BIOINFORMATICS & INTELLECTUAL PROPERTY RIGHTS: PROTECTION METHODS AND STRATEGIES UNDER INDIAN PERSPECTIVE

Rishabh Shukla*

Abstract

Intellectual Property (IP) for Bioinformatics plays a key role in accelerating developments of biological sciences and biotechnology industry. This paper will focus on current global position of IP protection in Bioinformatics and biological databases with special focus on India. The methods of protection which have been proposed after analyzing the characteristics of Bioinformatics and the biological databases and considering different database protection methods will be critically examined. Further, this paper seeks to critically analyze the diffusion process of information proposed by Junli Chang and Xuezhong Zhu in the practical world and will try to answer the question of whether the primary database in Bioinformatics should be made a public asset or whether it should be protected by IP laws to avoid misuse. While protecting the databases, the focus on balancing the economy of the country will also be looked upon along with subtle methods in IP protection for secondary database.

History

The concept of Bioinformatics was not known before the discovery of human cell. In 1953, when the structure of double helix model of DNA was analyzed and proposed by Watson and Crick, which was based on the x-ray data obtained by Franklin and Wilkins, there were intuitions that in future, these biological datas will be of utmost importance and hence needs to be preserved.Thus came the concept of Bioinformatics in 1960s with efforts of Margaret.

* 3rd year, BBA.L.L.B, New Law College BVDU.

O. Dayhoff, Walter M. Fitch, Russell F. Doolittle and others wherein the use of Computers, Information Technology and Biology come together to form and develop into a fully developed single discipline to analyze and understand biological information. In other words, it is an amalgamation of life science and Information Technology, with the term 'Bioinformatics' coined by Paulien Hogeweg and Ben Hasper in 1978.[1]

Bioinformatics is interpreted and defined by various people with areas of expertise in Biology, Computer Science, and Information Technology differently, but the official definition of Bioinformatics was given by NCBI stating that Bioinformatics is "*the field of science in which biology, computer science and information technology merge into a single discipline. There are three important sub-disciplines within Bioinformatics: the development of new algorithms and statistics with which to assess relationships among members of large data sets; the analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains and protein structure; and the development and implementation of tools that enable efficient access and management of different types of information."*[2]

Bioinformatics focuses on cellular and molecular level of biology that is used in genetics and for drug discovery, and focuses in areas of genomics involving large scale DNA sequencing. The field of Bioinformatics depends not only on the analysis and understanding of biological data, but also preservation of this biological information for use in the years to come. This biological information is stored in special database called 'Biological Database'. The very first databases of Bioinformatics, or the biological databases, was created few years after the first protein sequence was discovered and reported in 1956 called as bovine insulin. Almost a decade later, when the first nucleic acid sequence was found and reported, Dayhoff gathered all the information available at that time on sequence data and created the very first biological database.[3]

The tools which came in aid to the research, analysis and understanding of data were created shortly after the creation of database, because even if

---

[1]*Scope of Bioinformatics in Biological Sciences*, Biology-Today.com, http://www.biology-today.com/Bioinformatics-biostatistics/scope-of-Bioinformatics-in-biological-sciences/

[2]Sabu M. Thampi, BIOINFORMATICS (2009), https://arxiv.org/ftp/arxiv/papers/0911/0911.4230.pdf.

[3]*Ibid.*

database was created, there was need to still get data from the biological material available in the world efficiently and at a faster rate. Due to this, the NCBI (National Centre for Biotechnology Information) was established in 1988 by the then USA President Ronald Reagan and Congressman Claude Pepper and has been in function ever since.[4]

Bioinformatics & IPR: Indian Perspective

The field of Bioinformatics in India was slow at the start but now it has evolved rapidly as Information Technology companies have also started focusing on the Life Science and Biotechnology sector. A lot of companies have gained major achievements in this area, for example, TCS, HCL, Cognizant Technologies, Mphasis, and Infosys. The companies engaged in Bioinformatics in India have a good opportunity to garner a large share of the market all over the world in providing services in this field such as mapping and DNA sequencing, data mining, proteomics and molecule design simulation, functional genomics and similar areas. The drastic increase in the number of participants focusing work to companies and the ever-increasing volume of genomics data have inspired many IT, pharmaceutical, and Biotechnology companies to enter in the field of Bioinformatics resulting in India's biotech sector being in the top five in Asia-Pacific region.

