

# AI and Intellectual Property

Module 10 of a course on *Ethical Issues in AI*

*Prepared by*

**John Hooker**

*Emeritus Professor, Carnegie Mellon University*

CMU Osher, February 2024

# Plagiarism?

- Generative AI can produce images or text very similar to copyrighted material.
  - *Even if the prompt does not specifically request the material or its source.*

GPT-4 response to prompt  
“black armor with light sword.”

Source: G. Marcus and R. Southern,  
“Generative AI has a visual plagiarism  
problem,” *IEEE Spectrum*, 6 Jan 2024.



# Plagiarism?

- Generative AI can produce images or text very similar to copyrighted material.
  - *Even if the prompt does not specifically request the material or its source.*

GPT-4 response to prompt  
“animated toys.”

Source: G. Marcus and R. Southern,  
“Generative AI has a visual plagiarism  
problem,” *IEEE Spectrum*, 6 Jan 2024.



# Plagiarism?

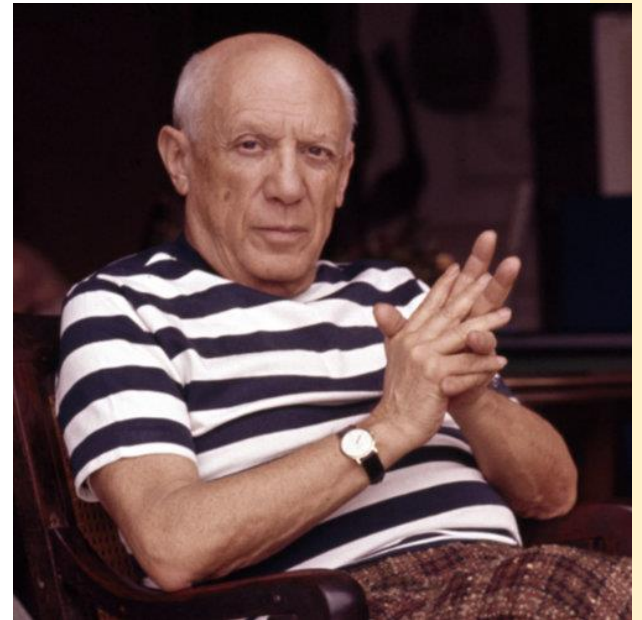
- Generative AI can produce images or text very similar to copyrighted material.
  - *Even if the prompt does not specifically request the material or its source.*
    - Because GPT etc. are trained by scraping online material.
  - *Some authors/artists claim this is **plagiarism** that threatens their **livelihood**.*
    - It allows content producers to “steal” ideas from truly creative people.
    - There have been a number of **lawsuits** (e.g. *NY Times*)
    - OpenAI has made agreements to **pay** some parties for use of copyrighted material in training set.

# Plagiarism?

- The **legal** issues are unsettled.
  - *IP law is complicated and confusing*
    - ...with little coherence since genetic engineering.
- We focus on the **ethical** angle.
  - *Does generative AI unethically use others' **intellectual property**?*

“Good artists borrow, great artists steal.”

--Attributed to Pablo Picasso

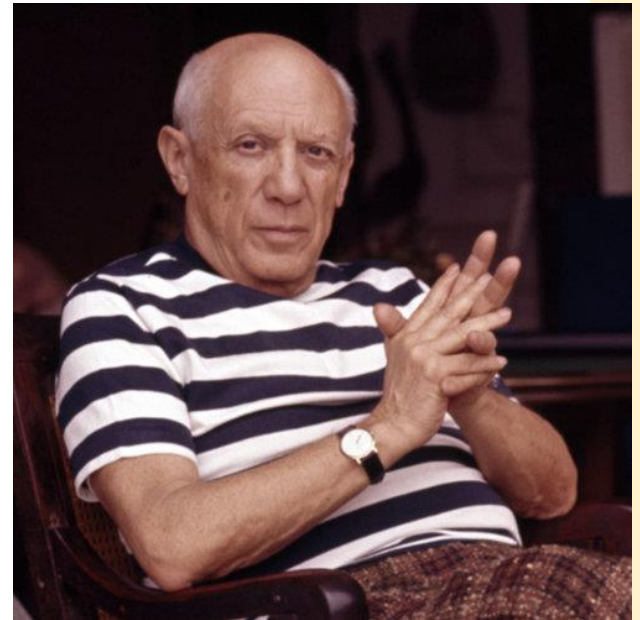


# Plagiarism?

- The **legal** issues are unsettled.
  - *IP law is complicated and confusing*
    - ...with little coherence since genetic engineering.
- We focus on the **ethical** angle.
  - *Does generative AI unethically use others' **intellectual property**?*
  - *We begin with an ethical analysis of **property**, then **IP**.*

“Good artists borrow, great artists steal.”

