

WIPO



SCCR/7/5

ORIGINAL: English

DATE: April 4, 2002

E

WORLD INTELLECTUAL PROPERTY ORGANIZATION
GENEVA

STANDING COMMITTEE ON COPYRIGHT AND RELATED RIGHTS

Seventh Session
Geneva, May 13 to 17, 2002

**A STUDY ON THE IMPACT OF PROTECTION OF UNORIGINAL DATABASES ON
DEVELOPING COUNTRIES: INDIAN EXPERIENCE**

*Study prepared by Mr. Phiroz Vandrevalla, Chairman
National Association of Software and Service Companies (NASSCOM)
New Delhi*

TABLE OF CONTENTS*

	<u>Page</u>
SUMMARY OF THE STUDY	2
Indian Database Industry.....	2
Legal Regime	3
Conclusion	4
STUDY	5
METHODOLOGY	5
INTRODUCTION.....	6
ISSUES.....	7
I. THE DATABASE INDUSTRY IN INDIA.....	8
A. GOVERNMENT DATABASES.....	8
(a) Scientific Databases	9
(b) National Informatics Centre (NIC)	9
(c) Remote Sensing Data	10
(d) Ancient India: “Traditional Knowledge” Data.....	11
(e) “Diverse India”–Genomic Databases.....	12
(f) Bio-Informatics	14
B. PRIVATE SECTOR.....	14
(a) Specific Examples of Private Databases	16
(i) Centre for Monitoring the Indian Economy (CMIE), Bombay	16
(ii) Tata Energy Research Institute (TERI) Publications.....	16
(b) Online Databases.....	17
(c) Data Entry	17
(d) Revenues	17
(e) Need for Legal Protection	17

*

At the request of its member States, WIPO commissioned, in 2001, five studies on the economic impact of the protection of non-original databases in developing countries and countries in transition. This study, one of those five, contains the research and opinions of only its author and does not in any way reflect the views or position of WIPO.

	<u>Page</u>
C. INDIA'S INFORMATION TECHNOLOGY CAPABILITY	18
(a) Data and Content Management	19
(b) Projections for the Indian Database Industry	21
II. LEGAL RESPONSE: INTERNATIONAL	22
III. INDIA: EXISTING LEGAL REMEDIES	25
(a) Copyright Protection.....	25
(b) Contractual Protection	26
(c) Information Technology Act	26
(d) Protection of Technological Measures	27
(e) Unfair Competition.....	28
CONCLUSION	29

SUMMARY OF THE STUDY

With the advent of the digital era and the consequent increase in the creation and dissemination of electronic databases, calls have been made for a regime that would protect even unoriginal databases, i.e. databases that lack the requisite originality/creativity to qualify for copyright protection, but that nonetheless, involve a substantial investment/effort.

It therefore becomes important to determine as to whether countries and in particular, developing nations such as India need to provide for some sort of *sui generis* protection in favor of “unoriginal databases.” This determination would, to a large extent, depend on the status of the database industry in India.

Indian Database Industry

An analysis of the Indian database industry reveals the following:

- (a) 80% of the databases in India still emanate from the domain of the Government.
- (b) Only a small portion of the above data (about 40%) is available to the general public. The remaining 60% of data is either internal data that is department specific and therefore holds no relevance for the general public or “classified,” as it includes sensitive information pertaining to atomic research, space research, defense research, etc.
- (c) The private database industry, though a mushrooming one is still fragmented at the moment. Owing to the fragmented nature of this industry, documented data/statistics relating to this industry are hard to come by.
- (d) As is usual with government data, most of it is not commercialized. Apart from the fact that the government intends to encourage scientific research through the wide dissemination of data, another reason for non-commercialization could be that most of the data generated by the government is raw and needs to be analyzed and classified further in order to be useful.
- (e) One area that has seen that some sort of commercialization is Remote Sensing Data. With its advanced capabilities in remote sensing, the National Remote Sensing Agency (NRSA) has begun to commercialize the various data generated by *inter alia*, selling it to various other countries.
- (f) The traditional argument that stronger Intellectual Property Regimes would facilitate the transfer of technology to the developing countries does not hold good in the case of the “database” industry, where the key subject matter is “data.” Therefore, absent a finding that the industry is either a well-developed one or has such potential, the need for a separate *sui generis* system in India to protect the industry may not be felt.
- (g) Currently, the range of database companies in India are the same as in any other part of the world, i.e. databases containing legal information, banking and financial information, market research information, stock market information, ticketing and travel information, etc. These sectors offer little competitive advantage for the Indian database industry and therefore it becomes important to look into sectors that could be potential niche areas, having tremendous commercial potential.

(h) The potential for data generation and consequent wealth creation is perhaps the strongest in three broad areas:

(i) Traditional Knowledge Data

Owing to India being one of the most ancient civilizations in the world, it has tremendous reserves of traditional knowledge such as traditional medicinal knowledge, folklore, art, etc. It is only now that the Government has felt the need to document this traditional knowledge. To give an example of one such initiative, the Government has initiated the Traditional Knowledge Digital Library (TKDL) program; this involves the procurement of various literature on traditional medicinal systems, such as Ayurveda and the scientific classification of this information as would enable an easy retrieval of the same. With this surge in documentation, traditional knowledge databases would gain prominence and command high revenues.

(ii) Genomic Data

Another sector that could see huge revenues is the genomic industry in India. India has perhaps the most diverse population in the world; consequently its diverse gene pool would be very attractive for medical researchers.

Besides human genomic data, India is rated as one of the “biodiversity” hotspots in the world; thereby making it a potentially rich source of animal/plant/microbial genomic data. The potential uses of this vast and varied genomic data could bring in substantial revenues for the country.

