

Considering Creatives, Copyright and AI: A Response to the UK Government Consultation

April 2025

Convocation Research + Design, DACS and PICSEL

CONVOCATION
RESEARCH + DESIGN

DACS



Introduction

Some of the great strengths of the United Kingdom are its art, fashion, films, tv shows, video games, and creative culture, and the impact that creative culture has on the entire world. Moreso, the creative industries, and its workers, are a core component of the UK's economy. The creative industries, in the United Kingdom, are a multi-billion dollar industry, having [contributed over 124 billion pounds to the UK's economy in 2022](#) alone. However, there have been heavy disputes on how the creative industries are being negatively impacted by the rise of generative AI, and its ingestion of copyrighted material from web scraping and large datasets, such as [LAION 5-B](#). Surveys conducted by the [Association of Photographers](#) and [Association of Illustrators](#) in 2025 found that 58.1% and 32.40% of respondents, photographers and illustrators respectively, have already lost commissioned work to generative AI. One of PICSEL's members reported their average product placement revenue dropped from £70,523 in 2022 to £18,864 in 2024 – a staggering 73% drop in revenue.

The goal of this paper is not to pit every instance of AI against creatives and artists, as the UK is home to some of the most world renowned [AI artists who make their own artistic datasets and train their AI models](#). Rather, this paper aims to outline policy and technology steps forward around transparency, rights reservation and labelling, based on input from workshops with 26 experts in arts, law and technology and analysis of pre-existing policy. In the first sentence of the UK Government's Consultation on AI and Copyright, the creative industries are described as a core sector of the UK, along with AI. The consultation document emphasises the current 'status quo' of mass scraping of digital content for AI models and the disputes related to that scraping cannot continue, as it "risks limiting investment, innovation, and growth in the AI sector, and in the wider economy. It effectively prevents creative industries from exercising their rights," which this research paper agrees with. The UK government has the time to decide and make new forms of policy to benefit both sectors, supporting both the creative industries, and AI. With that in mind, whatever decisions the UK government makes, it should ensure the steps are not arduous for users to engage with or riddled with [harmful design patterns that could cause friction, confusion, and user fatigue](#).

This paper aims to highlight that new forms of engagement are possible, that allow for creatives' work to be used with their consent, and still allow for AI innovation. Recently, the [Serpentine Galleries built a collaborative data choral data set](#) for an exhibition with AI artists Holly Herndon and Mat Dryhurst featuring the voices of 150 choirs across the UK. This choral and vocal based data set, however, allows for all participants to weigh in on the usage of the dataset, returning consent to the collective as a form of 'bottoms up data governance.' While an artistic project, this experiment exists as a proof of concept illustrating that new

forms of data governance, and data stewardship are possible, and that art making is not necessarily at odds with AI development. It's this kind of creative and innovative problem solving towards data governance, stewardship, and policy that the UK Government should explore.

Expert Engagement and Workshop Design

From 1 December 2024 to 20 February, 2025, Convocation Research + Design (CoRD Labs) was commissioned by collective rights management organisations, DACS and PICSEL, to create a research report in response to the government's consultation, to surface from stakeholders across the creative sectors and policy experts what considerations, frameworks, and policy should inform the consultation in regards to rights reservations, transparency and labelling of data.

For this project, CoRD Labs conducted ecosystem mapping and analysis on generative AI consumer facing products; auditing and analysing current opt-out systems in regards to AI and/or UK/EU regulation (such as GDPR); designed and facilitated two workshops to gather expert feedback and insights. Workshop One was in person on 30 January, 2025 with 18 participants at DACS office in London, and the second workshop was held online on 14 February, 2025 with 8 participants. Workshop participants included legal experts, educators, curators, funders, artists, policy experts, data scientists, and technologists. Participants included artists, members of DACS and PICSEL, British Council, the Open Data Institute, Tate Modern, National Portrait Gallery, the Serpentine Galleries' Art and Technology team, Spawning, and others.

The workshop's structures, surveys, and activities were designed around the following questions from the Government consultation. These questions serve as the main focal points of our research and this research report:

Rights reservation

6. What action should a developer take when a reservation has been applied to a copy of a work?
8. Do you agree that rights should be reserved in machine-readable formats? Where possible, please indicate what you anticipate the cost of introducing and/or complying with a rights reservation in machine-readable format would be.
9. Is there a need for greater standardisation of rights reservation protocols?
10. How can compliance with standards be encouraged?
11. Should the government have a role in ensuring this and, if so, what should that be?

Transparency

17. Do you agree that AI developers should disclose the sources of their training material?
18. If so, what level of granularity is sufficient and necessary for AI firms when providing transparency over the inputs to generative models?
19. What transparency should be required in relation to web crawlers?
20. What is a proportionate approach to ensuring appropriate transparency?
21. Where possible, please indicate what you anticipate the costs of introducing transparency measures on AI developers would be.
22. How can compliance with transparency requirements be encouraged, and does this require regulatory underpinning?
23. What are your views on the EU's approach to transparency?

labelling

40. Do you agree that generative AI outputs should be labelled as AI generated? If so, what is a proportionate approach, and is regulation required?
41. How can government support development of emerging tools and standards, reflecting the technical challenges associated with labelling tools?
42. What are your views on the EU's approach to AI output labelling?

Methodology

This consultation carried out a multi-faceted approach integrating ecosystem mapping, regulatory analysis, and stakeholder engagement. First, we conducted an ecosystem mapping and analysis of generative AI consumer-facing products, identifying key trends, risks, and regulatory gaps. In parallel, we audited and analysed existing opt-out mechanisms in relation to AI and UK/EU regulations, including GDPR, to assess their effectiveness and transparency.

To incorporate diverse expert perspectives, we designed and facilitated two workshops. The first workshop was held in person on 30 January, 2025, at DACS' London office with 18 participants, while the second took place online on 14 February, 2025. The workshops were structured around key questions from the Government consultation, with activities and discussions designed to surface considerations, frameworks, and policy recommendations related to rights reservation and data transparency. Insights from these engagements directly informed the findings and recommendations of this report.

Takeaways and Findings

“The more I reflect, the more frustrated I feel that the consultation is so stacked in favour of one obvious, not very imaginative or innovative route forward” – a curator with a background in data and technology, reflecting on the nature of the government consultation’s questions.

As this paper has established earlier, our research and workshops are not anti-AI or anti-innovation, and participants were similarly nuanced in their input, feelings, and suggestions. It is a false equivalency to assume AI innovation must come at the expense of creatives and their work, a sentiment that repeatedly appeared in our results. Some participants who are creatives and artists expressed a deep interest in AI development; others talked about the importance of institutions in supporting the rights of works and, then how those works are used as data within AI systems with consent and compensation from the artists. Towards the end of workshop 1, the conversation shifted to the importance of new opportunities and developing solutions for licensed AI for both the creative sector and the UK economy. Below are the major themes that emerged from our ecosystem mapping research, policy analysis, workshops and corresponding surveys.

Major Themes:

- **AI fatigue:** participants highlighted that conflating all of AI in this conversation is unhelpful. It is best to just focus on generative AI, and separate that type of AI from other AI innovations. One participant pointed out that “not all AI is generative AI. Some AI, like assistive AI is beneficial to consumers while Generative AI which is not as it is assimilating our work, style”, etc.
- **The Pros and Cons of Standardisation:** participants highlighted the need for standardisation to be able to create more ease of use across products and to hold platforms accountable. However, as participants also discussed, a one size fits all approach could be challenging to implement and result in a one size fits none result.
- **A need for nuanced control and granularity**– participants repeatedly stressed a need for nuance and granularity across any proposed opt out/opt in and transparency mechanisms. Creatives want the ability to make granular and nuanced decisions about their images and better understand how their work is being used, by whom, when, and where.
- **Accessibility and terminology:** Participants stressed the need for easily understandable information about where their work has been used, without requiring specialised expertise. The language choices and overall accessibility of mechanisms including opt out/opt in and transparency, must be prioritised.
- **Accountability of Institutions:** Participants emphasised the accountability of institutions is needed in ensuring the proper use of rights holders' works.

- **Need for Concrete Plans:** Participants expressed frustration with the lack of concrete plans from the government regarding transparency measures, and next steps from the consultation.
- **Fears that a decision has already been made:** The tone of the consultation's questions lead to multiple participants feeling as though the government has already decided what paths to take forward, specifically with opt out and rights reservation.
- **Questions over responsibility:** Participants raised questions over whom or what agencies, and organisations will be in charge of managing opt outs on behalf of creatives, and what agencies or regulators will be responsible for when rights of images, opt outs, etc are violated. Who is in charge and what responsibility they have deeply matters in this context.
- **Using AI for 'Good':** Participants raised questions around what roles can AI technology play to alleviate frictions with opt out and make it easier to opt out (image hashing, etc). Perhaps, this is a space where AI can be used for good.
- **Supporting Creativity through Regulation:** Participants want to ensure AI models allow more space for creativity and diverse aesthetics, and regulation should support that.

Further Themes Surfaced from Research

Other themes emerged in our research. Each section exists as concerns participants had, along with insights, suggestions, and provocations for next steps, further research, and considerations for government to centre in their future work in regards to generative AI and artists' copyright.

How Will This Work?

Almost every theme below will touch on this point, but this was a common sentiment expressed repeatedly by participants: how will "this" work? "This" refers to a myriad of topics, but under the general umbrella of once there is regulation or legislation, what are the plans, how will the plans unfold, and how will it impact artists who work with datasets, collective rights management organisations, lawyers representing cultural and creative institutions, curators working with artists' and their artifacts, artists who want to opt out, artists who want to opt in, artists who want to opt out and opt in some works, large institutions, academia, and technology companies. There is a need for clarity and transparency in how plans will unfold, and how next steps will be determined.

- **Need for concrete plans:** Participants expressed frustration with the lack of concrete plans from the government regarding transparency measures.

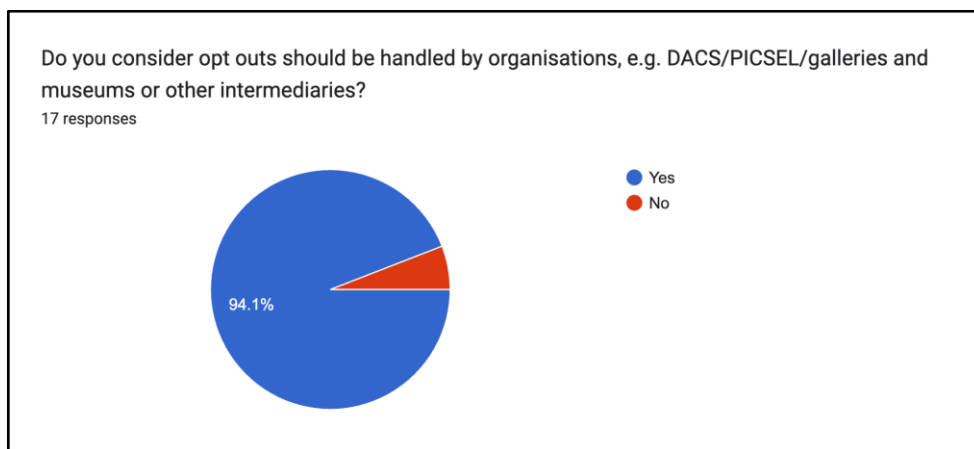
- Need for more investigation, research, engagement and time from the government spent on this topic and providing next steps. Participants seem worried a decision would be made quickly and there would not be enough 'time' to consider all kinds of solutions and alternative solutions.
- A feeling that the government has already made a decision: Participants felt as though the government is already leaning towards opting out as the only way for creatives to have consent and agency over the use of their works from the focus of the consultation's questions.

Who is Responsible?

Participants repeatedly questioned and wondered who *would* be responsible, and who *should* be responsible for managing the opt out/opt in, rightsholders and responsibility of stewarding the works as data sets.

- Centralised Agency or Collective Rights Management: Participants suggested that pre-existing CMOs could step in to handle this management, or that a new centralised agency or regulator could be created. However, as one participant noted, even a new agency or pre-existing CMOs would not be able to address those that forget, cannot, or are not aware that they can opt out, even though those creatives should be compensated or offer consent in some way shape or form. Further in this report, we address the shortcomings of opt out.
 - One participant suggested, "If one of the implications of the consultation is that it places a burden on artists to adjust their workload and think with the consumption of AI first, does this not give rise to a market opportunity to play that role to manage the data infrastructure on behalf of its customers or artists. Presumably, that was one of the reasons that gave rise to CMOs in the first place – managing their rights or licensing wasn't at the forefront of artists' mind or didn't fit into their workflows. This could be a natural evolution of the role of the collecting society of CMO that we're talking about. In 10 years' time we could see a lot of these preference signals managed by intermediaries rather than artists having to adjust metadata themselves."
 - Another participant stated, "I think the idea of a central registry that you could upload all your images to is relatively doable but not ideal. GenAI systems are only as good as the quality of the art they are trained on. The registry still would not address the issue of systems being trained on everyone's work that doesn't opt out, even though they should be paid. The registry is the best option I've heard of so far."

