ReTracker: an Open-source Plugin for Automated and Standardized Tracking of Retracted Scholarly Publications

Ly Dinh, Yi-Yun Cheng, Nikolaus Parulian
School of Information Sciences
University of Illinois at Urbana-Champaign
{dinh4;yiyunyc2;nnp2}@illinois.edu

ABSTRACT

In this paper we present *ReTracker*, an open-source plugin for tracking retracted scientific journal articles. *ReTracker* helps Zotero users detect any papers in their current libraries that were retracted on PubMed. At its first release, *ReTracker* plugin adds metadata for each entry on its "retraction status" as the plugin performs local cache checking and sends a query to the tool's database. If there is a match in the paper's full title between the query and *ReTracker*'s database, the retraction status metadata will be updated. We believe that *ReTracker*, as an open-source, free tool, is a necessary solution that is distinctive from existing tools (e.g. Open Retractions, CrossMark, etc.), in that it enables researchers to manage their existing citation libraries without extra efforts to repetitively search for and manually flag retracted papers from external database(s). With future releases, the tool aims to expand its functionalities in terms of standardizing retraction tags, adding retraction data for other item types, and providing pop-up notices for web browser plugins.

KEYWORDS

Scientific Retraction; Scholarly Communication; Metadata Improvement; Publishing Ethics

1 INTRODUCTION

With a considerable rise in the amount of journal articles and conference proceedings retracted in the past few years, research on scientific retractions becomes increasingly necessary. There is a myriad of reasons for a paper to be retracted, ranging from errors or study design to deliberate misconduct such as duplication of data, plagiarism of text, or fabrication of results. While progress in scientific reproducibility has been made to ensure these errors and misconducts are minimized, issues arise when retracted papers are still being cited and recognized as valid and reliable research [3]. At times, retracted works are even cited at a higher rate than before it was retracted (see [4]). Several studies have shown that there is yet to be a unified way to inform researchers of retracted papers [1,8].

Therefore, we sought to resolve these inconsistencies with a single unified tool *ReTracker* that can provide users with comprehensive and reliable information on any retracted papers from PubMed. The goal for this tool is not to point fingers at problematic research, but to encourage greater transparency and honest conversations around the types of errors and issues that were previously encountered.

2 RELATED WORKS

2.1 The Retraction Issue

Retracted papers are sometimes being cited long after retraction because scholars who cite these papers were not aware of the retraction. This is a recurring issue as flawed and erroneous works continue to be circulated as reliable prior works and at times directly used to validate future study methods and findings [1].

As a remedy to such problematic citations, Committee on Publication Ethics (COPE) [2] released guidelines for publishers and editors on the retraction process, but not on the citation format of retracted papers, nor on what to do when retracted papers continue to be cited [1, 9]. Moreover, the format for retraction notices should be standardized across platforms [1]. As of now, only PubMed and Scopus include retracted papers as a separate search category (e.g. labeled as "Retraction of Publication" or "Erratum"). On Google Scholar and Web of Science, users must manually find titles that contain "Retracted Article", "Article Retracted", or "Retraction Notice".

2.2 Current Tools for Identifying Retractions

A number of scholars have observed the growing need to have one database for "classifying retractions" [5, 8]. To meet such demand, resources such as Retraction Watch, Open Retractions, Crossmark, and Office of Research Integrity (ORI) were created. Each of these resources have their own database for searching retracted papers, either through the paper's DOI (Open Retractions), a paper's title (Retraction Watch), or by first author (ORI's case summaries), which may result in "inconsistent representations of retractions" [7] across these different databases.

