

ISSN :2582-6433



# INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary  
Peer Reviewed 6th Edition

VOLUME 2 ISSUE 6

[www.ijlra.com](http://www.ijlra.com)

## **DISCLAIMER**

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Managing Editor of IJLRA. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of IJLRA.

Though every effort has been made to ensure that the information in Volume 2 Issue 6 is accurate and appropriately cited/referenced, neither the Editorial Board nor IJLRA shall be held liable or responsible in any manner whatsoever for any consequences for any action taken by anyone on the basis of information in the Journal.

Copyright © International Journal for Legal Research & Analysis

# EDITORIAL TEAM

## EDITORS

### **Megha Middha**



*Megha Middha, Assistant Professor of Law in Mody University of Science and Technology, Lakshmangarh, Sikar*

*Megha Middha, is working as an Assistant Professor of Law in Mody University of Science and Technology, Lakshmangarh, Sikar (Rajasthan). She has an experience in the teaching of almost 3 years. She has completed her graduation in BBA LL.B (H) from Amity University, Rajasthan (Gold Medalist) and did her post-graduation (LL.M in Business Laws) from NLSIU, Bengaluru. Currently, she is enrolled in a Ph.D. course in the Department of Law at Mohanlal Sukhadia University, Udaipur (Rajasthan). She wishes to excel in academics and research and contribute as much as she can to society. Through her interactions with the students, she tries to inculcate a sense of deep thinking power in her students and enlighten and guide them to the fact how they can bring a change to the society*

### **Dr. Samrat Datta**

*Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board*



## **Dr. Namita Jain**



**Head & Associate Professor**

*School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC -NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.*

*Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India. India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time & Utkrisht Bharat Foundation, New Delhi.(2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020). Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019*

## **Mrs.S.Kalpana**

**Assistant professor of Law**

*Mrs.S.Kalpana, presently Assistant professor of Law, VelTech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Avadi. Formerly Assistant professor of Law, Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr. Ambedkar Law College, Pudupakkam. Published one book. Published 8 Articles in various reputed Law Journals. Conducted 1 Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration. 10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.*



## **Avinash Kumar**



*Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship. He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi. Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi. He has qualified UGC – NET examination and has been awarded ICSSR – Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and learning.*

## **ABOUT US**

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN

2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench.

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN

2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

# **BIG DATA PROTECTION IN IP**

**Authored By- Apsi Adithya Kumar  
LLM (IP) PHD, IUCIPRS**

## **Brief**

The interfaces between "Big Data" and IP matters both because of the impact of Intellectual Property (IP) rights in Big Data. This Article looks at both sides of the coin, focusing on several IP rights, namely copyright, patent, exclusivity and trade secret/confidential formation.

The term "Big Data" can be defined in a number of ways, but the four most important are volume, veracity, velocity and variety. Big Data corpora are often generated automatically, and the question of the quality or trustworthiness of the data is crucial. If all previous features are present, a Big Data corpus likely has significant "value".

In many cases, Big Data corpora are protected by protection that relies on trade secret law combined with technological protection from hacking, and contracts. A publicly available corpus, in contrast, must rely on erga omnes IP protection—if it deserves protection to begin with. Copyright protects collections of data; the sui generis database right (in the EU) might apply; and data exclusivity rights in clinical trial data may be relevant.

The human-written (AI) software used to collect (including search and social media apps), store and analyse Big Data corpora is considered a literary work eligible for copyright protection, subject to possible exclusions and limitations.<sup>18</sup> The analysis that follows focuses on the harder question of the protection of the *Big Data corpora* and of the *outputs* generated from the processing of such corpora

A Committee of Experts meeting under the auspices of the World Intellectual Property Organization (WIPO), which administers the Berne Convention, concluded that, the only mandatory requirement for a literary or artistic work to be protected by the Convention is that it must be "original". In its Article 2, when discussing the protection of "collections", it states that "collections of literary or artistic works such as encyclopedias and anthologies shall be protected *as such*, without prejudice to the copyright in each of the works forming part of such collections. There are two layers of copyright in an encyclopedia: the "organizational layer", granted to the maker of the collection based on the "selection or arrangement" of the individual entries,

photographs and illustrations; and the "collective layer", which is generally treated as a collective work. In a collection of this type, there are first, a right in each entry, and each illustration or photograph, which is either transferred or licensed by the maker or publisher to the person making or distributing it.

