Open science platforms fighting clandestine abuses of piracy and phishing: The Open Science Framework Case

Ayumi Ikeda[1,2*], Fumiya Yonemitsu[1,3*], Naoto Yoshimura[1,4], Kyoshiro Sasaki[5], and Yuki Yamada[2]

[1]Japan Society for the Promotion of Science
[2]Kyushu University
[3]Chuo University
[4]Ritsumeikan University
[5]Kansai University

*Corresponding authors
Ayumi Ikeda
Address: Institute for the Promotion of Business-Regional Collaboration, Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo, Hokkaido, 060-0812, Japan
E-mail: aikeda4444@gmail.com TEL
& FAX: +81-92-802-5250

Fumiya Yonemitsu
Address: Faculty of Letters, Chuo University, 742-1, Higashinakano, Hachioji-shi, Tokyo, 192-0393, Japan.
E-mail: y.fumiya.0408@gmail.com
TEL & FAX: +81-42-674-2624

Abstract

The Open Science Framework (OSF) is an important and useful platform for researchers to practice open science. However, OSF has recently been misused for criminal purposes, especially on search boards for watching pirated copyright works, leading to phishing sites. This misuse can negatively influence the OSF server function; therefore, it is important to take appropriate measures. To protect the sound base of open science in the future, this paper reports cases where OSF has been abused for illegal activities and discusses various measures, including those already implement by OSF management.

Keywords: open science, metascience, Open Science Framework (OSF), preregistration, open repositories

Open Science Framework

The Open Science Framework (OSF) is a free and open-source project management platform that supports research projects (Elliott et al., 2021). OSF is the flagship product of the Center for Open Science, Inc. (COS), whose mission is to increase the openness, integrity, and reproducibility of scientific research. This platform is useful for researchers to share data and materials and pre-register their research protocols. Indeed, OSF has played a key role in projects on the reproducibility of psychological research (e.g., Open Science Collaboration, 2015). At present, researchers in various research fields and psychologists use OSF to manage their research projects and open collaboration. Thus, the OSF is vitally necessary for good research practices; unfortunately, we report cases where the OSF has been abused for illegal activities. Note that we reported this issue to the COS before making this document public, and according to the COS, they were already aware of this issue and were taking swift and concrete actions to address it. The following is a summary of what has happened up to that point we discovered this incident; we present a necessary discussion to prevent similar crimes from being committed in the future, regardless of OSF involvement.

Accessing illegal copies of copyrighted works by abusing the OSF site

One of the authors was searching the name of the last author "Yuki Yamada" on the OSF site for purely academic purposes and coincidently encountered a movie production starring a popular Japanese actor of the same name. It was created as an individual project page with embedded links to illegally uploaded websites. We were surprised when we found that page, but we suspected that this was not the only such case. Thus, we began to search for other works illegally shared through the OSF site. Several such project pages were identified, and the details of the common methods employed are described below.

First, a person illegally uploads copyrighted works (e.g., commercial films) using social media as a storage site. Then, they set up a project on the OSF site and make it public with links to the illegally uploaded works in the "Description" section of the project. In other words, the OSF platform is abused as a search board to watch pirated copies, called a "leech site." For example, a search for "tumblr.com" in the OSF revealed that on January 11, 2022, when we first became aware of this problem, there were 324 projects and 140 users (Fig 1). One of them was the project name "ดูหนัง ▷ Spider-Man: No Way Home【2021】เต็มี เรื่อง HD พากย์ไทย THAI" (it results in "Watch Movie ▷ Spider-Man: No Way Home [2021] Full HD Thai Dub Thai" when google translated into English). The description in this project included a link to the illegally uploaded movie "Spider-Man: No Way Home," which is a popular movie of an American comic hero (Fig 2). This was just one example, but there were many other projects that had links to famous movies and paid content.

Moreover, a search on "tumblr.com" in the OSF showed that as of September 27, 2022, the number of projects had decreased to 263. By contrast, the number of users increased to 380. In addition, a search on "tumblr.com" in the OSF showed that as of May 28, 2023, the number of projects had decreased to 94. By contrast, the number of users increased to 394. Unfortunately, all 94 projects were abused in such a way that the project name contained the name of an anime, game title, or application, and the project had direct links to external websites. In the results of user counting from the first one displayed, there were only two general users with either academic or personal life posts on the "tumblr.com" link. Certainly, it is possible that some of the accounts classified as malicious were actually engaged in

legitimate commercial activities on sites with links to OSF. However, we chose not to make such distinctions as they would introduce arbitrariness on the part of the authors.

