The insurmountable problems with generative AI opt-outs

Ed Newton-Rex

November 2024

Introduction

Generative AI companies wish to train AI models on copyrighted works, but rights holders view doing so without a licence as unacceptable.[1] Some propose as a solution to this implementing an opt-out scheme, which would allow rights holders to opt out of generative AI training, but which would allow AI companies to train on any works not opted out.[2]

However, opt-out schemes of this nature are both hugely unfair to creators and rights holders, and woefully ineffective in practice. This essay outlines a number of the inherent issues with opt-out schemes for generative AI training.

1. It is impossible to opt out downstream copies of your work.

There are two approaches to opting content out of AI training, which have been described as location-based and unit-based. Location-based opt-outs indicate that certain URLs should not be used for training; unit-based opt-outs involve adding metadata to content, indicating that it is opted out.[3] Neither are remotely effective at controlling the usage of downstream copies of rights holders' works in generative AI training.

The most widely used opt-out schemes are location-based: for instance, robots.txt, which contains information about how the web domain where it's hosted can be accessed by web crawlers. These only work for web domains you control. But your work is likely to appear in many other places: for instance, a photographer's photo may feature in an ad on another website; a composer's song may be used in an online video; screenshots of a newspaper article may be shared on social media. You

[1] https://www.aitrainingstatement.org/
[2] https://www.ft.com/content/26bc3de1-af90-4c69-9f53-61814514aeaa
[3] https://openfuture.eu/wp-content/uploads/2024/05/240516considerations_of_opt-out_compliance_policies.pdf

have no control over whether these downstream uses of your work are opted out of generative AI training, since you do not control the URLs where they're hosted.

Unit-based opt-out schemes, on the other hand, add metadata to content that indicate the content's opt-out status. But metadata is easily removed (e.g. X removes EXIF metadata when images are uploaded)[4], and, besides, there is no way to add metadata to text. Unit-based opt-out schemes fail to opt out downstream copies where metadata cannot exist in the first place or does exist but is subsequently removed, which are common.

There are no effective opt-out schemes that reliably opt out *content itself* from training. Doing so is impossible. One theoretical solution could be maintaining a central directory of opted-out works and using automatic content recognition (ACR) to scan datasets for opted-out content; but even this does not work, as ACR cannot reliably identify opted-out works. Relying on ACR for opt-out information would allow modified versions of opted-out content to be trained on, as well as individual copyrighted elements that make up part of a larger copyrighted work (e.g. the composition in a sound recording, the lyrics in a song, or the dialogue in a film), bootlegged recordings of live performances, and works transposed from one medium to another (e.g. screenshots of text). None of these would be caught by automatic content recognition systems scanning for the opted-out works.

So opt-out schemes only work for web domains you control. It is telling that by far the most common method for opting out of generative AI training to date has been robots.txt, and this gives rights holders no control whatsoever over how downstream copies of their works are used. There are no known solutions to this issue: opt-out simply does not work.

2. Most people miss the chance to opt out.

The evidence suggests that the majority of people who have the option to opt out of generative AI training do not realise they have that option.

Take-up for opt-out schemes already in existence is low. Today, the most common way of opting out of generative AI training is by blocking AI crawlers in your website's robots.txt file. A Cloudflare study from July 2024 found that only 16% of the 100 top-visited sites hosted on Cloudflare blocked AI crawlers (and only 8.8% of the top 1,000).[5] Even among the most popular sites, opt-out take-up is relatively low: more than 40% of the top 100 English language news sites blocked no AI crawlers as of February 2024,[6] a full 15 months after the release of ChatGPT - and these are the companies you would expect to be the most likely to be aware of opt-out availability.

[4] https://us.norton.com/blog/how-to/how-to-remove-gps-and-other-metadata-locations-from-photos
[5] https://blog.cloudflare.com/declaring-your-aindependence-block-ai-bots-scrapers-and-crawlers-with-a-single-click/
[6] https://pressgazette.co.uk/platforms/news-sites-block-ai-web-crawlers-chatgpt-google/

When companies run opt-outs, they generally don't share opt-out numbers, but, when they do, we see similarly low take-up: for instance, when AudioSparx ran an opt-out of a training deal with Stability AI for its members, only 10% opted out.[7]