Some Indian IT companies like TCS (Tata Consultancy Services), Infosys and WIPRO have already engaged a part of their companies in the Bioinformatics sector. Pharmaceutical companies in India such as Biocon, Biological E, Nicholas Piramal and other such similar companies are making fast strides in the Bioinformatics area. Strand Genomics is one of the examples of emerging pure-play Bioinformatics companies that India has witnessed.[5]

Before 1955, the regime of patents in India could only protect the processes rather than the product itself. On January 1, 1955 when India became a member of TRIPs, an ordinance amending the Patent Act 1970 allowed applications in the fields of pharmaceuticals and agrochemicals and also awarded Exclusive Marketing Rights (EMR) to these products for a period of five years. The enactment of Patent (Amendment) Act 1999 took place in March 1999, which came into force with retrospective effect from January, 1995 with the ordinance's provisions. India immediately saw over a thousand of applications for patents on pharmaceutical products after it had

[4] *Ibid.*

[5] Pankaj M. Madhani, *Indian Bioinformatics: Growth Opportunities and Challenges*, 3 (2) PRERANA- JOURNAL OF MANAGEMNET THOUGHT AND PRATICE 7 (2011).

started receiving applications for patent of products. The applications for patents were opened when in April 2005 The Patents (Amendment) Act 2005 was enacted.[6]

Due to the large amount of patents which were filed and pending between 1995 and 2004 that became inconsiderable for patent by the Comptroller General after January 2005, a move later retracted, and all applications filed and pending before January2005 which contained product claims to be examined in accordance with the Patents Act 2005 as per the directions of the High Court of Kolkata[7] in the case of *Auguron Pharmaceuticals Inc v. Controller of Patents.*[8]

Under the new Patent regime in 2006, the first pharmaceutical company in India to gain product patent was the Swiss drug maker F-Hoffman Roche for the Pegasys, a new generation of Hepatitis therapy drug.[9] This kind of development has to be looked upon and cannot be neglected. Because sooner or later, the other field which involve information and patenting will emerge and the IP regime has to be capable enough to handle those areas effectively, but for this to happen, we must start working from now on.

The IP Laws and Bioinformatics

So why do we need IP Laws in the biological data when it is created by nature? Can we actually patent something which is not created by humans but only discovered? Can it be protected with IP Laws when it is not actually an invention? The answer to this question came in the case of *Diamond v. Diehr,*[10] where it was held that "*to qualify as patentable subject matter, the biological sequence has to be categorized as a process, machine or apparatus and idea of itself in not patentable and neither the principle in the abstract."*[11] This created new opportunities for people engaged in the field of Bioinformatics because methods for discovering the specimens, and the ways

---

[6] Amiya Kumar Patiand Ankit Patel, *Biotechnology and IPR: The Importance of Patent in Indian Perspective,*2 LawMantra 1(2015).

[7] *Ibid.* See also, Pankaj Musyuni & Swati Gupta, *Patent protection for combination/composition claims in India,* Lexorbis (July 18th, 2019), https://www.lexorbis.com/patent-protection-for-combination-composition-claims-in-india/.

[8] Auguron Pharmaceuticals Inc. v. Controller of Patents, MIPR 2009 (2) 345.

[9] *Id.*

[10] Diamond v. Diehr, 450 U.S. 175 (1981).

[11] *Supra* note 1.

of manufacturing the tools and the tools itself including products could not be patented because it now became clear as to what should be protected under IP law, but there were still issues regarding other aspects of Bioinformatics.