Two findings were indicated by observing changes in continuous search results over a year and a half regarding the same search term: First, OSF was actively engaged in removing malicious projects; however, malicious users continued to engage in their activities. Second, ordinary users experienced harm because their content was buried in the search results.

The percentage of abused projects registered in English increased compared to those registered in languages of specific cultures. Malicious users may have changed their approach to reach more widely within this short period.

Today, there are strict punishments not only for those who upload pirated works, but also for those who save them. In addition, leech site administrators may be punished if they leave the sites. In fact, the precedents of the European Court of Justice that provides links to illegal content constitutes copyright infringement (*GS Media v. Sanoma Media*, 2016). In addition, the German Federal Court of Justice clarified that online platform "YouTube" could be held liable if they fail to take appropriate action against copyright infringements committed using the service (*Bundesgerichtshof*, 2022). In 2021, Germany implemented the new EU Copyright Directive, which requires online services to ensure that infringing content is removed and not reuploaded. If platforms fail to do so, they can be held accountable for the damage. Even for academic search sites, if the operators neglect their responsibilities, they could be charged with a crime, and the site could be shut down. Thus, the loss of a major platform for open science, which implies the loss of a huge amount of information on study data, materials, and registrations, could have a serious negative impact on the scientific community.

OSF actively addresses this issue. First, our research results on "tumblr.com" serve as evidence of the diligent removal of malicious projects. In addition, we reached out to a COS member and inquired about specific measures against spam. They informed us that they had used multiple spam filters that considered a plethora of different factors (everything from odd titles to suspicious activity from a nearby IP address) when flagging spam. Simultaneously, they took steps to address the case of legitimate accounts among these flagged accounts. The following email was sent to flagged accounts to prevent false banning:

"Hello,

Thank you for reaching out to the OSF support team! We investigate reported cases of OSF users being accidentally flagged or deactivated on a case-by-case base. If you have an account or material that has been flagged as SPAM or deactivated please report your case here: Account Disabled or Falsely Flagged as SPAM reporting. Cases are reviewed within 48 to 72 hours.

If this is not the case please write back into OSF support with more details about your situation.

Thank you,"

Thus, COS members have attempted to do so. However, malicious users continue to use OSF beyond the measures to prevent it.