Further India’s information technology potential could be leveraged in the genomic industry to catapult India as a leader in the Bio-informatics sector.

(iii) Information Technology

Advances in information technology have helped unleash a wide variety of databases and consequently given a major fillip to the database industry. India’s status as an IT super-power could be leveraged to aid the growth of the Indian database industry.

The Government has recognized this fact and had, in one of its reports, identified “database creation” as one of its priorities. The IT task force, constituted by the Government, had in its IT action plan, included “databases” within the broad definition of the “content industry” and identified ways and means in which this industry could be developed.

Legal Regime

Although India does have a very strong copyright regime, which had till date been in *pari materia* with the UK “sweat of the brow” doctrine, a recent decision by the Delhi High Court seems to cast a doubt on whether the Indian Copyright regime protects “unoriginal” databases. Further, India has not recognized a separate tort of unfair competition/misappropriation and therefore this common law right may not be available to protect the efforts of the Indian database industry.

In this context and in order to attract more investment to the database industry in India, it is critical that India provide for a protection regime that would help protect unoriginal databases.

Conclusion

India is “ancient, diverse and mathematical.” The “ancient” quality of India translates into an immense wealth of “traditional knowledge.” The “diverse” quality translates into a varied gene/blood pool and consequently a wealth of genomic data. The “mathematical” quality translates into India’s IT capability, which would then be used to leverage the database industry in general.

The problem of potential non-access could be solved by carving out specific exceptions in favor of the academic and research community. The argument that government data cannot be commercialized does not hold much ground in the light of the fact that the Indian Government has now started commercializing its data, especially in areas where it is technologically advanced, such as remote sensing. More importantly, with the digital era churning out tons of data today, one should structure the regime in such a way that mere data is not protected; rather only databases that involve the expenditure of substantial investment/resources should be protected. One also needs to clearly define what constitutes a “database.”

STUDY

METHODOLOGY

As the topic itself suggests, this paper is a study of the Indian database industry in order to determine whether the creation/generation of unoriginal databases by the industry warrants protection.

Since the database industry in India is not organized at the moment, documented data/statistics relating to this industry were hard to come by. Therefore, industry figures and perspectives have largely been based on interviews with key personnel from the industry; the figures given by them are at best an approximation. Apart from these interviews, other data on the industry was procured from the Internet.

The paper is sub-divided into various chapters as follows:

- The Introduction gives a broad outline of the database industry in general and the various “protection” concerns that have arisen in today’s digital age.
- Chapter I then goes on to discuss the database industry in India, including both the government and private sectors. Since the issue of protection of unoriginal database has assumed tremendous significance in the light of the digital revolution, this chapter will also focus on India’s IT capability and the consequential potential to generate more databases.
- Chapter II details out the international legal response (in particular the European and the US response) to the concerns raised by the database industry in general.
- Chapter III charts out the Indian legal system, in so far as protection for the database industry is concerned.
- The Conclusion highlights the fact that for this burgeoning industry in India, additional protection is imperative. However, certain specific exceptions will have to be carved out, in order to ensure that genuine access to the scientific and research community is un-impeded.

INTRODUCTION

In an information society, databases containing and cross-referencing all types of information and presenting such information in an easily accessible form are of immense value. The databases are a vital element in the development of a global information infrastructure and an essential tool in promoting economic, cultural and technological advancement.¹

The overall landscape of the database industry has altered considerably over the years with the technological advancements providing new mediums for making the information goods available to the consumers. The convergence of digital computing and telecommunications technologies in the 1990's led to an unprecedented boom in the information industry.² Not only was the collection, compilation and arrangement of the information simplified but the dissemination and distribution of such information also made easy. This no doubt created an enormous capacity for the mobilization and expansion of global scientific and technical resources.³

Consequently the database industry has come to mean various commercial and non-commercial activities relating to the bibliographic, textual and statistical databases as well as the information, education and entertainment materials in the electronic form including audio, video and multimedia forms.

As the technological advancements are in a state of flux, the database owners are constantly required to make substantial investments to keep pace with the new developments. However it is argued that the traditional intellectual property regimes have failed to protect the interests of the database owners. Thus the database providers have lobbied with the law-makers across the world to provide *sui generis* right in the contents of their database so as to safeguard the investments that they have to make in the collection, compilation and management of the databases.⁴

¹ Preamble to the WIPO Draft Treaty on Intellectual Property in Respect of Databases available at http://www.wipo.org/eng/diplconf/6dc_all.htm, visited on January 3, 2002.

² See e.g. Robert. P Merges, "*The End of Friction? Property Rights and Contract in the "Newtonian" World of Online Commerce,*" 12 Berkeley. Tech L.J. (1997); Julie E. Cohen, "*Lochner in Cyberspace: The New Economic Orthodoxy of Rights Management*", 97 Mich L.R. 462 (1998).

³ "*The use of computers made it economically feasible to collect, store, manage and deliver huge amounts of data at a time when continuously expanding databases have become the building blocks of knowledge, especially in the observational sciences.*" See Chapter IV, "Bits of Power," c.f. J.H. Reichman and Pamela Samuelson, "*Intellectual Property Rights in Data*", 50 Vand. L.R., 51 (1997).

⁴ See J.H. Reichman and Pamela Samuelson, "*Intellectual Property Rights in Data*", 50 V and L.R., 51 (1997).

ISSUES

The main issues are highlighted as under:

1. Whether a developing country such as India requires a separate *sui generis* legislation for the protection of unoriginal databases?
2. If yes, how do we reconcile this right, with providing un-impeded access to information by the scientific and research community?

