https://www.lexology.com/library/detail.aspx?g=29839a72-1fdb-4fe5-9acc-7f97355d04ad

AI and Copyright Reform: A Tale of Two Cities

Haldanes Solicitors and Notaries

With the evolution of generative artificial intelligence (AI) constantly growing in pace, countries around the world are scrambling to update their laws to tackle new challenges and opportunities posed by this technology.

As development of AI involves the ingestion of huge quantities of data for training, much of which is protected by copyright laws, navigating between the competing interests of AI firms and copyright owners is proving to be something of a Gordian knot for lawmakers.

Regulations that are too loose risk undermining established rights and interests; whilst regulations that are too restrictive may cost a country its competitive edge internationally.

This article will examine two recent public consultations in 2024 concerning copyright and AI: one launched by the Hong Kong Government in July 2024[1]; the other launched by the UK Government in December 2024[2].

These two consultations are of interest not only because of the similarity between the underlying legislation in question (the UK Copyright, Designs and Patents Act 1988; and Hong Kong's Copyright Ordinance (Cap. 528); but also because both governments have clearly stated their preferred course of legislative action. The difference in the approach taken by the two highlights how governments can reach radically different conclusions from similar laws.

I will focus on the two main aspects which have received the most coverage: the potential introduction of a broad text & data mining exception to copyright infringement; and the status of copyright protection for AI-generated works.

Both consultations cover other topics, like deep fakes and transparency by AI firms, but these topics have received less coverage and there do not appear to be any immediate legislative plans related to them.

Introduction of a broad text & data mining exception ("TDM Exception")

The training of AI algorithms involves the large-scale collection, extraction, copying and analysis of data. Data is mainly collected by "scraping" information published on the Internet through the use of web-crawlers.

This involves the copying and use of data, a large portion of which is protected by copyright laws. Unless a relevant legal exception applies, permission from the copyright holders is required to use (or in legal terms "reproduce") this data.

Currently most leading AI modules (e.g. ChatGPT, Stable Diffusion) have been trained without obtaining permission, leading to copyright holders (e.g. artists, news publishers, music publishers) alleging the AI firms behind these modules have infringed their copyright by using their data for AI training. In turn this has led to multiple lawsuits mostly, but not exclusively, in the US. [link to https://www.haldanes.com/publications/artists-rage-against-the-machine/].

In response to these lawsuits, AI developers have argued, amongst other things, that their actions are protected by the "fair use" exemption under US law[3].

However, most jurisdictions around the world (including most European countries, China, Hong Kong, India, and Australia) do not operate under US "fair use" principles. Instead, most jurisdictions operate under a "closed" framework (usually labelled as "fair dealing") where each specific exemption must be individually legislated and spelt out in the relevant legislation. Any use that does not fall under a specific exemption will be an infringement.

Some jurisdictions such as Japan, Singapore, and the EU have introduced various text and data mining exceptions ("**TDM Exception**").

Japan and Singapore have introduced broad TDM Exceptions covering both commercial and non-commercial use; the EU provides a TDM Exception with an "opt-out" option; while the existing TDM Exception in the UK is confined to non-commercial academic research only.

Both the Hong Kong and UK governments have signalled their intention to introduce, in Hong Kong's case, or broaden in the UK's case, TDM Exceptions to cover commercial use. Both governments have justified these changes by the need to promote AI development and maintain competitiveness in the technology sector.

Both governments recognize the TDM Exception should be subject to conditions and safeguards to protect the rights of copyright holders. The devil, however, as always, is in the detail.

The Hong Kong Government highlighted in only one paragraph there should be "conditions" (using an opt-out option as an example) to the TDM Exception, but did not elaborate on further conditions it had in mind and how these conditions would operate in practice.

The UK Government, on the other hand, devoted 15 paragraphs to discussing the potential conditions and how these conditions would operate in practice. It is particularly laudable that the UK Government recognised the difficulty and uncertainty in determining what constitutes a valid rights reservation or opt-out.

The UK Government proposed that "right reservations" (i.e. opt-outs from AI training) should be in "*effective and accessible machine-readable formats, which should be standardised as far as possible*", pointing to existing measures such as deploying protocols that block web-crawlers, or flags in the metadata of a work which make clear it is not available for training.

