



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Legal dimension of artificial intelligence – intellectual property law perspective

dr Monika Szkarłat

Centre for Artificial Intelligence and Computational Modeling
Marii Curie-Skłodowska University in Lublin

OSLO, 3-4.10.2024

Who am I?



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Plan

1. AI in science - introductory remarks
 - a) Benefits
 - b) Risks
1. Copyright and AI in EU law and national legal regimes
 - a. EU AI Act
 - b. Directive 2019/790/EU on copyright in the Digital Single Market
 - c. Definition of authorship/ownership
 - d. Infringement of copyright by creators of AI, users

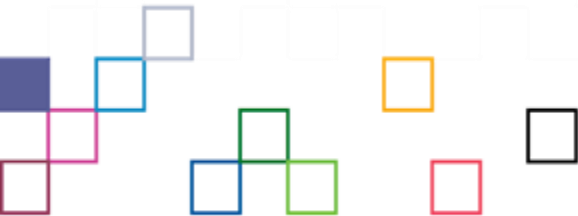


AI in science - introductory remarks



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

The most prominent AI techniques across STEM fields include artificial neural networks (ANNs); machine learning (ML) (including deep learning (DL)); natural language processing; and image recognition.





Advanced simulations and data representation

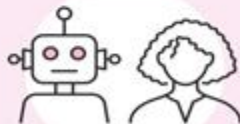
Computational microscope

Resource of inspiration

Agent of understanding



Acquiring new scientific understanding



Transferring scientific insights to a human expert



Identifying surprises in data and models



Identifying areas of interest from literature

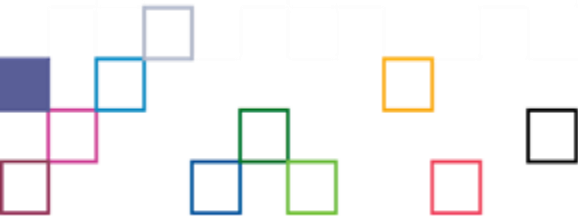
Source: The Royal Society, 2024.



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Benefits

- improve the quality of academic writing,
- assist with translation, or emulate specific writing styles (e.g. producing lay summaries),
- streamline administrative tasks and assist in drafting grant applications,
- accelerates the process of data collection and analysis,
- improve accessibility for researchers from diverse backgrounds.



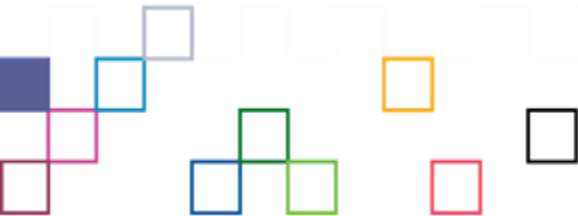
Risks

- 'knowledge workers' (Drucker, 1959) - what's the future for them in the age of AI?
- T. S. Eliot: 'good authors borrow, great authors steal'/'good models borrow, great models steal' (Chesterman, 2024);
- de-skilled scientists;
- principle of 'garbage in, garbage out';
- securitization/weaponization of scientific data;
- lack of understanding about how AI models work;



Risks

- ML models are commonly referred to as 'black-box models';
- AI models as proprietary systems;
- Private sector dominance and centralisation of AI-based science development;
- The private sector and open science;
- LLMs in a scientific project - increased exposure to 'hallucinations';



ARTIFICIAL INTELLIGENCE

Why Meta's latest large language model survived only three days online

Galactica was supposed to help scientists. Instead, it mindlessly spat out biased and incorrect nonsense.

By Will Douglas Heaven

November 18, 2022

Source: <https://www.technologyreview.com/2022/11/18/1063487/meta-large-language-model-ai-only-survived-three-days-gpt-3-science/>



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Copyright and AI in EU law and national legal regimes

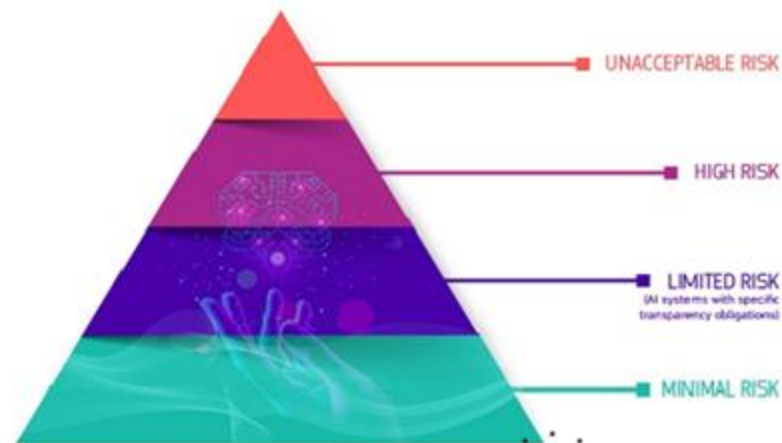


EU AI Act

Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence

A risk-based approach

The Regulatory Framework defines 4 levels of risk for AI systems:



Source: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

EU AI Act

How does it all work in practice for providers of high-risk AI systems?