--Attributed to Pablo Picasso



# Intellectual property?

- Does the concept of **intellectual property** make sense?
  - *Is there an ethical basis for it?*
- The concept of **intellectual property** is relatively **recent**.
  - *Talk of “IP” can leave the impression that IP is **like other property**.*



# Intellectual property?

- **Patents** were introduced to encourage **disclosure of ideas**, not to establish **property rights**.
  - For example, in the U.S. Constitution
- One cannot patent or copyright **ideas**.
- The more recent concept of “intellectual property” tends to **legitimize strong IP rights**.





# Intellectual property?

- IP is **not** like other property.
  - *One can use IP without denying others the use of it.*
- The generalization argument for property rights **breaks down...**

# Property rights

- **Generalization argument:** Violating property rights is cannot be generalized.
  - *The thief takes someone's property because he wants it to have **possession and use** of the property.*
  - *But if no one respects property rights, the thief **won't have** possession and use.*
  - *In effect, there will be **no property**.*



# Property rights

- But **millions of people** can possess and use IP at the same time.
  - *So taking IP to possess and use it is **generalizable**.*



# IP online

- None of the previous says that there is no reason to respect IP.
  - *Only that the generalization argument does not apply.*

# IP online

- None of the previous says that there is no reason to respect IP.
  - *Only that the usual generalization argument doesn't apply.*
- A **specific** generalization argument applies to **GPTs and online content...**

# IP online

- Scraping online content without permission or payment...
  - *...for purposes of training a GPT*
  - *...is **not generalizable** if the type of content in question **would not exist** if GPT developers **never paid for it**.*
- It doesn't matter whether the GPT output **resembles** the input.
  - *It's **not a question of plagiarism**.*

# **Appendix**

## **Intellectual Property Law**

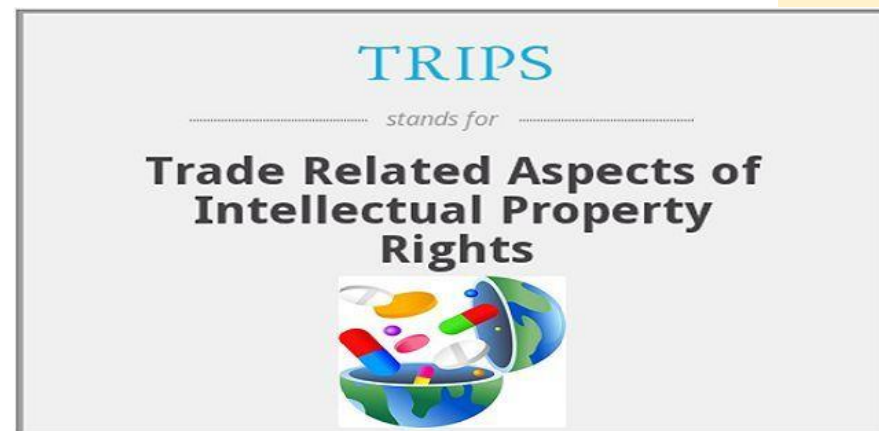
*A Brief Introduction*





# Intellectual property law

- The primary international treaty is the **TRIPS agreement**.
  - *Trade-Related Aspects of Intellectual Property Rights.*
  - *Added to **GATT** (General Agreement on Tariffs and Trade) at the Uruguay Round, 1994.*
  - *Amended at 2001 **WTO** Ministerial Conference in Doha.*
  - ***TRIPS-plus** consists of many bilateral agreements.*



# Intellectual property law

- **U.S. Constitution** provides for Federal regulation of patents and copyrights.
  - Administered by **U.S. Patent and Trademark Office & U.S. Copyright Office.**