- Creating a new regulator: participants raised the idea that a new regulator could be created specifically to address and investigate AI companies in future regulation, as well as tackle opt in/opt out for creatives and artists.
 - One participant wondered, “What are the tools regulators have to check against all this, this is where the pipeline of registries comes into place? Are auditors responsible for data science now?”
 - Another participant emphasised, “The most basic issue to me is that every company transparency has a corresponding auditor, and that might be a public expectation or some kind of expectation for a technical auditor. Identifying who that is and clarifying who that auditor is and their capacity to actually do the work is key. Underlining the importance between transparency of the system and training data, and explainability of models is also key.”
- New forms of data governance or licensing: participants also suggested the notion of moving beyond opt in or opt out, and exploring new forms of data governance and licensing. As mentioned earlier, this type of proof-of-concept has already been explored with the Serpentine Galleries’ Choral Data Set.
 - One participant explained the value of new models by sharing, “There’s many opportunities for case-by-case licensing for bespoke AI models, rather than the current practice of scraping all works available online and distilling the information.”



From our survey conducted on January 30, 2025

Responsibility Towards Infringement

Once responsibility is determined towards a series of institutions or a new regulator, participants highlighted a need for responsibility and measures against infringement to help

enforce regulation. Participants stressed the importance of holding companies accountable, and suggested fines as an example of that type of accountability.

- There must be a way to hold companies accountable for opt out infringements
 - One participant wrote in our survey that “there should be a way to report it to the company running the model, and then escalate to another organisation if necessary.”
- Other potential punishments, like fines, also surfaced in our survey. Below are answers from the workshops’ participants.
 - Compensation/fines
 - If unauthorised use exists, there should be similar financial penalties to GDPR
 - I think a central organisation, like the ICO, in charge of managing the opt outs and auditing use. There should be fines levelled if not complying with transparency. Fines can help pay for the opt out database.

Don’t Forget Creatives and Small Start Ups

Participants also highlighted the importance of the UK’s small but mighty and innovative start-ups (those at focus on AI and those that use data), new galleries, new institutions, and new creative companies started by creatives. Where will these important users sit in this wider conversation of regulation? Participants expressed concern that the conversation is only focusing on how larger companies, like Open AI, could implement regulation. Participants stressed that creatives and small startups should not get lost in future regulation or decision making.

- **Supporting Creativity through Regulation:** Participants want to ensure AI models allow more space for creativity and diverse aesthetics, and regulation should support
- **Creatives CAN help AI:** Creatives can help be a part of AI innovation, as paid data creators.
 - One participant stressed, “If we could produce work which we profit from that’s specifically created for machines, so that rather than feeding its own tail on works which we’ve already made there could be an opportunity to create new works for this.”
- **Where do small organisations and small data sets sit in this regulation?** As stated earlier, participants are pro-innovation, and experimental thinking, as that is at the heart of creativity. With this in mind, participants emphasised a need to address the question of “where will small datasets and small organisations fit into the forthcoming regulation.” [This is something the government must consider and address as new innovation in AI technology is championing the agility and accuracy of using small data sets.](#)

- o One participant explained, “There are clear alternatives that exist, some artists train their own models using their own work and creating value in this manner. This experience of smaller datasets.”
- o Another participant suggested, “Smaller model developers should be able to negotiate and design bespoke models, so if there’s a biased system that favours the largest companies.”

Standardisation: The Pros and Cons

Participants spoke about the importance of standardisation as needed for ease of use and to hold platforms accountable, but also highlighted how a one size fits all approach could be challenging to implement and result in a one size fits none result.

- One participant wrote in our survey, “There has to be a standardised way of doing this because platforms have to be accountable, and looking at every single use may be challenging.”
- Another stated in our survey that “...one size fits all also wouldn’t work...”
- A third survey respondent wrote that “an ambition towards standardisation is a good thing, but the practicality of it is challenging.”

Using AI and Tech for Good

Participants spent time also brainstorming, as many have a background in policy and technology, on how AI can also be utilised to help track and document unapproved uses of images in large AI datasets. Some participants who work on Haveibeen trained explained how that tool functions, and how the implementation of metadata in works, image hashing etc can be used to better document misuses of creative works. Others highlighted potential policy issues related to the technology of HOW to determine if a creative’s work is used. This conversation highlighted that all technology is imperfect, including computer vision and AI; what if image hashing can only determine a partial likelihood of 80% that a creative’s work is similar to a file in a large dataset? A participant wondered how government, a regulator, or a third-party auditor/investigator would respond to any technology determining likeness or likelihood? This raises questions including: what technical tools the government uses and how that becomes platform policy; and what percentage matches are ‘enough’ to determine an image is used if third party auditors/investigators/regulators/government are using AI tools and hashing to help someone opt out? It’s worth considering this scenario: what about a slightly changed artifact, such as the many variations of [Pepe the Frog](#) or [Adele Morse’s “Stoned Fox?”](#)

How will government reconcile things that are 'almost' matching? What's the threshold for similarity? It's not enough to opt out a particular image or artifact when there can be variations of it, such as different crops, colour correction, someone applying a filter, etc.

These considerations need to be addressed, otherwise there's the potential risk of the opt out and related investigatory or auditing tools not working as well as they could or should.

Regardless, the potential of using AI to help track AI misuses is a promising avenue to consider, as long as corresponding policy is also considered with it.

- **Can Image Hashing Help?**

- One participant explained how Haveibeen trained works on image identification: "Haveibeen trained does –also have a reverse image search, to that point the underlying technology which is able to identify that is also being used. We also determine similarity scores and an assorted ranking based on these, and if there's any exact duplicates, we'll automatically opt those out. If there's similar images individuals have the ability to sift through those and add them to the donottrain registry; if your image was on a book cover it would likely be."
- Another participant raised how many different kinds of pre-existing technology can be useful for image identification, auditing, and opting out by saying "This is a huge concern that tech companies brought up around the advocacy around option 3 in the consultation. There are techniques e.g. perceptual hashing which allows you to transform the data, assign a hash to it, assign visual similarity scores to data and remove data from datasets – and there's open standards being developed around this. Combined with strong provenance, watermarking and perceptual hashing – which only works for some data types and modalities of data – there are still ways to identify works despite metadata stripping."

- **How Will Policy Consider Technology's Limitations?**

- One participant raised this important point, "But there are questions around what data would be excluded, would a 95% match have to be excluded or would it have to be an exact match? It's a tricky platform policy question, as this could affect the scientific output of models. This is effectively what OpenAI is trying to build using techniques we brought to their attention (perceptual hashing) which can help identify data, but there are technical issues that make this challenging to scale across the web. The cost factor is also likely a reason we've not seen them going to market with this, particularly since they don't have a clear business model yet."

The Necessity of Transparency

Transparency is a major focus in this research report, and both of the workshops covered transparency, with one workshop focusing just on transparency. Prior to the workshops, DACS and CoRD Labs' adapted one of CoRD Labs' co-founders' definitions [on machine learning transparency](#), which outlined a need for 'meaningful transparency' for consumers that must consider: legibility, auditability and impact. Meaningful transparency calls for information and action to ensure transparency works for consumers, centring consumer agency and understanding; principles seen across both UK and EU technology regulation. Our goals for the government consultation on transparency would include:

- Allowing copyright holders to enforce rights
- Opening access to allow individuals and/or third parties to scrutinise datasets
- Facilitating the assessment of legal liabilities by AI users, encouraging uptake
- Ensuring measures are robust yet proportionate, especially for e.g. start ups

Building on these needs, for our workshops, we defined 'meaningful transparency' as:

Legibility: any transparency measures introduced by the Government should be easily understood by anyone, without the need for a lawyer. You should be able to understand what information you are entitled to receive and how to obtain it. This will allow you to determine whether works have been used for training

Auditability: the information provided by AI companies should enable copyright owners to determine whether a violation of their copyright / IP rights has taken place.

Impact: if a violation of their rights has taken place, the information provided by AI companies should enable rightsholders to seek redress – for example through a regulator or the court system.

This definition was designed thinking of creatives who do not have legal representation and are not part of a collective rights management organisation. This definition, coupled with the ten point transparency model, is the basis of our requirements to the consultation on transparency.

Ten Point Transparency Model Suggested by DACS, PICSEL and CoRD Labs:

1. Source and owners of each dataset, including any datasets used to generate Synthetic data. (This should include but not be limited to third-party data or datasets, online sources, publicly available online data or datasets, apps, licensors, offline purchases. For data publicly available online, information disclosed must include all relevant URLs and a description of the information gathered from the URL.)
2. Work-level information of all data ingested (i.e. image size, metadata, labels)
3. Timeframe of data collection
4. Source (location) of dataset
5. Specific legal basis on which each dataset was ingested and processed
6. Percentage/weighting of each data source to overall training/fine-tuning.
7. Region in which training took place.
8. Description of the measures taken to ensure compliance with UK copyright law, including measures taken to comply with any proposed opt outs including those in other jurisdictions.
9. List of third-party providers used to acquire data.
10. Where applicable, the types of automated data collection tools e.g. web crawlers used to acquire data, including by third parties.

In our two workshops, we presented the ten point transparency model, our definitions of meaningful transparency, as well as an overview of the EU AI Act's transparency requirements and California's laws on transparency to lead a guided discussion, and related survey questionnaires, to source feedback, opinions and recommendations from participants.

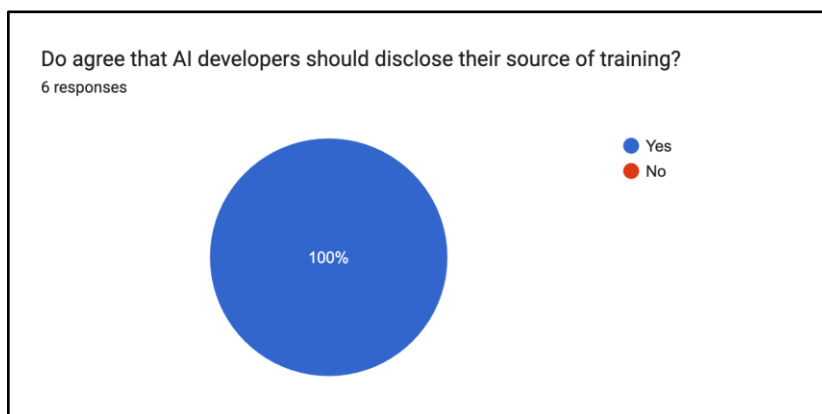
Transparency Themes

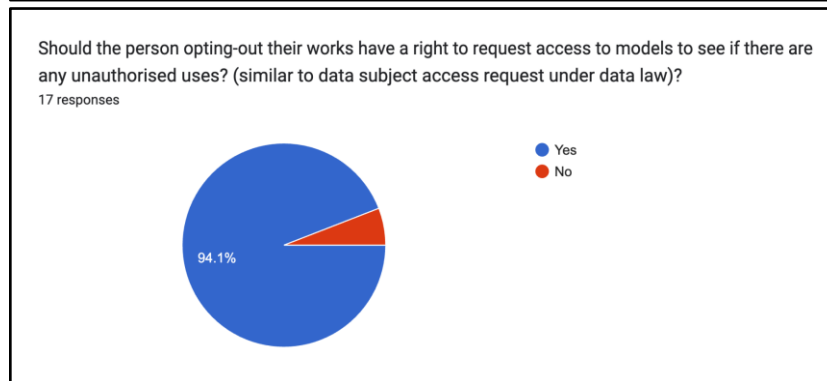
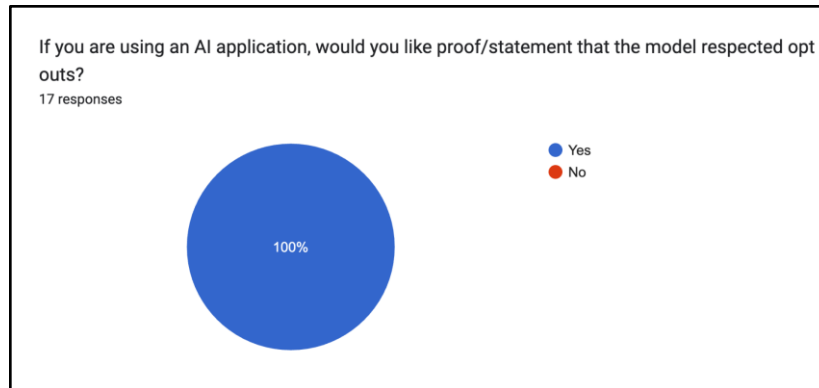
Participants raised a variety of concerns, questions and feedback, but overall stressed the need for transparency, along with wanting disclosures, granularity, and a desire for more concrete plans from the government. For example, what will be the level of disclosure related to transparency, and how will it be impactful and useful for consumers and companies alike?