Fig 1. A search result on "tumblr.com." Contributors' names are mosaicked by the authors.
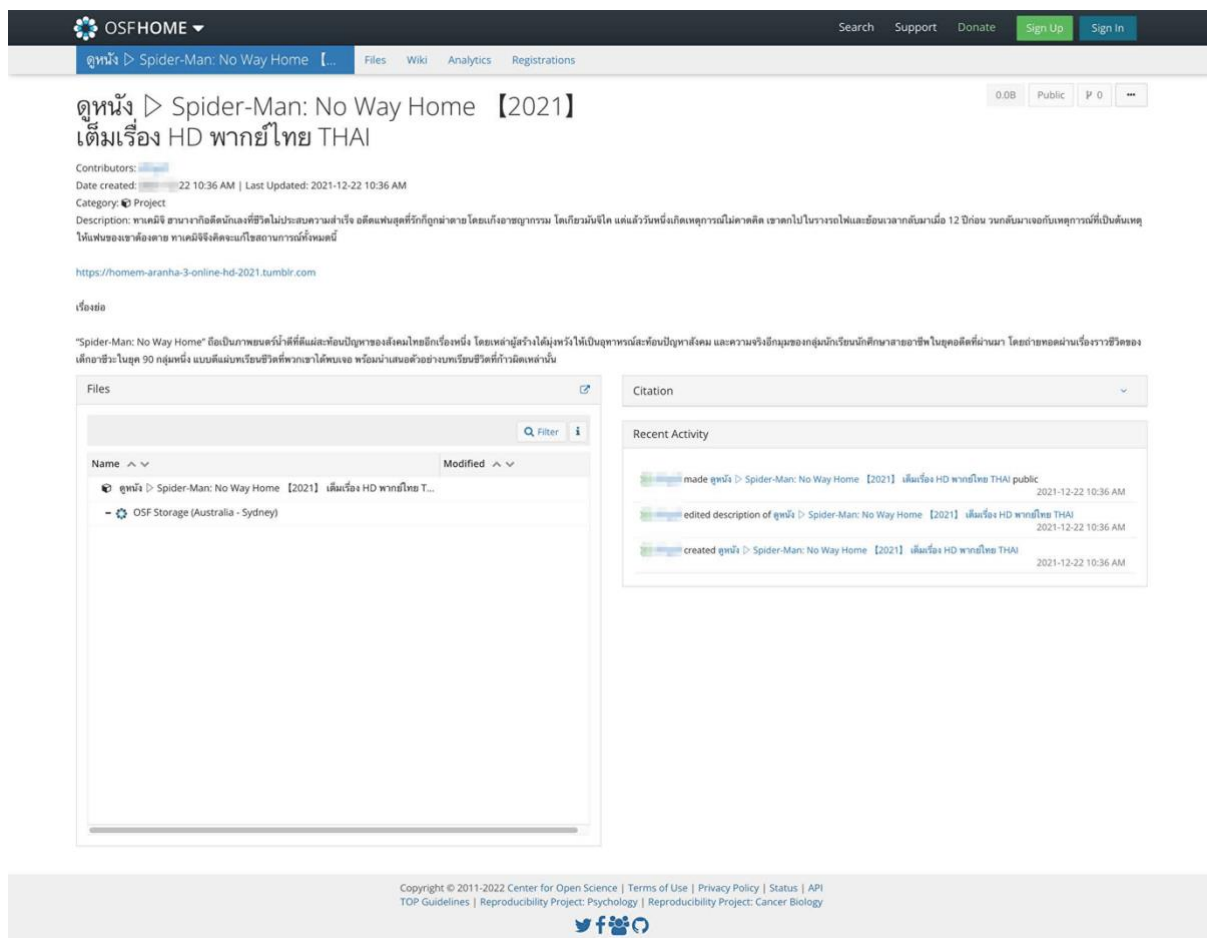
Fig 2. A project that has a link to an uploaded pirated movie. A contributor's name and the

link to an uploaded pirated movie are mosaicked by the authors.

The abuse of academic search platforms other than OSF

We found similar cases to OSF which linked to pirated films in widely used academic

research sites, such as figshare, Zenodo, Google Scholar, Academia.edu, and ORCID (search

words: "1080p," "download" and "Full HD"). Most illegal posts found in September 2022 were

deleted by May 2023, indicating that these academic research sites continue to clean up users'

illegal acts in the same manner as the OSF. However, there were some, though not many,

postings that have remained for more than six months without being deleted (Fig 3). In addition,

as of May 2023, new malicious posts were created (Fig 4). Researchers generally do not search

