

The Intellectual Property Office
Via email

Friday 7th January 2022

Dear Sir / Madam,

We thank the IPO for the opportunity to contribute to this important topic. By way of introduction to Infosys we are a next generation digital and technology services company focusing on AI, Automation, Cloud, and other cutting-edge services. Infosys maintains a presence in 45 countries, employing 225,000 people worldwide. The UK is our second largest market and we have approx. 3000 employees working at sites throughout the region.

We enclose a detailed submission with answers to the specific questions posed by this consultation however we would like to submit the following additional information.

Recent developments in the I.T industry do call for legislative intervention in copyright law owing to the typical scenarios in which AI systems are used. It is important to underscore that usage scenarios and value dynamics are different in the IT industry when compared to other industries where AI is adapted for creation of copyrightable works (such as for AI generated art, music, language processing etc).

AI is being actively adapted and used in the IT industry as a part of digital transformation initiatives of organisations, with the primarily usage being the running of both internal operations & business processes. In such scenarios, the data set used to 'train' A.I systems was typically proprietary operational, personal, financial, or technical data of the user organizations. The output generated by AI systems was usually business or control data.

The interest in such AI generated output was minimal from a copyright perspective and the nature of the output not suited for IP licensing or assignment. The primary examples are computer vision, voice recognition, product research, text classification, fraud detection, preventive diagnostics, industrial control systems etc.

However, organizations like OpenAI and EleutherAI have matured AI systems to the level where considerably useful software code generation is now possible. Such AI generated software code certainly requires copyright protection as it will have to be used, licensed, enforced and/or assigned in the course of business by both IT product and services companies alike.

In summary, we would like to submit that clear provisions are now required to address the following areas:

- i) copyright law regarding the ownership of AI generated software code;
- ii) the TDM exception to copyright while training AI systems using the software code of another entity, including issues of 'lawful access' and 'contractual opt out'; and
- iii) ownership of 'Enhanced AI Models' that are created through the training dataset.

INFOSYS LIMITED
14th and 15th Floor
10, Upper Bank Street
Canary Wharf
London E14 5 NP
T +44 20 77153300
F +44 20 77153301

Corporate Office: CIN:
L85110KA1981PLC013115

44, Infosys Avenue,
Electronics City, Hosur Road
Bengaluru 560 100, India
T 91 80 2852 0261
F 91 80 2852 0362
askus@infosys.com
www.infosys.com

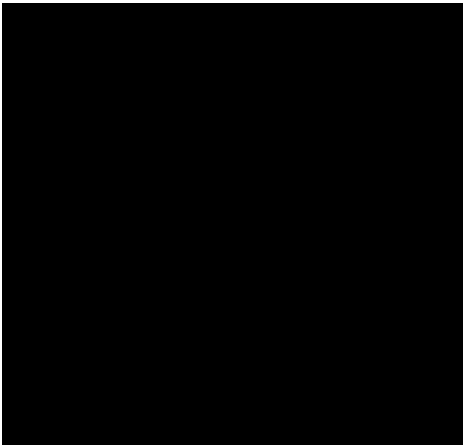
The ecosystem involved in the process of software code generation should be considered when addressing the various legal issues. We typically see the following AI usage scenarios in practice:

1. Training dataset (Owner: Client) + AI model (Owner: Infosys) = AI output [Owner: Client] + Enhanced AI Model
2. Training dataset (Owner: Client) + AI model (Owner: Open-Source) = AI output [Owner: Client] + Enhanced AI Model
3. Training dataset (Owner: Client) + AI model (Owner: Third Party Vendor) = AI output [Owner: Client] + Enhanced AI Model.
4. Training dataset (Owner: Infosys/Licensed) + AI model (Owner: Infosys/Vendor/Open Source) = AI output [Owner: Client] + Enhanced AI Model

Considering that the ownership of AI systems, training datasets and the output generated by AI systems may rest with different organizations, clear provisions in this area will help bring business certainty, accelerate innovation, and prevent downstream litigation.

We are happy to provide any additional information or assistance as the IPO may require.

Yours faithfully,



Section A

Copyright – computer generated works (CGW)

1. *Do you currently rely on the computer-generated works (CGW) provision? If so, please provide details of the types of works, the value of any rights you license and how the provision benefits your business. What approach do you take in territories that do not offer copyright protection for computer-generated works?*

Response:

It is noteworthy that the UK was the first country to introduce provisions in copyright law to recognize copyrightability of computer-generated works back in 1988 through section 9(3) of the CDPA. Countries like Japan, Singapore, India, and China have also begun to recognise copyrightability of CGW/Computer-Generated Works (which essentially is 'output' generated by an AI system).