Biology and IP rights go hand in hand, and so it is obvious that Bioinformatics and IP Laws will have a strong relation with each other. Not just because Bioinformatics involves biology and technology, but also to protect the biological information which has been discovered by people engaged in the Bioinformatics field and avoid the misuse of this information. So for Bioinformatics to survive in the market, IP laws are the key factors because nearly all aspects of Bioinformatics are patentable, and now that all aspects have become patentable, there is also an opportunity to generate IP value from investing in researches and tools.

The scope of protecting Bioinformatics can be divided into two areas. The first one being Bioinformatics tools in the form of databases and compilation of raw data and the second being tools which help in retrieval and analysis of data. Though the IP protection scope in Bioinformatics is divided into two parts, there are a lot of issues related to Bioinformatics in the IP front, some of which are computer software and biological research, databases of genetics and genomes, algorithms and computational techniques, concept of interoperability between the existing databases, biological molecules, but most importantly the research data itself. So how can they all be protected?

The various constituents of Bioinformatics can be protected by various means:

1. Patents can be used to gain monopoly over the technologies that have been created for the aid to research in Bioinformatics;
2. Trademarks can be used to protect trade names, domain names, product names, and service marks and slogan by companies;
3. Copyrights can be used to protect Bioinformatics related materials such as books, articles, software codes etc.;
4. Trade Secrets are used to protect compilations, databases and facts, formulas, and processes.

---

[77]Vallari Dubey, *Legal Nature of Bitcoins: The Encrypted Digital Currency,* (April 29th, 2017),http://vinodkothari.com/blog/legal-nature-of-Bitcoins-the-encrypted-digital-currency-by-vallari-dubey/.

[78]Reserve Bank of India Act, 1934.

[79]*Foreign Exchange Management (Manner of Receipt and Payment) Regulations,* Reserve Bank of India, (May 2nd, 2017), https://rbi.org.in/scripts/BS_FemaNotifications.aspx?Id=10392.

Protecting the tools, technologies, facts, formulas and all the other things created by humans in the field of Bioinformatics is easy, but what about protecting the biological databases?

The biological databases are the compilations of the data that have been gained by extensive research, analysis, and understanding of the biological specimens and life science. They are divided into two categories, the primary database and secondary database.

Primary databases are populated with experimentally-derived data such as nucleotide sequence, protein sequence or macro-molecular structure. Experimental results are submitted directly into the database by researchers, and the data are essentially archival in nature. By contrast, secondary databases comprise data derived from the results of analyzing primary data. They are often referred to as curated databases.

The first question was whether primary database should be protected with the IP laws or not, because it is something on which a lot of investments have been made, and a lot of effort has been put into the research to compile the database, thereby resulting in the protection of these databases with the IP laws in order to commercialize it and avoid misuse. The government wants it to be public so that it helps the researchers, but the companies have a different take on it. The battle is hard and long fought, seeming to never-end. Thus, primary databases should not be made public. This is due to three reasons–

1. A notable number of primary databases are funded by government and utilize the money they gain by the taxes incurred, and therefore should be made freely accessibleto the benefit the public;
2. IP protection for primary databases will gravely affect the foundations of research in life sciences, since these databases are essential to scientists in the field of research gene engineering, molecular biology and proteomics. The use of free databases will promote communication among research groups, and avoid tedious and expensive repetition;

---

[77]Vallari Dubey, *Legal Nature of Bitcoins: The Encrypted Digital Currency,* (April 29th, 2017),http://vinodkothari.com/blog/legal-nature-of-Bitcoins-the-encrypted-digital-currency-by-vallari-dubey/.

[78]Reserve Bank of India Act, 1934.

[79]*Foreign Exchange Management (Manner of Receipt and Payment) Regulations,* Reserve Bank of India, (May 2nd, 2017), https://rbi.org.in/scripts/BS_FemaNotifications.aspx?Id=10392.

3. It is difficult to accurately find the source of data in important primary databases because they are provided by not only the makers of the database but also by active researchers in the field of life science.

