Code Is Different: How the Developer Community Drives Content Moderation on Code Collaboration Platforms

Margaret Tucker, Policy Manager

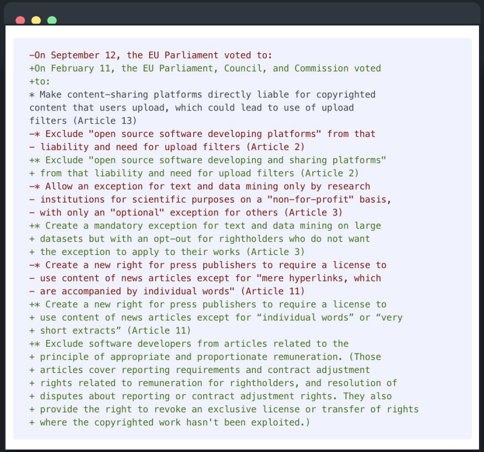
GitHub

Code is different from other content. Its functional purpose creates unique considerations for platform moderation.

- Copyright
- Dual-use
- Network effects of takedowns

Code has different copyright concerns.

- Functional purpose
- Independent duplication
- Open source sharing
- Filtering false positives



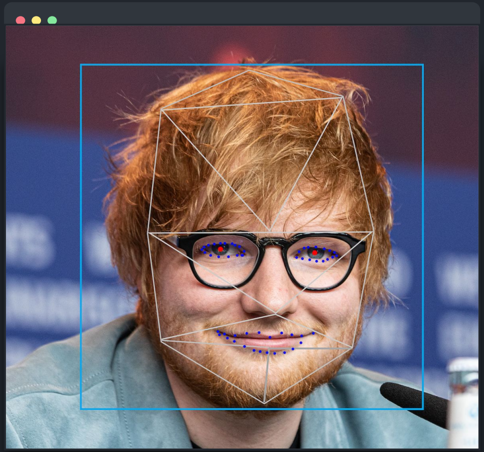
```diff
-On September 12, the EU Parliament voted to:
+On February 11, the EU Parliament, Council, and Commission voted
+to:
* Make content-sharing platforms directly liable for copyrighted
content that users upload, which could lead to use of upload
filters (Article 13)
-* Exclude "open source software developing platforms" from that
- liability and need for upload filters (Article 2)
+* Exclude "open source software developing and sharing platforms"
+ from that liability and need for upload filters (Article 2)
-* Allow an exception for text and data mining only by research
- institutions for scientific purposes on a "non-for-profit" basis,
- with only an "optional" exception for others (Article 3)
+* Create a mandatory exception for text and data mining on large
+ datasets but with an opt-out for rightholders who do not want
+ the exception to apply to their works (Article 3)
-* Create a new right for press publishers to require a license to
- use content of news articles except for "mere hyperlinks, which
- are accompanied by individual words" (Article 11)
+* Create a new right for press publishers to require a license to
+ use content of news articles except for “individual words” or “very
+ short extracts” (Article 11)
+* Exclude software developers from articles related to the
+ principle of appropriate and proportionate remuneration. (Those
+ articles cover reporting requirements and contract adjustment
+ rights related to remuneration for rightholders, and resolution of
+ disputes about reporting or contract adjustment rights. They also
+ provide the right to revoke an exclusive license or transfer of rights
+ where the copyrighted work hasn't been exploited.)
```

Case study: EU Copyright Directive

Code has dual-use applications.

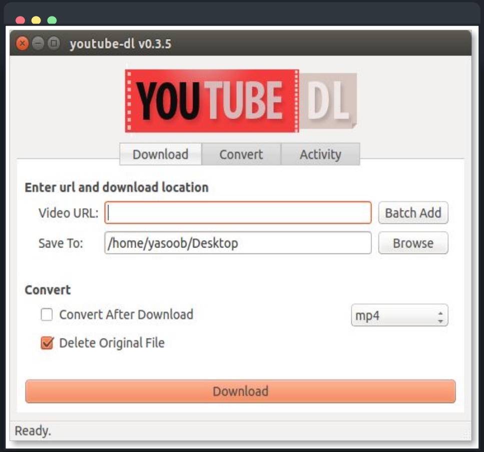
Supporting legitimate and beneficial research while disallowing harmful misuse

Case study:
Security research



Case study:
Deepfake NCII and disinformation policy

Moderating code requires careful deliberation of context and network effects.



Enter url and download location

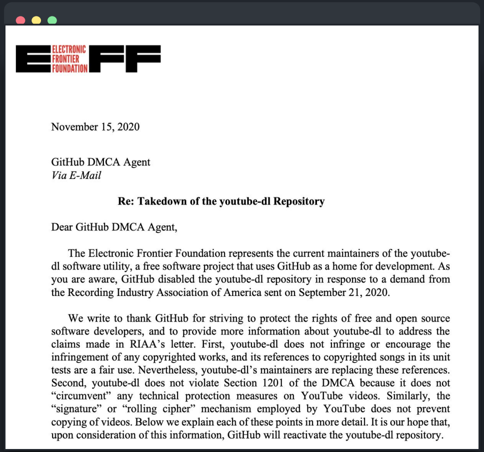
Video URL:

Save To: /home/yasoob/Desktop

Convert

□ Convert After Download mp4

☑ Delete Original File



ELECTRONIC FRONTIER FOUNDATION

November 15, 2020

GitHub DMCA Agent
Via E-Mail

Re: Takedown of the youtube-dl Repository

Dear GitHub DMCA Agent,

The Electronic Frontier Foundation represents the current maintainers of the youtube-dl software utility, a free software project that uses GitHub as a home for development. As you are aware, GitHub disabled the youtube-dl repository in response to a demand from the Recording Industry Association of America sent on September 21, 2020.

