Digital Society (2023) 2:25
https://doi.org/10.1007/s44206-023-00055-1

BRIEF COMMUNICATION

International Companies and Governments Need to be Alerted to a Fatal Flaw in DOI Management Policy to Prevent Security Breaches

Yoshiyasu Takefuji[1]

Received: 10 March 2023 / Accepted: 16 June 2023

Abstract
This paper will identify a fatal flaw in the current Digital Object Identifier (DOI) management policy regarding software reproducibility validation with service providers. The fatal flaw can cause security breaches for individuals and organizations over the Internet. Under the current DOI policy, once software code with known security vulnerability is published with unique DOI, no one can modify or delete it. This paper will also address how the DOI software policy should be fixed or updated for deleting harmful software DOI numbers. We must be aware of this fatal flaw on the DOI management policy for preventing security breaches.

Keywords Digital Object Identifier (DOI) · Software reproducibility · DOI policy · Security breach · Management policy

1 Introduction

Software reproducibility plays a key role in increasing productivity and reducing development costs. However, due to current DOI policy, this is the first reported security breach vulnerability in software productivity verification in the world. The current DOI policy on software should be corrected as soon as possible.

The International DOI Foundation (IDF) was established in 1997 to develop and manage the DOI (Digital Object Identifier) system. The Standard (ISO 26324, Digital

Highlights
- A fatal flaw in the DOI (Digital Object Identifier) policy was identified in software.
- Software with known vulnerability and published DOI may cause security breaches.
- The DOI policy flaw on software is harmful to individuals and organizations worldwide.
- All businesses need to be aware of the fatal flaws in their software DOI policies.

✉ Yoshiyasu Takefuji
takefuji@keio.jp

[1] Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, 135-8181 Tokyo, Japan

Published online: 02 July 2023

Object Identifier System) was approved in November 2010 and published in May 2012. According to IDF, (1) A DOI name is an identifier (not a location) of an entity on digital networks. (2) It provides a system for persistent and actionable identification and interoperable exchange of managed information on digital networks. (3) A DOI name can be assigned to any entity — physical, digital or abstract — primarily for sharing with an interested user community or managing as intellectual property. (4) The DOI system is designed for interoperability; that is to use, or work with, existing identifier and metadata schemes. (5) DOI names may also be expressed as URLs (URIs).

According to DOI FAQ, approximately 275 million DOIs have been assigned through a worldwide federation of registrars. Many documents in academia and industry are documented by DOIs. In other words, The DOI number is a unique URL for accessing the document on the Internet. In other words, DOIs are used in many publications on all areas of science, technology, and business worldwide.

The DOI for a document remains fixed over the lifetime of the document. This DOI policy may be fine with many documents, but it may be harmful in software DOI numbers with known source code vulnerabilities. This means that malicious fraudsters can exploit known vulnerabilities to cause security breaches that can cause significant financial damage. This fatal flaw in the DOI policy should be disclosed and published and the problem fixed as soon as possible.

This paper shows a fatal flaw in the current DOI policy on software reproducibility with service providers. The fatal flaw can cause security breaches for individuals and organizations over the Internet, which can have a significant impact on academia, industry, and business to cause big financial damages. According to statista.com, as of 2022, the average cost of a data breach in the USA amounted to 9.44 million US dollars.

Without an updated current DOI policy, software code with a DOI number could cause security breaches and other harm to society. The current DOI policy states that once a DOI document has been assigned a DOI number, it cannot be updated or deleted forever. In other words, the current DOI policy is valid for many documents other than software source code. Companies and governments in all sectors need to be alerted to the fatal flaws in their software DOI policies to prevent security breaches.

The reproducibility crisis in science, technology, and business refers to the fact that the results of many scientific studies have been found to be difficult or even impossible to reproduce (Laraway et al., 2019).

In 2016, Baker wrote an article entitled “1,500 scientists lift the lid on reproducibility” (Baker, 2016). However, Olavo B. A. et al. reported that reproducibility is costly (Amaral & Neves, 2021).

This reproducibility crisis can be found in computer software. Software reproducibility validation is essential for scientists and engineers to utilize claimed open-source codes for rebuilding new applications. The current DOI policy of software reproducibility validation has the potential to harm our society.