In the past few years, AI has been actively adapted and used in the IT industry as a part of digital transformation initiatives of client organizations. Typical current usage is for running internal operations, business processes etc. In such scenarios, the input needed is training dataset to train the AI models are typically proprietary data that are operational, personal, financial, or technical data of the user organizations.

On the other end, the output generated by the AI systems is usually business or control data. The interest in such AI generated output was historically low from a copyright perspective. The nature of outputs was such that it was mostly not suited for IP licensing or assignment (as the output was consumed operationally). The primary use case examples are computer vision, voice recognition, product research, text classification, fraud detection, preventive diagnostics, industrial control systems etc.

It is expected that with the emergence of new capabilities of AI in software development, a reasonable portion of software will be AI-generated in the near future. It is likely that user organizations will hesitate in adopting AI for software development if its copyright ownership is uncertain.

Organizations like OpenAI, EleutherAI, and others have matured the AI systems to the level where considerably useful software code generation is possible. Such AI generated software code certainly requires copyright protection as it will have to be used, licensed, enforced and/or assigned in the course of business by the IT companies.

Companies providing IT services typically develop software for their clients and are usually under a contractual obligation to transfer copyright/IP ownership of such software to the client. In the absence of copyright for such Computer-Generated works (CGW), this core contractual obligation will continue to remain unmet and on uncertain legal grounds.

Unlike other forms of literary works, software is a 'functional' work. Typically, it is licensed to users for business and technical consumption and can be fundamental to the work of many organizations. Inability to enforce copyright licenses & ownership may slow down or disrupt the workings of such organisations. Absence of protection for CGW may also have a negative impact on innovation and development of new software solutions in the IT industry.

On a related note, it is important not to assess the issue of market-failure against absence of copyright for CGW in other jurisdictions. Software development through AI is a relatively recent innovation and the impact of the absence of copyright ownership may not be realised for some time. In this respect there are limited downsides in continuing with the current provisions of Section 9(3) of the CDPA.

Finally, in our view, the issues around confusion between human and AI works, and the risk of false attribution are not substantial.

2. Please rank these options in order of preference (most to least preferred) and explain why.

Response:

- i) **Option 0: Make no legal change (Most preferred)** – As mentioned in our response to Question 1 above, the current statutory language of the CGW provision is sufficient to address the authorship issue in AI generated works. Most of the arguments listed in our response to Question 1 should be considered. We contend that no changes should be made to current provisions for CGW.
 - ii) **Option 2: Replace the current protection with a new right of reduced scope/duration** – The scope of copyright protection with regards to copying, selling, distributing, or making derivative works is appropriate for CGW. Other industries may argue for the reduction of the copyright term from 50 years (from the year of creation of CGW) however the term of copyright does not significantly impact the IT industry.
 - iii) **Option 1: Remove protection for computer-generated works (Least preferred)**- As the AI model and AI output is generated by the computer itself, removing protection does not serve anyone's purpose in the IT industry. Absence of any legal protection of AI models and AI output might reduce the adoption of this technology and innovation.
- 3. If we introduce a related right for computer-generated works, as per option 2, what scope and term of protection do you think it should have? Please explain how you think this scope and term is justified in terms of encouraging investment in AI-generated works and technology.*

Response:

We do not believe that any significant change should be made by changing the scope and term of copyright protection.

4. *What are your views of the implications of the policy options and of AI technology for the designs system? –*

5. *For each option, what are your views on the risk that AI generated works may be falsely attributed to a person? –*

Copyright – text and data mining (TDM)

Training of AI models through TDM results in ‘enhancement’ of the AI model. This essentially means modification to the neural network or to the traditional machine learning models of the AI under the influence of the training data set.

Such trained and enhanced AI models are of significant business value. It is therefore important that there is no legal ambiguity in terms of IP ownership of such enhancements that happen as part of training. Progress in AI is largely attributed to open-source AI models, data and frameworks. It is limited by restrictions in utilizing proprietary source code as training data (particularly for certain legacy technologies like COBOL or ABAP). For some of these technologies barely any open-source code is available, that can potentially be used for training.

There are several technical scenarios where TDM will be the differentiator. Multiple organizations with legacy systems need documentation for existing code (ii) multiple organizations requiring translation of code for the same set of software languages (ie translating VB to .NET), (iii) to train and then generate test cases of already existing code (iv) it will allow saving the infrastructure, effort & time costs for repeating the training effort multiple times over.