Big Data is sometimes defined in *direct contrast* to the notion of SQL database and reflected in the TRIPS Agreement. Big Data software is unlikely to "select or arrange" the data in a way that would meet the originality criterion and trigger copyright protection

Data generated by AI-based TDM systems that have initially high but fast declining value, such as financial information relevant to stock market transactions, could be subject to copyright protection in some jurisdictions. In the US, the tort of misappropriation is applicable to "hot news" in US law, or the protection against parasitic behaviour available in a number of European systems.

The use of noSQL technologies may mean that Big Data corpora are not protected by the sui generis right in the Database Directive. The Directive refers to the database maker's investment in "obtaining, verification or presentation of the contents" and then provides a right "to prevent extraction and/or re-utilization" of that data. The Court of Justice of the European Union defined "investment" in obtaining the data as "resources used to seek out existing materials and collect them in the database" - but does not cover the resources used for the creation of materials which make up a database. The main argument for this distinction is that the Database Directive's economic rationale is to promote and reward investment in database production, not in generating new data.

An AI-capable TDM system might be used in enhancing the use of patent information. The "patent bargain" is basically a fair disclosure of an invention in exchange for a limited monopoly on its use, especially on a commercial basis. AI applications in this field already go further, and the trajectory of their development leads to some potentially remarkable conclusions.

There is a concern that TDM tools might prevent the use of clinical trial data, which is seen as a negative development. This is because it is the collected clinical trials, and their ability to provide a large and comprehensive dataset, that make them valuable. It is not the specific health and safety outcomes proven by those data that are so valuable.

Patents may become more difficult to obtain due to massive Big Data –based AI disclosures of possibly new incremental innovations. Such a system could conceivably disclose new molecules

and predict their efficacy. In such a case, it would be near impossible to patent the drug unless patented by the AI "inventor". The data exclusivity right might fill that void.

Application of trade secret law to Big Data. Trade secret and confidential information law could be used to protect data acquired for purposes of TDM. Trade secret law typically works far better for business information than private data. One might expect the default contracts may not adequately protect users or consumers, though privacy or consumer protection laws may impose limits on contractual freedoms. The protection of confidential information could apply to "data coming from a machine-to-machine process", as well as the use of such data by companies in the so-called "collaborative economy 3.0" - where they share their Big Data with each other. Possibilities of welfare gains by third parties, since this regime applying to knowledge commons such as the IoT enables spillovers, and therefore its presence may not necessarily be perceived as a bad thing.

Excessive restrictions on access to lock-in effects by major data gathering entities might have negative welfare impacts warranting governmental intervention in "data--driven platform markets characterized by strong network and lock--in effects--and in new technological contexts that might otherwise be ripe for competitive innovation

In sum, the interfaces between Big Data and IP are about finding ways to adapt IP rights to allow and set proper parameters for the generation, processing and use of Big Data. This includes an analysis of how Big Data may infringe IP rights. There is also an issue of rights *in* Big Data, however. Courts and legislators have years of questions to answer on both constraints in and protection of Big Data.

## Introduction

Big data is currently a hot topic in many fields, including management and marketing, scientific research, national security, government transparency, and open data. Both the public and private sectors are increasingly utilising big data analytics. This study aims to provide an overview of the issues as we see them and to contribute to the big data discussion.

In this subject, technological capabilities and the range of possible applications are quickly developing, and there is ongoing debate regarding the consequences of big data. Our goal is to balance the many privacy hazards connected with big data with the benefits that big data provides to organisations, individuals, and society as a whole. We believe that adhering to essential data protection rules and measures will aid in the long-term sustainability of big data's developing benefits. The benefits cannot be simply traded for the right to privacy.

This poses a tension in terms of intellectual property because there are continuing attempts to better protect authors' rights in the digital age; efforts that could be perceived as incompatible with the needs of big data.

Put simply, big data refers to the unlimited use of data, whereas traditional IP protection aims to prevent this. An IP attorney can help you solve the complexities of this situation, but at first glance, it seems to be going in the opposite direction.

In order to proceed with a big data project without violating the Copyright Act, the project manager should theoretically contact the individual authors represented in the dataset and obtain permission.