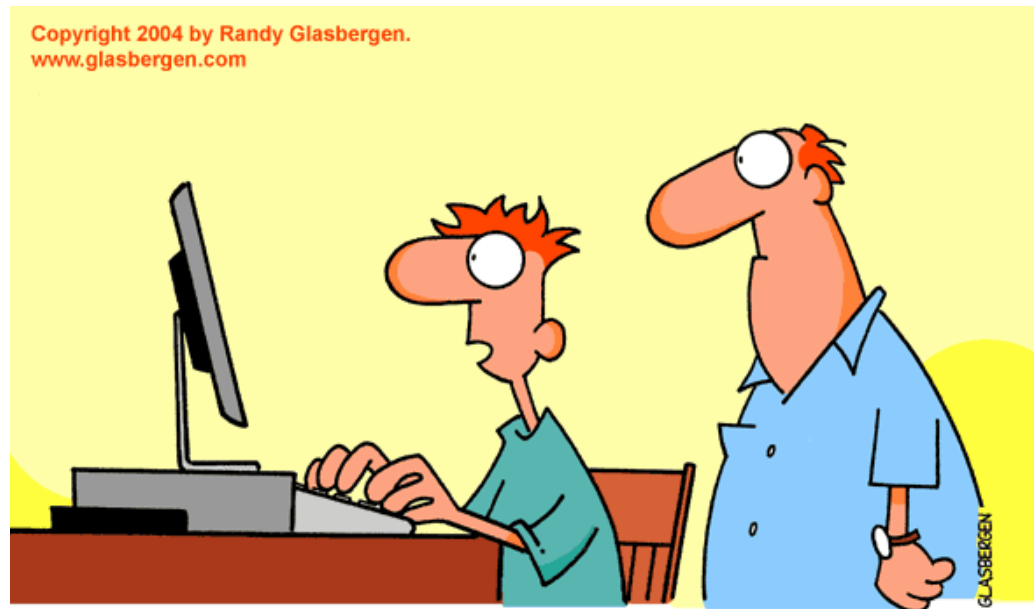
# Intellectual property law

- Main types of intellectual property:
  - **Copyright.**
  - **Trademark.**
  - **Patent.**
  - **Trade secret.**



# Copyright

- A **copyright** limits the number of copies others can make of a document or work of art without permission.



"I KNOW IT'S ILLEGAL TO DOWNLOAD MUSIC FROM  
THE INTERNET... BUT YOU SAID THE STUFF  
I LISTEN TO ISN'T MUSIC!"

# Copyright

- A **copyright** limits the number of copies others can make of a document or work of art without permission.
- **Duration** of copyright:
  - *In TRIPS:*
    - **Life + 50 years**
  - *In U.S.:*
    - Individual: **Life + 70 years**
    - Corporate-owned works: **95 years** from publication, or **120 years** from creation, whichever is shorter.