As per reports of certain analysts, widespread adoption of AI may create a double digit increase to the UK’s GDP in the next ten years. For robust AI systems to become available, it is critical that the AI models are trained with most relevant dataset, which is mostly possible through meaningful TDM exemptions.

6. *If you license works for TDM, or purchase such licences, can you provide information on the costs and benefits of these? For example, availability, price-point, whether additional services are included or available, number and types of works covered by the licence etc.*

Response:

Infosys typically use its own proprietary AI models that are trained on its own training datasets, open-source code, or licensed datasets. There is no specific insight about licensing, price points etc for works for TDM.

7. *Is there a specific approach the government should adopt in relation to licensing?*

Response:

It will create a level playing field if the UK's approach is aligned with the approach taken by the EU, Japan, and Singapore. Interestingly, organizations in the US are operating under a fair use exception (though, it is still not legally tested).

However, if the UK IPO decides to give an opt-out option, then it will be helpful for a "code of practice" to be created which will be followed by the rights holders and AI model users during negotiations of license agreements. The "code of practice" should include the following:

- a) Rights holders should not unreasonably withhold or refuse to provide the license on the copyrighted content for TDM activity.
- b) The conditions under which the rights holder can refuse to provide the license for TDM.

8. *Please rank the options in order of preference (most to least preferred) and explain why.*

Response:

Order of Preference: Option 4, Option 3, Option 2, Option 1

1. **Option 4: Adopt a TDM exception for any use, which does not allow rights holders to opt out (Most Preferred) –**

- a) The work generated by TDM (i.e., the enhanced AI model), is different from the original copyrighted works used for training purposes. In other words, the enhanced AI model is not a substitute of the original copyrighted work and there are no technical means for the end user to copy or reverse engineer the trained AI model to obtain the original work. This kind of use is equivalent to a 'transformative work' under the 'Fair use' concept in United States.
- b) If the UK IPO decides to go with Option 4, it will be important that there is sufficient clarity about the meaning of 'lawful access' to the training dataset to avoid potential disputes between organizations. It is critical that lawful access is defined in a way that it includes explicit as well as implicit licenses granted for a certain purpose (unrelated to TDM). In the absence of such clarity, the usage of training dataset may attract litigation by the dataset owners in several business-to-business contractual scenarios.
- c) It is equally important to note that the following additional provisions will be required as part of the proposed TDM exception: (i) the enhanced AI model should be explicitly recognized as 'derivative work' of

the AI model, and (ii) the person who ‘makes the necessary arrangements’ is recognized as the ‘author’ [it needs to be like authorship provisions under section 9(3)]. Ability to reuse, license, enforce the enhanced AI model is also a central issue to the development of AI. At present, these issues are open to interpretation and clear legal provisions will help in avoiding ambiguity.

- d) Considering that it is likely that ownership of AI systems, training datasets and the output generated by AI systems may rest with different organizations, the need to have clear law on the issues is higher than ever before. In that context, this consultation process initiated by the UK Government is very timely and appreciable.
 - e) It is also noteworthy that Japan, and Singapore have already proposed this kind of TDM exception for machine learning purposes.
2. **Option 3: Adopt a TDM exception for any use, with a rights holders opt out:** Providing an unfettered right to rights holders to opt out from providing a license for TDM may have different implications for different kinds of industry use cases. The opt out option will serve greater purpose in such scenarios where the result of the TDM will generate a work which is competitive to the original work. Furthermore, instead of a complete opt out option, it would be reasonable to have a compulsory licensing regime that allows a choice to the rights holder to charge/not charge the license fees for their work which will be used for TDM.
 3. **Option 2: Extend the existing TDM exception to cover commercial research and databases** – Extending current provisions of Section 29 is needed, but will require the safeguards featured in our response to Option 4 above.
 4. **Option 1: Improve licensing environment for the purposes of TDM** – In case the UK Government does not extend the TDM exception for commercial purposes, then the Government’s intervention will be required to strike a balance between the interest of the copyright holder and the AI Industry. The Government is requested to come up with “code of practice” to be followed by copyright holder and AI industries while negotiating the license agreement. The “code of practice” should include the following aspects:
 - c) Rights holders should not unreasonably withhold or refuse to provide a license;
 - d) The conditions under which right holders can refuse to provide the license;
 - e) Once a license is provided for TDM, the rights holder should not further restrict the purpose of TDM.
 5. **Option 0: Make no legal change** - As mentioned above, where the outcome of the TDM (i.e. the AI model) is different from the original work to the extent that it is no longer required to retain the copy of the original work, we do not believe it is justified in not allowing a TDM exception for commercial purposes. We request the UK Government to adopt a clear approach to address the TDM exception.