- **Accessibility and Terminology:** Participants stressed the need for easily understandable information about where their work has been used, without requiring specialised expertise. Participants asked that the language choice and overall accessibility be duly considered.

- **Explainability Matters:** explainability, or legibility, is deeply related to accessibility but goes a step further; it also requires in depth information that a consumer can understand without having to be a data scientist. Consumers deserve to know how a system works in plain language, with accurate, relevant and pertinent information.
 - One user stated, "There's a few distinctions in machine learning about what learning is, explainability is often the first step – which is a human thing to ask for."
- **Granularity in Transparency Requests:** Participants noted the need for tools that enable granularity in transparency requests. Users responded well to our ten points and definitions of transparency, citing a need for granularity. A one size fits all transparency request, or transparency disclosure is not going to work for all creatives or companies.
- **Global Challenges to Transparency:** Participants pointed out the challenges of achieving transparency on a global scale will be difficult due to differing regulations and safeguards in different regions. What the UK requires might be different from the EU, and different from Canada, Australia, New Zealand, the US, and the list goes on.
 - One user emphasised, "transparency on a global scale is also challenging, if Europe provides for certain safeguards but another region does not – how can transparency across borders be achieved."
- **What is the Level of Disclosures Related to Transparency?** Participants emphasised that the government must also think beyond the transparency definitions and clauses, and also focus on what disclosures will look like and to whom will this information be disclosed? Disclosure can be similar to auditability and impact; creatives want to know how these systems work, and who is using their works.
 - 17 out of 18 participants voted that AI developers should disclose sources of training, and also wanted proof that the model respected opt outs. 94% of the 17 users surveyed also wanted the right to request access to models to check if unauthorized users were occurring.
 - One participant wondered, "What level of disclosure, to whom are we disclosing? The tech companies are going to do everything within their power to disclose as little as possible, and the middle ground – which I consider fair – would be to mandate disclosure to a regulator."
 - Another participant highlighted disclosures as information finding related to opt outs, stating, "Artists would also like to understand how opt-out systems interact with other stacks in the software stack – and other systems + further actors in the arts ecosystem. Interoperability from the opt-out to other possibilities. It's a design and a technical question."

- **Government Should Invest in Infrastructure and Data Science:** Participants highlighted that government has the opportunity now to invest in technology infrastructure related to AI and AI auditing, and should see it as a long term investment by supporting data scientists in government or regulatory roles and supporting and funding open source tools for AI investigations and audits.
 - One participant suggested, “I see it similarly in this being the early stages of the significant infrastructure we’re going to need around training data and around the web to ensure that AI models are built in a compliant manner, safely and responsibly. There’s going to be all sorts of data infrastructure we’ll have in 10–15 years, e.g. databases of opted out works, venues for the publication of training data.”
 - Another participant stated “If you’re going to push for this level of granular transparency which is necessary to form the basis of a licensing regime, I don’t think you’ll get it if you say this information has to be open at this level of granularity. So, pushing for at least one part of government to be responsible for creating a repository of transparency information that can be queries by rights holders or organisations on their behalf.”
 - An additional participant emphasised, “I think the public sector should employ data scientists to identify unsafe data which should not be trained on, which is a larger issue of funds being unlocked for regulatory competency on data science. Funding open-source tools to do this auditing and codify this gold standard into something that can be generalised and extend beyond national boundaries is crucial.”





The Elephant in the Room: The Pros, Cons and Considerations of Opt Out

Similar to transparency, opting in/opting out was a large focal point of our research and workshops. When discussing opt out/opt in, a variety of themes and concerns emerged, branching out to issues related to transparency, such as disclosures and evidence of opt out working. However, our discussion surfaced many other issues related to opt out, including the design of opt out, the ease of use of opt out systems, and the prominence of opt out, along with the arduous nature that opt out can occur. For example, would an opt out system require a creative to opt out every single instance (and copy) of a piece of work? One image of an art piece might be shared hundreds of times, potentially with a slight cropping, different colours, etc. Would a creative need to track down all of those different instances or would just registering the image one time suffice? Answering this issue is paramount to ensuring fairness and agency for creatives. Participants, along with CoRD Labs, DACS and PICSEL highlighted the importance of any system being easy to use, easy to understand and easy to find.

Harmful Design Patterns in Opt Out/Opt In

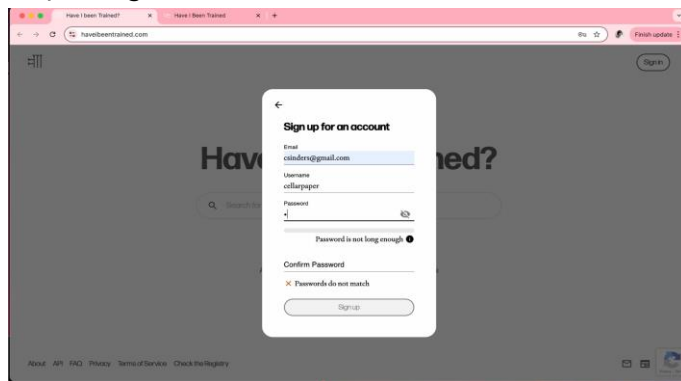
Participants shared fears of opting out being difficult to find, arduous or full of frictions. This fear is not unfounded when reflecting on the phenomena of “dark patterns” or harmful

design patterns. “Dark” patterns, often called manipulative design patterns, harmful design patterns or deceptive design patterns, are design patterns that unintentionally or intentionally confuse, manipulate, design or obstruct users from making their intended choices, choices not in their best interests, or choices that benefit the company over the user. Harmful design patterns (or dark patterns) have [varying impacts on consumers](#), including privacy, economic, and cognitive burden that these harmful patterns place on users’ decision-making online. Both the Information Commissioner’s Office and the Competition and Markets Authority have highlighted harmful design patterns as practices that can influence consumers’ decision making and impact markets’ outcomes. One of the authors of this report was a [co-author of a joint regulatory paper with the ICO and the CMA](#) on harmful design patterns and online choice architecture practices; this paper explored how harmful design patterns can impact and violate current UK competition law and GDPR.

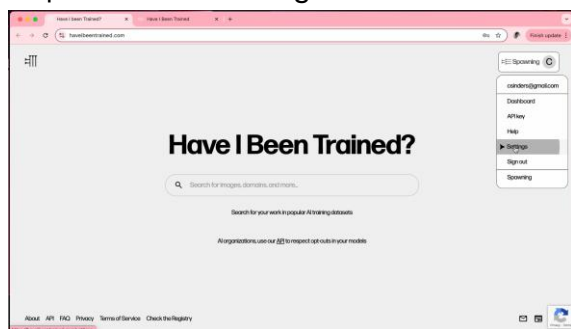
Spawning and Haveibeentrained

For this workshop series and related paper, CoRD Labs analysed three different products focusing on opt out, with four different flows in total. It’s important to highlight how different these products are and what they aim to achieve for consumers. We looked at Spawning’s Haveibeentrained website and web browser extension (two different product types under the same project), Microsoft’s Copilot opt out, and Instagram’s GDPR related AI opt out. Spawning’s Haveibeentrained project exists as a tool of user agency to allow creatives to opt out their images from a series of different public AI datasets Spawning has collected. The machine learning company Hugging Face uses Haveibeentrained in their models, tools, and related software. Haveibeentrained is relatively easy to use, but it does require users to track down different instances of the same image when first using the service. After a while, Haveibeentrained ‘learns’ which images the user has opted out of and tries to surface other instances of those images for the user to further opt out.

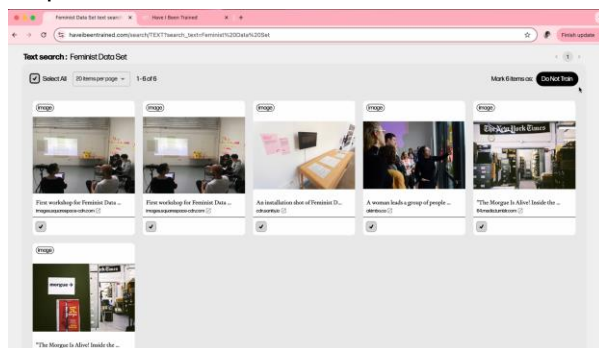
Step 1: Sign in



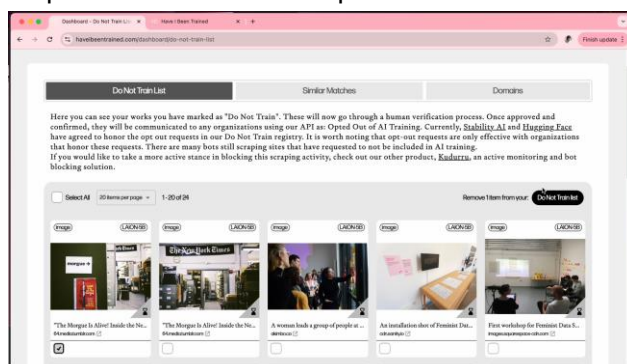
Step 2: Start Searching via the search bar



Step 3: Search for works



Step 4: Select works to opt out



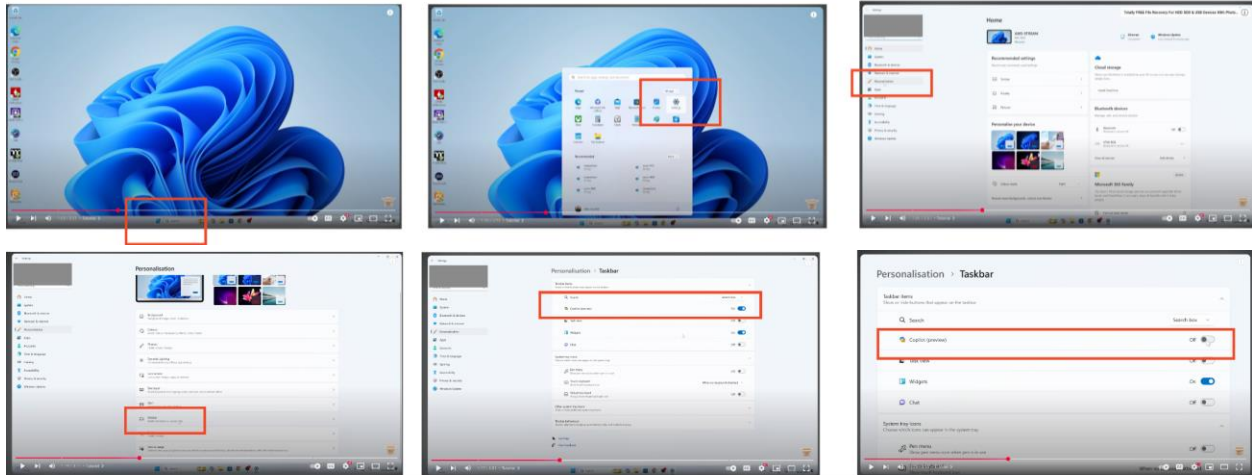
We analysed Haveibeenentrained to highlight that, while it is currently a fantastic tool to empower consent and control for creatives of their work, it does create extra work for users to opt out their images. When first using it, a user has to think about or consider how other websites might have named, stored or labelled their images. For example, one of our co-authors searched for their own work, and used variations of their own name, the work's name, etc. Some work only appeared when the co-author's last name was used and buried deep in the search results. This is not a Spawning issue but an issue of how much information metadata can have, and how different websites, platforms, galleries, magazines, and general media will label and mislabel images they are using. There's no one standard way to label an image online, so simply searching for "ARTIST NAME + TITLE OF WORK" will not necessarily surface all instances of that artistic artifact online. The Haveibeenentrained example is not shared to critique the tool, but because it provides a necessary service that no large generative AI companies provide, namely allowing creatives to reclaim control of their works. However, this example is used to illustrate how the government might need to rethink opt out to avoid the time-consuming activity of a creative individual opting out of each instance of an image of their work.