The article references several examples of big data and cites reports and other publications. Information is taken from publicly available sources and links are provided in the footnotes.

## IP Protection and Big Data

The concept of secrecy, a form of trade secret law-based security coupled with technical hacking prevention and contracting, involves an enormous amount of information. Therefore, when figuring out which IP rights can be followed, it is important to distinguish between public and large amounts of undisclosed information (such as the Google database that powers search engines and advertisements). Secret Legions are often virtually protected from competition by secrecy. In other words, your competitor may create an aggressive legion to gain market share. Public data must rely on the security of intellectual property for evaluation.

The proposed EU General Data Protection Regulation<sup>1</sup> contains a number of provisions that would have a bearing on the use of personal data in big data analytic.

Data minimization and data anonymization – burden of proof on data controllers. The need for transparency; A shift in the balance of data protection forces designed by default; Possibility of extending data protection responsibilities to organizations outside the EU.

Personal data must be "restricted to the minimum necessary with respect to the purpose of processing" and must be processed "only in cases and during periods when the purpose cannot be achieved" by processing information that does not contain personal data" (Art. 5 (c) EU GDPR)<sup>2</sup>

Furthermore (Article 17 EU GDPR)<sup>3</sup> "right to be forgotten" means that the data subject may request the deletion of personal data if it is no longer necessary for the purpose for which it was collected or processed. A recent decision of the European Court of Justice under the current directive (which is being implemented by the DPA) also supports this direction<sup>4</sup>. Under the proposed regulation, data controllers not only require a "transparent and easily accessible policy" for the processing of personal data, but also communicate with data subjects "in an understandable manner using clear and simple language appropriate to the data subject" ask you to do .

---

<sup>1</sup>(<http://curia.europa.eu/juris/document/document.jsf?text=&docid=152065&pageI ndex=0&doclang=en&mode=lst&dir=&occ=first&=1&cid=243691>)

<sup>2</sup>[EU GDPR 1998, Art. 5 (c)]

<sup>3</sup>[EU GDPR, Art. 17]

<sup>4</sup>*Google Spain SL and Google Inc v Agencia Española de Protección de Datos and González* C-131/12

Data controllers would have to put in place methods to ensure that only the bare minimum of personal data is utilised and that it is stored for no longer than is necessary for the processing. Big data is frequently described as a power dynamic that benefits corporations and governments. The Regulation implies a desire to change the power balance in favour of the individual by giving them more explicit rights over their personal data processing<sup>5</sup>. While the Commissioner supports the Regulation's protections in general, it is critical to ensure that the provisions are effective in reality, which requires more thought about what this degree of prescription would accomplish. The Regulation clearly aims to address some of the most pressing data protection concerns raised by big data analytics, but whether it will be implemented in its current form remains to be seen.

## **Challenges To Patent System Posed By Big Data**

Although big data is not patentable in and of itself, the algorithm and software programme may be covered by the law. Furthermore, while big data content cannot be patented in general, it may be protected as a patent if it may provide an economic advantage by articulating it as an innovation that is inherent unique and can be utilised for industrial use by the company seeking to assert its rights<sup>6</sup>. Consider the scenario of a computer that was said to be used to assess the qualifications of candidates for a vocational training programme<sup>7</sup>. The Court distinguished two types of computer use: the first is using a computer to carry out a scheme or plan in which the computer only acts as an intermediary, and the second is using a computer to improve its functionality or solve a technical problem that is outside of the computer's normal use. "Putting a business process or strategy into a computer is not patentable unless the computer performs the scheme or method in an inventive manner," the Court stated. However, finding patents for these computer-generated inventions can be a tricky equation, as tasks generated by uncontrolled artificial intelligence are not patentable. This can challenge traditional notions of intellectual property (IPR).

---

<sup>5</sup>Neil M. Ricahards and Johnathan H. King, "Three Paradoxes Of Big Data" Vol. 66, September 2013.

