



YouTube's Content ID System

Elly Lee, Shunya Kodaira, Samuel Hsu

We present this evaluation as external auditors hired to assess YouTube's Content ID system for liability, fairness, and operational risks.

Introduction: Why Content ID Exists

- YouTube receives 500+ hours of video per minute
 - ◆ Impossible for human review!
- YouTube built **Content ID** in 2007 to scan uploads for copyrighted material
- Goal:
 - ◆ Protect copyright owners
 - ◆ Reduce legal liability
 - ◆ Manage copyright at scale



Definitions

- **Copyright:** People who make creative works (e.g. music, art, writing, videos) control how their works are used
 - Example: If you record a song, a YouTuber generally can't reuse your song without permission
- **Fair use:** Use of copyrighted works is sometimes allowed without prior permission in limited, transformative ways (Harvard Office of the General Counsel, 2023)
 - Note: Fair use is *contextual*, meaning that algorithms cannot determine it detect reliably
- **DMCA:** online platforms can't be sued for their users' copyright infringement if they provide a way for copyright owners to take down infringing content (Electronic Frontier Foundation, n.d.; US Copyright Office, n.d.)

Youtube and Content ID

- In most cases, using copyrighted media requires permission unless the use qualifies as fair use
- **Content ID:** data-driven decision-making system for detecting the use of copyrighted material in YouTube videos
- Copyright owners generally **earn money** off of infringing videos (Google, n.d.; Scott, 2020)
 - They can also block videos, strike them down via DMCA, or sue
- “Content ID means that video creators... generally don’t get sued and don’t get DMCA takedowns” (Scott, 2020)





How Content ID Works

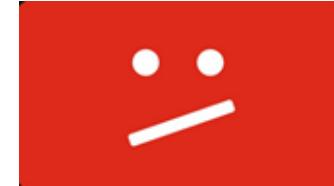
1. Copyright owners upload “*reference files*” of their work to YouTube’s database, converting into digital fingerprints of unique audio and video patterns
2. When a user uploads a new video, Content ID scans for matches, then applies copyright owner’s predefined policy (Google, n.d.)
3. When the system notices a match, a copyright owner’s policy takes place
 - a. *Monetize*: Has the video run ads, where the revenue goes to the content owner
 - b. *Track*: Leaves the video alone, but analytics of the video are sent to the content owner
 - c. *Block*: Prevents the video from being viewed
 - d. *Mute*: Mute the audio segment from a video

Benefits

- Handles **immense volumes** of copyright checks **daily** that would otherwise be extremely difficult, if not impossible, for humans
- Detects most straightforward copyright matches accurately
- Enables **automatic revenue sharing** for copyright owners
- **Reduces legal risks** for small creators
 - “Content ID means creators generally don’t get sued...” (Scott, 2020)
- Helps maintain **platform compliance** with DMCA



Issues



- **False identification:** wrongfully flags or takes down original works
 - British composer *Simon Wilkinson* - one of **his** popular tracks taken down as copyrighted material owned by **someone else** (Israel, 2016)
- **Wrongful penalties**
 - Canadian YouTuber *Corey Vidal* - his YouTube channel **gone** due to his Star Wars a capella tribute, **lost** thousands of revenue and months of appeals (Israel, 2016)
- **No fair use recognition:** cannot detect fair use of copyrighted works
(Trendacosta, 2020)
 - Video producer *Shane Luis* - **removed** one of his videos in fear of earning a strike, despite believing his use was **legal** (Israel, 2016)
 - Reviewer *Doug Walker*: “I have never had a day where I felt safe posting one of my videos even though the law states I should be safe” (Harmon, 2016)

Issues



- **Power imbalance** - Large media companies have more influence than small creators
 - In 2016, **Fox**'s TV show, *Family Guy*, used a 2009 footage of gameplay of *Double Dribble* by another YouTube user, **sw1tched**, then YouTube took down sw1tched's video, claiming it contains content from Fox (Despotov, 2016)
- **Opaque system:** No transparency on match confidence, thresholds, or reference files

Result: *Chilling effect on creativity and speech*

Current Solutions

- While not perfect, youtube has taken actions to prevent false claims
 - **Transparency:** Copyright claim shows the exact timestamps on where the system detected a copyright match
 - **Creator tools:** An option where a video can be edited to avoid copyright issues (remove audio during a specific section)
 - **Dispute/Appeal System:** Creators can appeal to the copyright owner if they believe they have been falsely accused
- The solutions above have worked as an alternative/solution to false claims
- Limitation:
 - Doesn't prevent false claims from taking place in the first place
 - These solutions still rely heavily on automation and do not address biases built into the matching process.