To elaborate:

1. The main concern of the database industry while soliciting enhanced protection has been an effective remedy against the problem of infringement. It is most vehemently contended that infringement/piracy could become a potential disincentive to the creation of new value-added databases. The advent of electronic databases has amplified the problem of piracy with the databases being expensive to make but cheap to copy.⁵ It also argued that whereas with effective legal protection, database providers would have the confidence to willingly disseminate data and thus make the information more readily accessible.
2. The calls by the database industry for additional protection have sparked off a vociferous debate on the level of protection justified so that the access to information is not unduly hampered. The scientific and academic community has all along countered the demands of the database industry for additional protection. It is argued by them that the facts and ideas are building blocks of intellectual discourse and should not be removed from the public domain. These communities fear that the extent of protection sought by the database industry would inevitably lead to mini-monopolies over information.

The most potent of the arguments put forward by the scientific community against attempts at commodification of data is that it would jeopardize basic and applied scientific research. It is felt that the recognition of the legitimate concerns of and incentives for commercial database producers should not lead to undue restrictions for the scientific community in their legitimate pursuits which are of immense value to society. In the words of Paul A David:⁶

“We really do not know how much further the current rush towards privatization of scientific and technological knowledge can go before it starts to seriously undermine the inherited structure of fragile conventions and institutions that support cop-operative research activities, thereby setting in motion the contraction of the global domain of open scientific enquiry.”

The scientific community argues that there is no justification for additional avenues of protection for databases, as the existing legal remedies are sufficient for the effective

⁵ In fact, the European Union had initially restricted its draft *sui generis* directive to the electronic media. It was only later that this was extended to the print media as well. See Reichman, *supra* note 4.

⁶ Paul A. David, *“The Digital Technology Boomerang: New Intellectual Property rights threaten global “open science.”* (See www-econ.stanford.edu/faculty/workp/swp01005.pdf, visited on January 7, 2002).

protection of database providers. It is alleged that the assertion that investors are deterred by a perceived lack of incentives remains anecdotal and unsubstantiated.

I. THE DATABASE INDUSTRY IN INDIA

As to whether there exists a specific “database industry” at all in India is a debatable proposition. However, although the industry may appear to be fragmented at the moment, there is no denying fact that it has great potential for growth.

Most databases in India still continue to emanate from the domain of the government. A conservative estimate of the percentage of government databases in India pegs it at 80% of the total databases created/generated in India.⁷

In fact, the mushrooming database industry in the private sector relies heavily on the data generation by the government agencies, which data are then processed according to the specific requirements of the particular company/industry. It is estimated that approximately 80% of the data in databases created by the private enterprises are procured from government databases/agencies.⁸ However, though 80% of the data generated in India is by the government, only a small portion (about 40%) is available to the general public.⁹ The remaining 60% of data is either internal data that is department specific and therefore holds no relevance for the general public or “classified,” as it includes sensitive information pertaining to atomic research, space research, defense research etc.

A. GOVERNMENT DATABASES

Today, various government agencies are engaged in the collection/generation of large amounts of data including *economic, demographic, socio-demographic audit, agriculture, commerce, communication, energy, environment, finance, food and civil supplies, health and family welfare, home affairs, human resources development, industry, information and broadcasting law & justice, parliamentary affairs & tourism, personnel, public grievances and pension, petroleum and natural gas, planning, program, implementation, science and technology, steel and mines, surface transport, and urban development* data.

As is usual with government data, most of it is not commercialized. Apart from the fact that the government intends to encourage scientific research through the wide dissemination of data, another reason for non-commercialization could be that most of the data generated by the government is raw and needs to be analyzed and classified further in order to be useful.

⁷ Per statement of Dr. Chandra Gupta, Head of Computerized Data Collection, CSIR (these figures are mere approximations).

⁸ Id.

⁹ Id.

(a) Scientific Databases

Amongst the government databases, the most prominent are the scientific databases.¹⁰ In fact, it is estimated that there are more than forty research institutes in India under the aegis of the Council of Scientific and Industrial Research (CSIR) engaged in the collection and generation of data/databases, marketed through publications or on CD-ROMs.

For example:

- The National Institute of Oceanography at Goa has valuable data relating to pipeline routes, submarine routes, shallow seismic surveys, seabed surveys and magnetic surveys;
- The National Botanical Research Institute at Lucknow would have important data on things like lentils;
- The Indian Institute of Petroleum at Dehradun would have important data on crude and petroleum products in India;
- The Geographical Survey of India would have important data relating to seismic activity, mineral resources, land use and environmental impact monitoring;
- The Spice Board at Cochin has valuable data on Indian spices which is used to promote their exports; and
- Various other bodies are collecting data on bio-resources such as medicinal plants—their names, occurrences, characteristics and uses.

(b) National Informatics Centre (NIC)

Amongst the government agencies, the most active creators/developers of database is the National Informatics Centre (NIC). NIC caters to academicians, researchers, professionals and businessmen by providing a variety of information. Some of them are highlighted as under:

– *District Information System (DISNIC)*: DISNIC program facilitates the development and analysis of databases (both spatial and non-spatial) for decision support in 27 sectors. These include education, industry, agriculture, cooperative, microlevel planning etc. This program is operational in about 460 districts in the country. Implementation of DISNIC program facilitates development planning and responsive administration right from grass root level.

– *CourtNIC*: CourttNIC is a user friendly information system designed to provide information on pending cases before the Supreme Court of India. The information made available on COURTNIC includes, *inter alia*, the name of the case, and its status and position.

¹⁰ According to one estimate there are approximately 400 to 500 databases in the scientific field in India.