STEP 1



A high-risk AI system is developed.

STEP 2



It needs to undergo the conformity assessment and comply with AI requirements.*

*For some systems a notified body is involved too.

STEP 3



Registration of stand-alone AI systems in an EU database.

STEP 4



A declaration of conformity needs to be signed and the AI system should bear the CE marking.
The system can be placed on the market.

If substantial changes happen in the AI system's lifecycle

GO BACK TO STEP 2

Source: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

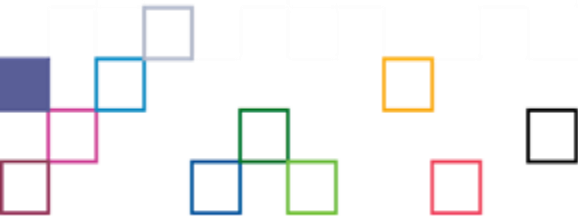
AI system means “software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate output such as content, predictions, recommendations, or decisions influencing the environments they interact with”.

“(…) current EU copyright law, as interpreted by the CJEU, leaves room for the protection of **AI-assisted output** in a wide range of creative fields. As long as the output reflects creative choices by a human being at any stage of the production process, AI-assisted output is likely to qualify for copyright protection as a ‘**work**’.”



- AI assisted output vs. AI generated output;
- Concept of co-authorship, but with whom?

DSM Directive - the legal definition of **data mining** was introduced along with several exceptions or limitations, which only apply when the data/copyright works are used for the training of the AI systems;



Copyright and AI (Definition of authorship/ownership)

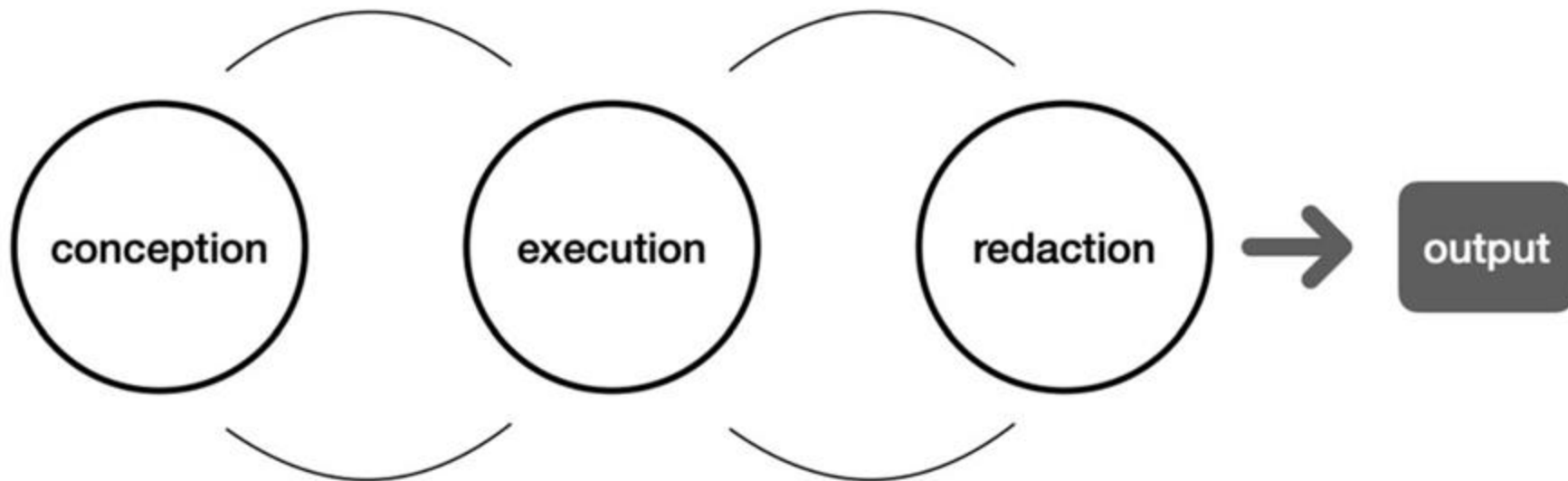
- A work protected by copyright law is the result of human labour;
- A work is the result of human labour, is original, possesses individual character, and is fixed;
- There is ongoing global debate as to whether a book written by AI will be protected by copyright law;
- Machines as right-holders and duty bearers;
- Creative contribution of a human being and AI model - case-by-case approach;



Diagram of an iterative creative process

Idea

Expression



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Copyright and AI (Definition of authorship/ownership)

Two approaches under consideration regarding the copyright protection of AI-generated works:

1. The AI-generated creation does not belong to anyone, and thus, anyone may use it.
2. The AI-generated creation is part of the public domain, meaning it is subject to legal protection, but anyone may use it.