for such words, thus preventing the problem from coming to light easily. This shows that these cases are common and widespread in academic tools, not just OSF. This presents a challenge in open software management.
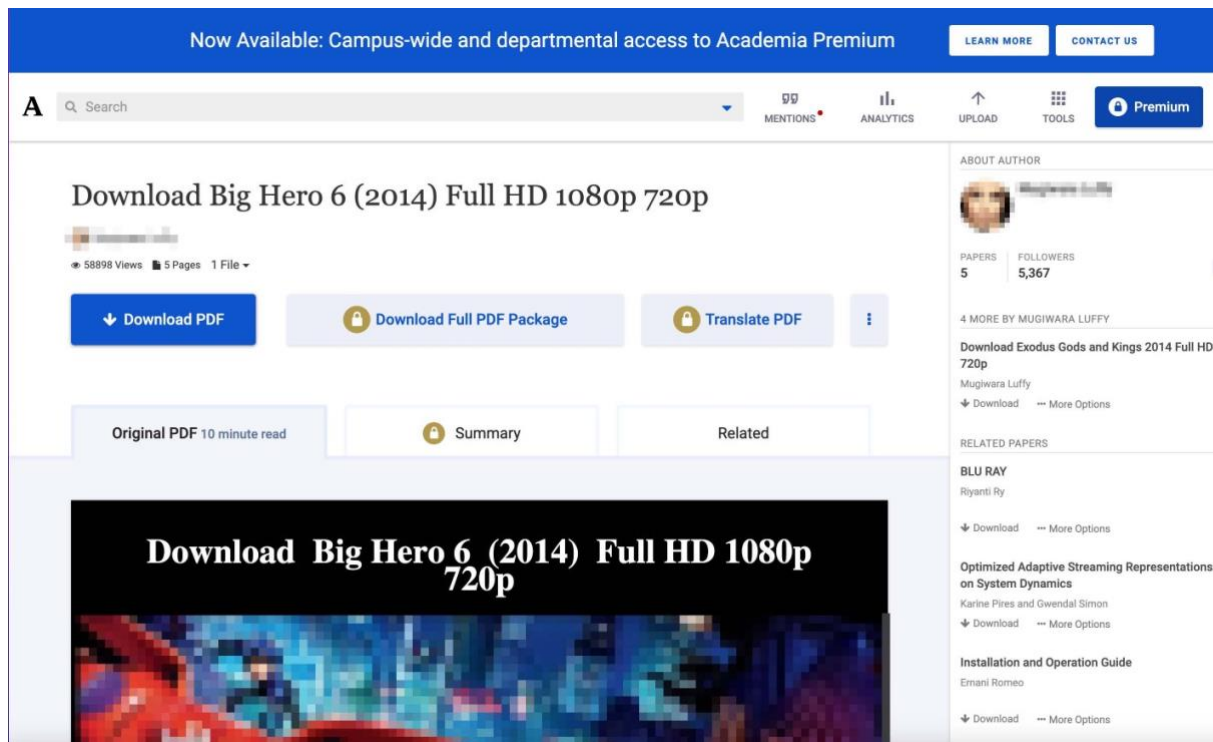
Fig 3. An example of abused postings on academic research sites other than OSF that have stayed for more than six months (The site is Academia).
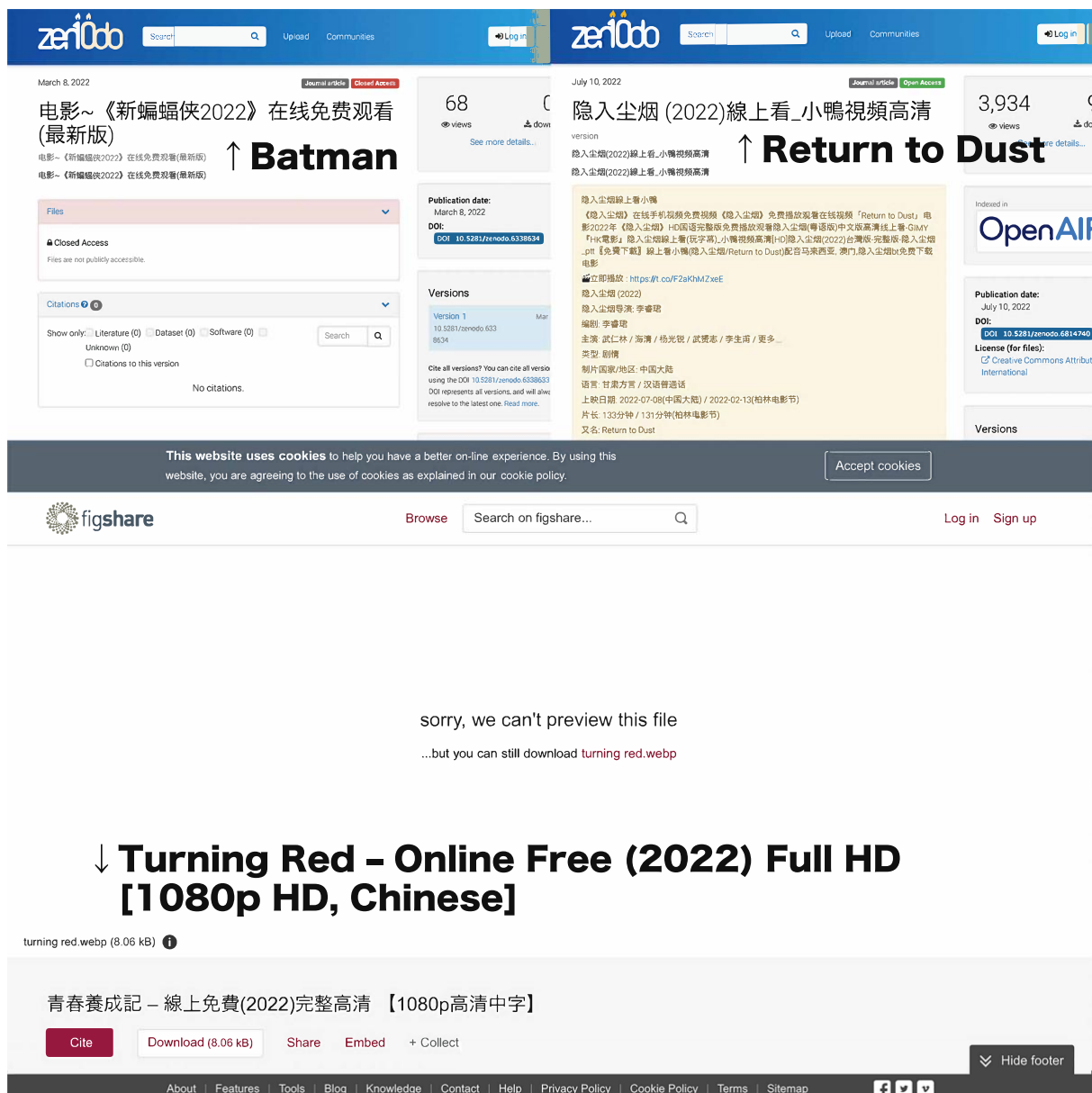
Fig 4. Examples of abusive postings on academic websites. Above, Zenodo; below, Figshare.