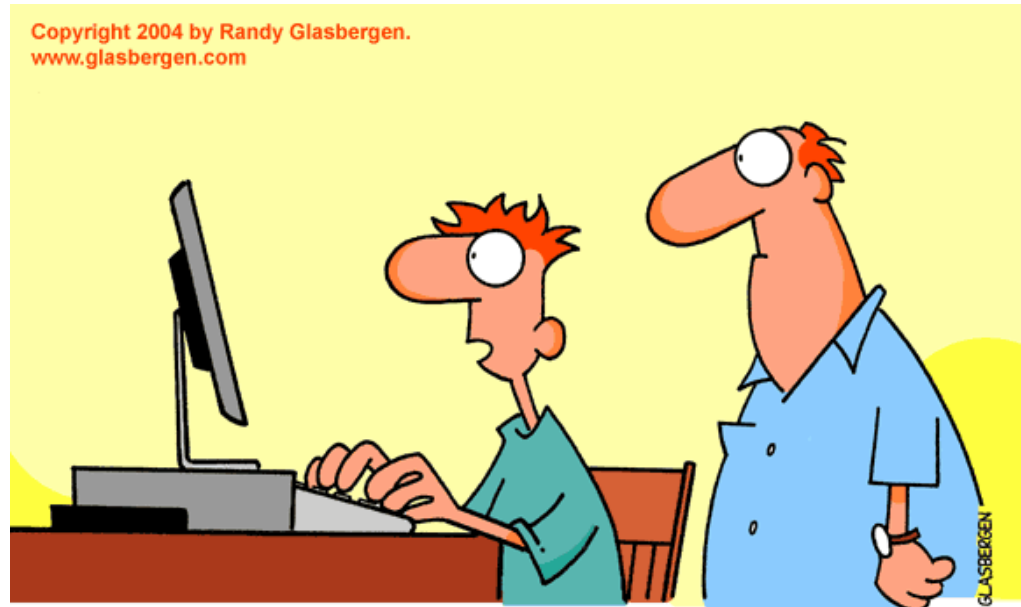
# Copyright

- A **copyright** limits the number of copies others can make of a document or work of art without permission.
- **Ideas** cannot be copyrighted.
  - *Only a particular expression of ideas.*
  - *Such as a literary work or musical composition.*



# Copyright

- Fair use.
  - *Is it legal to download this cartoon?*



"I KNOW IT'S ILLEGAL TO DOWNLOAD MUSIC FROM  
THE INTERNET... BUT YOU SAID THE STUFF  
I LISTEN TO ISN'T MUSIC!"

# Copyright

- From Section 107 of U.S. Copyright Act (emphasis added):

*The **fair use** of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, **teaching** (including multiple copies for classroom use), scholarship, or research, is **not** an infringement of copyright.*

*In determining whether the use made of a work in any particular case is a **fair use** the factors to be considered shall include—the purpose and character of the use, including whether such use is of a commercial nature or is for **nonprofit educational purposes**.*



# Copyright

- Can one copyright **software**?
- Generally, **yes**.
  - The TRIPS agreement explicitly says yes.
  - But one **cannot** copyright the underlying idea or algorithm.

```
#include <stdio.h>
int main(void)
{
    printf("Hello World!\n");
    return 0;
}
```

**C/C++**

# Trademark

- A **trademark** is a word, phrase, symbol, or design that **distinguishes** the maker or source of goods.
  - A *service mark* serves the same function for a service.



# Trademark

- A trademark can be **registered**.
  - *Or can be established in common law by commercial use.*
- Legal rights to a trademark can last **forever**.
  - *As long as it is in commercial use.*
    - *Cannot have legal rights to a trademark **already registered or in use** by another company in the same product class.*



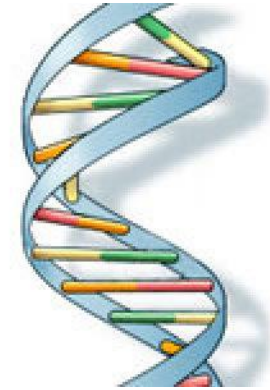
# Patent

- Designed to encourage **disclosure** of ideas in exchange for limited period of exclusive use.
  - **20 years** in TRIPS
  - In U.S.: **20 years**, except **14 years** for a design patent
- Can patent:
  - A **new** method, device, or material.



# Patent

- Cannot patent:
  - *A pure idea, such as a theorem.*
  - *Anything that occurs in nature.*
  - *A “way of doing business,” even if automated by computer.*
  - *“Look and feel,” e.g. spreadsheet.*



# Patent

- Can **software** be patented?
  - *Controversial, varies by country.*
  - *TRIPS agreement is **vague**.*
  - *In the U.S. one can patent:*
    - **A software-based invention**
    - Software that improves **computer functionality**
    - Software that solves a problem that arises only in the context of **computer technology**
  - *Ask a **patent lawyer** (a recognized specialty)*



# Patent

- Can an **algorithm** be patented?
  - *In principle, no.*
    - Technically, it must be embodied in some device.
  - *In practice, yes.*
    - Many algorithms have been patented.
    - Patent is not issued for source code.
    - It is issued for a “series of steps.”
    - ...which is an algorithm
    - However, **bubble sort** was denied a patent.



Muhammad Ibn Musa  
**Al-Khwarizmi**  
The Father of Algebra and Algorithms

# Patent

- A patent is **registered** in the name of the inventor.
- The **owner** may be someone else, or a company.
  - An employer normally **owns** any idea conceived by someone working **for hire**.