<sup>6</sup>Kashishipr, "The Interplay of Big Data and Intellectual Property Rights" [BlogCopyrightIntellectual Property RightsPatent](https://www.kashishipr.com/blog/the-interplay-of-big-data-and-intellectual-property-rights/), 23 June 2021, available at <https://www.kashishipr.com/blog/the-interplay-of-big-data-and-intellectual-property-rights/> (accessed on 29/04/2022)

<sup>7</sup>*Commissioner of Patents v RPL Central Pty Ltd* [2015] FCAFC 177

## Patent challenges

### Document Authentication and Management

The concerns examined under this group are record falsification, the arrangement of clever agreements, and the treatment of agreements after everything is said and done.

The following firms have successfully handled this issue:

- IBM — Handling many types of agreement layouts stored in a blockchain, where the type of format open is regulated by the event type and event records entered.
- Coin-plug – Checking the validity of a bank's exchange records by comparing the initial and subsequent records issued in accordance with a client.
- Bank of America – Using a private blockchain framework to improve and streamline the acceptance of reports moving between two different stockpiling devices.
- Alibaba – Validation and verification of records between clients by at least one user checking the record and then transferring it to a central server, which distributes the validated report to all other blockchain clients.

It is associated with patent families that are attempting to confirm data exchange through the system.

### Data Sharing and Consistency

The following firms have successfully solved this issue: -

- IBM – Method of securing a supplier's media material by storing it on a server and only transferring it to a media player application after verification.
- Coinplug — Registering only the Merkle tree model's root estimates that speak to the entire blockchain rather than the complete blockchain to a local PC.
- Bank of America - Using extremely complex hashes for each mutual information record to aid the framework in distinguishing between offer and other information records.

- Alibaba – Choosing an accord hub for information sharing and consistency using a democratic framework to streamline the handling steps required to check the blockchain

### Security and Secrecy

It identifies with patent families that are aiming to address challenges of information and motion protection and encryption.

- IBM – A method of securing a provider's media material by storing it on a server and only delivering it to a media player application after authentication.
- Coin-plug — A verification framework based on a blockchain-based electronic wallet.
- Bank of America – A framework that keeps track of asset accessibility and converts non-secure instruments into secure instruments that require customer and mark approval before access is granted.
- Alibaba – Using a blockchain stage to confirm exchange requests before transferring assets to the client.

### Transactions in General

It includes arrangements for information exchanges and trade, as well as tracking and determining the types of transfers.

- IBM – blockchain to identify parties with outstanding transactions.
- Coin-plug — Using blockchain to track transactions between parties without requiring the perception of an open location or the use of QR codes.
- Bank of America – Using a blockchain framework to allow customers to migrate data from one bank to the next without the need for an aggregator.
- Alibaba – A clever blockchain agreement framework that determines the optimum agreement arrangement to use based on the business exchanges required.

## **Text Mining And Copyright**

In text and data mining, enormous amounts of copyrighted material are frequently copied. To 'mine' books and other content, researchers must utilise computer programmes to access, copy, and process them. Even if researchers have legal access to and can read the content, such as through their university library, copying a significant percentage of those works may be illegal.

Copyright, on the other hand, was never meant to restrict the use of a work's ideas, facts, or information. In a recent case involving internet browsing, the UK Supreme Court reaffirmed this principle, saying, "Broadly speaking, producing or distributing copies or adaptations of a protected work is an infringement." Simply looking at or reading it is not an infringement<sup>8</sup>. Text and data mining might be considered a technology that merely substitutes for human sight and reading. As a result, copying in the context of a text mining process could be seen as a byproduct of the technology's operation rather than an activity aimed at exploitation of copyright-protected content. In this regard, copyright owners (publishers) have traditionally been ready to allow academics to 'mine' works in their catalogues, particularly if the research might result in mutually advantageous outcomes, such as the development of software tools that increase the value of their catalogues. Instead of being competitors, readers and researchers are partners of copyright owners.

## **Exception For Text And Data Analysis**

Copyright laws in the United Kingdom allow academics to make copies of works "for text and data analysis."

This means that if a person has legal access to a work, they can create a duplicate of it in order to do a computational analysis of the information contained within it.

The exception is subject to the following conditions:

- 1) The conceptions and lies must be for the goal of non-commercial study.
- 2) The copy is accompanied by an appropriate acknowledgement (unless this is practically impossible)

---

<sup>8</sup>(*The Newspaper Licensing Agency Ltd v Public Relations Consultants Association Ltd*, [2013] UKSC 18)

Copyright is also violated if a copy is transferred to a third party or used for a purpose other than those permitted under exceptions, according to the requirements (although the researcher could ask the owner for the permission to do either of these things).

