work more on genuine and reliable research. Journal editors and reviewers can save time and effort by easily removing submissions from paper mills. Publishers can maintain the quality and prestige of their journals, allowing the scientific community to rebuild credibility and disseminate extant knowledge, rather than misinformation to the research community and wider public.