YNZ GROUP

CAN AI TRULY "CREATE"! COPY"RIGHT" OR WRONG?

WEDNESDAY WISDOM
01-01-2025

Generation of a verse[1], letters or even a report in a matter of minutes has become easy and convenient at the click of the keys, give appropriate commands on the artificial intelligence platforms like OpenAI's ChatGPT, Google's Gemini and Microsoft Copilot and there comes the verse: "Flowing through the woods...". Ever wondered what if the same verses are presented by poets at poetry reading, either poet claiming the Authorship[2] of the poem.

1. What is Generative Artificial Intelligence ("Generative AI")?

Generative AI has not been defined under any legislation across the world. Generative AI is the general-purpose AI model which generates content like text, images, videos or audio, etc.

However, Generative AI may be considered under the General Model AI. General-purpose AI model under European Union Artificial Intelligence Act, 2024 means an AI model which is trained with large amounts of data using self-supervision at scale, which displays significant generality and is capable of competently performing a wide range of distinct tasks[3] .

2. How does Generative AI get a human touch?

The answer to our question lies in the Language Learning Model (LLM).

LLM is a type of artificial intelligence that excels at processing, understanding and generating human language. The work generated through LLM will depend on the quantum of training and resources available for such LLM. [4]

The training of LLMs require vast amounts of textual data from books, articles, websites so as to form a sufficient database for the processing. Such training will include genre like scientific, financial, medical, legal, artistic etc. to ensure a complete perspective of any issue to LLM.

[1] The article reflects the general work of the authors and the views expressed are personal. No reader should act on any statement contained herein without seeking detailed professional advice.
[2]Use you can also read for more information on: https://www.ynzgroup.co.in/articles/IP/Copyright,%20Authorship%20And%20Originality%20A%20Digital%20Trilogy.pdf
[3] Article 3 (63) of the European Union Artificial Intelligence Act, 2024
[4] https://sap.com/resources/what-is-large-language-model

For an Artificial Intelligence tool to generate an image on the command: "Generate a beautiful image with a hut and a background of mountains, the scenery should consist of a river flowing peacefully through the valley. The scene should consist of a farm of strawberries between the river and the hut and season of strawberries is in full bloom".

For an AI tool to process commands consisting of adjectives and phrases used by humans such as "peacefully" and "season is in full bloom", it needs relevant training wherein it can analyze and produce the required output. To comprehend the commands, the LLM are often trained with books or poems which describe in detail how does the season in full bloom or how does peaceful is perceived by humans to produce such output.

3. Are there any legal issues with the usage of training material by LLM?
Yes, legal issues of copyright infringement are coming up.
The creators of original books and media agencies are alleging that artificial Intelligence tools like OpenAI have been using the content generated by various authors, content creators and media agencies without giving them due credit and without obtaining proper licenses to such materials.
This is giving a rise to conflicts for the infringement of the Intellectual Property of such material by the Artificial Intelligence tools.

4. Who is an Author?

An "author" in relation to Copyright, which is a distinctive Intellectual Property for any creative work such as —

i. in relation to a literary or dramatic work: the author of the work;
ii. in relation to a musical work: the composer;
iii. in relation to an artistic work other than a photograph: the artist;
iv. in relation to a photograph: the person taking the photograph;
v. in relation to a cinematograph film or sound recording: the producer; and
vi. in relation to any literary, dramatic, musical or artistic work which is computer-generated: the person who causes the work to be created;[5]

5. How does one determine the authorship in work generated by Artificial Intelligence?

The definition of Author under the Copyrights Act, 1957 clearly states that the author for computer generated work shall be the person who created such work.

As quoted by one of the Generative AI Tool with regards to ownership of the content generated, *"To the extent permitted by applicable law, you (a) retain your ownership rights in Input and (b) own the Output. We hereby assign to you all our right, title, and interest, if any, in and to Output"*[6]

However, another Generative AI tool stated *"Some of our Services allow you to generate original content. We won't claim ownership over that content. You acknowledge that we may generate the same or similar content for others and that we reserve all rights to do so."*[7]

The determination of the authorship still remains grey area. Generative AI produces the content based on the commands and prompts by the user, but it does not assure that similar content shall not be regenerated.