These low take-up numbers don't tally with the data we have on creators' and rights holders' opinions of how their content should be used by AI companies. Creators and rights holders overwhelmingly believe permission and payment should be required for generative AI training. In a poll of 4,274 songwriters, composers, and publishers in Australia & New Zealand, 95% said copyright holders must be asked for permission before their works are used as input for AI systems, and 93% called for remuneration.[8] Polls of authors[9] and artists[10] show roughly the same numbers. Cloudflare reports that, when Cloudflare users take action on AI crawlers, 85.2% block them, while only 10% allow them (and they conclude that "we hear clearly that customers don't want AI bots visiting their websites").[11]

Why this huge discrepancy between rejecting unlicensed training in principle and actual opt-out take-up? Because people don't realise they can opt out. For example, when Udemy ran an opt-out scheme, 404 Media reported that many Udemy creators missed the window to opt out, and only heard about it after it had closed.[12]

It is virtually impossible to picture an opt-out scheme that becomes well-known among the majority of the people who would be able to use it. Even the most widely-used AI opt-out method to date, websites' robots.txt file, is little-known and poorly understood. And even those who do hear about a given opt-out scheme need to prioritise opting out against competing priorities, leading some who would like to opt out to delay, possibly indefinitely. This makes opt-out take-up even less likely.

3. Opt-out is binary, which harms rights holders.

Large language models are becoming a common feature of search engines, with Google now providing AI summaries and with new products entering the market like Perplexity and SearchGPT. Appearance in search engines is critical to many publishers, since search engines are many internet users' entry point to the web.

AI companies that provide search functionality crawl the web for data to train on. Existing opt-out schemes are binary: you allow training or you don't. For example, you can provide a binary allow/disallow for certain crawlers via robots.txt; and you can assign a binary value to the

[7] https://musically.com/2023/09/13/stable-diffusion-maker-launches-stable-audio-text-to-music-ai/
[8] https://www.apraamcos.com.au/about/supporting-the-industry/research-papers/aiandmusic
[9] https://societyofauthors.org/2024/05/02/egm-resolutions-and-next-steps/
[10] https://cdn.dacs.org.uk/uploads/documents/News/Artificial-Intelligence-and-Artists-Work-DACS.pdf
[11] https://blog.cloudflare.com/declaring-your-aindependence-block-ai-bots-scrapers-and-crawlers-with-a-single-click/
[12] https://www.404media.co/massive-e-learning-platform-udemy-gave-teachers-a-gen-ai-opt-out-window-its-already-over

c2pa.ai_generative_training entry of the Coalition for Content Provenance and Authenticity (C2PA) metadata[13]. This means rights holders cannot distinguish between allowing their content to be referenced in search results (in order to direct users to their websites) and allowing their content to be trained on (which enables the resulting trained model to output material based on, and competing with, their content).

So, if opt-out schemes are adopted, publishers and copyright holders have only the illusion of choice. If they opt out of AI training, they opt out of being findable on the internet.

4. Emerging technologies like smart glasses render existing opt-out schemes useless.

Location-based and unit-based opt-out schemes, already woefully deficient for the reasons outlined above, completely fall apart in a world where training data is not gathered by crawling URLs or accessing content via the web. This world is already upon us.

Meta sells augmented reality glasses called Ray-Ban Meta Smart Glasses. These are selling extremely well.[14] They are equipped with cameras and a microphone, and Meta has confirmed that it may train AI models on data the glasses capture.[15] This means copyrighted works, captured via smart glasses, will be used to train AI models. There is no opt-out scheme in existence that can give rights holders control over which of their works can be used for AI training in this context: works ingested by smart glasses are neither accessed via a URL, nor are they accessed in a way that preserves metadata.

It is widely expected that smart glasses will see major consumer adoption over the coming years: the market is projected to grow at a CAGR of 53% from 2023 to 2030.[16] Any opt-out scheme brought into law now would totally fail to address AI training on content gathered by these types of devices, which has already started and which is expected to grow rapidly.

