



A study on IPR in the era of AI and big data

Dr Anu Antony

Associate Professor, Department of Commerce, Krsitu Jyoti College of Management and Technology
Autonomous, Changanassery, Keralam, India

DOI: <https://doi.org/10.66856/ijcmr.2026.12.4.12297>

Abstract

The rapid development of Artificial Intelligence (AI) and Big Data has transformed the creation, use, and dissemination of knowledge and information, creating significant opportunities as well as challenges for Intellectual Property Rights (IPR). Traditional IPR frameworks, which were primarily designed to protect human-created works and inventions, are increasingly being tested by AI-generated content, machine learning systems, large-scale data processing, and automated innovation. This study examines the relationship between IPR and emerging AI and Big Data technologies, with particular focus on copyright, patents, trademarks, trade secrets, and data protection. It explores key legal and ethical questions concerning ownership of AI-generated works, inventorship of AI-assisted inventions, the use of copyrighted materials for training AI models, and the protection and commercialisation of data. The study also analyses the difficulties faced by existing IPR regimes in balancing innovation, public access, privacy, and the rights of creators and technology developers. By examining emerging legal approaches and policy developments, the study aims to identify potential reforms that can make IPR frameworks more responsive to technological advancement while ensuring fair protection for human creativity and innovation. Ultimately, the study highlights the need for a balanced, flexible, and technology-neutral IPR framework capable of addressing the complexities of the AI and Big Data era.

Keywords: Intellectual property rights (IPR), artificial intelligence (AI), big data, copyright, patent law, ai-generated works, data protection, machine learning, digital innovation, technology and law

Introduction

Intellectual Property Rights (IPR) are designed to protect human creativity, innovation, inventions, artistic works, designs, brands and commercially valuable information.

The rapid development of Artificial Intelligence (AI), Generative AI and Big Data, however, is challenging many traditional assumptions underlying intellectual property law. Modern AI systems depend heavily on large datasets for training. These datasets may contain copyrighted books, photographs, music, software, articles, designs and other protected materials. At the same time, AI systems can generate new text, images, music, software and inventions, raising questions about authorship, ownership, infringement and liability.

The World Intellectual Property Organization (WIPO) identifies several major areas of interaction between AI and IP, including authorship, ownership, remuneration for training data, AI-assisted innovation and protection of AI-generated outputs

Objectives of the Study

The major objectives of this research are

1. To examine the relationship between IPR, AI and Big Data.
2. To analyse the challenges AI creates for copyright, patents, trademarks and trade secrets.
3. To examine the legal status of AI-generated works and inventions.
4. To analyse the copyright implications of AI training data.
5. To study the challenges associated with ownership and liability.
6. To identify possible legal and policy approaches for balancing innovation with protection of creators.

Research Methodology

This research adopts a doctrinal and analytical methodology.

The study is based primarily on

- International IP principles and policy documents.
- Reports and publications of WIPO and OECD.
- Academic literature on AI and intellectual property.
- Emerging judicial and regulatory developments.
- Comparative analysis of approaches to AI-generated content and training data.

The research is primarily qualitative and examines existing legal principles in light of technological developments.

Understanding AI and Big Data

Artificial Intelligence

AI refers to technologies capable of performing tasks that traditionally require human intelligence, such as

- Learning
- Reasoning
- Pattern recognition
- Prediction
- Language processing
- Image generation
- Decision-making

Big Data

Big Data refers to extremely large and complex datasets that can be collected, processed and analysed using advanced computational technologies.

AI and Big Data are closely connected because data provides the foundation on which many AI systems learn and operate.

WIPO notes that training data is essential to AI systems and that large amounts of publicly available internet data may include copyright-protected text, images, designs and music.

Why AI Creates a New IPR Challenge

Traditional IP law generally assumes a human creator or inventor.

AI changes this assumption.

For example

Human → AI system → Output

The central legal question becomes

Who owns the intellectual property in the output?