However, I believe the UK Government's stance that rights reservations or opt-outs should be in "*effective and accessible machine-readable formats*" is not good policy.

While deploying machine-readable measures should be feasible for corporate rights holders, like news outlets, or publishing companies, it may be beyond the technological and financial capabilities of most individual right holders.

For example, a majority of artists simply insert watermarks, labels or notes to show they do not agree with their work being used in AI training, which are most likely not machine-readable and thus "ineffective" under the UK Government's proposed standard.

It should also be highlighted that imposing an "opting out" requirement places an additional burden on copyright owners.

As a matter of principle, the burden should be on AI firms to ensure they are only using "legitimate" data for training and they should bear the cost and risk of ensuring their data collection methods do not accidentally capture data belonging to copyright owners who have opted-out. I believe any clear statement of intention to opt out should be sufficient.

I also believe governments should address the inherent disparity in power between copyright holders and platforms.

The majority of individual right holders (artists, content creators, indie musicians) rely on platforms such as X (formerly Twitter), YouTube, DeviantArt to publish their work and are subject to the terms and conditions of such platforms.

This gives platforms considerable power over the work published there. For example, in November 2024, X changed its terms of service to state all content published on its platform could be used for AI training for X's in-house AI algorithm Grok, prompting an outcry from users[4]. Governments should address whether such blanket licences are valid and if they can be overridden by users' individual opt-outs.

Copyright protection for AI-generated works

Another hot issue is whether work "created" by AI (e.g. an article created by ChatGPT, or an image generated by Stable Diffusion) could and should be protected by copyright.

The answer is "no" in many jurisdictions like the US and most European countries, as the work doesn't have the necessary human authorship to qualify for copyright protection. As mentioned

in my previous article (Link to the article in https://www.lexology.com/library/detail.aspx?g=92277358-605b-43fc-b8b0-93db1707f37d), the UK's Copyright, Designs and Patents Act 1988 contains a rather unique provision (section 9(3)) that could, in theory, qualify AI-generated work for copyright protection by attributing human authorship of AI-generated work to "*the person by whom the arrangements necessary for the creation of the work are undertaken*" (the "**Arrangement Provision**").

The Arrangement Provision has been adopted in several countries modelled after the UK common law system, such as Hong Kong, Singapore, India and New Zealand.

In its consultation paper published in July 2024, the Hong Kong Government pointed to the Arrangement Provision already present in the Copyright Ordinance (section 11(3)) as a justification not to make any legislative change for the time being, opting instead for the Courts to further develop jurisprudence surrounding the Arrangement Provision organically.

However, as discussed in my earlier article, the Arrangement Provision is filled with difficulties and uncertainty. In the UK the provision has only been applied once by a court in over 30-years of recorded cases, , and that application was not in relation to an AI work.

In the consultation paper published in December 2024, the UK Government acknowledged the problems associated with the Arrangement Provision: 1) it is a logical contradiction to have an "*original*" work that is "*computer-generated without a human author*"; and 2) there is no justification to provide a separate class of protection for computer-generated works.

The UK Government also concluded that the Arrangement Provision would have little positive effect as it would not lead to increased AI investment in the UK, and in fact most leading AI nations do not provide copyright protection to AI-generated work.

It leads to the surprising conclusion that, should the public consultation phase show insufficient evidence of any positive effect from the Arrangement Provision, the UK Government's preference would be to **remove the Arrangement Provision altogether**.

This does still leave open the possibility of reforming the Arrangement Provision "*to clarify its scope,*" but that would also require compelling evidence to persuade the UK Government to take this course.

I agree removing the Arrangement Provision is the correct move as the provision creates more uncertainty than it resolves. I also believe there is no economic justification to accord copyright protection to AI-generated work.

It remains to be seen whether the Hong Kong Government will change its course in light of this new development from the UK.

There is, however, slightly ironic that the Arrangement Provision, which was drafted in the 90s to specifically cater for AI technology, is now at risk of being removed at the very time when AI technology is becoming a reality.