Microsoft Co-Pilot

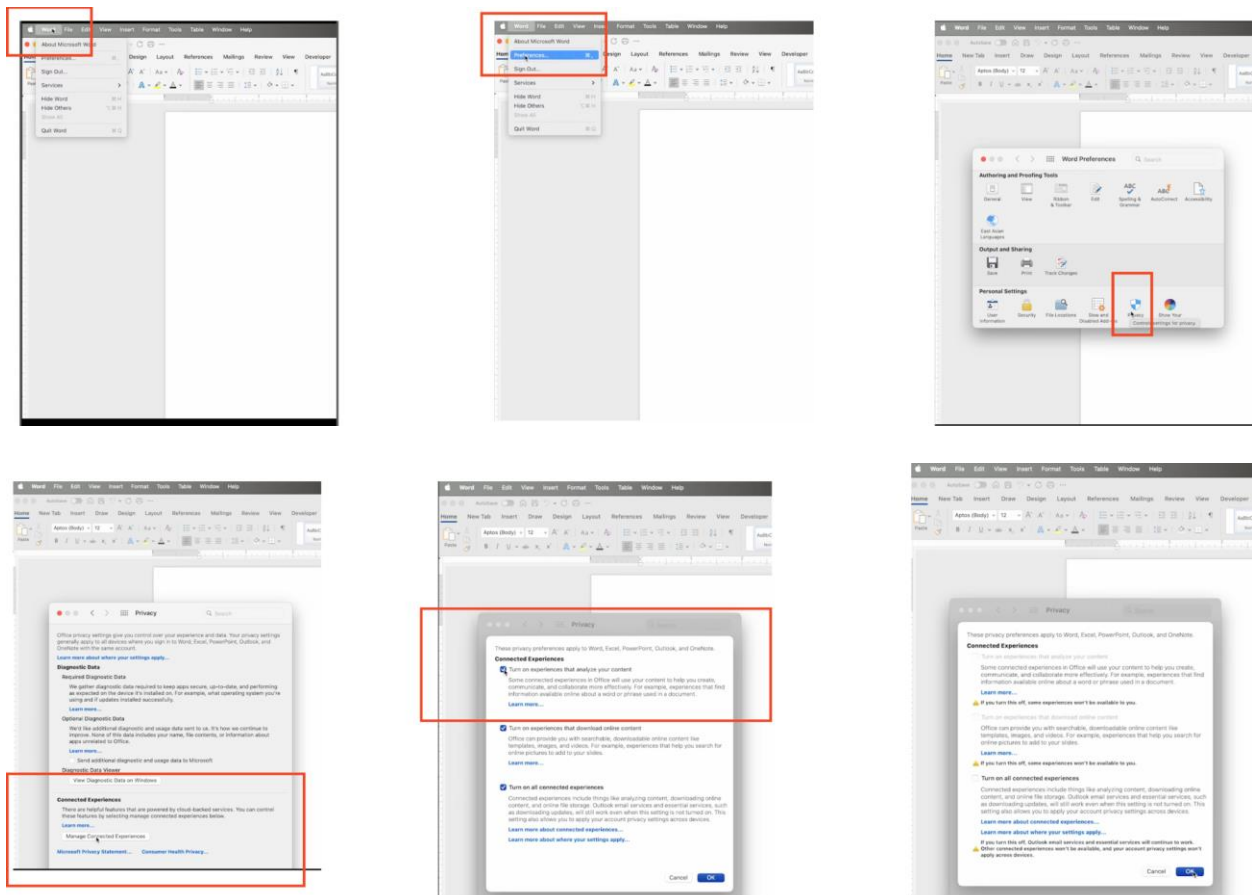
Microsoft's Co-Pilot is an AI powered chatbot and tool to assist in a variety of tasks related to Microsoft 365, including Microsoft Office and other programs used for day-to-day work. Microsoft has created two ways to engage with Co-Pilot in terms of administrative controls. It's technically very simple to 'turn off' Co-Pilot as a feature, and as illustrated below, it's accessible via the main navigation bar. However, this flow only 'turns off' Co-Pilot but it does not opt the user out of Co-Pilot or Microsoft tracking the user's data online to train AI programs.

None of the Co-Pilot UI mentions tracking or using the consumer's data. Microsoft describes this Co-Pilot and other data tracking as "experiences that analyse your content."

Turning Off Co-Pilot



Disabling Co-Pilot



However, to actually disable Co-Pilot and opt out of 'experiences that analyse your content', e.g. tracking, a user has to undertake a slightly more arduous and confusing design

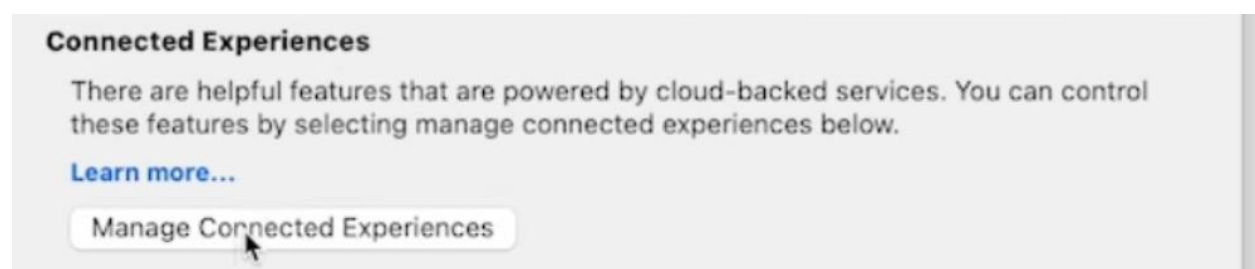
flow. Based on our expertise, we feel confident in describing this flow as an example of harmful design patterns.

Consumers might even conflate the use of Co-Pilot as learning from them and think that by disabling Co-Pilot, they are opting out of all AI systems or systems that analyse their content. However, the only way to opt out is via, to our knowledge, each different product inside of Microsoft 365. In this example, opting out of Microsoft Office. To fully opt out of tracking, it appears a consumer would need to opt out within each individual Microsoft 365 product.

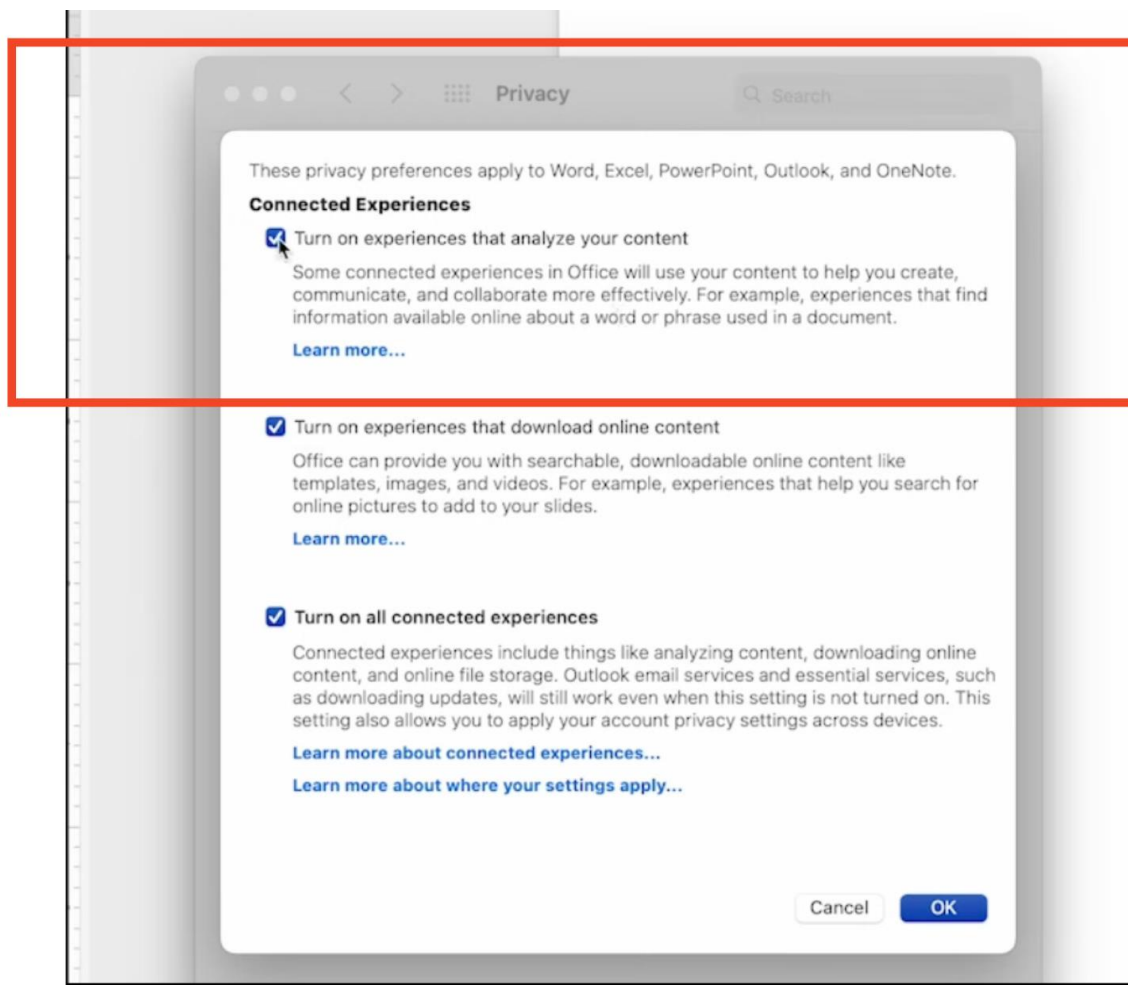
Questions we considered in our analysis:

- Is it intuitive and clear that the consumer knows where or how to opt out?
- Are the descriptions of the task clear? Does 'manage connected experiences' sound like the 'right' setting to disable one's content for AI training?
- Does 'turn on experiences that analyse your content' adequately explain to a user what that setting does or does not do?
- How would a user know this is the setting to go to?
- How would a user know that Microsoft is analysing their content?
- How were consumers notified of this new feature?

What makes this a harmful design pattern is the friction in opting out of training for Co-Pilot and other Microsoft products, since users have to do this in each individual product, creating a more arduous process. The [UK's Competition and Markets Authority \(CMA\)](#) would define this as a 'sludge' or designs, actions or choice architecture "[that] makes it hard for consumers to act in their interests (such as adding friction to cancellation processes)." Furthermore, looking into the steps, and names of settings to opt out of training, the names of the settings as "Connected Experiences" and burying connected experiences at the bottom of the Privacy settings, is a form of "hindering or obstruction." While it makes sense to have privacy related settings, like a user's data being used to train AI, in privacy, the issue here is the name and descriptions of "Connected Experiences." A close up of the setting reveals that Connected Experiences does not mention AI, data or training at all.



Clicking on “Manage Connected Experiences” also does not explain or mention data, tracking or AI. However, this is the setting a user must turn off to turn off tracking and experiences or data used to train Microsoft’s AI. The labelling and descriptions in products matter, and it’s a form of design called “UX copywriting.” How products describe what they are doing should accurately reflect what they do. If clicking to opt out of the feature that says: “Turn on experiences that analyse your content” is the precise feature to turn off AI training or anything related to AI; the descriptions and labels should reflect that. In this context, how would a user understand what a ‘connected experience’ is and that it pertains to data, AI and training? They would know if Microsoft explicitly, accurately and legibly shared what this feature does.



[Gray et al., defines obstruction](#) as “impeding a task flow, making an interaction more difficult than it inherently needs to be with the intent to dissuade an action.” Relatedly, [the European Data Protection Board](#) refers to this particular type of action(s) as ‘hindering’,

which is “an obstruction or blocking of users in their process of getting informed or managing their data by making the action hard or impossible to achieve.” Throughout their report, the EDPB gives examples of hindering such as ‘taking longer than necessary’, giving ‘misleading information’, and ‘creating a dead end’.

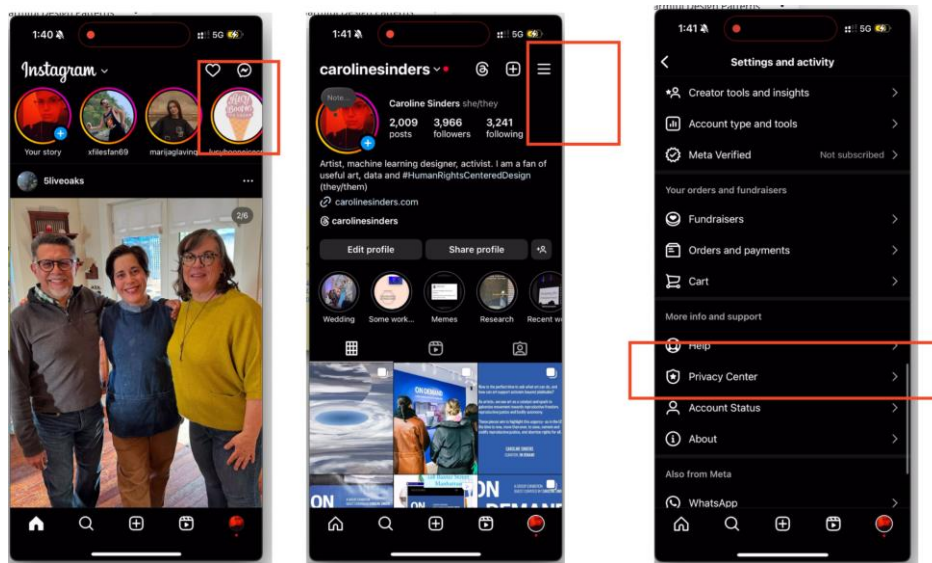
Additionally, that the setting is automatically turned on, would be a harmful design pattern called a ‘default’ which the UK’s Competition and Markets Authority (CMA) as the designs, action or choice architecture “[applying] a predefined setting that the consumer must take active steps to change.”