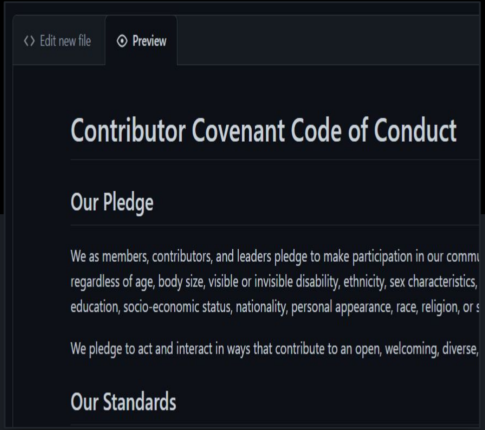
We write to thank GitHub for striving to protect the rights of free and open source software developers, and to provide more information about youtube-dl to address the claims made in RIAA's letter. First, youtube-dl does not infringe or encourage the infringement of any copyrighted works, and its references to copyrighted songs in its unit tests are a fair use. Nevertheless, youtube-dl's maintainers are replacing these references. Second, youtube-dl does not violate Section 1201 of the DMCA because it does not "circumvent" any technical protection measures on YouTube videos. Similarly, the "signature" or "rolling cipher" mechanism employed by YouTube does not prevent copying of videos. Below we explain each of these points in more detail. It is our hope that, upon consideration of this information, GitHub will reactivate the youtube-dl repository.

Case study: youtube-dl

Moderating a developer community

Developer community and open source norms

Platform moderation beyond takedowns



Contributor Covenant Code of Conduct

Our Pledge

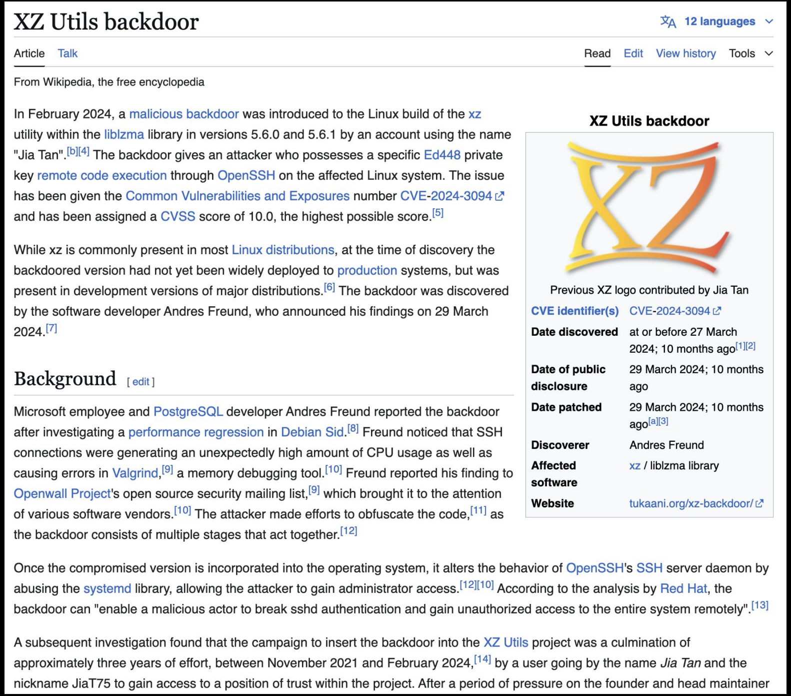
We as members, contributors, and leaders pledge to make participation in our commu
regardless of age, body size, visible or invisible disability, ethnicity, sex characteristics,
education, socio-economic status, nationality, personal appearance, race, religion, or s

We pledge to act and interact in ways that contribute to an open, welcoming, diverse,

Our Standards

Encouraging community content moderation

Case study: xz backdoor



XZ Utils backdoor

From Wikipedia, the free encyclopedia

In February 2024, a malicious backdoor was introduced to the Linux build of the xz utility within the liblzma library in versions 5.6.0 and 5.6.1 by an account using the name "Jia Tan".[b][4] The backdoor gives an attacker who possesses a specific Ed448 private key remote code execution through OpenSSH on the affected Linux system. The issue has been given the Common Vulnerabilities and Exposures number CVE-2024-3094 and has been assigned a CVSS score of 10.0, the highest possible score.[5]