– *GISTNIC*: GISTNIC (General Information Service Terminal–National Informatics Centre) is an on-line database of NIC that offers on-line information retrieval services from its wide ranging static and quasi-static areas such as Indian Economy Monitor Population Census Database, Village Amenities Database, Rural Technology Database, Tourist Guide of India, District Profiles, University and College directory of India, Database on Traditional Sciences and Technologies of India etc.

(c) Remote Sensing Data

Perhaps one are that is seen some sort of commercialization is Remote Sensing Data. The key governmental agency which is involved in this data generation is the National Remote Sensing Agency (NRSA), an autonomous institution supported by the Indian Space Research Organization (ISRO), responsible for acquisition, processing and supply of data from remote sensing satellites.¹¹

Data from Indian Remote Sensing (IRS) Satellites is used for various applications of resource survey and management such as:

- Pre-harvest crop acreage and production estimation of major crops
- Flood risk zone mapping and flood damage assessment
- Snow-melt run-off estimates for planning water use in down stream projects
- Forest survey
- Coastal studies
- Cartographic applications

It is believed that the data so collected by Indian satellites such as IRS 1C and IRS 1D have very high quality resolution¹² and is of great commercial value.¹³ The commercial tie-ups of the ISRO/NRSA include the following:

- S1-EOSAT, a USA Company receives and markets IRS data worldwide.
- NRSA recently entered into a satellite data product distributorship agreement with BPL Innovision Technologies Pvt. Ltd., India. As per this agreement, BPL is authorized to distribute satellite data acquired by NRSA to India users in the private sector only. BPL's access and involvement with corporate entities engaged in high technology and infrastructure development areas is expected to significantly enlarge the utilization of remote sensing for infrastructure projects in the area of telecommunications, roadways, power, ports, pipelines, etc.
- The clients of ISRO/NRSA include several governments such as South Africa, Nepal, Bangladesh, Sri Lanka, Indonesia etc.

¹¹ The NRSA also runs the Indian Institute of Remote Sensing at Dehra Dun, which currently supplies data from four Indian Remote Sensing satellites (the IRS series) to the users.

¹² India recently launched Technology Experiment Satellite (TES) which provides a 1m imagery. See "Overview of Indian GIS Industry: Year 2001," See <http://www.gisdevelopment.net/news/overview/overview.htm> (visited on January 3, 2002)

¹³ Telecom Agencies of various countries use remote sensing data for laying their cables.

Currently ISRO/ NRSA has an average sale of 18,000 data products per annum¹⁴ ISRO/NRSA is now set on capturing a substantial part of the growing world market of raw data (approximately US\$500Million); it plans to fly Cartosat-1 (with 2.5-metre resolution), and Cartosat-2 (1 m resolution) in 2002 and 2003.¹⁵

Therefore, one is witnessing the gradual trend of the government to capitalize on its data generation capabilities, in areas of technological superiority.

(d) Ancient India: “Traditional Knowledge” Data

Another sector that has tremendous financial potential is “traditional knowledge”. Owing to India being one of the most ancient civilizations in the world, it has tremendous reserves of traditional knowledge such as traditional medicinal knowledge, folklore, art etc. It is only now that the government has felt the need to document this traditional knowledge. With this surge in documentation, traditional knowledge databases would gain prominence and command high revenues. To give an example of one such initiative, the Government has initiated the Traditional Knowledge Digital Library (TKDL) program.

This program, conceptualized by the Department of Indian Systems on Medicinal Homeopathy (ISMH) on October 1999 is currently being undertaken by the National Institute of Science Communication (NISCOM). It involves the procurement of various literature on traditional medicinal systems, such as Ayurveda and the scientific classification of this information as would enable an easy retrieval of the same.¹⁶ The genesis of this initiative lay in the public outcry against the patenting of Indian traditional knowledge systems in the USA.¹⁷

The TKDL project seeks to classify traditional knowledge on the basis of a novel classification system (known as traditional knowledge resource classification, TKRC)¹⁸. The

¹⁴ See <http://202.54.32.164/test/main/DD.HTML> (visited on January 8, 2002)]. However the revenue figures are still not optimal. In the 1998 CAG (Comptroller and Auditor General, India) report, it was stated “*Though there has been continuous increase in the collection of data year after year, the sale/utilization of data had not kept pace with the collection.*” See http://www.cagindia.org/reports/scientific/1999_book1/overview.htm, visited on January 3, 2001.

¹⁵ See http://www.cagindia.org/reports/scientific/1999_book1/overview.htm, visited on January 3, 2001.

¹⁶ See Dr. R.A. Mashelkar, Director General, CSIR, “*Intellectual Property Rights and the third world,*” CSIR, New Delhi.

¹⁷ For example, two US based Indian were granted a US patent 40,504 on March 28, 1995, for the “use of turmeric in wound healing.” This patent which was assigned to university of Mississippi Medicinal Centre, USA claimed “the administration of an effective amount of turmeric through local and oral roots to enhance the wound healing” process as a novel finding. CSIR (Council of Scientific and Industrial Research, India) opposed the patent and located 32 references (some of them have more than 100 years old) which showed that the claimed novelty of Turmeric was part of the traditional knowledge of India. See supra note 16.

¹⁸ See Report on the Task Force on Traditional Knowledge Digital Library (TKDL), May 2000, New Delhi.

classification will be more or less on the lines of the classification followed by International patent classification (IPC).¹⁹

Although the TKDL began as a reactive measure to the fact that Indian traditional knowledge was being commercialized through patenting, the government has now realized the value of the database so created and is now contemplating commercializing it.²⁰ In fact, the fact that TKDL would involve the translation of the database into 4 international languages would greatly enhance its international market and bring in revenues for the country.