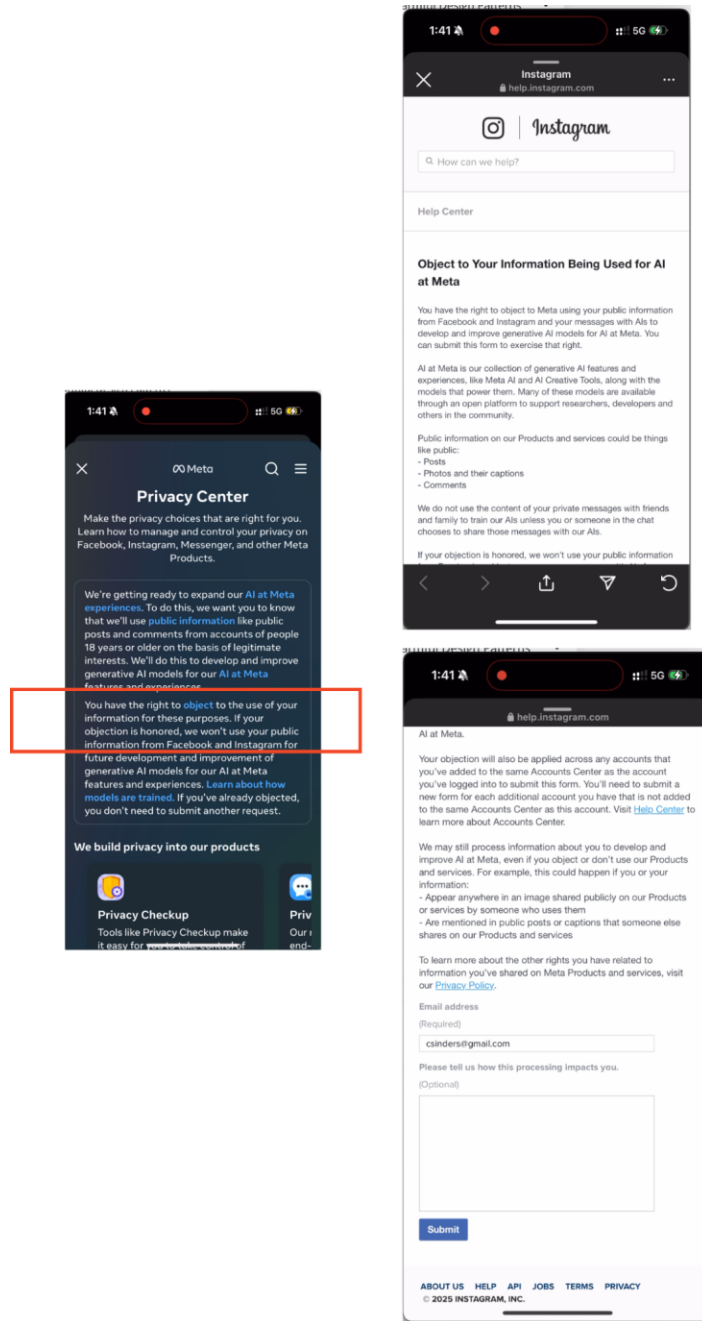
Instagram AI Opt Out

The last product we analysed under the lens of ‘harmful design patterns’ and to inform opt out/opt in design was Instagram’s AI Opt Out. The same kinds of harmful design patterns also appear here, with the same kinds of considerations. While it makes sense that the opt out for AI training would be under privacy centre, it’s not under privacy settings. Accessing the Privacy Centre requires the user to click on the ‘settings’ hamburger menu and scroll all the way to the bottom of the settings page. Long scrolls, and multiple clicks, can be arduous for users, as there is much more information they need to sift through to find the right setting. Once clicking on “Privacy Centre”, the user is then sent to a setting full of varying texts, some unbolded and some text **bolded in the Meta blue**. Within the red square, we have highlighted “You have the right to object” with the word “object” in that bolded blue. Clicking on the bolded blue ‘object’ sends a user to a new pop-up page on their device. What makes this a harmful design pattern is that a bolded word is not necessarily indicative of a hyperlink of underlined text, or a setting where a user would go to. However, even a hyperlink would not be sufficient enough for a user to understand this is a form of choice architecture, and is the right ‘feature’ to click on to start the opt out process. The joint regulatory paper [by the ICO and CMA on harmful design features](#) a similar UI example, present on a cookie banner, and highlights how the design(s) “must be presented with equal prominence; the user must understand what they mean and must not be nudged towards one over the other.” In this case, it’s not clear that the bolded word ‘object’ is UI of choice architecture to lead them to the action of objecting; users would usually expect a button present. A hyperlink to send them to a new page or a bolded word is not enough of a call to action to meaningfully inform a consumer of choice, particularly in this case to object.

Furthermore, once a user clicks on “object” they are sent to a brand-new pop-up page and they must scroll all the way to the bottom to enter their email address and hit submit. This long scrolling adds additional friction, and at first glance, it seems as though the user must

write something to object. It's only when looking much more closely at the UI, does one notice in small, grey letters the sentence "Please Tell Us How This Processing Impacts You" and then in much lighter grey colour further below, the word (Optional) appears. It would not be surprising if users ended up making it all the way to stage and then stopping because the presence of entering text first looks like a requirement. A user might worry that they could write the wrong sentence and reverse their opt out. Additionally, the word (Optional) is so light that it violates accessibility design principles, and those with low vision, most likely cannot see the word optional, further impacting disabled users. This begs the question- why not make opt out or object a button or a setting, and much easier to find in the privacy settings or another place within Instagram? The use of sludges and hindering deeply impacts a user's ability to easily and fairly opt out.





Why Design Matters

How opting in and opting out is presented and visually illustrated to users deeply matters in terms of protecting their privacy, competitive markets, and their rights over their images. While not all of the examples are directly related to image opt out, these examples exist to show how different companies can and will interpret regulation that involves user choice, agency and/or consent, and warp it to their benefit. This can also be seen in the many examples of how companies have misinterpreted cookie banners related to GDPR and attempt to violate user's choices in regards to their privacy. It's imperative for the

government to consider the design and user experience of whichever systems or solutions they decide on. Design is an integral part of the user's experience online, and an integral part of ensuring understanding, consent and transparency.

Opt Out: Findings from the Workshop

As stated earlier within this report, opt outs and opt ins were one of the largest themes and talking points throughout the workshops and our research. Even as the workshops would change topics, to transparency or rights holders, the conversation would inevitably return to opt out and opt in, as that seems to be the mechanism the government is leaning towards as a solution. Thus, transparency measures and opt out/opt in systems will go hand in hand. Our workshops surfaced tensions between opt out and opt in, with a fear of overly arduous or difficult to find systems to opt in/opt out of work, and how so much of the burden of this solution will fall to individual rights holders. Participants highlighted a need for new systems, wondering why opt in could not be considered first? Why not have all work automatically opted out, and artists could select which works to opt in?

- **Opt-in vs. Opt-out:** A key discussion point was whether the default should be opting in rather than opting out, with some participants feeling the opt-out approach was too restrictive and linear. While standardisation of opt-out mechanisms is desirable, its practical implementation poses significant challenges.
 - One participant noted, "I would rather the Q be should we be invited to opt-in rather than out? It's interesting [that] these discussions started from that opt-out point."
 - Another stated, "We're in an interesting moment now, where opting out would be the default situation – and where does this leave AI companies."
- **Tactical Opt-Out:** Opting out was viewed as a way to resist market capture by large corporations and empower rights holders.
 - One participant emphasised, "The opt-out for me is about saying no, putting a doorstop to the giants. There should be a sense of discretion for rights holders to determine which works they want to grant access to, and to whom."
- **Ease of Use/Burden of Opt-Outs:** Participants expressed concern about the time and effort required for individual opt-outs, suggesting that AI companies should bear more responsibility in identifying copyrighted works. This point was particularly emphasised during workshop 1, session 2 which focused on harmful design patterns and designing new opt out systems.

- **Granularity:** Participants emphasised a need for 'granularity', especially in terms of understanding the current limitations of Opt Out and it being represented as a binary. Participants suggested flexible and additional settings, considerations and variations on IP in regards to a singular piece of work. Presently, opt out is presented as a very binary and static setting of one decision being made for one piece of work; however, a singular piece of work could have many licenses attached to it. Other users suggested thinking through more nuanced settings of opt in and opt out together.
 - One participant emphasised, "It's completely binary, and what about if I want to opt-out of some works and not others? "
 - Another participant stated, "Opt-out en masse and more granular opt-in? Could the opt-in be focused on individual licensing schemes?"
- **Design Prominence:** as evidenced in the harmful design sections, design deeply matters. Participants emphasised that the prominence of opt out/opt in should be given 'main' real estate or be one of the first things a user can see and access, especially if placed into a product or platform. As discussed in the harmful design patterns category, often user choice and opt out, in regards to services or regulation, can be incredibly buried or riddled with harmful design patterns.
 - One participant emphasised, "Copyright opt-out should be given same prominence as GDPR "
 - Another participant suggested a specific design element by saying, "If you consider the example of the VPN, it used to be that you had to download it onto the computer, then it became an add-on and now it's at the top of the preferences of IOS. Could the opt-out be dictated at the device level, cf Apple ID?"
 - Another participant highlighted disclosures as information finding related to opt outs, stating how this impacts design, "Artists would also like to understand how opt-out systems interact with other stacks in the software stack – and other systems + further actors in the arts ecosystem. Interoperability from the opt-out to other possibilities. It's a design and a technical question."

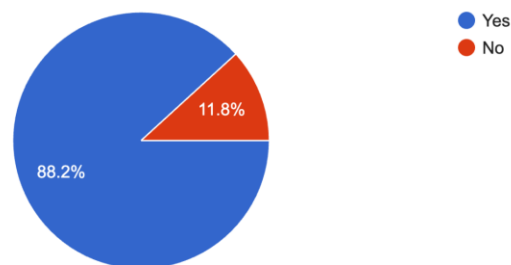
Recommendations and Conclusion

We encourage the government to consider all of the themes in this research report as specific research to further engage with and policy to consider. Specifically, government should:

- Explore the feasibility of **opt-in** systems and alternative licensing models
- Require AI developers to demonstrate the feasibility of standardised, yet flexible, opt-out mechanisms that allow for nuanced control over the use of copyright works
- Prioritise **transparency and accountability** by implementing accessible tools and frameworks for tracking the use of copyrighted material in AI training, while also explaining how these systems work
- Develop expert working groups, including UX and design experts, to work on the practical application of protocols and measures including transparency, labelling and any proposed opt-in/opt-out systems
- Ensure granularity, nuance and user consent across all systems
- Ensure ongoing dialogue and consultation with stakeholders to address the evolving challenges of AI and copyright
- Ensure that small models, small companies and artists, including artist made models and datasets, are reflected in policy decisions
- Think through the ripple effects of policy choices; for example, creatives and artists need the ability to share their work and have it shared online. Systems or policies that do not allow for the sharing of artworks will deeply impact the creative industries in the UK and globally.

Would you be concerned if an opt-out also made the work unsearchable? (whether your work or someone else's)

17 responses



Lastly, many participants emphasised the benefit of time that the government has. Now is the time to deeply consider not just the past, but the present and the future, and think

outside the box to ensure creatives' works are honoured, while also allowing for AI innovation. Government has the time and ability to explore new ways of thinking through ownership, labelling, and metadata images and how that relates to opt in/opt out. The UK has an expansive wealth of experts within it, on AI and creativity, on the arts, and on copyright law and IP. Government should use this time to engage with all kinds of experts and researchers, including everyday working creatives and artists themselves. Some of the co-authors of this report are working artists, and artists who work deeply with AI, [making their own models, and photographic datasets](#).

Creatives are experts, especially about the nuances, provenance, labelling, archiving, and sharing of creative work. Consider all of these solutions our participants came up with for: "how should a model demonstrate compliance with opt outs?"