While xz is commonly present in most Linux distributions, at the time of discovery the backdoored version had not yet been widely deployed to production systems, but was present in development versions of major distributions.[6] The backdoor was discovered by the software developer Andres Freund, who announced his findings on 29 March 2024.[7]

XZ Utils backdoor

Previous XZ logo contributed by Jia Tan

CVE identifier(s)	CVE-2024-3094
Date discovered	at or before 27 March 2024; 10 months ago[1][2]
Date of public disclosure	29 March 2024; 10 months ago
Date patched	29 March 2024; 10 months ago[a][3]
Discoverer	Andres Freund
Affected software	xz / liblzma library
Website	tukaani.org/xz-backdoor/

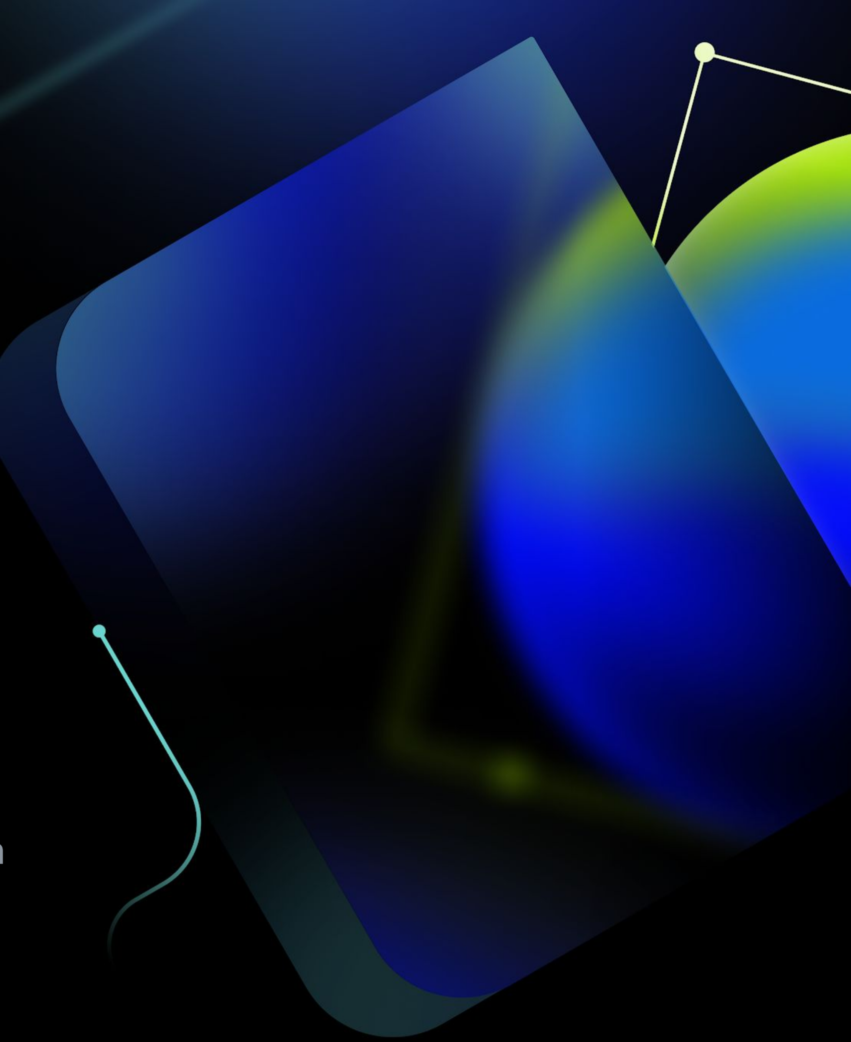
Background [edit]

Microsoft employee and PostgreSQL developer Andres Freund reported the backdoor after investigating a performance regression in Debian Sid.[8] Freund noticed that SSH connections were generating an unexpectedly high amount of CPU usage as well as causing errors in Valgrind,[9] a memory debugging tool.[10] Freund reported his finding to Openwall Project's open source security mailing list,[9] which brought it to the attention of various software vendors.[10] The attacker made efforts to obfuscate the code,[11] as the backdoor consists of multiple stages that act together.[12]

Once the compromised version is incorporated into the operating system, it alters the behavior of OpenSSH's SSH server daemon by abusing the systemd library, allowing the attacker to gain administrator access.[12][10] According to the analysis by Red Hat, the backdoor can "enable a malicious actor to break sshd authentication and gain unauthorized access to the entire system remotely".[13]

A subsequent investigation found that the campaign to insert the backdoor into the XZ Utils project was a culmination of approximately three years of effort, between November 2021 and February 2024,[14] by a user going by the name *Jia Tan* and the nickname JiaT75 to gain access to a position of trust within the project. After a period of pressure on the founder and head maintainer

New Frontiers for Content Moderation



AI

AI developer tools

Model hosting

Scaling moderation

One billion developers on GitHub by 2030

Get in touch with us

Collaborate on site policies

- 30 day notice-and-comment period for substantive site policy changes
- Provide feedback via pull request or opening an issue in our site-policy repo

Open an issue in our developer-policy repo

Share public policy issues of concern to developers

Email us policy@github.com

Thank you