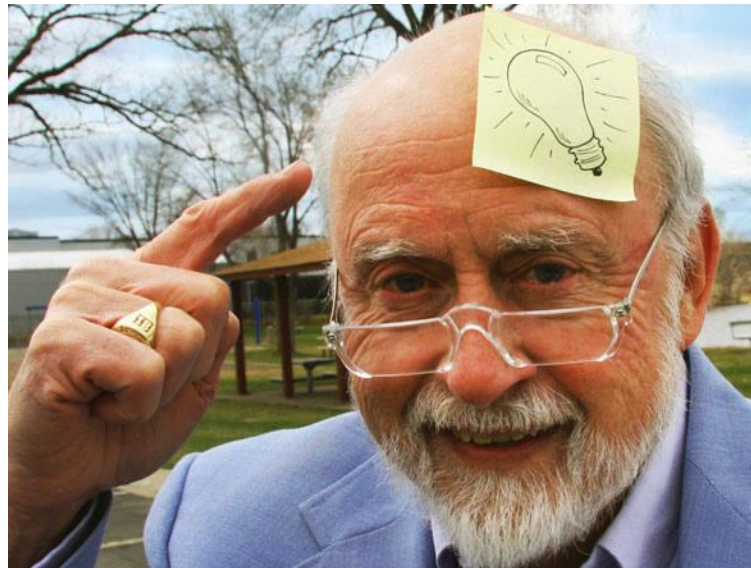
# Patent

- Who works **for hire**?
  - Normally, **full-time employees** work for hire and do not retain IP rights.
  - Normally, **consultants** do not work for hire, depending on contract



# Patent

- Case study: Post-it Notes
  - *Invented by Arthur Fry in 1974*
    - To make notes in choral music.



# Patent

- Case study: Post-it Notes
  - *Invented by Arthur Fry in 1974*
    - To make notes in choral music.
  - *Fry was employed by 3M.*
    - 3M marketed Post-it Notes in 1980.
    - Today, over \$2 billion annual revenue.
    - Fry received **zero royalties** for his invention.



# Patent

- Case study: Post-it Notes
  - *Invented by Arthur Fry in 1974*
    - To make notes in choral music.
  - *Fry was employed by 3M.*
    - 3M marketed Post-it Notes in 1980.
    - Today, over \$2 billion annual revenue.
    - Fry received **zero royalties** for his invention.
    - But he was promoted to “Corporate Researcher”
    - National Inventors Hall of Fame, 2010.



# Trade secret

- A **trade secret** is a secret formula or device that provides a commercial advantage.
  - *It can be bought, sold and licensed.*
- It remains intellectual property forever, or until the secret gets out.
  - *For example, the formula for Coca-Cola.*



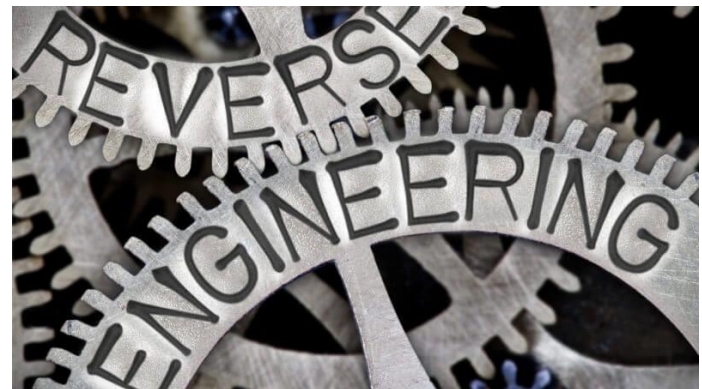
# Trade secret

- *There is one copy of the Coca-Cola formula.*
- *Kept in a vault in Atlanta, USA.*
  - Two people have access to the vault.
  - Their identities are secret.
  - Not allowed to travel together
  - When one is about to die, he/she chooses a successor
- *Syrup shipped as 9 components, combined at factory.*



# Trade secret

- The law does not prohibit **use** of a trade secret.
  - *It only prohibits others from **stealing** a trade secret.*
- It is normally legal for another company to conceive the idea independently and use it.
  - ***Reverse engineering** is not theft (the idea was not really secret).*



# NDAs

- A **nondisclosure agreement** prevents disclosure of proprietary information.
  - *The **duration** of the NDA is usually stated (1-5 years).*
    - If not stated, it may be forever.
  - *“**Proprietary**” is defined in the NDA.*
  - *Disclosure of **trade secrets** at any time is **illegal**.*
    - With or without an NDA.
    - Generally the company identifies what it considers to be trade secrets.