---

[5] Section 2(d) of the Indian Copyrights Act, 1957
[6] https://openai.com/policies/row-terms-of-use/
[7] https://ai.google.dev/gemini-api/terms

YNZ GROUP

To understand how Generative AI functions and whether it produces similar content, we conducted an experiment wherein a few people gave a same prompt to generate a dialogue between 2 school friends meeting after a long time. The dialogue conversations generated were not identical however certain common traits were found in the generated work such as names of the characters or setting of the place where the characters meet, etc. Following are the excerpts of the results obtained:

Setting: A bustling coffee shop near a university.
Characters:
- **Alex**: A cheerful and energetic individual, always curious about others' lives.
- **Sam**: More reserved, but warm and sincere, with a thoughtful demeanor.

Alex: [Waving excitedly] Sam? Is that you?! Oh my gosh, it's been ages!

Sam: [Turning around, eyes widening] Alex! No way! What are the odds of running into you here?

Alex: Seriously! What's it been, like... five years since high school graduation?

Sam: More like six. Time flies, huh?

Alex: [Pulling out a chair] Mind if I sit?

Sam: Of course not! Sit down. So, what's new? Tell me everything. Last I heard, you were off chasing some big dreams.

Alex: [walking into the coffee shop] "Is that really you, Jamie? I can't believe it!"

Jamie: [looks up from their phone, surprised] "Alex? Oh my god, it's been ages! I didn't expect to see you here! How long has it been? Ten years?"

Alex: "At least! The last time we saw each other was... graduation day, right? Wow, time flies."

Jamie: "Tell me about it! I almost didn't recognize you at first. You look... different, but in a good way. What's been going on with you?"

Alex: [laughs] "You too! I guess we've both changed a lot. Well, I moved to another city for work right after graduation. Been settling in, you know, the usual adult stuff—job, apartment, figuring life out."

Jamie: "Same here, honestly. After school, I went to college, then traveled for a bit. Ended up back here for a job in marketing. But I think I've changed the most in terms of my hobbies. Remember how I used to be obsessed with playing guitar? Now, I've picked up photography."

But who shall be eligible to claim the authorship in case two people generate same content with similar type of prompt using Generative AI, does the first person to generate the content be considered as a crucial factor in determining the authorship?

6. What are various perspectives to these copyright issues?

Fair Use exception: As per Section 52 of the Copyrights Act, states the acts which shall not constitute an infringement of copyright, such acts are as follows:

- Content or information used for private or personal use including research;
- Criticism or review of any work;
- Reporting of current affairs and events;
- Transient or incidental storage of work in the technical process of electronic transmission or communication to the public;
- Reproduction of any work in accordance with any law;
- Reading or recitation in public from a published literacy or dramatic work
- Reproduction of any work for academic purposes, etc.

Determining fair use of copyrighted material is crucial in any copyright infringement claim. Exemptions from copyright infringement include private or personal use, such as research, criticism, review and reporting of current affairs[8].

[8] Section 52(1)(a) of the Indian Copyrights Act, 1957

7. What is the ongoing dispute about Copyright Infringement?

ANI (Asian News International), an international news agency that licenses data to news and media organizations has filed a suit against OpenAI, the company behind one of the most popular Artificial Intelligence (AI), claiming copyright infringement with false attribution and improper use of their content. [9]

ANI Allegations:	V/S	OpenAI Defences:
Use of ANI content to train the LLM of OpenAI without proper licensing and permissions		OpenAI stated that it uses the publicly available information published by ANI and in case if ANI wishes to block use of their published information, it can block the access to such information by the commonly used methods for restricting web crawlers
Use, publishing and storage of ANI material which is made accessible by ANI only to subscribers.		OpenAI clarified that it does not have an access of the subscription-based or licensed content of ANI and has only utilised the publicly available information for the training of its LLM

ANI Allegations: **V/S**   **OpenAI Defences:**

> **False attributions and publication of such work not produced by ANI**

> **OpenAI also denied the accusations relating to false attributions stating that no grievances have been raised so far stating the inaccuracies in the attributions.**

> **Similarly, 13 legal proceedings around the world have been filed by authors, content creators across the world against OpenAI for illegal use of content for training of the LLM.**

> **OpenAI also admitted that they are facing multiple legal actions across US, Canada and Germany and no court so far has granted an injunction against Copyright infringement on the part of OpenAI.**

OpenAI has also upheld a tough stance and argues that the Court does not have territorial jurisdiction over the said matter as the OpenAI does not have physical servers in India where the training of such LLM take place .[10]

8. What is the current stage of the ongoing dispute between ANI vs. OpenAI in India?

The Hon'ble High Court of Delhi has passed an interim order issuing summons to OpenAI wherein the jurisdiction aspects of the case shall be addressed, and the Court has also appointed Amicus Curiae to assist in legal proceedings. The matter is sub-judice and the next hearing is scheduled for January 28, 2025.