- Similar to a GDPR cookie
- standardised options for all – an extension of copyright law
- Full crawl logs of the system, including url for what it collected into the data set v. what was found that got bounced back by an opt out (as text/interface so an individual could reference without technical skill). This would then need to be verifiable by cross referencing with the actual dataset used for the training (which maybe could be done by an independent expert with the technical capacity?)
- A digital certificate stored and accessible online on a permanent basis that I can view, print, amend as necessary with secure access, but similar in mechanics and essence to the blockchain system.
- A formal regulated and transparent certificate that the model respects the agreed licences or rights.
- Ideally opt in choice, otherwise a legally signed document
- Something like the Fairtrade certificate or secure sign for banking transactions – some icon to show the system can be trusted
- A statement saying that the model conforms to the law, when it is put into legislation
- Evidence of sources
- A statement to that extent. But also it would be important for me to know that the AI developer works with licensed training data.
- A clear watermark or logo and detailed description of terms in plain English (and plain other languages)
- To understand which datasets have been licensed, and licensing practices used, and remuneration strategies

It's important to remember that not all AI is generative AI, as this report has previously argued. But it's necessary to emphasise that again and again. Not all AI is generative AI; not all AI is useful AI, and not all AI is impactful AI. In the past decade, the UK has repeatedly passed landmark legislation such as GDPR and the Online Safety Bill. It's worth remembering the positive impacts of that regulation, but also how many people are impacted by that regulation.

Generative AI is not a core internet infrastructure or key product to that many companies. It is unclear how life changing Generative AI will be for general consumers and companies, but it is clear how it's impacting creatives currently. It is too early to assume that Generative AI will be as groundbreaking as the search engine, or infrastructure like [Amazon Web Services \(AWS\)](#), which powers over 30% of the web. This is why it's so imperative to separate how varying kinds of AI use data, and misuse data, especially 'data' that are works from creatives, who own and hold their own copyright and IP. Right now is precisely the time to be grappling with and understanding all of the different facets of AI innovation, including how artists and creatives can help foster AI innovation, but this should not come at the detriment of creatives and artists. Artists and creatives know how to make creative work, and a lot know how to make images, sound, and moving image datasets for AI. It's imperative to include creatives in shaping this policy.

Annex:

Workshop Design and Structure

Workshop 1 – AI Consultation – 30/01/2025

The following section will present the workshop structure, objectives and insights to further explore the response to the consultation with a variety of experts from the field. This workshop had 18 participants in person. This workshop looked to invite various stakeholders and engage in critical discussions covering the themes of: UK consultation on AI (session 1), harmful design patterns (session 2) and transparency (session 3):

Session 1 looked at the consultation questions and objective of the session.

The session followed a conversation guided by the survey questions and sections. The sections in the survey were organised following the consultation questions. Follow-up questions were asked to the participants to further dwell on the subject matter and were also asked to give their opinion and observations.

Session 2 looked at harmful design patterns and opt-out.

The session presented an overview of what are harmful design patterns, where to spot them and how to report them. Followed by a hands-on design challenge for the participants to design their own opt-out mechanism.

Session 3 looked at transparency and accountability.

The session presented a guided conversation by the survey and the participants within their groups, responding to several statements about transparency and accountability. Follow-up questions were asked to participants where their different perspectives rose up in a multi-scale conversation.

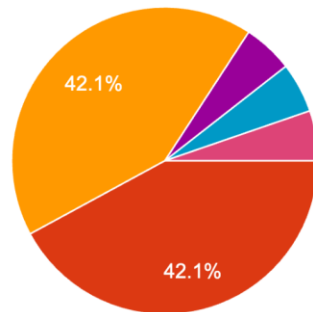
Survey Questions and Answers:

Session 1

What is your professional background? Are you a

What is your professional background? Are you a

19 responses

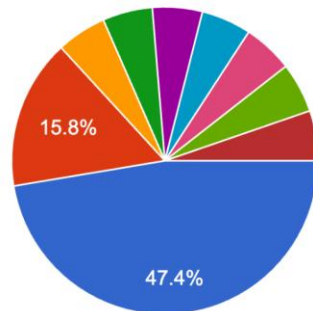


- Lawyer or policymaker
- A working artist or creative practitioner
- Arts related professional (a curator, arts administrator, staff, etc)
- Technologist or working in or related to technology (this includes open source...)
- CMO (collective management organisation – e.g. DACS, PICSEL)
- Hybrid 2, 3 & 4
- Copyright professional at DACS, as w...

Is the notion of opt-out clear to you?

Is the notion of opt-out clear to you?

19 responses



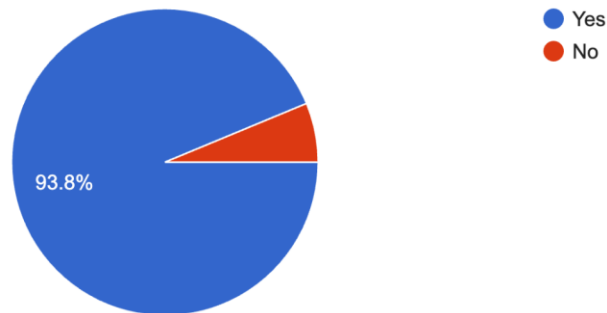
- Yes
- No
- Conceptually yes, but not at all at the level implementation
- I would like the option to "opt-in" rather...
- Too ambiguous
- Should be opt in as all images are cov...
- The legal concept is clear but the 'bina...
- I don't think how artists can opt out in...
- The concept is clear but the how, why...

Consultation Question 6: What action should a developer take when a reservation has been applied to a copy of a work?

If works have been opted out for training, do you think this needs to apply to all copies of that work, including ones made by other people? Example: "The OG stoned fox" (OG) by Adele Morse

If works have been opted out for training, do you think this needs to apply to all copies of that work, including ones made by other people? Example: "The OG stoned fox" (OG) by Adele Morse

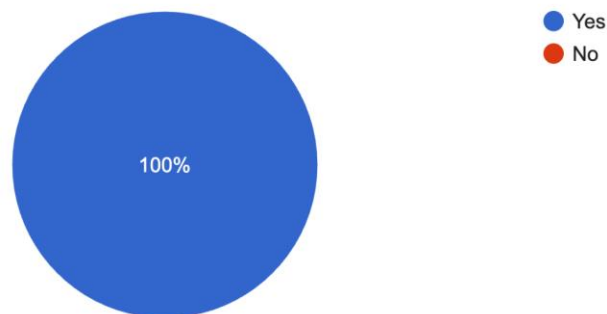
16 responses



If the developer wants to use this work, should they be able to with a licence?

If the developer wants to use this work, should they be able to with a licence?

17 responses

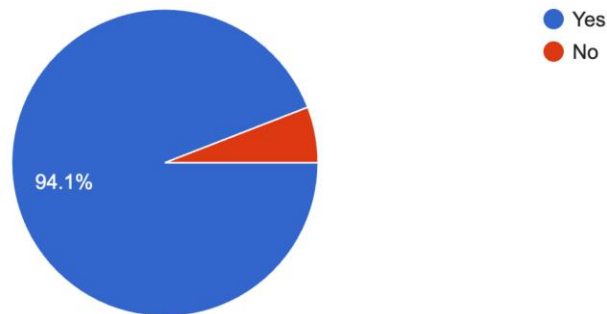


Consultation Question 8: Do you agree that rights should be reserved in machine-readable formats? Where possible, please indicate what you anticipate the cost of introducing and/or complying with a rights reservation in machine-readable format would be.

Should an opt out be an online technical tool?

Should an opt out be an online technical tool?

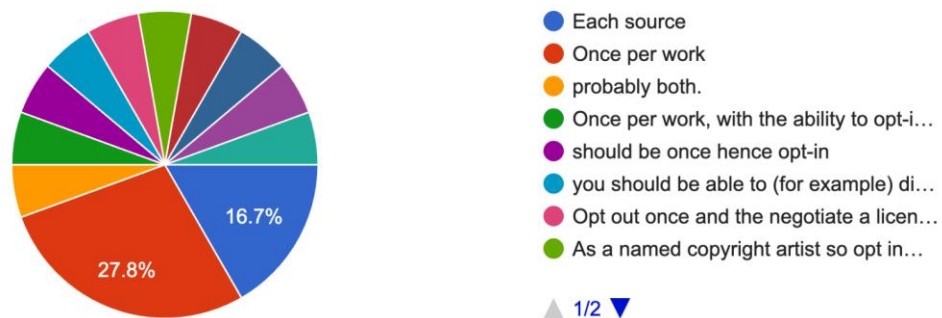
17 responses



Should you opt out for each model/training source or only once per work?

Should you opt out for each model/training source or only once per work?

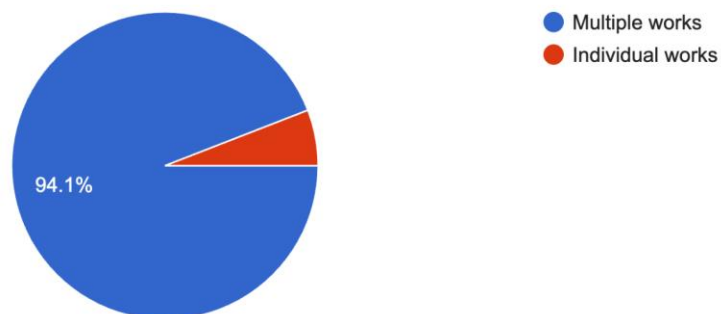
18 responses



Should you be able to opt out multiple works at once, or individually?

Should you be able to opt out multiple works at once, or individually?

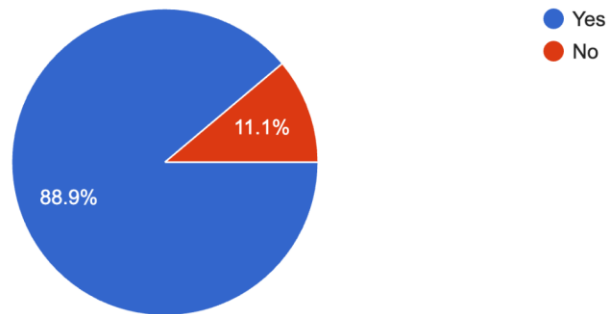
17 responses



Would you be concerned if an opt-out also made the work unsearchable? (whether your work or someone else's)

Would you be concerned if an opt-out also made the work unsearchable? (whether your work or someone else's)

18 responses



What do you understand by 'machine readable format'?

A format that is possible for a machine/computer to process

that there will be a technical layer to media that will function similar to watermarks but only vis-a-vis training

An open technical standard providing publishers, creators, and consumers the ability to trace and inform the use of their data

don't understand

metadata

Can be understood by a computer

Readable by a non-human element; e.g.: software, code, robot crawlers, spyware,

Can be regulated within an automated pipeline

Would have to research it

in a technical format which can be read and understood by different operating systems without input from the user

A way for computer systems and AI models to interpret it

Automatically 'harvestable' by technology etc.

That the choice/decision on opt out (or not) is interpretable by computer.

technical mechanisms that block crawlers from scraping content when opt-out is applied by a user; the below question re individuals

a format computer systems can read between each other which should be easily searchable, scalable and repeatable

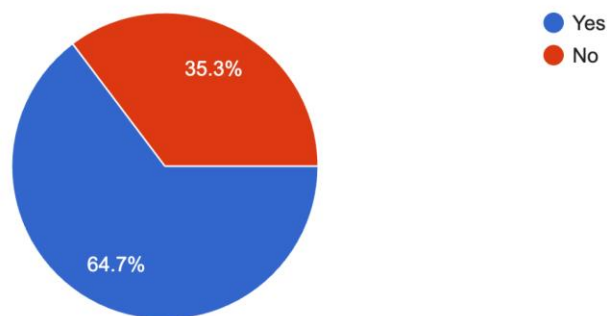
An automated or embedded (ie. through metadata or inherent code) method for work to be read by scrapers etc.

By this is I understand that the person wanting to use copyrighted works are able to read online whether or not we have opted out and automatically avoid touching these works.

Do you consider applying opt outs should be done by individuals?

Do you consider applying opt outs should be done by individuals?

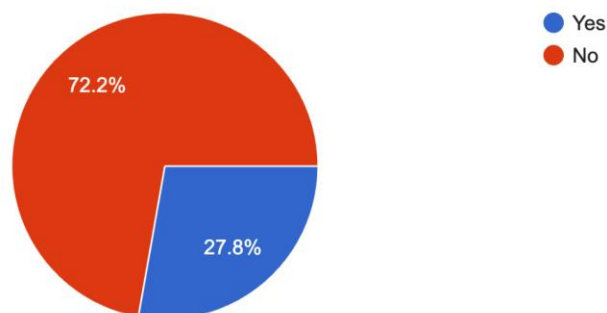
17 responses



Do you consider individuals have the means/time to apply opt outs?

Do you consider individuals have the means/time to apply opt outs?

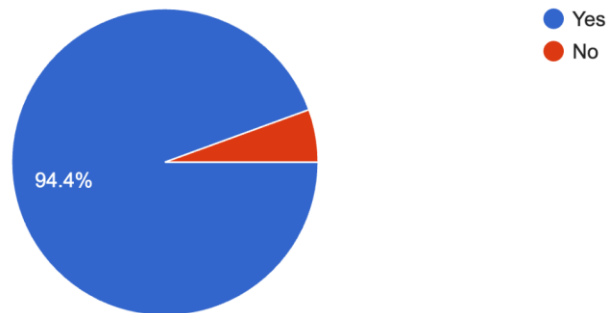
18 responses



Do you consider opt outs should be handled by organisations, e.g. DACS/PICSEL/galleries and museums or other intermediaries?