Intellectual Property, Bioinformatics Databases and Government Funds

The main objective of government in the field of biology and technology is to enhance the reach of biotechnology, thereby promoting the development and welfare of society in general. To achieve this aim, government should provide financial subsidies to the makers of Bioinformatics databases. As a result of these financial subsidies, primary database can be made freely available to users involved in active research and can further reduce the cost of secondary databases in research. However because financial resources of government are limited, there is a method proposed by Junli Chang and Xuezhong Zhu which is analyzed as follows: -

The Diffusion of Biological Information[12]

Providing data or the biological information is the diffusion process in this technique. The purpose of creating Bioinformatics database is to give a platform for the researchers to further obtain newer biological information. However there is a pattern observed in the diffusion and spread of biological information.

In the first phase researchers need to examine utilities for the diffusion of new biological information and explore ways of its possible application. So the cost is high and users are less, which leads to slow diffusion.

In the second phase, more users get involved and the diffusion rate becomes faster and faster where the information is mostly used.

In the third phase, the biological information has been utilized sufficiently and most researchers have mastered them. Meanwhile, because of high diffusion, the ways to obtain the biological information will also increase, while the research on the biological information will decrease. The rate of diffusion of information will become stable and not change much.

Considering the above model, it will be appropriate for the government to give varying financial subsidies to biological databases depending on the diffusion

[12]Junli Chang and Xuezhing Zhu, *Bioinformatics Databases: Intellectual Property and Protection Strategy,* 15 JOURNAL OF INTELLECTUAL PROPERTY RIGHTS 447 (2010).

phase of information. For the first stage, which is the slow stage, there is a lot of research on application which is a sensitive indicator of database cost, because high cost will lead to slower diffusion of biological information and hinder the progress of life science research, and so the government should provide subsidies to make up for the high cost of application based research.

In second phase, due to earlier research breakthroughs, a number of commercial applications involved in the use of biological information have already been formalized. Since the cost factor is now low, the government should provide partial support to the Bioinformatics database which is equal to not providing any subsidies to the users of secondary databases but still supporting primary database users.

In third phase, the diffusion and application of biological information are relatively mature and government subsidies are no longer required so the government can cancel subsidies fully and operation of Bioinformatics database can be regulated by the market which will reduce financial burden of government.

On the other hand, the protection of secondary database is somewhat different. As it is not complex, the various methods of protecting them are as follows: -

1. Copyright law can be used where the database is considered as a compilation according to Trips agreement;
2. Trademark law and trade secrets can be used where database is known to public;
3. *Sui Generis* law can be used where to qualify for database law protection, it needs to meet the requirements of qualitative and quantitative investment of substantive nature while obtaining the data, verifying the data or presentation of contents of database. This law can adapt to the development of database industry and protect the interest of database makers and the content itself but is only available to the European Economic Area and not outside, which leaves nearly one option to tackle the problem;
4. Contract Law can be used by makers to prevent breach off aith and infringement.[13]

[13]A. T. Jagdish, *Law of Intellectual Property and Bioinformatics,* 3(1) INTERNATIONAL JOURNAL OF SCINTIFIC RESEARCH PUBLICATIONS 1 (2013).

Challenges and Difficulties

The scope of Bioinformatics in India is evolving at an exponential rate. The need for tools and databases is increasing day by day with a felt need in medical, pharmaceutica land biotechnology field coming up fast. But to protect the Bioinformatics database and its tools and products, the right set of skills and knowledge are required by the patent agents in India to file, and handle cases related to the field of Bioinformatics. The number of agents who have areas of expertise in subjects like these are very less in India. This has not only made it difficult for the IP regime to get a stronghold on the field of Bioinformatics, Biotechnology, and lifesciences due to lack of knowledge and skills in the aforementioned areas, but has also hindered an opportunity to step up the economy by convincing the companies outside the country to set up in India. Due to less patent agents being able to handle these kinds of cases, the workload on the existing patent agents increases so much that it becomes difficult for them to focus on the issues at hand, which eventually leads to pendency in filing of patents, etc.