Possible candidates include:

- The person who created the prompt
- The person who operated the AI
- The AI developer
- The owner of the training data
- The organisation using the AI
- Nobody, if the law requires human authorship

This creates uncertainty about ownership and protection.

Copyright and AI

Copyright is one of the areas most significantly affected by AI.

AI can

- Generate articles
- Create paintings
- Compose music
- Produce videos
- Write computer code
- Modify existing creative works

This creates two major questions

Question 1: Can AI-generated content receive copyright protection?

Many copyright systems are traditionally based on human authorship.

Therefore, purely machine-generated content creates uncertainty regarding:

- Authorship
- Originality
- Ownership
- Duration of protection
- Enforcement

WIPO describes this as the “output problem”—whether AI-generated content should be eligible for copyright protection.

Question 2: Who owns AI-assisted content?

A different situation occurs when a human uses AI as a creative tool.

For example

Human creativity + AI assistance = AI-assisted work

Here, the degree of human creative contribution may become important in determining whether copyright protection is available.

AI Training Data and Copyright

One of the most controversial issues is the use of copyrighted material to train AI models.

AI developers may collect enormous quantities of information from the internet.

Training datasets can contain

- Books
- News articles
- Photographs
- Artwork
- Music
- Films
- Software code
- Academic publications

The legal question is

Can copyrighted material be copied and processed for AI training without the copyright owner's permission?

There is no single universally accepted answer.

WIPO notes that copyright owners and AI developers have competing arguments concerning the use of protected works in training datasets, while litigation and policy discussions continue in multiple jurisdictions.

The OECD similarly identifies data scraping for AI training as an important emerging IP issue and examines different legal and policy responses.

Big Data and Intellectual Property

Big Data creates another fundamental question:

Can data itself be protected as intellectual property?

Individual facts or raw data may not always qualify for copyright protection.

However, legal protection may arise from

- Copyright in databases
- Database rights in certain jurisdictions
- Trade secrets
- Contractual restrictions
- Confidentiality obligations
- Protection of associated software and technology

The economic value of data has increased substantially because data can be used to train AI systems and generate commercially valuable insights.

Therefore, ownership and lawful access to datasets have become strategically important.

Patents and AI

AI also challenges traditional patent law.

Traditionally, a patent is granted for a qualifying invention attributable to an inventor who satisfies the applicable legal requirements.

AI can now assist humans in

- Discovering new compounds
- Designing products
- Optimising engineering systems
- Identifying technical solutions
- Generating potential inventions

This creates an important question

Can an AI system itself be recognised as an inventor?

WIPO's discussions on AI inventions highlight questions concerning the role of AI in the inventive process and the standards required for patent protection.

A practical approach increasingly focuses on identifying the human contribution to the inventive process rather than simply treating AI as an independent inventor.

Trade Secrets and AI

Trade secrets are particularly important in the AI industry.

Companies may protect

- Algorithms
- Model architectures
- Training methodologies
- Proprietary datasets
- Source code
- Business strategies
- Confidential prompts and processes

Trade secret protection can be attractive where disclosure of an invention through patent registration would reveal commercially sensitive information.

However, the use of third-party AI tools creates a significant risk:

Employees may unintentionally disclose confidential company information while interacting with an AI system.

Therefore, organisations need clear internal policies governing what information may be entered into AI systems.

WIPO identifies differences in contractual treatment of trade secrets, confidential information and ownership of outputs as important risks when organisations adopt generative AI tools.

Trademarks and AI

AI also affects trademark law.

Generative AI can create

- Logos
- Brand names
- Advertising materials
- Product images
- Synthetic brand personalities

Potential problems include

- AI-generated logos resembling existing trademarks
- Unauthorised use of famous brands
- Consumer confusion
- AI-generated advertisements containing protected marks
- Deepfakes involving brands or public figures

Businesses therefore need to conduct appropriate trademark and content checks before commercially using AI-generated material.

AI and Infringement Liability

Another important question is:

Who is responsible when AI produces infringing content?