When software code is published with a unique DOI number, the fatal flaw is that it is impossible to fix or remove content or software code, even if a known software vulnerability exists.

Many peer-reviewed publishers such as Springer Nature (Editorial, 2018, 2019a, b; Jeffrey, 2019) and Science (Science Translational Medicine Editorial Policies, https://www.science.org/content/page/stm-editorial-policies) have been making partners with software reproducibility service providers such as Code Ocean. Many peer-reviewed publishers such as Elsevier, IEEE, Black & Wiley, Taylor & Francis, Springer Nature, SAGE, and Cambridge University Press emphasize the importance of software reproducibility (Editorial, 2018, 2019a, b; Jeffrey, 2019; Pérignon et al., 2019; Joppa et al., 2013; Stodden et al., 2016; Science Translational Medicine Editorial Policies, https://www.science.org/content/page/stm-editorial-policies).

However, publishers have the narrow goal such that reviewers are allowed to evaluate and verify the authors' proposed claims in submitted papers using given software codes. Authors must submit their software codes to Code Ocean and Code Ocean will verify the submitted codes, if the verification is successful then Code Ocean will send a message "Acceptance for publication".

However, the current DOI policy is that once software code was published with a unique DOI number, no one can change or delete it. This DOI policy may seem perfectly fine at first glance for many documents, but if there are security issues in software source codes, such as user names and passwords being publicly exposed, you will want to modify or remove the software code.

According to IBM report in 2022 IBM (IBM, 2022), the average cost of a data breach is at an all-time high of US$4.35 million. We should avoid data breach as much as possible. This paper will discover and identify why the current DOI policy is harmful to our society using the recent my published DOI experience.

1.1 Document DOI vs. Software DOI

Scientific papers are tentative products of the pursuit of truth, logically assembled from facts and evidence. Since these are written based on the findings at the time of writing and agreed upon after review by peer reviewers and others, there is little incentive to change or delete them. In scientific papers, there is a mechanism for correcting rebuttals and matters arising corrections when problems arise.

The release of software is, as the author states, a trade-off between providing reproducibility at the time of the study and providing the opportunity to change or remove vulnerabilities. The current DOI policy emphasizes the former, providing reproducibility at the time of the study, and sustainability. We know every day that we do not have perfect software code, as evidenced by updates to the OS and libraries.

On the other hand, there is a GitHub site, for example, which assumes that software will be improved and updated to solve the problems. The GitHub site is designed to release new software version to fix the problems. In other words, fixed DOI software is harmful in general and in the future.

This paper shows a concrete example of the fatal flaw in the DOI policy on software reproducibility validation. Software reproducibility validation is recommended in many top journals, including *Science*, *Nature*, *NEJM*, and *The Lancet*. This paper will be essential for scientists, engineers, and business people around the world to prevent unnecessary financial losses.

2 DOI Flaw Experience in Software

The author published a dynamic DNS updater, dyDNS with Code Ocean (Takefuji, 2022). The dynamic DNS updater is to update the dynamic IP for freedns.afraid.org. freedns.afraid.org is one of the largest free dynamic DNS providers. Via wired, Wi-Fi, or mobile SIM with dynamic IPs, dyDNS allows users to access the changed-IP machine with the same domain name which is very convenient and useful with a variety of mobile services and servers.

However, current free dynamic DNS providers use plain text for usernames, passwords, and domain names, which is not secure. Therefore, the author developed a new dyDNS application for enhancing security. dyDNS allows users to change unsecured plain text to secured crypted file with OpenSSL.

However, the author found that the key file published on Code Ocean is publicly readable. The key file was supposed to be readable only to a user (author), but it should not be publicly readable or disclosed. The key file is used to decrypt crypted document. In other words, publicly exposing the key file means that anyone can decrypt the crypted file and can read the plain text such as username, password, and domain name. In other words, important private information on dynamic DNS was exposed in public.

However, the current DOI policy for software reproducibility validation does not allow the author to alter the publicly readable key file to unreadable file in public. The key file should not be readable to public.

In other words, although security vulnerabilities are known in published software codes with the unique DOI number, you cannot remove or alter them forever with the current DOI policy.

This is a fatal flaw in the current DOI policy on software or software reproducibility validation with many publishers. You can't leave the vulnerability as it is?