The copyright status of AI-generated works is currently a complex and evolving legal issue without clear consensus. Some key points to consider:

1. In many jurisdictions, copyright law traditionally requires human authorship. This has led some legal experts to argue that AI-generated works may not be eligible for copyright protection at all.
2. For AI-assisted works where there is significant human involvement in guiding the creative process, the human creator may be considered the copyright holder.
3. Some argue that the companies or individuals who created and trained the AI models could potentially claim copyright, though this is debated.
4. A few countries have started to address AI-generated works in their copyright laws, but there's no international standard yet.
5. In the US, the Copyright Office has stated that it will not register works produced by an AI system without substantial human involvement.
6. There are ongoing discussions and legal cases that may help clarify this issue in the future.

Given the complexity and ongoing developments in this area, it's best to consult with a legal expert for specific situations. If you'd like more details on any aspect of this issue, I'd be happy to elaborate.



potentially altering the current landscape.

6. Practical Considerations:

- **Documentation of Input:** To strengthen claims of authorship, users should maintain records of their inputs, prompts, and any creative decisions made during the generation process.
- **Consultation with Legal Professionals:** Given the intricacies and regional differences in intellectual property law, it is advisable to consult with legal experts when dealing with AI-generated content, especially for commercial or high-stakes applications.

Conclusion:

Currently, the ownership of copyright for AI-generated works predominantly rests with the human user who provides creative input and directs the AI's output, or with the entity that owns the AI system, depending on the jurisdiction and specific circumstances. In instances where AI-generated content lacks sufficient human authorship, such works may not receive copyright protection and could fall into the public domain. However, this is a rapidly evolving legal area, and ongoing legislative developments may further clarify or redefine these principles. Users and creators should stay informed about the latest legal standards in their respective regions and seek professional legal advice when necessary.



Copyright and AI (Definition of authorship/ownership)

ChatGPT (previous answers):

“I do not have the ability to own intellectual property or any other legal rights”

“Any text or other content that I generate is the property of OpenAI, as the creator and owner of the tool that I am.”

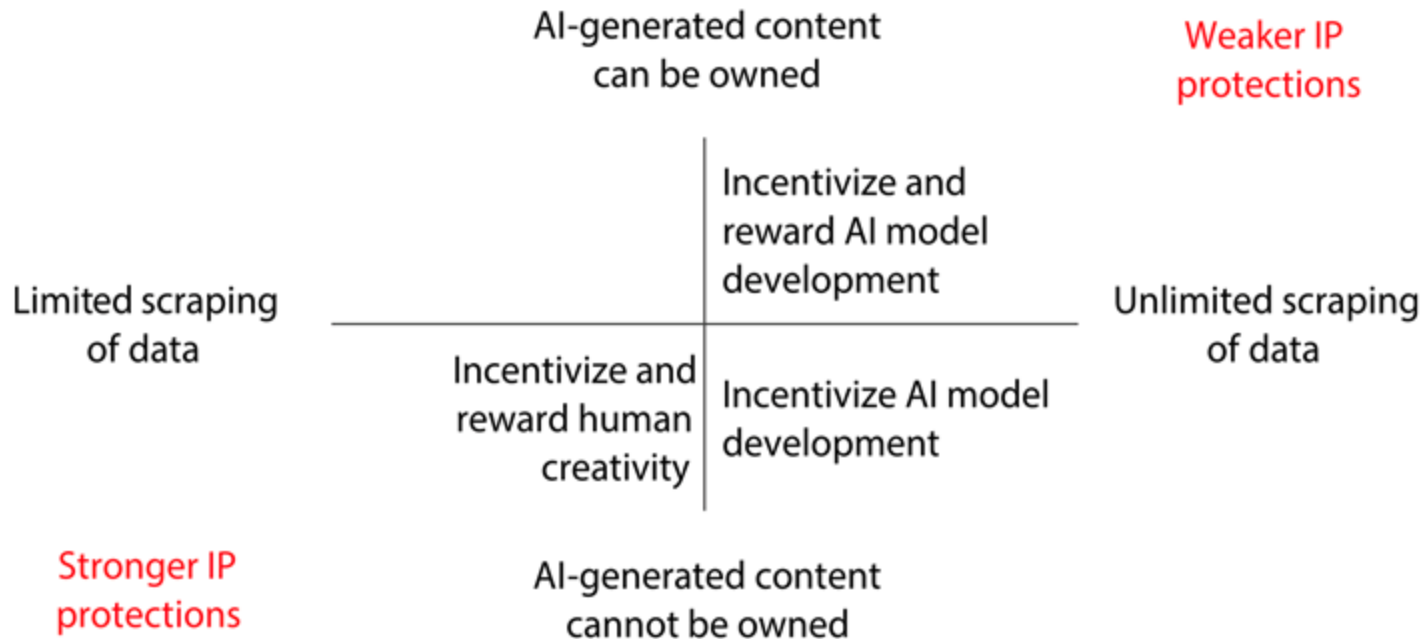
Revised answer: *“The text generated is not the intellectual property of the model itself. Instead, the intellectual property rights belong to the person or entity who has commissioned the model to generate the text.”*