---

[13] https://c2pa.org/specifications/specifications/1.3/specs/_attachments/C2PA_Specification.pdf

[14] https://techcrunch.com/2024/10/21/metas-smart-glasses-outsell-traditional-ray-bans-in-some-stores-even-before-ai-features-roll-out/

[15] https://techcrunch.com/2024/10/02/meta-confirms-it-may-train-its-ai-on-any-image-you-ask-ray-ban-meta-ai-to-analyze/

[16] https://www.globenewswire.com/news-release/2024/07/09/2910156/28124/en/Global-Smart-Augmented-Reality-AR-Glasses-Business-Analysis-Report-2023-2030-Growing-Technology-Commercialization-Activity-to-Benefit-Market-Developments-and-Expansion.html

5. The changing landscape of web crawlers makes opt-outs impossible to keep up with.

Many rights holders will want to allow some AI companies to train on their work, but not others. So they need a way to control who can and can't access their work for training. But, under an opt-out scheme that gives rights holders this control, there will always be periods of time where companies they don't want to train on their works are able to do so.

The landscape of web crawlers is confusing and continuously changing. When new AI companies are founded, they often develop new crawlers; and established companies sometimes replace their crawlers.[17] Many rights holders block the crawlers they've heard of, but fail to block many others.[18] Even the best-resourced rights holders fail to block all crawlers.[19] And a hypothetical, well-resourced rights holder that managed to block all crawlers would still find there was a delay between new crawlers being launched and its ability to identify and restrict them, during which time their works could be accessed for training.

This problem cannot be solved under an opt-out scheme that seeks to give rights holders any control greater than simply opting out of all generative AI training. Opt-in is the only way to give rights holders true control over which companies can train on their works.

6. Introducing opt-outs implies exoneration of historical infringement, and even models and companies that 'respect opt-outs' often still benefit from opted-out works.

Most generative AI opt-out schemes are introduced after the AI companies that commit to respect them have started training and releasing models. But no opt-out scheme has, on its introduction, required these companies to retrain their models from scratch with no trace of the opted-out works in their training data. This means that the introduction of opt-out schemes implicitly exonerates historical infringements potentially spanning years, in which time the AI companies benefited from unlicensed copying of copyrighted works while the rights holders had no chance to opt out.

Even if AI companies are required to retrain their models without directly training on opted-out works, they will likely still use previously-generated synthetic data to train future models. Synthetic data is data generated by another AI model, and is widely used already in generative AI training.[20] The models used to generate synthetic data are often themselves trained on copyrighted works. Unless the AI company also recreates their synthetic data from scratch (which has never happened

[17] https://www.404media.co/websites-are-blocking-the-wrong-ai-scrapers-because-ai-companies-keep-making-new-ones/
[18] https://blog.cloudflare.com/declaring-your-aindependence-block-ai-bots-scrapers-and-crawlers-with-a-single-click/
[19] https://pressgazette.co.uk/platforms/news-sites-block-ai-web-crawlers-chatgpt-google/
[20] https://www.ft.com/content/053ee253-820e-453a-a1d5-0f24985258de

as a result of an opt-out scheme), then even after the introduction of an opt-out scheme they will continue to use synthetic data that was likely created using models trained on the opted-out works.

Further, opt-outs do not take immediate effect. Many models are trained well in advance of being released (for example, GPT-4 finished training in August 2022[21], but wasn't released until March 2023 and was only made generally available to all developers in July 2023[22]), so, after a rights holder has opted out, models that were trained on their works will continue to be released; models are rarely replaced immediately upon a replacement becoming available (as of November 2024, GPT-4 is still live, despite OpenAI having released newer models[23]), and certainly not as a result of an opt-out being introduced; and, in the case of open source models, models that have been released cannot be rolled back (i.e. they will be available forever). So there is always a delay between a rights holder opting out of training and models that use their works ceasing to be available, sometimes a significant one. The rights holder has no control over the length of this delay. (Notably, this is in stark contrast to copyright, which, in countries like the UK, arises automatically when a work is created and fixed in a material form.)

Once again, opt-out schemes fail to give rights holders true control over how their works are used.

7. The administrative burden of opting out all your works is huge.

Creators generally publish and license large numbers of works across a variety of platforms and over a long period of time. As an example, I have published music in audio format on MySpace, Bebo, SoundCloud, Facebook, X, Spotify (via two different distributors), Apple Music, and my website; I have published sheet music on MuseScore, Medium, on my website, and via two music publishing companies; I have published writing in various newspapers and journals, academic papers, blogs and websites. On top of this, I have licensed music to various companies and events. (And no doubt there are places my work has surfaced that I've forgotten.) If I want to opt out of all generative AI training on my work - which I do - it would be incredibly difficult to do so, and a significant investment of time (if indeed it's even possible). And not only is an opt-out scheme asking creators and rights holders to take on this burden; it is also asking them to go through the same opt-out process for every future work they publish. Again, this is in stark contrast to copyright, which in countries like the UK arises immediately and automatically.