Of course the database created by the TKDL project itself would easily qualify for protection as an original database under the existing copyright regime²¹. However, in the light of the fact that this traditional knowledge initiative would spur others also to documents such data, it is quite possible that there would emerge several unoriginal databases as well.²² In this context, it becomes critical that the existing copyright regime be supplemented by a *sui generis* system, so that all traditional knowledge databases are protected; this would facilitate the commercialization and trading in of such data.

Over and above the fact that TKDL database would be protected as an original database under the Copyright Act 1957, the said database would also be protected through deployment of watermarks and other technological measures.

(e) “Diverse India”–Genomic Databases:

The advent of the Human Genome Project²³ caused a large amount of genomic data to be generated. Although this data needed to be analyzed further before one could derive some functional information, the generation of this raw data began assuming commercial significance.²⁴ Having completed “hand in hand” the first draft of the human genome this

¹⁹ Most patent examiners around the world use the IPC during patent examination. Although the IPC has more than 1,00,000 sub-groups for retrieving information on modern scientific inventions, it has only one such sub-group for retrieving information on “medicinal plants.” Since the Indian TKDL will have information on 5000 subgroups, their inclusion in the IPC will enhance the quality of patent examination substantially.

²⁰ The National Institute of Science Communication (NISCOM), which is the main body developing this traditional knowledge library, is trying to work out a mechanism wherein this database would be freely accessible to patent examiners but would have to be paid for by individual users—per statement of Mr. V.K. Gupta, Director, NISCOM.

²¹ The Indian Copyright Act, 1957 protects original compilations and databases as “literary works.” Since the TKDL will be generated using a complex classification mechanism known as TKRC, it is evident that the database so generated would merit copyright protection. Indian courts had, in the first ever instance of online copyright infringement, restrained an Italian party from copying the herbal database of an Indian pharmaceutical company, The Himalaya Drug Company.

²² In fact, NISCOM itself has in the past documented several Indian medicinal plants in various databases, etc.

²³ The Human Genome Project was a fifteen-year, three billion dollar project to identify and map every gene in the human genome. A first draft of the human genome was recently released.

²⁴ Although the HGP project began as a collaborative research project, with the discovery of large amounts of genomic data that had the potential of yielding commercially valuable therapeutic

spring, the International Human Genome Sequencing Consortium²⁵ and Celera²⁶ are still split on the issue of whether Celera should be entitled to intellectual property protection for the results of its sequencing efforts. Whereas the Consortium makes the sequences available for free on the Internet,²⁷ Celera keeps them in the private domain and only provides access thereto on commercial terms, albeit that free, limited access is provided to qualified academic users.²⁸

The Indian government, in keeping with the HGP norm of facilitating scientific research through open sharing of data, began making its genomic data public.²⁹

Although the government has not yet thought about commercializing genomic data, it is evident that India specific genomic data would command a high commercial value in the market. The reasons for this are as below:

- Since India has one fifth of the world's population, India's genomic data would perhaps be the most diverse and therefore ideal for genetic analysis.
- Considerable number of cases of rare genetic disorders are likely to be found merely because of the population size.
- It is possible that half of the world's genetic mutations are in India.
- India has a large number of very competent medical practitioners, modern clinical expertise and a vast network of hospitals, clinics and health centers, unlike in many other developing nations, thereby facilitating quick sampling and analyzing of blood/tissue samples.³⁰

Beside human genomic data, India is rated as one of the "biodiversity" hotspots in the world; thereby making it a potentially rich source of animal/plant/microbial genomic data. The potential uses of this vast and varied genomic data could bring in substantial revenues for the country.

[Footnote continued from previous page]

products, private researchers began attempting to stake proprietary claims in such information, including even applying for patents.

²⁵ see <http://www.nature.com/genomics/human/index.html>, visited on January 5, 2002.

²⁶ see <http://www.sciencemag.org/cgi/content/full/291/5507/1304> visited on January 5, 2002.

²⁷ see, <http://www.ncbi.nlm.nih.gov/Unigene/Hs.Home.html> visited on January 5, 2002.

²⁸ See the terms of the Celera Free Public Access Click-On Agreement, available at: <http://publications.celera.com/pubsite/terms.cfm>., visited on January 5, 2002.

²⁹ These databases were set up at the Indian Institute of Science (IISc) at Bangalore, University of Pune, Jawaharlal Nehru University (JNU), New Delhi and Institute of Microbiological Technology (IMTech), Chandigarh. These databases included mirror sites of international development databases as well.

³⁰ Samir K. Brahmachari, Suparna Gupta and Mitali Mukerji, "Genomics: Predictive Medicine and Drugs of Tomorrow," New Millennium Lectures: Celebrations of Science—National Institute of Science Commission, 1999.

(f) Bio-informatics:

Further, India's information technology potential could be leveraged in the genomic industry to catapult India as a leader in the Bioinformatics sector. As discussed earlier, India has rich source of genetic and other clinical data, both from its large population and from the rich animal and plant diversity that India possesses.

"Bioinformatics" basically entails the use of information technology to determine the functional relevance of sequence data.³¹

Bioinformatics can be used to identify proteins and therefore develop new drugs. The benefits of providing greater protection to bioinformatics data could be two fold:

- Indian genomic companies appoint such bioinformatics service providers to create databases from which they generate revenues (either through the development of drugs or through the sale of such data).
- Even if the bioinformatics service provider is engaged by a foreign company, a strong protection regime would create further incentives for the foreign company to supply such data to the bioinformatics company in the first place, thereby making it more competitive in the world market.