Copyright and AI (Definition of authorship/ownership)

- Another challenge pertains to the development of bespoke AI models tailored to specific clients, such as an algorithm created for the purposes of particular scientific or implementation-based research;
- From the perspective of academic work, it is important to remember that copyright law does not protect scientific ideas;
- Since it is not possible to designate an author for an AI-generated creation, there is no subject for copyright protection. Consequently, such a creation becomes part of the public domain, and its utilisation is not subject to any restrictions.





Source: Chesterman, 2024.



UMCS
CENTRUM SZTUCZNEJ INTELIGENCJI
I MODELOWANIA KOMPUTEROWEGO

Infringement of copyright by creators of AI, users

- AI models trained on copyright protected resources, protected databases and industrial property;



World ▾ US Election Business ▾ Markets ▾ Sustainability ▾ Legal ▾ Breakingviews ▾ Technology ▾

Litigation | Copyright | Intellectual Property | Data Privacy | Litigation

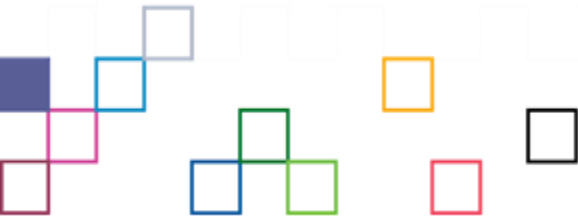
Getty Images lawsuit says Stability AI misused photos to train AI

By Blake Brittain

February 6, 2023 6:32 PM GMT+1 · Updated 2 years ago

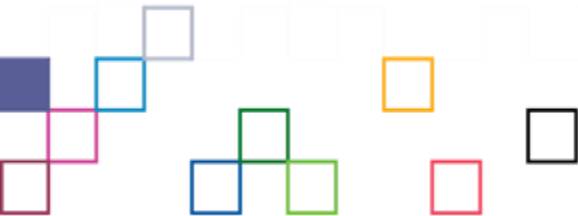


REUTERS/Wolfgang Rattay Purchase Licensing Rights ↗



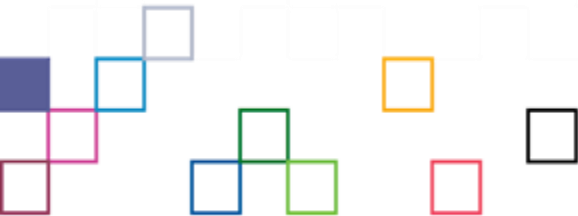
Infringement of copyright by creators of AI, users

- The reliability and verifiability of the data/sources on which the AI model was trained are crucial. Querying the AI model about this will not yield satisfactory answers, as the model only has access to metadata.
- It is essential to review the terms and conditions or terms of use of these tools. Analyses indicate that, in most cases, liability for the outcomes of their use is transferred to the user.
- Responsibility will always lie with those who publish content containing false scientific conclusions, statements infringing on others' rights, etc.
- A potential solution may involve the utilisation of technologies designed to flag toxic or false content.



Recommendations

- 1) Review the terms and conditions or terms of use of the AI tools before you decide to use them.
- 2) Reproducibility checklists and protocols e.g. Machine Learning Reproducibility Checklist, Checklist for AI in Medical Imaging (CLAIM), REFORMS checklist.
- 3) Follow the community standards for documentation (domain-specific community standards).
- 4) Follow the laws of the relevant jurisdiction that govern the AI and the creation of the work.



Thank you