There have been suggestions that patent agents can eventually evolve and learn about these areas with time when they will deal with these kind of cases and filings, but the problem which can arise in this situation is that it will take a lot of time for them to learn the ways, methods and strategies to finally be able to tackle these fields effectively in no time. The more effective way of going forward in these areas is that the law should be such that it is able to identify the people having knowledge in the field of Biology, Information Technology and IP laws along with the patent agents who can come together to make an alliance to advise, draft and file patents and even hold clarifications and oppositions effectively and thus Bioinformatics will be given more consideration in the Indian legal system. Others methods include not only increasing the number of patent examiners in the field of Bioinformatics but also improving the efficiency of the patent examiners, by having more training programs at regular intervals for patent examiners particularly in new areas of technology like Bioinformatics which should be a major focus in countries like India.

The DIT schemes called as SIP-EIP (Support International Patent Protection in Electronics and IT) can help in refund of 50% of the cost of filing patent to save money and support international patent protection. The scheme will reimburse the costs incurred by SMEs and Technology Start-up Units in filing international patent applications (in the field of Electronics and ICT)

fortheirindigenousinventions.[14] The highlights are as follows: -

1. The funds are given as a grant, which means, no refunds can be expected.
2. There is no license fee or any other benefits which is asked by DIT.
3. Applicants are free to hire any lawyers, which opens up opportunities for the ones who are interested in handling these cases and want to learn more in the areas of biotechnology, and Bioinformatics and similar fields of technology and IP laws.
4. 50% of all the expenses, including lawyer's fee is refunded by DIT. But, to enjoy this benefit, the criteria of "50%" must not exceed Rs. 15 Lakhs, which is approximately $37,000 from an international perspective.
5. The Start-up/SMEs need not be an ICT company; it can be a biotech company or any other company. The only requirement is that the invention sought to be protected relate to the broad area of Electronics/ IT. Thus, Biotech with Bioinformatics R&D can also avail of this scheme.

The methods aforementioned will contribute to the growth of the economy of the country by various means such as more efficient ways of patent filing with the SIP-EIP scheme and strategies of diffusion theory. This will also make a statement to the worldwide IP dominion that the Indian IP regime is effective enough to let companies outside of the subcontinent establish themselves and make investments without any worries and hesitations.

Conclusion

The field of Bioinformatics and IP laws are intertwined when it comes to protecting one's hard work and income. Due to such rapid growth in the field, there are a lot of opportunities to attract more patent filings and give a boost to the Indian based biotech companies as well as companies situated outside the continent to invest in the Indian economy. This will not only give more exposure to the Indian biotechnology regime but will also help in the growth of the economy, thus leading to India's development and also the development of effectiveness of patent agents to file and handle cases of the IPR and Biotechnology.

---

[14]Mrinali Kochupillai, *Promoting DIT Scheme*, spicyip.com (July 22nd, 2008), https://spicyip.com/2008/07/promoting-ip-protection-dit-scheme-for.html.

CALL FOR ARTICLES FOR VOLUME 2

VIT Law Review being a scholarly publication strives to publish leading research articles, notes and comments authored by the legal fraternity from across the world. The VIT Law Review has been serving as a research arm for practitioners, academicians and the broader intellectual community.

Submissions are invited for the following:

- Articles (5000-10000 words)
- Essays (3000-6000 words)
- Notes & Comments (2000-5000 words)
- Book Review (1000-2000 words)

All submissions must adhere to the following guidelines:

- The submitted manuscript should not be published previously or under consideration for publication.
- Footnoting must be mandatorily done according to the 20th Edition Harvard Bluebook Citation.
- The font must be Times New Roman with size 14, for headings, and 12 for the remaining body with footnotes in Times New Roman, size 10.
- Line Spacing -1.5 and 1.0 (footnotes) and a margin of 1" on all sides.
- The submission must contain an abstract of no more than 300 words.
- Co-Authorship is permitted to a maximum of two authors.
- Plagiarism will be viewed with zero tolerance.

Submissions shall be mailed to chennai.lawreview@vit.ac.in with the Subject line as **'SUBMISSION FOR VIT LAW REVIEW'.** Articles will be considered on rolling basis

Contact details:

Faculty Editors
Prof. Tania Sebastian, +919810237136.
Prof. Ankit Kumar Padhy, + 919407910104