Furthermore, text and data analysis copies are not for sale or hire.

Contracting out of the activities covered by the exception is not an option, according to the regulation.

Contractual terms that purport to limit or prevent the performance of the exception-authorized acts are unenforceable. The exclusion applies to all sorts of copyright works, as well as recordings of performances, even though text and data analysis is primarily focused with mining literary works. Policymakers considering enacting an explicit TDM exceptions or limitations should consider the following questions:

Whether the exception applies to only one type of right (reproduction) or all types of rights (adaptation/derivation);

Is it possible to have contractual overrides?

Whether the content should come from a legal source?

What kind of data dissemination, if any, is possible?

Whether TDM is used for a non-commercial purpose.?

In response to the main question, if allowing TDM is considered a normatively valid goal, the right holder should no longer be able to block it by using one right fragment from the bundle of copyright rights. Irini Stamatoudi concluded from an analysis of the rights involved that right fragments beyond reproduction and variation were far less applicable<sup>9</sup>. Nonetheless, it seems safer to phrase the exception or difficulty as a non-infringing use, as in the USA Copyright Act's Section 107 (fair use).<sup>10</sup>

Second, for the same reason, contraction overrides should be prohibited.

Unless there was only one TDM provider for a given type of job, it's difficult to see how they might

---

<sup>9</sup>Irini A. Stamatoudi, 'Text and Data Mining', in Irini A. Stamatoudi (ed.), *New Developments in EU and International Copyright Law* (Wolters Kluwer, 2016) 262

<sup>10</sup>The US Copyright Act (17 USC s 107)

be effective. Even if a clause barring contractual overrides is not mentioned in the text of the statute, contract law principles may render the restriction invalid.<sup>11</sup>

On the surface, the legal source factor in French law appears to be compelling. It appears to be difficult to argue against requiring the data's source to be legitimate. However, putting it into practise poses certain difficulties.

To begin with, a human user may not always be able to determine whether or not a source is legal; the issue may be even more unclear for a machine.

Second, determining the legality of a foreign source may entail an examination of the law of the country of origin, because copyright infringement is found using the *lex loci delicti*, which necessitates first determining the source's origin. Perhaps a requirement focusing on sources that the user is aware of or would be grossly negligent in ignoring *in not knowing* were illegal might be more appropriate.<sup>12</sup>

The final two tasks are a little more challenging. It may be important to communicate the information to others who are interested in the project if the data comprises copyrighted content. German legislation exempts a "restricted circle of people for cooperative scientific research," as well as "third parties for the purpose of checking the quality of scientific research." This represents the scientific exceptions, which include project-based work by a small group of scientists under the supervision of peer reviewers. This would make it impossible for TDM to scan libraries of books and make snippets available to the public, as Google Books does.<sup>13</sup>

---

<sup>11</sup>Lucie M.C.R. Guibault, *Copyright Limitations and Contracts: An Analysis of the Contractual Overridability of Limitations on Copyright* (Kluwer Law International, 2002) 224-225

<sup>12</sup>TRIPS Agreement (n 1), Art. 39(2)

<sup>13</sup>Copyright Act of 9 September 1965 (Federal Law Gazette I, p. 1273), as last amended by Article 1 of the Act of 1 September 2017 (Federal Law Gazette I p. 3346), art. 60d

## **Conclusion**

We understand that big data analytics can assist society in a variety of ways, including scientific and medical research. These advantages come on top of better products and services for consumers and business advantages. Nonetheless, these benefits should not come at the expense of an unjustifiable violation of privacy. Data protection principles should not be regarded as a stumbling block to progress, but rather as a foundation for promoting privacy rights and encouraging the development of innovative ways to inform and involve the public. Transparency regarding the aim and impact of analytics is not just required by law; it may also boost people's confidence as "digital citizens" in the age of big data.

To summarise, the interfaces between Big Data and IP are all about finding ways to adapt IP rights to allow and set proper parameters for the generation, processing, and use of Big Data. This includes a look at how Big Data might violate intellectual property rights. However, Big Data poses an issue in terms of rights. Courts and politicians have been questioned for years about the restrictions and protection of Big Data.