Links for phishing scams also appear on the OSF

In addition to copyright infringement, OSF has been abused in other, more serious cybercrimes. We found many accounts unrelated to academic research and attempted to direct users to sites where money was exchanged (Fig 5). We again conducted a search in the OSF using "tumblr.com" as a search word, which was used as a destination for illegal uploads, as

introduced above. Consequently, several suspicious accounts were identified. Some accounts had links in their personal website sections that directed users to malicious websites claiming to be online casinos. These online casino sites encourage visitors to register as members and enter their phone and credit card numbers. A website that withdraws important personal information from site visitors is similar to the characteristics of scams that are rampant on the Internet. For example, we searched for the site link name in the personal website section (i.e., "happyluke withdraw money") using the Google Chrome browser. Chrome then warned us that "passwords, messages, credit cards, and other information may be stolen by malicious users." and blocked our access. This fact indicates that OSF is abused to reach dangerous sites that are blocked by Internet security systems.

These illegal accounts and project pages are shared by SHARE (https://share.osf.io/). The harmful effects of these abuses spread outside the OSF. SHARE is an open API tool that helps bring together scholarships distributed across research ecosystems for greater discoverability. SHARE usually facilitates studies through an easy search, but the downstream effects of inappropriate content being picked up may spoil these merits. That is, a malicious user would use SHARE to accelerate the search for illegal projects and parasites and facilitate the pollution of open science platforms. We found the above cases of abuse by chance, and these findings could be the tip of the iceberg. Researchers should anticipate hidden abuse projects that are not assumed by authors. The vulnerability of OSF and some academic research sites to invasive use, such as illegal uploads and phishing pages, indicates that the site can also be used for more dangerous criminal activities, such as trading illegal items or personal information.
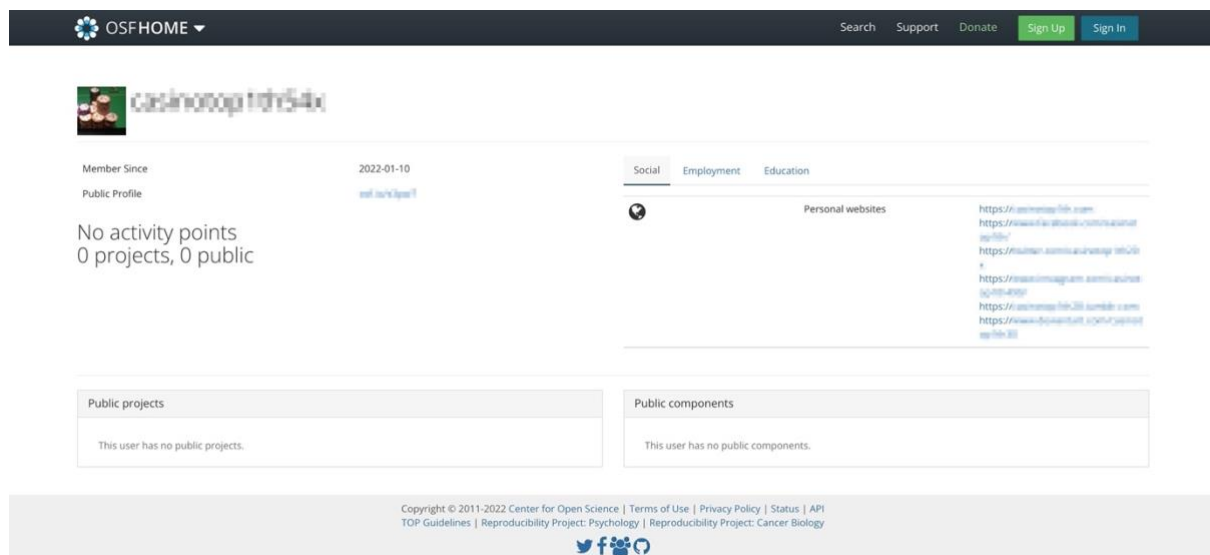
Fig 5. An OSF personal page of a person posting phishing sites. A contributor's name, OSF ID, and links to phishing sites are mosaicked by the authors.

