

Copyright and Privacy JP Essey

One of the fundamental attributes in a democratic, capitalist state is the concept of ownership and private property, which can take on many forms from physical objects to intellectual ideas and original, creative expressions in a variety of modalities. Private ownership enhances a person's sense of responsibility in a rippling effect. An individual is incentivized to work harder to attain, use and keep their property as well as to create and innovate. This in turn can have an advantageous value to the individual as well as to a country's economy and stability. As our society has become ever more connected and digitized in the past several decades, the concept of copyright has been under stress and is in need of evaluation and clarification.

In its most fundamental form, copyright is a legal framework that gives the creator of an original work, the exclusive right to show, reproduce, sell, or distribute their work for a set number of years. A copyrighted piece of work may not be used without the permission of the creator. There are a few exceptions to this such as fair use or when included in certain instances such as an educational critic or a critical review. This is done to ensure that a creative environment is fostered in the society.

With LLMs and the method of the way they are trained, fault lines as to what is considered as a copyright infringement have come to the fore. Many authors of published works state that their copyrighted work has been infringed upon without financial compensation due to an LLM ingesting their work and including it in some output, at times with a financial reward paid to them. From the creator's point of view, they did not authorize the use of their work nor did they receive any compensation from it, and they should be entitled to a financial reward as well as having to ask to allow use. Besides those concepts there is the fundamental idea of loss of control over their work. This would disincentivize individuals from producing original work.

The response to that perspective is that an LLM is not reproducing an author's work but their idea, which is not copyrighted. There's a strong possibility that no copyright was infringed upon as the law protects a particular expression of an idea or fact; it does not give exclusive ownership to that idea. And an LLM 'reads' a copyrighted text to 'learn' a particular subject and not to memorize an author's particular turn of phrase. An analogy is similar to a student or researcher. (Rahman and Santacana, 2023)

The topic of privacy is more of an immediate concern as it impacts not only authors of created, original work but any individual who uses the internet. Although a person who uses the internet has most all of their digital footprint marked somewhere, current LLMs gather it all and make it easily accessible in one easy location. This leakage of personal identifiable information (PII) is of a serious concern. The consequences can be extremely detrimental to an individual. An individual can have a whole array of negative consequences from diminished personal liberty by fear of being observed or followed, to financial fraud, to being stigmatized for their expression or POV of having said something, to potential career damage, to distrust of digital systems, and much more. This is an extremely pressing issue.

Both copyright and the issue of privacy need the focus and urgency commensurate with the capacity for harm that LLMs can cause. Those issues should be at the head of organizational to-do lists if LLMs are to be productive tools in our culture to enhance our capacities rather than diminish them.