Do you consider opt outs should be handled by organisations, e.g. DACS/PICSEL/galleries and museums or other intermediaries?

18 responses



If you had to put a price tag on individuals' time to carry out opt outs, what would that be?

at least £100 per hour as this shouldn't be a job and it should be a system where people would opt-in rather than opt-out of involuntary inclusion

it really depends on what the proposed technical system is going to be

£5k per year

would be endless and fruit less exercise – hence selective opt in.

living wage

Universal basic income

hard to say without the technical specifics – if it happened at interface level maybe it's super cheap (baked into i.e. uploading) – if it's quite manual then you'd want an (expensive) expert

I attempted inroads into identifying and 'opting-uit' those works of mine that had been used for training. After a full day of doing this, I'd identified five works but had by no means exhausted the search for every occurrence of that work, and certainly I became quickly aware that there were a potentially enormous quantity of my other works that had been scraped. The cost of that time – to seek out and 'opt-out' every instance of my work – is not possible to quantify.

5-10days per year potentially

Wouldn't be needed if Opt In was an option

Not sure without seeing a more fully-fleshed out opt-out system

Approx one month per year.

I think this would also depend on the mechanisms put in place to manage the opt out.

This is a very hard question as for some large artist studios, this would make sense, but this might be a huge burden on most artists.

This is a commercial service: £100 per hour.

I don't know but money would be better spent to pay into an intermediary organisation as a member.

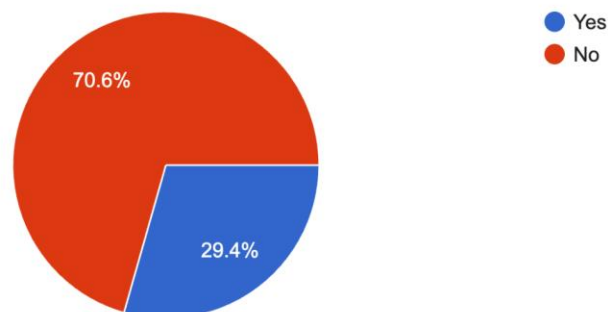
20–30K, namely an extra person

Consultation Question 9: Is there a need for greater standardisation of rights reservation protocols?

Have you used any tools like robots.txt to block web-crawlers?

Have you used any tools like robots.txt to block web-crawlers?

17 responses



If yes, do you have any comments on its efficacy?

easily site stepped and not real option

There is no reason to believe that robots.txt is totally effective. There are plenty of cases in which robots has been ignored

I've never actually verified ...

No; they're part of the site-building software I use so can't comment on their efficacy. I do know my work continues to be plagiarised and has been using in training also.

Can be ignored, no transparency currently, different jurisdictions

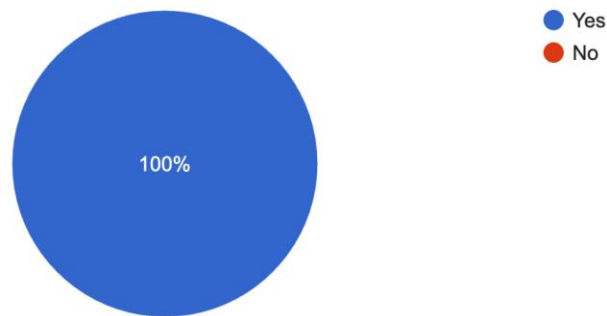
n/a

It does catch web crawlers that we are aware of

Should there be a standardised way to opt out?

Should there be a standardised way to opt out?

18 responses

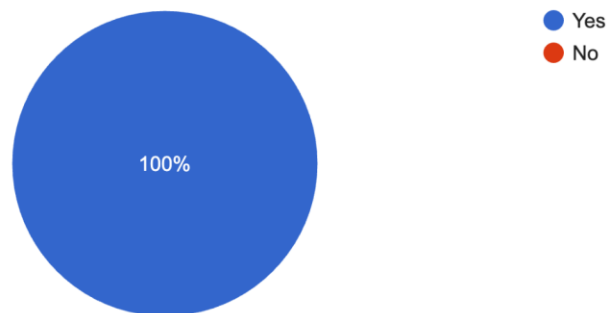


Consultation Question 10: How can compliance with standards be encouraged?

If you are using an AI application, would you like proof/statement that the model respected opt outs?

If you are using an AI application, would you like proof/statement that the model respected opt outs?

18 responses



If yes, what would you expect to see?

transparency around training data

Similar to a GDPR cookie

standardised options for all – an extension of copyright law

a simple statement

A statement or fancy logo

Full crawl logs of the system, including url for what it collected into the data set v. what was found that got bounced back by an opt out (as text/interface so an individual could reference without technical skill). This would then need to be verifiable by cross referencing with the actual dataset used for the training (which maybe could be done by an independent expert with the technical capacity?)

A digital certificate stored and accessible online on a permanent basis that I can view, print, amend as necessary with secure access, but similar in mechanics and essence to the blockchain system.

A formal regulated and transparent certificate that the model respects the agreed licences or rights.

Ideally opt in choice, otherwise a legally signed document

Something like the Fairtrade certificate or secure sign for banking transactions – some icon to show the system can be trusted

A statement saying that the model conforms to the law, when it is put into legislation

Some kind of encrypted PDF, documentation, blockchain tech etc.

Evidence of sources

A statement to that extent. But also it would be important for me to know that the AI developer works with licenced training data.

A clear watermark or logo and detailed description of terms in plain English (and plain other languages)

To understand which datasets have been licenced, and licensing practices used, and remuneration strategies.

An email at the very least confirming that opt out has taken place and confirming what has been agreed.

How should a model demonstrate compliance with opt outs?

transparency re training data where that data is querable

Similar o a GDPR cookie

A watermark or logo

Something like standardised / unit tests could be good. But these have the potential to be gamed if too standard/public (VW emissions scandal style)

I don't feel I have enough technical knowledge to answer this. If the tech bros can create AI image generation models, they are capable of creating the accountability that would come with opting out.

Regulation by government

Wouldn't be able to use it? (coded?)

needs to be hand in hand with transparency – so can investigate to show opted out. Also can have audit to test if opt outs work

With a statement, and it should demonstrate this under testing (ChatGPT already does this with Bacon, for example)

See above!

Dataset transparency

As above A clear watermark or logo and detailed description of terms in plain English (and plain other languages)

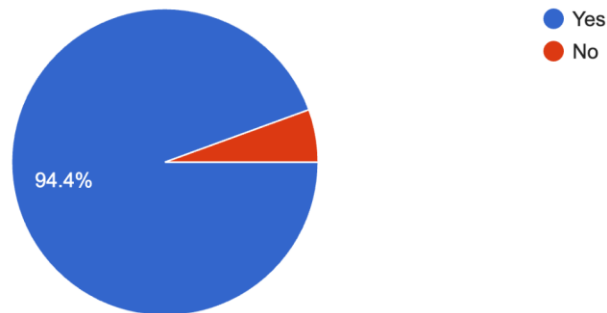
Code of conduct on the one hand. Some way of evidencing the licences of the data.

We would need to know that a model was stopped from taking works, perhaps with a report or notification from our systems. They could also publish a list of websites/rights holders that they have access to for us to check.

Should the person opting-out their works have a right to request access to models to see if there are any unauthorised uses? (similar to data subject access request under data law)?

Should the person opting-out their works have a right to request access to models to see if there are any unauthorised uses? (similar to data subject access request under data law)?

18 responses



What would be consequences of this situation, if any come to mind?

yes, but realistically this also needs to be automated.

mine field of admin

Compensation

ideally retraining without the content – seems very unlikely (plus what if open source) – so maybe heavy compensation/fines?

Only a greater awareness, and empowerment, and possibly a lot of anxiety! Providing there was an immediately accessible route to redress and remove those unauthorised uses, that scenario can only be a good thing.

If unauthorised use exists, there should be similar financial penalties to GDPR

Royalties should apply to any copyright images

I think a central organisation, like the ICO, in charge of managing the opt outs and auditing use.

There should be fines levelled if not complying with transparency. Fines can help pay for the opt out database.

There should be a way to report it to the company running the model, and then escalate to another organisation if necessary

If evidence of unauthorised use is found? Redress should be standard, repeatable and swift with financial compensation

inefficient. there needs to be a streamlined technical way of doing this that is both for data contributors and AI developers.

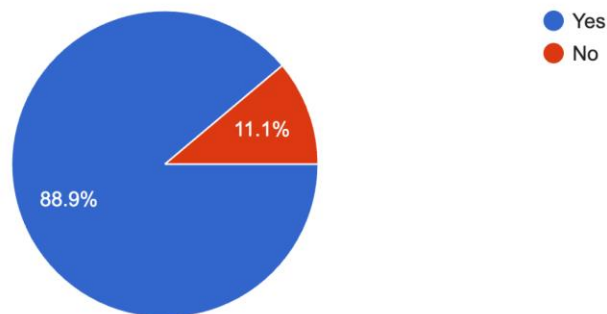
there should be penalties and requirement to cease and desist immediately

There would be greater transparency and accountability. As a gallery we would be answerable to rights holders and if accused of allowing AI training we would also be able to defend ourselves.

Do you think opt outs need to be standardised internationally?

Do you think opt outs need to be standardised internationally?

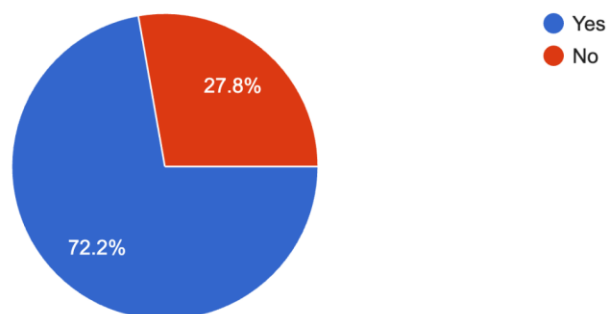
18 responses



If there becomes a database of opted out works, should the government / international governments have a role in managing this?

If there becomes a database of opted out works, should the government / international governments have a role in managing this?

18 responses

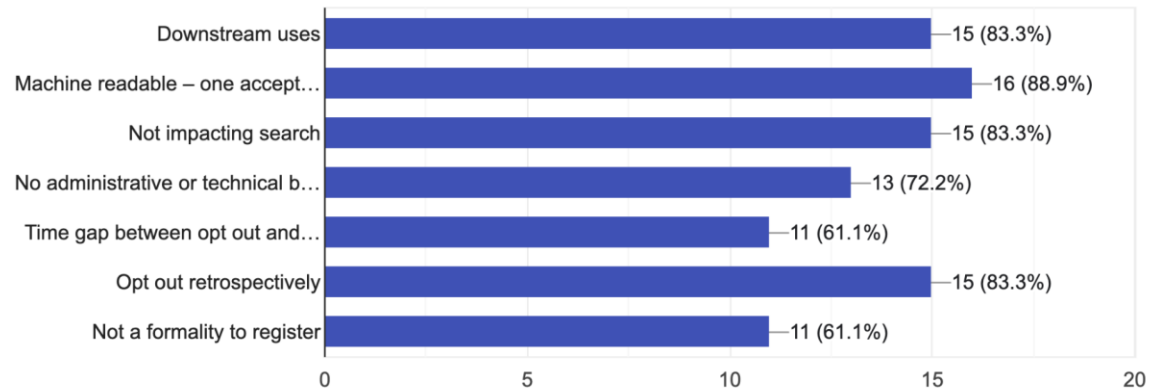


Overview of opt-out

Which of the following do you believe an opt-out must achieve? (select as many as apply)

Which of the following do you believe an opt-out must achieve? (select as many as apply)

18 responses

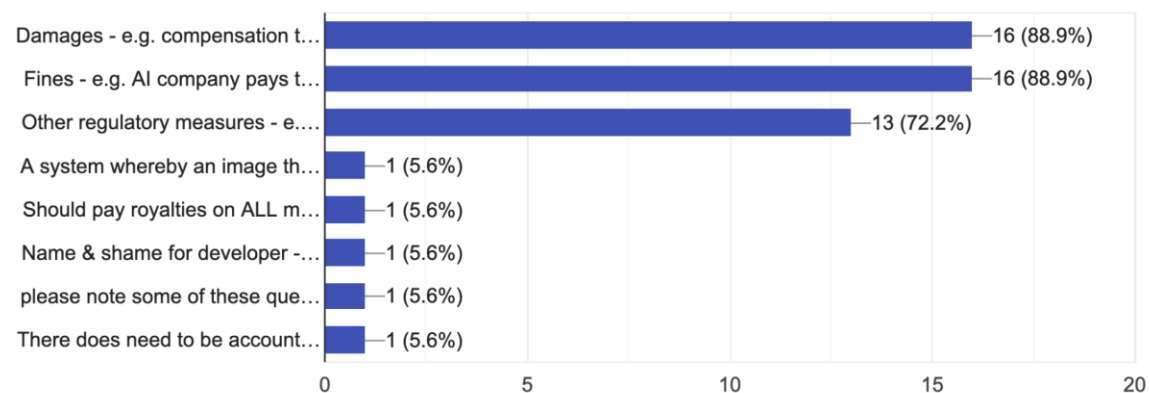


Consultation Question 7: What should be the legal consequences if a reservation is ignored?