In addition to legal issues, illegal activities conducted on the research platform can hinder general research activities. For example, researchers studying films, gambling, and social media possibly use words such as a specific movie title, casino, and Tumblr in the "Description" section, and then illegal projects which contain piracy or phishing sites will become noise in search for their sane projects. Thus, accounts and projects involving criminal activities prevent researchers from engaging in comfortable search activities and undermine the important functions of the OSF[1]. Based on our search for tumblr.com in January 2022, the oldest project with a link to the storage site was built on December 28, 2021. Furthermore, we searched again for the oldest malicious project on May 28, 2023, and found that the oldest

---

[1] The OSF HOME page states: "Discover projects, data, materials, and collaborators on OSF that might be helpful to your own research."

project was created on June 16, 2021, and the second oldest was created on December 5, 2021. Fortunately, to the best of our knowledge, there were no abuse projects finding loopholes in the automated systems of OSF before 2021. Since the preprint of January 2022, when we reported this problem, some measures have already been taken by server administrators; nevertheless, we are devising the following measures:

Possible solutions for illegal activities at OSF

How can the soundness of an OSF site be recovered? Here, we discuss some of the mitigation strategies for this issue, which are the same as some of the COS' responses to spam[2]. First, the COS responsible for the OSF has defined its terms of use (https://github.com/CenterForOpenScience/cos.io/blob/master/TERMS_OF_USE.md), and the kind of content we reported here clearly violates its Article 9 ("PERMITTED AND PROHIBITED USES"), so any account that does such activities will be banned and the content will be forcibly removed under Article 10 ("LIMITATION OF ACCESS AND REMOVAL OF CONTENT"). Obtaining help from rational users may be a way to eliminate illegal activities on OSF sites. For example, when volunteers find criminal accounts, they report them to the COS, which then bans them. However, there are limitations to labor and manpower. Thus, the COS team may need to hire additional dedicated staff or install additional automated detectors in the future.

Restricting the sign-up method may also be effective. The above types of illegal activities are primarily due to the advantage of allowing anyone to create accounts. Thus, it will be effective to limit and identify those who can create the accounts by registering ID

---

[2] This point has been confirmed by one of the COS' members.

information that can identify researchers such as an email address that has a domain for academic institutions (e.g., ". ac" and ".edu") or ORCID[3]. However, this may hinder the involvement of the public in science (i.e., citizen science) and cause inconvenience to early career researchers whose affiliations tend to change frequently.

Moreover, regarding piracy, attention should be paid to projects aimed at non-English-speaking audiences, especially those in Asian languages. A questionnaire investigation conducted by the Agency for Cultural Affairs in Japan showed that copyright damage in Asian countries was the most common (Agency for Cultural Affairs, 2010).

