



DRAFT Generative AI fourth call for evidence: engineering individual rights into generative AI models

10th June 2024

Introduction

The Creators' Rights Alliance (CRA) is a coalition that exists to promote, protect and further the interests of creators through policy, advocacy, and campaigning work. We speak on behalf of 23 major creator led groups, trade associations, and unions, between them representing over 500,000 creator members, working in the UK's creative industries – from authors, artists, photographers and illustrators to translators, performers, musicians, and journalists – on policy issues as diverse as fairer contract terms and working conditions to copyright and intellectual property.

Creators' Rights Alliance members bring knowledge, insights and perspectives from several UK, European and international forums; including the European Federation of Journalists Authors' Rights Expert Group; The International Authors Forum and the European Illustrators Forum.

In 2023, the CRA was invited to join the Creative Industries Council (CIC). The Council's focus is addressing the challenges and opportunities facing the UK's creative industries to help drive forward progress on key areas of growth for the sector, including access to finance, skills, export markets, innovation and intellectual property (IP).

2. Do you agree with the analysis presented in the call for evidence document (As linked here: Generative AI fourth call for evidence: Individual rights)? - required

Where personal data is collected from other sources:

The analysis is clear, though there are a couple of changes that could help to support Creators, as follows:

- Publishing specific, accessible information must include the dates of the sources, for example, the data set includes data on x platform from January 2001 to March 2023. This would provide Creators with a better understanding of if their data is likely to be included or not.
- If developers applied privacy enhancing technologies or other pseudonymisation, this could make it more difficult for Creators to identify if their work has been used without permission.

The right of access:

Rights to erasure, to rectification, to restriction of processing and to object to processing, should extend to include a method of enforcement to enable the take down of content, and to secure compensation for financial loss or other damage (moral rights).

Moral Rights Include:

- The right to attribution
- The right to object to derogatory treatment of a work
- The right to object to false attribution
- The right to privacy of certain photographs and films

Damage to individual cultural and social identity should be considered from an economic and social point of view. We have already seen cases of poor use of AI models that have directly and negatively impacted individuals, such as:

Actor: Stephen Fry had his voice cloned to read a documentary without his prior knowledge¹.

Actor: Emma Watson had her voice used to read offensive messages².

Definitions:

¹ [Stephen Fry shocked by cloning. The Guardian](#)

² [Extra safeguards coming after AI generator used to make celebrity voices read offensive messages](#)

Personal data is defined in the UK GDPR as:

“‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the **physical, physiological, genetic, mental, economic, cultural or social identity of that natural person**”.

Collecting and using creators’ personal data can only be done lawfully by agreeing consent in advance. Seeking permission in advance through a licensing system would allow compliance with copyright law and with data protection. Otherwise, we will see a continual stream of legal cases (see below, re Meta’s new ‘privacy policy’).

GDPR rules:

Creators’ personal data must be collected and processed in line with GDPR rules, as well as in compliance with all other pre-existing laws, such as intellectual property laws – and this should be the case both at the input and output stages.

As regards Article 5 (1) (d) of the GDPR, the ICO should consider how to ensure accuracy within personal data including personal data in the form of personal creative output into music, books, writing, films, animations, illustrations, photographs and drawings.

Considerations

- Labelling of data
- Content origin definition
- Dates from / to
- Transparency of data publishing
- Accountability for breaches in transparency data publishing
- Clear guidance from GDPR to developers, being clear and warning that using cultural data (including language, music, films, images, physical, physiological, genetic, mental, economic, cultural or social identity of that natural person) without permission is breaching copyright laws as well as data protection

Permissions:

Controllers/developers are already regularly using data without prior permission or agreement³. For many Creators' Rights Alliance members, the over-reach potential of 'legitimate interest' is adding additional concern to an already worrying new reality.

There are routes that would allow AI developers to lawfully access and process data to fulfil their training purposes. We are fortunate in the UK to have a sophisticated, world-leading licensing system which already works well, and ensures the collective management of rights and subsequent remuneration of creators for the exploitation of those rights. These systems are open and transparent, and subject to government regulation. AI developers should work with creators to consider models to extend such systems to machine learning. Collective licensing would be respectful of copyright legislation and would allow rights holders to the opportunity to agree or decline requests. Creators should always have opt-out rights.

Meta

Creators' Rights Alliance would welcome hearing the ICO position on the updated privacy policy from Meta, which gives Meta sole discretion to decide whether to grant user's request to object to have their information used⁴. Does ICO believe such discretionary decisions would conform with GDPR?

3. Where training or fine-tuning data is web scraped or collected in other ways, what measures do you think are effective to inform individuals about how, why and by whom their personal data is being processed? (As described in the call for evidence document: Generative AI fourth call for evidence: Individual rights) - required

Software development companies would be outside of copyright law and data protection if they undertook web scraping of personal data, without gaining permission in advance.

In the event that a creator suspects that their work has been used without permission outside of copyright law and data protection:

- Companies should respond to requests for information within 28 days
- Companies should provide examples of the methods they've used to try to respond to the request for information

³ See paragraph 230, [House of Lords - Large language models and generative AI - Communications and Digital Committee \(parliament.uk\)](#)

⁴ <https://www.dacs.org.uk/news-events/what-artists-and-their-beneficiaries-need-to-know-about-metas-new-privacy-policy>

- Companies must comply with the law and not use data without prior permission. Where there are clear and obvious methods of contact, which the company has ignored, they should have to explain why they did not take this option and face additional, including but not limited to financial, legal penalties for having done so
- The impact of previous rights breaches must be recognised, and future breaches avoided with stringent deterrents in place

It is the responsibility of platforms to ensure consent has been given and robust processes are in place to provide transparency and accountability. Although many individuals may be unaware that their data might be being used, we do not consider possible difficulty in tracing original sources of data as sufficient to the continued use of member data and works, and instead call on developers to recognise the multiple impacts on creators of having works and data used in this way - these include but are not limited to rights breaches, reputational harm and impacts on remuneration.

4. What kind of information do individuals need in relation to their data in the context of generative AI so they can exercise their rights?

Creators need to know if their work has been used in training data, for example, if a photographer shared an image online, the photographer needs to be able to find out if that image has been used as an input, to build an AI machine, to be able to control the output. The best way to manage this is to agree use in advance. It would be impossible for creators to monitor all uses of their work, whereas an unmonitored use would have an exponential detrimental impact on earnings.

There are routes that would allow AI developers to lawfully access and process data to fulfil their training purposes. We are fortunate in the UK to have a sophisticated, world-leading licensing system which already works well, and ensures the collective management of rights and subsequent remuneration of creators for the exploitation of those rights. These systems are open and transparent, and subject to government regulation. AI developers should work with creators to consider models to extend such systems to machine learning.

Collective licensing would be respectful of copyright legislation and, by extension, rights holders' rights, and would allow the provision of permission, or not, in advance.

Jen Hunter, Policy and Public Affairs, Creators' Rights Alliance, contact@creatorsrightsalliance.org