Figure 1: Zotero Library interface with *ReTracker* metadata

978-1-7281-1547-4/19/$31.00 ©2019 IEEE
DOI 10.1109/JCDL.2019.00092

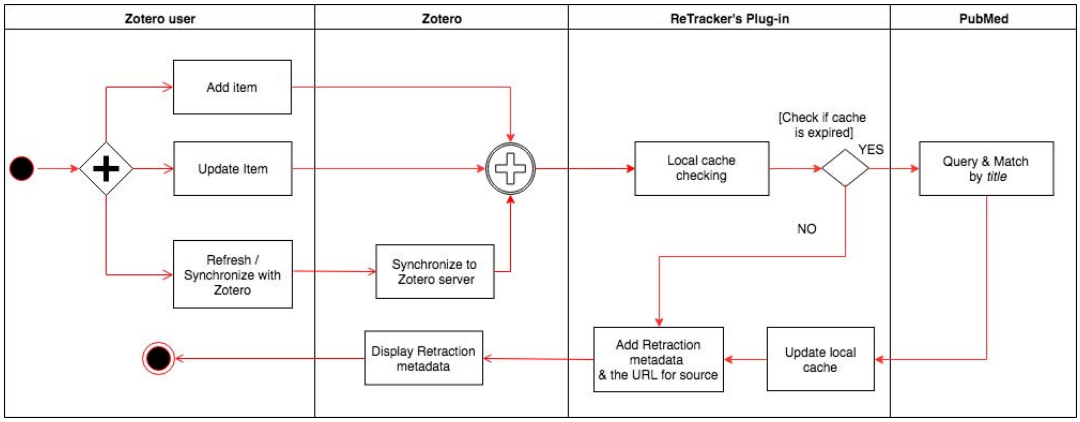

Figure 2: Detailed events for ReTracker

3 RETRACKER: AN UPDATEDSOLUTION

Recognizing all the prior efforts in identifying and tracking retractions, we propose a novel tool that checks within a researcher's current citation library for updates and notifies the researcher if there is a retracted paper in his/her library. Our tool, *ReTracker* is designed as a Zotero plugin, instead of a standalone widget, so that researchers can manage their existing citation libraries without extra efforts to repetitively search for and manually flag retracted papers from external databases.

ReTracker sends online query to PubMed to match and retrieve retraction metadata. We believe this model is more sustainable and automated in long terms. Target users of *ReTracker* are those who actively use Zotero and save articles in their personal Zotero libraries. Current release of the tool and instructions on how to install *ReTracker* can be found on [6].

3.1 Main Features

Currently, *ReTracker* has two main features: (1) Updating retracted metadata. Whenever a user adds or updates an item, or refresh their Zotero library (ZL), it will trigger the checking of PubMed. Then the metadata in user's ZL will be updated with "this article has been retracted" if an retracted paper is found; (2) Syncing Retraction Notices. Articles that already contain a retraction notice (for example, the title of the article usually will begin with "Retracted", "Retract", or similar) and are downloaded to Zotero library will also be flagged with the retracted metadata as well.

3.2 Design

ReTracker's design consists of two parts: (1) the plugin that can be installed in Zotero; (2) the query request that *ReTracker* sends to PubMed to match the retraction information. Core events are described in Figure 2:

(1) Whenever a user adds or updates an item, or refresh their Zotero Library (ZL), the checking of user's local cache will be triggered.

(2) The triggered events then will enforce local cache checking from the *ReTracker* plug-in. Then a query will be send to PubMed to seek for exact matches of retracted articles if the local cache is expired.

(3) To ensure that the *ReTracker* plugin does not overload requests to PubMed (e.g. by checking the titles too frequently), we store the result in the local cache with an expiration date.

(4) If exact matches for retracted articles have been found, then the metadata on user's ZL will be updated to include a noticeable label "This document has been Retracted".

4 FUTURE WORK

ReTracker is our first attempt to standardize retraction notices via Zotero plugin and libraries. Our tool has shown promising results to display retraction metadata actively to users. Our next steps are to provide a more holistic matching mechanism (e.g. matching titles+DOIs) that can fetch retracted articles with higher precision. We also plan to work more in depth with Zotero's full functionalities such as standardize retraction tags, develop full capabilities with other item types (such as books), and provide pop-up notices for web browser plugins. Ultimately, *ReTracker* is not a panacea for solving the issues with retraction, but rather, a way to unify incongruences and make errors more transparent.

REFERENCES

[1] Judit Bar-Ilan and Gali Halevi. 2017. Post retraction citations in context: A case study. *Scientometrics* 113, 1 (2017), 547–565.

[2] COPE. 2015. COPE-Committee on Publication Ethics flowchart. (2015).

[3] Peng Dong, Marie Loh, and Adrian Mondry. 2005. The" impact factor" revisited. *Biomedical digital libraries* 2, 1 (2005), 7.

[4] Atsunori Fukuhara, Morihiro Matsuda, Masako Nishizawa, Katsumori Segawa, Masaki Tanaka, Kae Kishimoto, Yasushi Matsuki, Mirei Murakami, Tomoko Ichisaka, Hiroko Murakami, et al. 2005. [RETRACTED] Visfatin: a protein secreted by visceral fat that mimics the effects of insulin. *Science* 307, 5708 (2005), 426–430.

[5] Felicitas Hesselmann, Verena Graf, Marion Schmidt, and Martin Reinhart. 2017. The visibility of scientific misconduct: A review of the literature on retracted journal articles. *Current sociology* 65, 6 (2017), 814–845.

[6] Nikolaus Parulian, Yi-Yun Cheng, and Ly Dinh. [n. d.]. ReTracker source code. https://github.com/nikolausn/ReTrackers.

[7] Amy Riegelman and Caitlin Bakker. 2018. Understanding the complexities of retractions: Recommended resources. *College & Research Libraries News* 79, 1 (2018), 38.

[8] Richard Van Noorden. 2011. Science publishing: The trouble with retractions. *Nature News* 478, 7367 (2011), 26–28.

[9] Elizabeth Wager, Virginia Barbour, Steven Yentis, Sabine Kleinert, et al. 2009. Retractions: guidance from the Committee on Publication Ethics (COPE). *Maturitas* 64, 4 (2009), 201–203.