Consistent with this trend, a number of illegal projects used Asian languages in the OSF from January 2022 to May 2023. Such projects had decreased to some extent by October 2022 but had not been completely eradicated. Therefore, crime projects and accounts are not easily noticed by the COS, which is probably composed mainly of English-speaking staff; and although cooperation from non-English speakers is essential, it is both costly and difficult. Therefore, we suggest introducing an automatic detection system that operates in linkage with a machine translation device to automatically suspend accounts with criminal risk. If the account holder launches a protest against suspension, the COS can examine the content and determine whether to restore the account or not. This approach can be used in languages other than English at a relatively low cost. In conclusion, not only the COS but also the entire scientific community should develop good ideas and measures to prevent OSF and other similar services and entities from being abused by criminals.

---

[3] Note that there are also problems of empty or "ghost" ORCID accounts or the possible abuse of ORCIDs to register potentially fake elements (Teixeira da Silva, 2020a, b).

activities.

Competing interests

The authors have no competing interests to declare.

Authors' contributions

Ayumi Ikeda: Conceptualization (lead); Funding Acquisition (lead); Investigation (equal); Writing – Original Draft Preparation (equal); Writing – Review & Editing (lead). Fumiya Yonemitsu: Conceptualization (lead); Funding Acquisition (equal); Investigation (equal); Writing – Original Draft Preparation (equal); Writing – Review & Editing (supporting). Naoto Yoshimura: Conceptualization (supporting); Funding Acquisition (equal); Investigation (equal); Writing – Original Draft Preparation (equal); Writing – Review & Editing (supporting). Kyoshiro Sasaki: Funding Acquisition (equal); Investigation (equal); Writing – Original Draft Preparation (equal); Writing – Review & Editing (supporting). Yuki Yamada: Funding Acquisition (equal); Investigation (equal); Supervision(lead); Writing – Original Draft Preparation (equal); Writing – Review & Editing (supporting).

References

Agency for Cultural Affairs (2010). [Questionnaire Survey on Piracy: Summary of Results] Kaizokuban higai nado ni kansuru anketo chousa shukei kekka gaiyo (in Japanese).

Retrieved January 12, 2022, from https://www.bunka.go.jp/tokei_hakusho_shuppan/tokeichosa/kaizokuban/pdf/higaicho sa_houkokusho.pdf

Elliott, F., Errington, T. M., Bagby, C., Frazier, M., Geiger, B. J., Liu, Y., … & Chen, L. (2021). OSF. Retrieved from osf.io/4znzp

Kyodo News (2021, Jun 2). *Ex-illegal manga site operator slapped with three-year prison sentence*. The Japan Times. Retrieved January 12, 2022, from https://www.japantimes.co.jp/news/2021/06/02/national/illegal-manga-website/

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349. doi:10.1126/ science.aac4716

Teixeira da Silva, J. A. (2020a). Failure of ORCID: 57 academics named "Beatriz". *Update Dental College Journal*, 10, 3-5. doi:10.3329/updcj.v10i2.50172

Teixeira da Silva, J. A. (2020b). ORCID: Issues and concerns about its use for academic purposes and research integrity. *Annals of Library and Information Studies*, 67, 246-250.

GS Media BV v. Sanoma Media Netherlands BV and Others, CJEU, C-160/15 (2016), https://curia.europa.eu/juris/fiche.jsf?id=C%3B160%3B15%3BRP%3B1%3BP%3B1 %3BC2015%2F0160%2FJ&oqp=&for=&mat=or&lgrec=en&jge=&td=%3BALL&ju r=C%2CT%2CF&num=c-160%252F15&dates=&pcs=Oor&lg=&pro=&nat=or&cit=none%252CC%252CCJ%2 52CR%252C2008E%252C%252C%252C%252C%252C%252C%252C%252C%252

C%252Ctrue%252Cfalse%252Cfalse&language=en&avg=&cid=1694400#section_do cuments (January 12, 2022)

Der Bundesgerichtshof (2023, June 9). *Zur Haftung von "YouTube" und "uploaded" für Urheberrechtsverletzungen*. Retrieved June 10, 2023, from https://www.bundesgerichtshof.de/SharedDocs/Pressemitteilungen/DE/2022/2022080. html