If a developer ignores opt out, what should be the consequence?

If developer ignores opt out, what should be the consequence?

18 responses

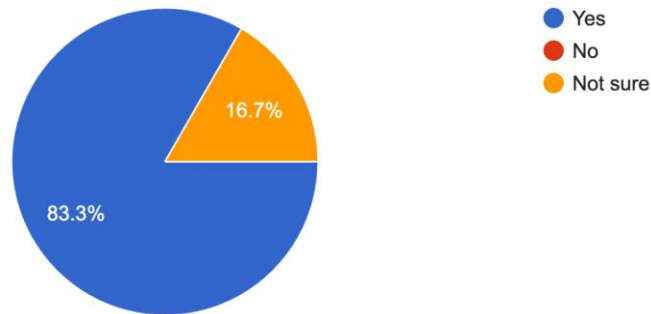


Session 3 – Transparency

Do you agree with DACS and PICSEL's view on 'meaningful transparency'?

Do you agree with DACS and PICSEL's view on 'meaningful transparency'?

6 responses



Do you have any comments on the concepts of 'legibility' 'auditability' and 'impact'?

They should follow copyright law

Arguably, there should be 'trusted data intermediaries', or existing/new entities that have greater privileges vis-a-vis disclosure than individual users. Additionally, transparency measures should:

Be ongoing rather than one-time disclosures

Include both technical transparency (what was used) and governance transparency (how decisions are made)

Be monitored and enforced by trusted intermediaries

Enable collective bargaining and negotiation

I agree that "legibility" should apply to ordinary users (i.e. non-lawyers, non-experts etc).

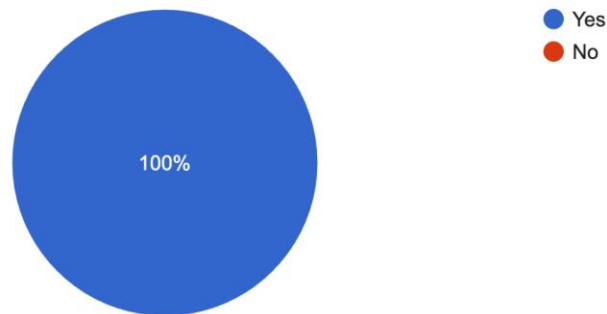
These are all very important, there should not be a shift in viewing copyright as something that is less of a right than others that doesn't need to be protected or made accountable. Perhaps a searchable register? Accountability isn't just on the other side. Institutions that have assets used in AI may need to report back on what and how it was used.

Consultation Question 17. Do you agree that AI developers should disclose the sources of their training material?

Do agree that AI developers should disclose their source of training?

Do agree that AI developers should disclose their source of training?

7 responses



Please give any further comments

As far as is possible, yes. But this is a simple question which would be quite complex to enact. 'The source' is shorthand for 'online', and given that billions of images are used for data training, I don't yet see how specific such a disclosure could be.

Since my imagery is deeply online and has been for almost 30 years, exists in hundreds of locations and is not necessarily attached to my name, or even to the title of the work (if there is one), "where" the work was sourced might be a difficult and potentially pointless question to want answering.

'Whether' feels like the more urgent but obviously more broad question to answer.

copyright/size/timescale used/royalty amount/territory

There should also be transparency about any fine-tuning or prompt engineering to control the reproduction of recognisable copyrighted material. Eg. Preventing a model from reproducing an identifiable style or not

With the development of zero-knowledge proofs, there is technical possibility to query datasets and to not reveal their contents. But nevertheless trusted intermediaries should have access to auditing to ensure compliance.

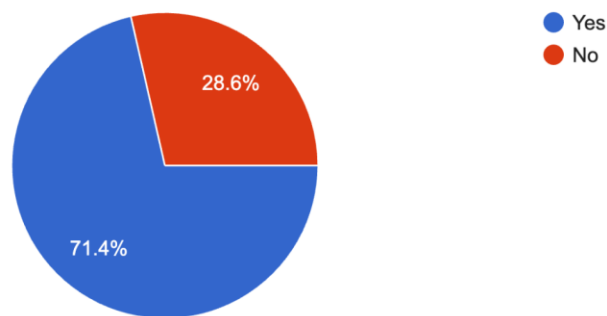
Absolutely, they would need to show what they have used and where they got it from. This is also useful if a rights holder has opted out but found their works has been used, the source can be identified.

Consultation Question 18. If so, what level of granularity is sufficient and necessary for AI firms when providing transparency over the inputs to generative models?

Looking at the proposed transparency model, do you think these measures are sufficient for transparency?

Looking at the proposed transparency model, do you think these measures are sufficient for transparency?

7 responses



Looking at the proposed transparency model, is there anything you would remove, or change?

I'd be interested to learn how the data referenced in the transparency model would be communicated to the recipient at the point of disclosure – whether as a downloadable document, an online document, whether that document is visual, verbal, a list of URLs, a stream of metadata or all of those and more. It would need to be interpretable by everyone and anyone.

Not remove, add! copyright. Notify copyright infringement

I mostly like them but think some more transparency about the model and fine-tuning would be good to add

Work-level information – like musical works proposal – should absolutely be as detailed as possible.

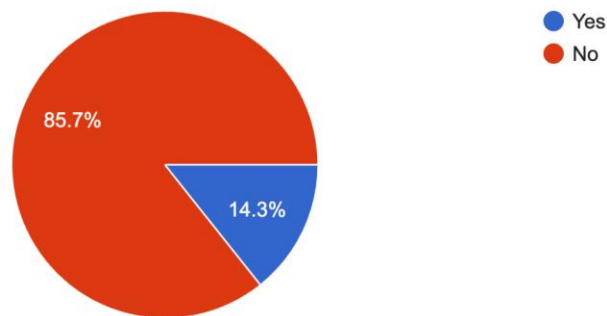
I would ask how many works have been used in the data sets.

Consultation Question 23: What are your views on the EU's approach to transparency?

Do you consider the EU's approach to transparency is sufficient?

Do you consider the EU's approach to transparency is sufficient?

7 responses



Please provide further thoughts

uncertain of the difference

No, but it is a start. To be legally enforceable would require an acceptance that had an 'opt-out' been violated, this would amount to an offence. I struggle to visualise this becoming an achievable scenario right now, but it may be in the future. The EU's approach does feel realistic and achievable despite my initial reaction that it doesn't go far enough.

To many grey areas that contradict copyright infringement

We need more legal enforcement rights

My understanding that in the Code of Conduct that is currently being developed the more hardline provisions from the AI Act will be watered down.

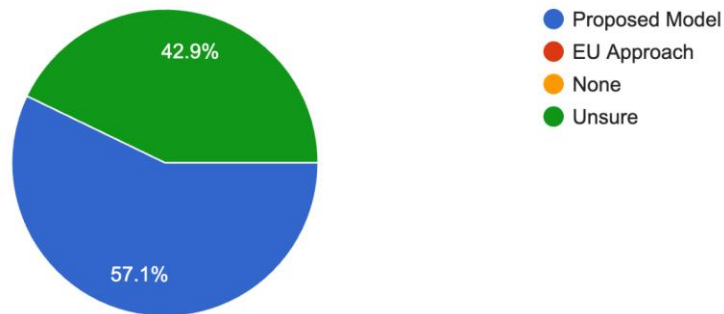
Yes, but proposed DACS / PICSEL model is even better...

It is not legally enforceable and doesn't give very specific details

Of the three models of transparency (proposed model, EU approach), which do you prefer?

Of the three models of transparency (proposed model, EU approach), which do you prefer?

7 responses



Could you explain your answer?

think the EU won't specifically tell you what has been violated and would expect this

The proposed model is more thorough and ambitious and as such, in the same way that you can price your car at £6000 knowing full well you will be knocked down by at least £1000, this model has space for a reduced version 'getting approved' which, despite being negotiated down, is still a robust system that can be built on.

The government should have a creative background team to recognise creators and understanding of the importance of copyright as it is our livelihood

I think 'transparency' only makes sense if it is full transparency on the training data and the weights.

...With the addition of an emphasis on output (which is the contentious part for rights holders) as opposed to the initial AI training.

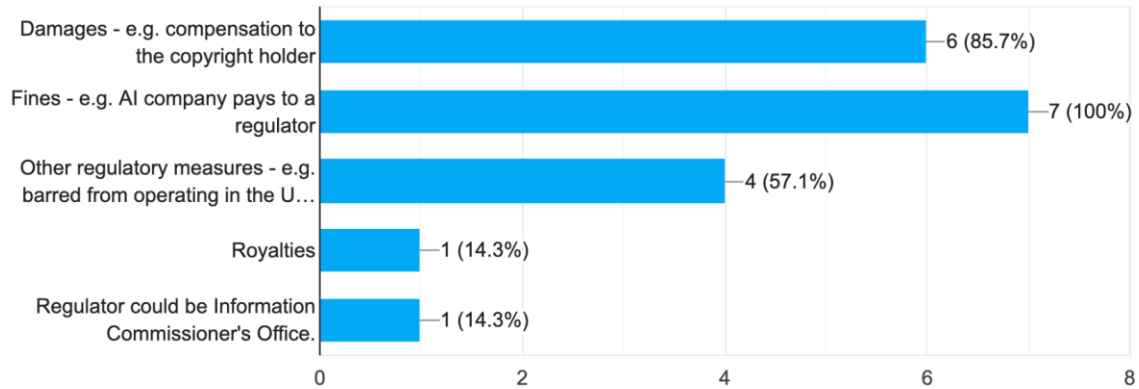
This is most useful for rights holders

Consultation Question 22: How can compliance with transparency requirements be encouraged, and does this require regulatory underpinning?

What measures do you think could be put in place to ensure AI companies comply with transparency requirements?

What measures do you think could be put in place to ensure AI companies comply with transparency requirements?

7 responses

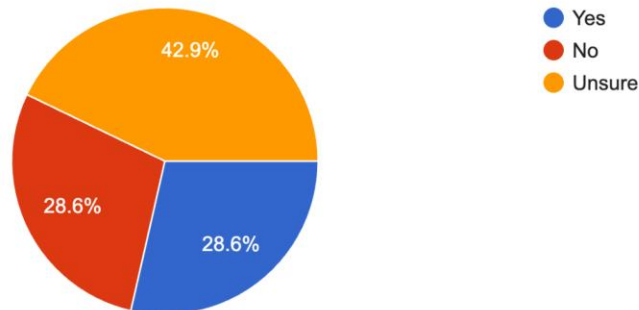


Final perspectives

Do you support a text and data mining exception with an opt out?

Do you support a text and data mining exception with an opt out?

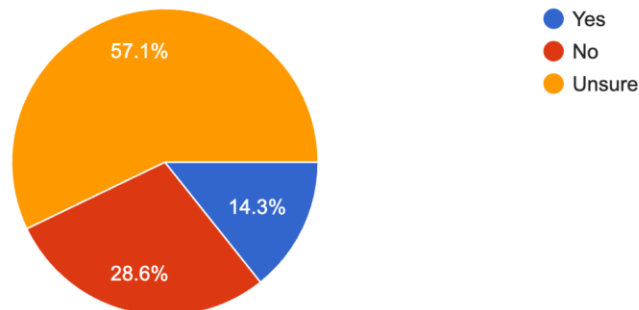
7 responses



Do you think the Government's 'option 3' meets the objectives of giving copyright owners more control over their work?

Do you think the Government's 'option 3' meets the objectives of giving copyright owners more control over their work?

7 responses



What influence, positive or negative, would the introduction of an exception along these lines have?

would make sure that it is encompassing and not easily circumnavigated.

I wish I completed survey at DACS as now feel to many grey areas so hard to give an opinion

An opt-in would be the preferable choice for rights holders. Opt-out would take time, resource, reduce creators' creative time, and upend the balance in favour of the tech companies.

I think it will be very unclear for a long time as to how to opt out properly and many tech companies will take advantage of this as well as make opt out as complicated as possible

Workshop 2 – Transparency Consultation – 14/02/2025

The entirety of workshop 2 focused on workshop 1's session 3: on transparency. This workshop was a guided conversation utilizing specific questions from our transparency survey, and allowing for participant input and response to several statements about transparency and accountability. Follow-up questions were asked to participants where their different perspectives rose up in a multi-scale conversation. This workshop had 8 participants.