There is no solution that would successfully opt out all of a rights holder's works from all generative AI training with a single decision. So the introduction of an opt-out scheme inevitably puts the huge burden of opting out each individual work, including historical works, on the rights holder.

---

[21] As described in the video at https://openai.com/index/gpt-4/

[22] https://openai.com/index/new-models-and-developer-products-announced-at-devday/

[23] https://platform.openai.com/docs/models/o1

8. Opt-out deadlines put undue pressure on rights holders and limit real control.

By their nature, opt-out schemes impose a deadline on rights holders to make a decision about how their work is used. AI companies follow ad hoc schedules for training their models, and if the rights holder has not opted out before the training run begins, they have missed the chance to opt-out until at least the next training run, which may be months or a year or more away. Whether an opt-out scheme names a specific deadline or not, there is time pressure on rights holders to make a decision, and if they miss the chance then their work will be included in generative AI training for some period of time against their will.

This inherent time pressure is detrimental to rights holders' ability to weigh up the pros and cons of opting out. Further, it limits the real control they have: any training runs that happen before they have made a decision will likely use their work, and the AI company in question will then benefit from exploiting their work until that model is retired.

9. The implications of opting out are poorly understood, lowering take-up.

Low opt-out take-up is partly due to hesitation among rights holders driven by incomplete information. In some opt-out schemes, the chance to opt out of training may be the first time a rights holder has ever seriously heard of or considered generative AI and its implications. Rights holders may be worried about the implications of opting out (e.g. how it will affect their inclusion in search, as discussed above), or they may simply prefer a wait-and-see approach in which temporary inaction is favourable to deciding one way or the other. Requiring rights holders to opt out of generative AI training inevitably results in some who would like to opt out to hesitate in doing so, due to lack of information

10. Smaller rights holders are even more likely to miss the chance to opt out.

While opt-out schemes are unfair to all rights holders, they are more unfair to smaller rights holders and individuals, such as creators. This is because smaller rights holders, without the resources to devote to paying attention to opt-out schemes, are less likely to exercise their right to opt out.

When Cloudflare analysed the top 1 million websites by number of visits, they found that the percentage blocking AI crawlers increased with increasing website visits.[24] Similarly, the Data Provenance Initiative found that while 25% of data from the 'highest-quality sources' in some AI training sets had been restricted, only 5% of all data in those training sets had been restricted.[25] These findings support the logical hypothesis that smaller rights holders such as creators are less

[24] https://blog.cloudflare.com/declaring-your-aindependence-block-ai-bots-scrapers-and-crawlers-with-a-single-click/
[25] https://www.nytimes.com/2024/07/19/technology/ai-data-restrictions.html

likely to make use of opt-out schemes, and means they can be highly disadvantaged by such schemes.

Top N Internet properties by number of visitors seen by Cloudflare	% accessed by AI bots	% blocking AI bots
10	80.0%	40.0%
100	63.0%	16.0%
1,000	53.2%	8.8%
10,000	47.99%	8.92%
100,000	44.53%	6.36%
1,000,000	38.73%	2.98%

Source: Cloudflare blog, 2024-07-03

Conclusion

Opt-out schemes for generative AI training don't work. Under existing opt-out schemes and any hypothetical future opt-out schemes, it is impossible for rights holders to successfully opt their works out of training. Location-based and unit-based opt-outs are woefully inadequate, and there are no better solutions proposed. Opt-outs give rights holders the illusion of control - nothing more.

Even if an opt-out scheme could be conceived of that truly gave rights holders control, putting the burden of opting out on rights holders is evidently unfair. Opt-out schemes inevitably lead to many works being used in training whose owners don't want them to be used: whether because rights holders don't realise they can opt out, or because the administrative burden of opting works out is far too great, or because the changing landscape of web scrapers means people fail to opt out successfully, or because the binary choice and lack of information leads to delay and indecision.

The only way to effectively ensure that rights holders' works are not used for generative AI training against their wishes, in a way that is fair to both rights holders and AI companies, is for training to be based on opt-in consent.