

Copyright and Content Platforms Lab

[Note: formatting with a markdown document was quite messy, so I have kept the write-up as a .pdf file. The photos are, however, uploaded separately.]

Task I.

1. For this assignment, I have chosen to work with investigating the copyright policies on YouTube. YouTube uses an identification system, Content ID, that can be used by copyright owners. Content ID is behind the look-up process that determines if uploaded content infringes on copyright. Potential copyright owners can have their content stored in a database of copyrighted material. Then, the automated detection system scans every uploaded video against this database. In the process, YouTube essentially creates digital ‘fingerprints’ of copyrighted audio and video and compares uploads against the reference fingerprints. Therefore, flagging is carried out automatically (within just seconds or minutes of upload). Additionally, copyright owners can also manually submit DMCA takedown notices for content not in the Content ID system.
2. When Content ID detects a match, the copyright owner has chosen in advance how to handle it. They can either block the video, making it unavailable across the platform. Alternatively, they can monetize the video by running ads on it and collecting the revenue. The owner also has the option to just monitor the viewing statistics without taking action or can simply mute the copyrighted audio portion. The uploader receives a copyright claim notice explaining what was matched, which copyright owner claimed it, what action was taken, and the time codes of the matched content. A Content ID claim is, however, not a copyright strike. The infringing account standing is not affected, but the video could, of course, be blocked.
3. YouTube offers several options to appeal a copyright claim. First, the infringer can remove the claimed content by trimming out the copyrighted portion. There is also the option to replace the audio or to dispute the claim if the infringer believes they have the right to use the content.

4. Copyright owners can choose to monetize videos that use their content and revenue goes to the copyright owner, not the uploader. Multiple copyright claims means revenue is split among claimants. The uploader cannot monetize videos with copyright claims (unless they have special agreements).
5. YouTube does have certain partner program opportunities that require a clean copyright record and prohibit monetizing content you do not have rights to. One example program is a copyright match tool. It is available to creators with 100,000+ subscribers and finds re-uploads of original content. Creators can request removal of unauthorized copies. However, this is of course less sophisticated than Content ID. Additionally, there are music licensing programs. YouTube's Audio Library provides free music for creators and involves partnerships with music publishers for cover song licensing. Finally, premium subscribers' fees are distributed to creators whose content they watch, meaning that copyright claims affect revenue distribution.

Task II.

1. [Raw Copyrighted Clip] For the first piece of content, I used a clip from the Tour de France (held by FloBikes): <https://www.youtube.com/watch?v=JTCpcFNS2-E>. I tried first uploading a five-second clip (just the first five seconds of the video). This went through successfully. Then, I tried uploading the entire five-minute video without any alterations. This also succeeded and, surprisingly, was not taken down the last time I checked this:

[See 'Task 2, 1.jpg' and 'Task 2, 2.jpg']

[See 'Task 2, 3.jpg' and 'Task 2, 4.jpg']

2. [Remix/Mashup] Here, I used the same Tour de France clip (held by FloBikes) combined with an alternative commentary from the same Tour de France stage: <https://www.youtube.com/watch?v=BdZDAD4NnMo>. I uploaded the full video overlaid with the new commentary and again did not find any issues.

[See 'Task 2, 5.jpg']

At the last time I checked all potentially-problematic videos were still on the platform:

[See 'Task 2, 6.jpg']

Task III.

1. [Direct Reference] For the first piece of AI content, I used ChatGPT to try to create direct reference images to various cartoon characters (including Mickey Mouse, Spider-Man, and SpongeBob) but ChatGPT refused these requests with a response of the form: "I'm sorry — I can't generate images of copyrighted characters like [cartoon character]."
2. [Style Mimicry] For the first piece of AI content, I used ChatGPT for image generation. I used the following prompt: "create an image in the style of Studio Ghibli." ChatGPT successfully generated this image without copyright warning, but when I tried the same prompt with Banksy, I was faced with the output that such a request violated guidelines. This was the corresponding AI output:

[See 'Task 3, 1.jpg']

It was also successfully uploaded to YouTube without being flagged or blocked:

[See 'Task 3, 2.jpg' and 'Task 3, 3.jpg']

3. [Original Content] For the second piece of AI content, I used ChatGPT for original content generation. I used the following prompt: "create an image of a totally fictional cartoon character." ChatGPT successfully generated the image without a copyright warning:

[See 'Task 3, 4.jpg']

It was also successfully uploaded to YouTube without being flagged or blocked:

[See 'Task 3, 5.jpg' and 'Task 3, 6.jpg']

Task IV.