In this case, the published password in Code Ocean could not be changed with the current DOI policy so that the registered password in freedns.afraid.org was forced to be changed in order to avoid public exposure of the private username, password, and domain name in public respectively.

The current DOI policy has a fatal flaw in software or software reproducibility verification: even if you know there is a security problem in published software code with unique DOI number, you cannot change or remove it.

This paper recommends that software codes with unique DOIs for software reproducibility validation with known security vulnerabilities should be modifiable or deleted. The current DOI policy on software should be updated or changed as soon as possible for preventing security breaches for individuals and organizations without significant financial damages.

3 Future Work

Digital objects covered by the DOI include research papers, databases/datasets, software, patents, and various other objects. This paper has raised the issue of software through the author's own experience. Future work is to examine the issues of DOI management policies for other objects in accordance with their characteristics.

For example, in the case of patents, even if the scope of rights at the time of patent registration is registered in the DOI, some countries have adopted a system in which the scope of rights is reduced by subsequent invalidation trials, etc. Therefore, as this paper pointed out for software in this study, if updates and changes are not made, third parties will be misled as to the scope of rights.

In addition, for databases and datasets, it is possible that after the release of a DOI, there may be circumstances that cannot be assumed at the time of release that make it necessary to withhold some data or datasets.

As described above, some digital objects may be motivated to be changed or deleted at the discretion of the creator after the DOI is released.

4 Conclusion

The current DOI policy on published software codes with DOI numbers does not allow users to modify or delete DOI documents. This DOI policy may be fine for many documents with unique DOIs, but must be modified or changed when DOI software source code with known vulnerabilities. This means that if there is a known vulnerability in the source code of software with a published DOI number, users should be able to modify or remove the offending DOI document or number for preventing privacy exposure and financial damages. International companies and governments in the world need to recognize this fatal flaw in the software DOI policy for preventing security breaches and financial damages for individuals and organizations.

Data Availability Not applicable.

Declarations

Conflict of Interest The author declares no competing interests.

References

Amaral, O. B., & Neves, K. (2021). Reproducibility: Expect less of the scientific paper. *Nature, 597*(7876), 329–331. https://doi.org/10.1038/d41586-021-02486-7. PMID: 34526702.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 533*, 452–454. https://doi.org/10.1038/533452a

Editorial. (2018). Easing the burden of code review. *Nature Methods*, *15*, 641. https://doi.org/10.1038/s41592-018-0137-5

Editorial. (2019a). Changing coding culture. *Nature Biotechnol*, *37*, 485. https://doi.org/10.1038/s41587-019-0136-9

Editorial. (2019b). Sharing high expectations. *Nature Machine Intelligence*, *1*, 329. https://doi.org/10.1038/s42256-019-0092-6

IBM. (2022). How much does a data breach cost in 2022?. https://www.ibm.com/security/data-breach

Jeffrey, M. (2019). Perkel, Make code accessible with these cloud services. *Nature, 575*, 247–248. https://doi.org/10.1038/d41586-019-03366-x

Joppa, L. N., McInerny, G., Harper, R., et al. (2013). Computational science. Troubling trends in scientific software use. *Science, 340*(6134), 814–815. https://doi.org/10.1126/science.1231535

Laraway, S., Snycerski, S., Pradhan, S., & Huitema, B. E. (2019). An overview of scientific reproducibility: Consideration of relevant issues for behavior science/analysis. *Perspectives on behavior science, 42*(1), 33–57. https://doi.org/10.1007/s40614-019-00193-3

Pérignon, C., Gadouche, K., Hurlin, C., Silberman, R., & Debonnel, E. (2019). Certify reproducibility with confidential data. *Science, 365*(6449), 127–128. https://doi.org/10.1126/science.aaw2825

Science Translational Medicine Editorial Policies. Retrieved June 24, 2023, from https://www.science.org/content/page/stm-editorial-policies

Stodden, V., McNutt, M., Bailey, D. H., et al. (2016). Enhancing reproducibility for computational methods. *Science, 354*(6317), 1240–1241. https://doi.org/10.1126/science.aah6168

Takefuji, Y. (2022). dyDNS for dynamic DNS updater in freedns.afraid.org [Source Code]. https://doi.org/10.24433/CO.2993693.v1