Possible parties include

- AI developer
- AI platform provider
- AI user
- Content distributor

For example, an AI image generator might produce an image substantially similar to a protected artwork.

Determining liability may depend on:

- How the model was trained
- The user's instructions
- The degree of similarity
- The platform's safeguards
- Applicable copyright exceptions
- National law

This demonstrates why AI creates not only an ownership problem, but also a liability problem.

Deepfakes and Personality Rights

Generative AI can create realistic:

- Faces
- Voices
- Videos
- Photographs
- Digital performances

This creates risks involving

- Personality rights
- Publicity rights
- Privacy
- Copyright
- Trademark rights
- Defamation

A person's voice or likeness can potentially be reproduced without consent.

Therefore, the regulation of synthetic media requires coordination between traditional IP law and other areas of law.

WIPO has identified voice, image and deepfake issues as part of the broader IP questions created by generative AI.

Major Legal Challenges

The major challenges can be summarised as follows:

1. Authorship

Who is the author of AI-generated content?

2. Ownership

Who owns AI-generated works?

3. Training Data

Can copyrighted works be used to train AI?

4. Infringement

Who is liable when AI generates infringing material?

5. Inventorship

Can AI be recognised as an inventor?

6. Data Ownership

Who controls commercially valuable datasets?

7. Transparency

Should AI developers disclose what data was used for training?

8. Compensation

Should creators receive payment when their works are used to train AI?

9. International Jurisdiction

Which country's law applies when AI development, training data, users and outputs cross borders?

International Legal Landscape

There is currently no single global legal framework that completely resolves AI and IP questions.

Different jurisdictions are experimenting with different approaches.

The major areas of international debate include

- Copyright exceptions
- Text and data mining
- Licensing of training data
- Transparency requirements
- AI-generated works
- AI-assisted inventions
- Protection of creators
- Compensation mechanisms

WIPO's ongoing work demonstrates that AI and IP policy is still developing internationally, with particular attention to copyright infrastructure, transparency, consent, compensation and AI innovation.

Emerging Regulatory Approach

A balanced regulatory approach should attempt to achieve two objectives:

Protect creators

Creators should have mechanisms to:

- Control use of their works
- Receive appropriate compensation
- Identify unauthorised uses
- Challenge infringement

Promote innovation

AI developers should be able to:

- Access legitimate datasets
- Conduct research
- Develop innovative products
- Experiment within reasonable legal boundaries

Therefore, regulation should avoid both extremes

Over-regulation → reduced innovation

Under-regulation → exploitation of creators

Proposed Solutions

The following measures could help address emerging AI-IP challenges:

1. Greater Transparency

AI developers should provide appropriate information about training-data practices.

2. Licensing Mechanisms

Licensing systems could allow creators to authorise the use of their works for AI training.

3. Opt-Out Mechanisms

Creators could be given practical mechanisms to indicate that their works should not be used for certain AI training purposes.

4. Human-Centred Copyright

Copyright law should clearly distinguish between purely machine-generated content and works containing meaningful human creative contribution.

5. AI-Assisted Patent Guidelines

Patent offices should develop clearer standards for inventions created with significant AI assistance.

6. Stronger Contractual Protection

Businesses should establish rules preventing employees from submitting confidential information to inappropriate AI tools.

7. International Cooperation

Because AI and datasets operate across borders, international cooperation is essential.

WIPO has emphasised the need for scalable copyright infrastructure capable of supporting transparency, consent and compensation across jurisdictions.

India: Emerging Importance

India faces a particularly important challenge because it has:

- A large digital economy
- A growing technology sectors
- A substantial creative industry
- Increasing use of generative AI
- Large amounts of multilingual data
- A rapidly expanding AI ecosystem

Indian IP law will need to address questions concerning

- AI-generated works
- Copyright ownership
- Training data
- Text and data mining
- AI-assisted inventions
- Digital piracy
- Deepfakes
- Personality rights

The Indian approach must balance innovation, access to technology and protection of creators.