Experiment 1: Raw Copyrighted Clip

Purpose and character: This upload is non-transformative and non-commercial. I simply re-uploaded FloBikes' copyrighted content without modification, commentary, or any added creative elements. The purpose appears to replicate the original work rather than transform it.

Nature of the copyrighted work: The Tour de France footage is highly creative content involving professional sports broadcasting with commentary, camera work, and production value. While it documents factual events (the race itself), the presentation is creative expression.

Amount and substantiality used: For the 5-second clip, I used a minimal portion of the original work. For the 5-minute upload, I used the entirety of the copyrighted segment. The first clip would probably pass as fair use for this reason.

Effect on the market: My re-upload could serve as a substitute for the original, potentially reducing views on FloBikes' official content and undermining their licensing agreements with Tour de France organizers. This weighs strongly against fair use.

YouTube did not automatically take this down, which certainly reflects a gap between legal standards and the efficacy of automated enforcement.

Experiment 2: Remix with Alternative Commentary

Purpose and character: This experiment combines two copyrighted sources (FloBikes video footage with alternative commentary audio), only introducing moderate sampling. While remixing can be transformative, it is unclear if this would be considered fair use or if Content ID simply failed to pick up a match.

Nature of the copyrighted work: Both sources are creative, professionally produced content, generally weighing against fair use.

Amount and substantiality used: Substantial portions of both copyrighted works were used in their entirety.

Effect on the market: This remix could substitute for either original work and potentially harm both copyright holders' markets.

Without receiving licensing permission from a network with Tour de France rights, this may not be considered fair use. Then again, a number of channels do add commentary over a sports event, potentially weighing in favor of fair use.

Some Relevant Case Law

Google v. Oracle: Even though this case dealt with software APIs, it set the precedent that transformativeness and purpose matter significantly. Google's use of Java APIs was fair use because it served a different function than the original. My re-uploads serve essentially the same purpose as the originals (viewing of Tour de France content) undermining any fair use claim. As a result, it seems to be *in spite* of the relevant law that they were able to stay up.

Lenz v. Universal Music: This case established that copyright holders must consider fair use before issuing DMCA takedowns. This seems relevant to the context provided at the outset: understanding YouTube's Content ID system, which allows copyright holders to make monetization/blocking decisions in advance without case-by-case fair use consideration.

Some Observed Gaps

Gap 1: Detection vs. Legal Standards

Copyright infringement requires substantial similarity and access, but fair use provides an affirmative defense based on the four-factor test.

YouTube's Content ID automatically detects matches against reference files and allows copyright holders to preemptively choose how to handle matches.

Neither of my clearly infringing uploads were detected or flagged by Content ID, despite using copyrighted sports content from professional sources. This suggests that either FloBikes may not have registered this specific content in Content ID's database or potentially that Content ID fingerprinting may not have matched due to quality differences. After all, I downloaded these at pretty low resolution.

Gap 2: Scope of Protection

Copyright law protects specific expressions but not facts, ideas, or functional elements. Sports broadcasts receive copyright protection for their creative camera work, commentary, and production choices.

YouTube relies on copyright holders to accurately claim only content they own and to consider fair use.

The lack of automated detection for regional sports content shows that smaller copyright holders have less practical protection than major studios and record labels.

Gap 3: AI-Generated Content

Copyright law is unclear about AI-generated content. Current U.S. Copyright Office guidance suggests that purely AI-generated works may not be copyrightable because they lack human authorship, though works with significant human creative input may qualify. YouTube's policies fail to specifically address AI-generated content differently from human-created content. Content ID would only flag AI outputs if they match existing copyrighted reference files.

My AI-generated images uploaded successfully without flags, even when created "in the style of Studio Ghibli." ChatGPT itself enforced stricter boundaries than YouTube by refusing to generate images of specific copyrighted characters.

Attached are all of the links to the uploaded videos (none of which were, perhaps surprisingly, taken down):

5-Second Clip: <https://youtu.be/X8sXGMgP2-Y>

Full Clip: <https://youtu.be/DiK43l9Q0y4>

Mashup Clip: <https://youtu.be/XLMsq-cN2Js>

Studio Ghibli-Inspired: <https://youtu.be/6BanGxeCXTA>

Original Content: <https://youtu.be/L-1BlkjeX4E>