Solution A: Reduce Creator Burden (Evidence Handling)

- Problem: Creators cannot easily provide proof of legitimate use or licensing
- **Create a Creator Evidence Submission Portal**
 - Allow upload of:
 - Screenshots of permission
 - License receipts
 - Proof of originality
- Why this helps
 - Reduces false positives
 - Gives creators a structured way to defend themselves
 - Improves accuracy and fairness in early stages
- Limitation
 - Requires partial human review
 - Increased operational costs
 - Does *not* resolve fair use (requires context)

Solution B: Increase Platform Accountability (YouTube Handles Appeals)

- Problem: Appeals currently go to the copyright claimant, who has financial incentive to reject appeals
- **Shift appeals to a YouTube-run independent review team**
 - Trained reviewers evaluate:
 - Accuracy of claim
 - Context of the content
 - Fair use considerations
 - Potential abuse by rights holders
- Why this helps
 - Reduces power imbalance
 - Ensures neutral review
 - Increases trust and transparency
 - Aligns with legal standards around fair use
- Limitation
 - Requires staff training and hiring
 - Appeals still may not prevent external DMCA lawsuits
 - Fair use remains a complex legal question

Auditor Summary: Final Recommendations

1. Improve transparency
 - a. Show match confidence scores
 - b. Disclose reference file owners
 - c. Provide timestamp of exact match
2. Reform appeals process
 - a. YouTube handles appeals -> not rights holders
 - b. Independent human review
3. Protect creators from false claims
 - a. Evidence submission portal
 - b. Penalties for abusive claimants
4. Redesign fair use workflow
 - a. Human-first review
 - b. Prioritize commentary/transformation tags
5. Invest in creator education
 - a. Clear resources on licensing + fair use

References



Despotov, A. (2016, May 31). *Fox Took a Video From a YouTube User's Account for Its Show and Later Claimed Infringement Over It*. The Fried Firm. <https://www.friedfirm.com/blog/fox-took-a-video-from-a-youtube-users-account-for-its-show-and-later-claimed-infringement-over-it/>

Electronic Frontier Foundation. (n.d.). *DMCA*. Retrieved December 1, 2025, from <https://www.eff.org/issues/dmca>

Google. (n.d.). *How Content ID works*. YouTube Help. Retrieved December 1, 2025, from <https://support.google.com/youtube/answer/2797370?hl=en>

Harmon, E. (2016, February 26). *Content ID and the Rise of the Machines*. Electronic Frontier Foundation. <https://www.eff.org/deeplinks/2016/02/content-id-and-rise-machines>

Harvard Office of the General Counsel. (2023, July 11). *Copyright and Fair Use*. <https://ogc.harvard.edu/pages/copyright-and-fair-use>

Israel, S. (2016, May 8). *YouTube to Update "Content ID," Bringing Hope to Professional YouTubers*. CBC News. <https://www.cbc.ca/news/science/youtube-content-id-1.3569899>

Marshall, A. (2025, August 21). *Fix a YouTube copyright claim fast (What you NEED to know)*. Free Music for YouTube Videos & Social Media • No Copyright. <https://helothematic.com/so-you-received-a-copyright-claim-on-your-youtube-video/>

Scott, T. (Director). (2020, March 23). *YouTube's copyright system isn't broken. The world's is.* [Video recording]. <https://www.youtube.com/watch?v=1Jwo5qc78QU>

Trendacosta, K. (2020, December 10). *Unfiltered: How YouTube's Content ID Discourages Fair Use and Dictates What We See Online*. Electronic Frontier Foundation. <https://www.eff.org/wp/unfiltered-how-youtubes-content-id-discourages-fair-use-and-dictates-what-we-see-online>

U.S. Copyright Office. (n.d.). *The Digital Millennium Copyright Act*. Retrieved December 1, 2025, from <https://www.copyright.gov/dmca/>