8. *If you have experience of the EU exception with opt out for rights holders, how has this affected you?*

9. How would any of the exception options positively or negatively affect you?
Please quantify this if possible.

Response:

All the exceptions articulated in Option 4 and Option 3 will make a positive impact in terms of availability of wider datasets for training. This is critical for accurate and efficient AI systems whilst having a minimal negative impact on copyright owners of the dataset.

Patents

10. Please rank these options in order of preference (most to least preferred) and explain why?
11. Would the changes proposed under Options 1, 2 and 3 have any consequential effects on the patent system, for example on other patentability criteria?

For options 1 and 2:

12. If UK patents were to protect AI-devised inventions, how should the inventor be identified, and who should be the patent owner? What effects does this have on incentivising and rewarding AI-devised inventions?
13. In considering the differences between options 1 and 2, how important is it that the use of AI to devise inventions is transparent in the patent system?
14. Would the UK adopting option 2 affect your global patent filing strategy, if so, how?

For option 3:

15. What term and scope of protection should a new right offer?
16. What should the criteria for grant of a new right be and why? Particularly should it:
 - a) Replicate the current requirements for a patent?
 - b) Set a different bar for inventive step?
 - c) Be an automatic or registered right?

General

17. What role does the IP system play in the decision of firms to invest in AI?

Infosys Response:

With the emergence of AI in software development, the role of the IP system is significantly enhanced in decision making process within the organizations.

18. Does the first mover advantage and winner-take-all effect prevail in industries adopting AI? How would this affect the impact of the policy options proposed on innovation and competition?

Response:

Our preliminary view is that AI is not inherently monopolistic for the first mover. It has a very broad range of use cases for everyone to participate.

19. How does AI adoption by firms affect the economy? Does the use of AI in R&D lead to a higher productivity?

Response:

Use of AI most certainly improves productivity.

20. Do the proposed policy options have an impact on civil society organisations? If so, what types of impacts?

Section B: Respondent information

A: Please give your name (name of individual, business or organisation).

[REDACTED]

B: Are you responding as an individual, business or on behalf of an organisation?

- 1) Business – please provide the name of your business : **Infosys Ltd.**
- 2) Organisation – please provide the name of the organisation
- 3) Individual – please provide your name

C: If you are responding on behalf of an organisation, please give a summary of who you represent. **NA**

D: If you are an individual, are you? **NA**

- 1) General public
- 2) An academic
- 3) A law professional
- 4) A professional in another sector – please specify
- 5) Other – please specify

E: If you are responding on behalf of an organisation, are you? **NA**

- 1) An academic institution
- 2) An industry body
- 3) A licensing body
- 4) A rights holder organisation
- 5) Any other type of organisation - please specify

F: If you are responding on behalf of a business or organisation, in which sector(s) do you operate? (choose all that apply)

- 1) Agriculture, forestry and fishing
- 2) Mining and quarrying
- 3) Manufacturing – Pharmaceutical products

- 4) Manufacturing – Computer, electronic and optical products
- 5) Manufacturing – Electrical equipment
- 6) Manufacturing – Transport equipment
- 7) Other manufacturing
- 8) Construction
- 9) Wholesale and retail trade; repair of motor vehicles and motorcycles
- 10) Transportation and storage
- 11) Information and communication – Publishing, audio-visual and broadcasting
- 12) Information and communication – Telecommunication
- 13) Information and communication – IT and another Information Services**
- 14) Financial and insurance activities
- 15) Real estate activities
- 16) Scientific and technical activities
- 17) Legal activities
- 18) Administrative and support service activities
- 19) Public administration and defence
- 20) Education
- 21) Human health and social work activities
- 22) Arts, entertainment and recreation
- 23) Other activities – please specify

G: How many people work for your business or organisation across the UK as a whole? Please estimate if you are unsure.

- 1) Fewer than 10 people
- 2) 10–49
- 3) 50–249
- 4) 250–999
- 5) 1,000 or more**

H: The Intellectual Property Office may wish to contact you to discuss your response. Would you be happy to be contacted to discuss your response? **YES**

I: If you are happy to be contacted by the Intellectual Property Office, please provide a contact email address.

[Redacted]

J: Would you like an acknowledgement of receipt of your response? **Yes/No**