B. PRIVATE SECTOR

Although the database industry in India is a fragmented one,³² there is no denying the fact that this is a growing industry. Contrary to the earlier scenario where information activities were mainly funded by the Government, one now finds an aggressive participation from the private sector.³³

The range of database companies in India are the same as in any other part of the world,³⁴ i.e. databases containing legal information, banking and financial information, market research information, stock market information, ticketing and travel information, etc.³⁵ The range of database companies in India would be broadly as under:

³¹ See "A white Paper on Bioinformatics in India," Dr. Vijay Chandru, and P. Hari, January 15, 2002

³² As per the statement of Mr. Pradeep Gupta, Managing Director, Cyber Media (India) Limited (a company that publishes a variety of information technology magazines and journals), "there is no concept of database industry in India."

³³ See BG Surinder Singh, "Indigenous Databases and Information Products/Services," NISSAT, New Delhi, 1997 (article available at <http://itt.nissat.tripod.com/itt9702/infoprod.htm>; last visited January 5, 2002), where he states that "While the indigenous information industry has to grow and become globally relevant, there is a need to nurture, help develop some of the potential databases and facilitate their increased access."

³⁴ According to our estimate, there were as many as 1000 database companies in India.

³⁵ As mentioned earlier, no informational documents relating to the database industry has been documented. Therefore, the author relies on interviews with key personnel from database companies.

- Credit card companies like Diners Club and banks like Citibank and Standard Chartered, have a database on their current customers and on prospective customers.³⁶ These databases are continually updated and refined, and apart from names and addresses, various other parameters such as incomes, credit ratings, the types of goods usually purchased and so on can be ascertained therefrom.
 - There are the databases on company profiles of the kind maintained by Dun & Bradstreet (D&B) which contain important information used by investors, importers, exporters, shareholders, banks and financial institutions. For example a typical database catering to the specific needs of importers/exporters would contain information including buyer's names and addresses, supplier's name and addresses and the description, weight, date of shipment compiling or value of the product. Infodrive India is one such agency, which maintains reports on all exports and imports to the United States of America from the rest of the world and also from the six major ports of India.
 - Another example is free-to-air television. Companies use people-meters to assess how frequently a particular television serial is viewed by a family. These figures are regularly compiled and the advertising revenues depend entirely on these ratings.
 - Databases are also maintained by newspaper houses, pertaining amongst other things, to current events such as elections, past trends in voting patterns, traffic accidents, crime rates, etc. This data, while used by the newspapers/magazines for their publications, is also immensely useful to researchers and psephologists.
 - The Yellow Pages are important databases maintained by companies like TATA DONNELLEY Ltd.
- ③ Mail order companies like Otto Burlington maintain databases of customers who are accustomed to purchasing through mail order.

There are also market survey agencies, which may either work for a client compiling a specific type of data or run an actual business of collecting data. At times data collection can be very specialized such as:

- interacting with Road & Traffic Authorities to collect data on the number of two-wheeler and four-wheeler vehicles on the road;
- interacting with hospitals to procure data on HIV patients or other health records;
- interacting with the police to obtain crime statistics etc.

³⁶

Since some of these databases contain information that is collected covertly from customers and would therefore be in violation of "privacy" norms, statistics/figures on the volume of databases created and traded in by these companies is hard to come by.

(a) Specific Examples of Private Databases

Amongst the private database collectors, the prominent ones are highlighted as under:

(i) Centre for Monitoring the Indian Economy (CMIE), Bombay

CMIE has created perhaps India's first and only integrated database on the country's economic and business sectors. This organization has a team of researchers working around the clock to monitor the Indian economy and track all data releases, prominent events and news items.³⁷ CMIE research and databases are offered in the form of well defined services titled "business beacon" (provides reliable longtime series data on over 7000 economic indicators covering population and democracy, banking and interest rates, foreign trade, etc), "Industry Analysis Service," etc. CMIE also has extensive licensing arrangements with third parties to disseminate their databases.

(ii) Tata Energy Research Institute (TERI)³⁸ Publications

This organization based in New Delhi provides the following databases

– *Energy-environment database*: It provides ready-to-use information on different segments (energy and environment) of the Indian economy and some aspects of international economy. It provides time series and cross-sectional data on these subjects. The data is compiled by various research groups at TERI and is collected mostly from authentic government and recognized private publications. The exact ambit of the database is provided as under:

Scope	National
Type	Statistical/Numerical
Subject coverage	Coal, Power, Oil and Gas, Renewable, Forest, Agriculture, Industry, Transport, environment, Economic indicators
Coverage available from	1971 to 1997
No. of data tables	900
Interface	Web based
Updation	Once a year

– *Rural Energy Database*: This database has data compiled from more than 700 rural energy surveys, covering villages spread over all major states in India.

– *Periodical Abstracts*: Database covers references of journal articles, reports and conference papers on energy, environment, and sustainable development.

– *Environmental Information System (ENVIS)*: Databases were built from core journals and newsletters subscribed to by ENVIS Centre at TERI.

³⁷ See http://www.cmie.com/aboutus/cont_act.htm

³⁸ See <http://www.isro.org/anounce03.htm>—visited on January 7, 2002.

(b) Online databases

In the context of the Internet, there have now emerged several Internet databases dealing with a wide variety of data.³⁹

(c) Data Entry

Generally, data in the private sector is collected through street-by-street service or phone service and the cost for this may typically be Rs.5 per record. The data is then entered and this task is usually outsourced to Data Entry Operators. Finally, the data is formatted in a readily accessible manner and handed over to the client. In fact, some Indian data entry operators are so adept in their field that major clients like USA Telephone Directories commission them for data entry (India's IT capability in this area will be discussed more elaborately below). Unfortunately the data entry stage is often a source of leakage, as companies have not yet started to develop good confidentiality agreements with their Data Entry Operators or at least have not yet developed a culture of enforcing such agreements.