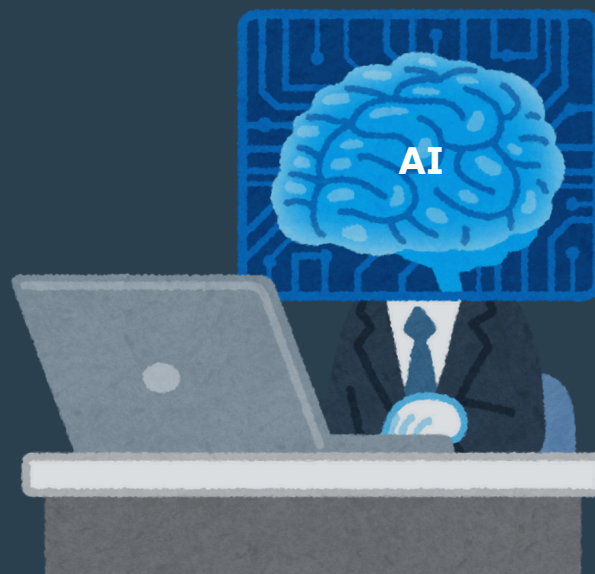[10] ANI v OpenAI: A copyright, AI training and false attribution dispute

9. What are the emerging trends adopted by certain Artificial Intelligence platforms to safeguards the rights of the end users against any Copyright infringement claims?

Platform like Microsoft Co-pilot on May 1st, 2024 declared an expansion of Co-pilot Copyright Commitment, which shall ensure to cover the outputs generated by the Generative AI platform Microsoft Copilot. The Commitment program by Microsoft Co-pilot shall defend the commercial customers against any Copyright infringement claims and cover any litigation costs for adverse judgements or for settlements arising from the use of Microsoft's Copilot generated works.

However, such claims shall be addressed by Microsoft if the commercial customers comply with the key conditions while generating such work, which are as follows:

i. Use of guardrails and content filters – Microsoft Copilot has some inbuilt filters and technologies which are likely to reduce infringement of third-party content which include guardrails like classifiers, metaprompts, content filtering, etc.

ii. No attempt to generate infringing works – The end users should not intentionally generate or recreate similar works causing infringement of Intellectual Property Rights for which they do not possess any rights. [11]

[11]https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/

Conclusion:

The evolution of Artificial Intelligence has the potential to transform any industry to its fullest potential, even beyond the human intellect. However, with its rapid advancement, there is an absence of legal framework to keep pace raising questions regarding copyright infringement, data protection and potential for misuse.

However, a small advisory has been issued by the Cyber Law and Data Governance Group, Ministry of Electronics and Information Technology, which states:

- To ensure the due diligence of content has been conducted prior to the content being posted on AI platforms and Intermediaries under the IT Act 2000 and as per the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.
- The advisory requires that intermediaries and platforms to obtain explicit permission from the Government of India before using or making available any under-tested or unreliable Artificial Intelligence (AI) models, Large Language Models (LLM), Generative AI models, software, or algorithms to Indian users. [12]

Therefore, a robust and adaptable legal framework is crucial to be implemented towards the responsible development and use of Generative AI while addressing the concerns of copyright infringement, data protection and misuse of the content.

[12]https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf

YNZ GROUP

For any feedback or response on this article, the authors can be reached on riddhi.bhosale@ynzgroup.co.in and ashvini.kandalgaonkar@ynzgroup.co.in

Author: Riddhi Bhosale

Riddhi is an Associate at YNZ Legal. By qualification she has completed her Bachelor of Science in Biotechnology and Bachelor of Law from University of Mumbai.

Co-author: Ashvini Kandalgaonkar

Ashvini is a Partner- Corporate Legal Advisory, She is experienced in corporate litigation and non- litigation. She provides training required under POSH Act.