Ethical Dimensions

The AI-IP debate is not purely legal.

It also involves ethical questions

- Is it fair to train AI on creators' works without compensation?
- Should AI-generated content be clearly labelled?
- Should creators have the right to refuse AI training?
- How should society distinguish human creativity from machine-generated content?
- How can innovation occur without undermining creative professions?

Therefore, future IP policy should consider both economic incentives and social justice.

Findings of the Study

The research identifies the following key findings:

1. AI is challenging traditional human-centred concepts of authorship and inventor ship.
2. Big Data has become a valuable economic resource for AI development.
3. The use of copyrighted material in AI training remains one of the most contested issues.
4. AI-generated outputs create uncertainty regarding copyright ownership.
5. AI-assisted inventions create challenges for traditional patent concepts.
6. Trade secrets are increasingly important for protecting AI technology and proprietary datasets.

7. Deepfakes create new intersections between IP, privacy and personality rights.
8. Existing IP laws can address some AI-related issues, but significant uncertainties remain.
9. International cooperation is increasingly necessary because AI development and data flows are global.
10. The future legal framework should balance innovation with fair protection and remuneration for creators.
15. World Intellectual Property Organization (WIPO). Berne Convention for the Protection of Literary and Artistic Works. Paris Act, as amended, 1971.
16. World Trade Organization (WTO). Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement). Marrakesh, 1994.
17. World Intellectual Property Organization (WIPO). WIPO Copyright Treaty (WCT). Geneva, 1996.

Reference

1. World Intellectual Property Organization (WIPO). Artificial Intelligence and Intellectual Property. Available at: <https://www.wipo.int/en/web/frontier-technologies/artificial-intelligence>
2. World Intellectual Property Organization (WIPO). Frequently Asked Questions: AI and IP Policy. Available at: <https://www.wipo.int/en/web/frontier-technologies/artificial-intelligence/faq>
3. Organisation for Economic Co-operation and Development (OECD). Intellectual Property Issues in Artificial Intelligence Trained on Scraped Data. OECD Artificial Intelligence Papers, No. 33. Paris: OECD Publishing, 2025. <https://doi.org/10.1787/d5241a23-en>
4. Organisation for Economic Co-operation and Development (OECD). OECD Digital Economy Outlook 2024, Volume 1. Paris: OECD Publishing, 2024. Available at: https://www.oecd.org/en/publications/oecd-digital-economy-outlook-2024-volume-1_a1689dc5-en.html
5. United States Copyright Office. Copyright and Artificial Intelligence, Part 1: Digital Replicas. Available at: <https://www.copyright.gov/ai/>
6. United States Copyright Office. Copyright and Artificial Intelligence, Part 2: Copyrightability. Washington, D.C.: U.S. Copyright Office, 2025. Available at: <https://www.copyright.gov/ai/>
7. Government of India. The Copyright Act, 1957. Available at: <https://www.indiacode.nic.in/indiacode/handle/123456789/1367>
8. Government of India. The Patents Act, 1970. India Code. Available at: <https://www.indiacode.nic.in/>
9. Government of India. The Trade Marks Act, 1999. India Code. Available at: <https://www.indiacode.nic.in/>
10. Government of India. The Information Technology Act, 2000. India Code. Available at: <https://www.indiacode.nic.in/>
11. Government of India. The Digital Personal Data Protection Act, 2023. Available through the official Gazette of India and India Code: <https://www.indiacode.nic.in/>
12. Thaler v. Vidal, 43 F.4th 1207 (Fed. Cir. 2022). United States Court of Appeals for the Federal Circuit. Available at: https://www.ca9.uscourts.gov/08-05-2022-21-2347-thaler-v-vidal-opinion-21-2347-opinion-8-5-2022_1988142/
13. Thaler v. Perlmutter, No. 22-1564 (D.D.C. Aug. 18, 2023). United States District Court for the District of Columbia.
14. World Intellectual Property Organization (WIPO). Convention Establishing the World Intellectual Property Organization. Stockholm, 1967.