(d) Revenues

Since the Industry itself is not an organized one, revenue figures for the same have not been documented. Market and Research India (MRI), a private database organization situated at New Delhi was set up in December 2000 and had revenues touching approximately Rs.1 crore (approximately US\$200,000) by the end of the year.⁴⁰ This perhaps would give some general indication of the lucrative nature of the database industry in India.

As regards the sales of databases, the figures vary. For example while MRI charges about Rs.130/- (approximately US\$2.5) per day for every unit of data that it supplies, other suppliers are known to charge much lesser sums (sometimes even as low as 50 paise (approximately US\$.01) for a unit of database).⁴¹

(e) Need for Legal Protection

Most database companies are of the view that in order to succeed, they need powerful software, hi-tech computer systems and strong laws to protect against the theft of data. With India's strong IT capabilities, providing powerful software and high tech computers at affordable rates may not pose much of a problem; however the lack of effective legal protection may operate as a disincentive to invest in this industry. To give an illustration of the kind of costs involved in creating databases, consider the following:

³⁹ For a comprehensive listing of the prominent online databases, visit <http://www.internets.com/sindia.htm>. These databases are diverse in nature and deal with a variety of subject matter such as hostels (www.hostels.com/in.html), software training institutes (www.braintrustindia.com/directory_databases/IT_India/comptrg.htm), marriage (www.youthsindia.com), homeopathic (www.indiaspace.com/homoeopathy/list.htm), etc.

⁴⁰ Per statement of Mr. Pandey, Tarkeshwar–Sr. Manager, Strategy & Planning, MRI, India.

⁴¹ Id.

The National Foundation of Engineering and the National Information System for Science & Technology (NISSAT) are working towards creating a database of engineering faculty from over 600 colleges in India. This would include various details such as the name of the faculty, his/her respective college, experience, projects handled, etc. The rough estimate for creating this simple database would be approximately Rs.10 lacs (approximately US\$20,000), apart from the considerable time and effort involved.⁴²

C. INDIA'S INFORMATION TECHNOLOGY CAPABILITY

In the context of the digital revolution, databases and their protection have assumed tremendous significance. As has already been discussed earlier, the advent of PC's,⁴³ the Internet and a host of other factors have given the much needed zest to the database industry and provided it with tremendous opportunities, to not only create databases but also to market them effectively.⁴⁴ With more people using computers, the economic value of electronic databases has increased manifold.

The digital revolution has thrown up several opportunities in the database industry for India. India's information technology (IT) capability should be leveraged to create and trade in a wide variety of commercially valuable databases. India's IT capability can be gleaned from the figures⁴⁵ below:

1. In 1999-2000, the Indian IT industry was estimated to have earned revenues of Rs.37,080 crore or US\$8.67 billion—a growth of almost 50% as compared to Rs.24,781 crore (US\$6.05 billion) in 1998-99.
2. In the last five years (1995-2000), the Indian IT Industry has recorded a C.A.G.R. (Compounded Annual Growth Rate) of more than 42.4%, which is almost double the growth rate of IT industries in many of the developed countries.
3. In India, IT spending as a percentage of the GDP is currently about 1.68%. In the USA, however, the IT spending as a percentage of GDP is more than 6%. With Government of India's resolve to increase IT spending, it is forecasted that by 2008, IT spending in India could touch 3% of its GDP.
4. The IT manufacturing sector is growing at an average rate of 30 to 35% annually over the past decade. And the industry has over 150 major hardware players supported by over

⁴² Per statement of B.J. Sunder Singh, National Information System for Science & Technology (NISSAT), DSIR, New Delhi.

⁴³ The introduction of commercial computers resulted in the development and use of computer databases, commonly referred to as electronic databases. The use of electronic databases allows information to be accessed, manipulated and used much faster than in paper based industries.

⁴⁴ The advent of the Internet, has enabled access to information residing in remote servers and consequently the creation of a universal market for this information.

⁴⁵ http://www.nasscom.org/it_industry/indic_statistics.asp visited on January 8, 2002. NASSCOM (National Association of Software and Service Companies) is not only a chamber of commerce; it is a single point reference on any information on IT software and services industry in India.

800 ancillary units and small time vendors engaged in sub-assemblies and equipment manufacturing.

5. Software continues to contribute a major portion of the Indian IT industry's revenues. During 1999-2000, the software industry's revenues constituted over 65% of the Indian IT industry's annual revenues.

6. More than 185 of the *Fortune 500* companies outsourced their software requirements to Indian software houses.

7. During 1999-2000, more than one million PCs were sold in India. This took the PC penetration in India to 4.3 PCs per 1000 people by the end of 1999-2000 (March 31, 2000). However, since then the penetration rate has gone up to 5 PCs per 1,000 people (as of December 31, 2000).

8. The Internet revolution took off in a big way—with the number of subscribers increasing to 9 lakh (0.9 million) by March 2000. This translates to almost 28 lakh (2.8 million) Internet users in the country. However, by December 31, 2000, the Internet subscribers went up to 18 lakh (55 lakh users).

9. More than 460 licenses were issued to private Internet Service Providers (ISPs) during the year and massive price reductions were announced by ISPs to combat the excessive competition. International gateway by private ISPs also became operational.

Amongst other things, these statistics also show the increasing levels of computer usage and Internet penetration in the country. No doubt, this might seem low when compared with the developed nations. However, amongst the developing nations, this is a substantial figure and augurs well for the database industry in India.

(a) Data and Content Management

A study undertaken by NASSCOM and McKinsey⁴⁶ (hereinafter referred to as the NASSCOM McKinsey report) showed that IT enabled services could yield revenues worth US\$17 billion in the year 2008. A key component of IT enabled services is “data management.” In the age of scientific and technological revolution, we face the problem of being overwhelmed by increasing volumes of knowledge and information, which flows from different disciplines.⁴⁷ Hence, it becomes critical to develop solutions that would help categorize and make sense of this data. India's IT skills in data management⁴⁸ would go a long way in bolstering the database industry.

⁴⁶ See “The Indian IT strategy,” 17 December 1999, New Delhi.

⁴⁷ The NASSCOM McKinsey study had quoted the Baldrige survey, which stated that “knowledge management” has emerged as one of the topmost CEO priorities. A key component of knowledge management is the creation of knowledge repositories such as research reports and databases.

⁴⁸ A number of Indian companies are engaged in data-entry, data-mining and data-warehousing operations. Even the IT majors in India such as WIPRO and SATYAM are actively involved in these activities.

Of course, one may argue that with India's focus on IT outsourcing, the data or the databases would continue to belong to the clients (most of whom are from the USA/European) and therefore India will have no say in the commercialization of these database. Nevertheless, a strong protection regime for databases would encourage foreign companies to supply more data while engaging the services of Indian companies.

The IT action plan of the IT Task Force had dealt with content creation and management as a separate subhead in its long term, National IT policy, where it described the "content industry" as follows:

"The Content industry has come to mean various commercial and non-commercial activities relating to the bibliographic, textual and statistical databases as well as the information, education and entertainment materials in the electronic form including audio, video and multimedia forms."

From the above definition, it is clear that the government is looking to now encourage the database industry. The action plan went on to state that the Government of India would give priority to the promotion of the content industry (which as defined earlier would include databases). Some of the key points mentioned in this policy are extracted as under:

- *Considering the vast global and domestic market, the content industry has a potential to grow to a size comparable to the computer software industry with enormous opportunities for income and employment generation. Content development being an integral part of any information system, the content industry is recognized as a priority sector for lending and according the same benefits for purposes of promotion and development as given to the software sector.*
- *It will be maximally ensured that any information or report collected by the government will be in electronic form at the entry point itself. Similarly, all the information to be made available to the public by the government, such as budget documents, customs and excise rules, railway time tables, telephone directory, maps including public domain digitized maps, etc. will be made available in the electronic form.*
- *India is known for its rich and diverse cultural heritage. It also possesses a vast wealth of traditional knowledge. These are mostly in Indian languages and should be promoted and preserved for posterity. The Government will, therefore, take initiatives, through appropriate projects, to create electronic images of the information on the Indian arts and culture, for wider dissemination and research.*
- *An effective Copy Right protection system is a prerequisite for development of creative works in the electronic medium. Therefore, the Indian copyright law will be strengthened in this direction. Further, there is a need for global harmonization of copyright laws. The conclusion of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement) and of the two World Intellectual Property Organization (WIPO) treaties, will be adopted for such harmonization.*

- *There is a need to promote and encourage hosting of non-commercial materials related to linguistic, social and cultural aspects of the people by the public or private organizations. The Government will take initiative for providing Websites, free of cost, for such purposes.*
- *Government will encourage and promote Indian companies and organizations to host their contents only on Web servers located in India with Indian domain addresses and these will be made available at internationally competitive prices. Any information hosted on these sites will follow the guidelines evolved by the Government in association with the Industry.*
- *For enabling Inter-operability between equipment, data, practices and procedures, Standards will be evolved to integrate hardware, software and communication systems and to exchange information across boundaries of different systems.*
- *Creation of knowledge bases requires trained manpower for collection, compilation, analysis and production of value added information products and services. Specialized training programmes, through existing institutions, will be initiated to meet the requirement of trained professionals in these areas. Traditional curriculum being offered by the universities and educational institutions in various fields related to content industry will be suitably modified, such as library science, journalism and mass communication.*
- *Multimedia and Internet are the foundations of the future content industry. Creation of multimedia products require specialists in multimedia designing, editing, programming, etc. Training programmes for these core professionals in multimedia, will be initiated in Engineering Colleges, Polytechnics and other concerned institutions.*

(b) Projections for the Indian Database Industry:

It is important to note that the government, in its strategic policy for the IT industry, included “database services” within the ambit of IT-enabled services and projected that with adequate incentives being given to this sector, this sector could yield revenues of about 6500 crores by 2008 and employ approximately 100,000 people.⁴⁹ This clearly shows that the industry has tremendous potential.⁵⁰

⁴⁹ see <http://it-taskforce.nic.in/vsit-taskforce/bbr3/chap1.htm>–visited on January 9, 2002.

⁵⁰ NISSAT, a government agency is involved in compiling an inventory of indigenous database efforts. This effort titled INDAB (Indigenous database) had in 1997 built a database containing 398 records providing information on 202, indexing/abstracting services and 113 indigenous databases. The analysis of this database index shows that though the major contributors of databases are government departments and R & D units, the participants from the private sectors are steadily increasing. The figures must have increased considerably since then. See B.G. Sunder Singh, National Information System for Science & Technology, DSIR, New Delhi, “*Indigenous Databases and Information Products/Services*,” See <http://itt.nissat.tripod.com/itt9702/infoprod.htm>.

II. LEGAL RESPONSE: INTERNATIONAL

The international legal community is on the issue of the quantum and quality of protection that should be accorded to the database industry. The European Union (EU) with its aim to bolster its database industry, which had been lagging behind its counter-part in the United States of America (USA) has assumed a protectionist stance with its Directive.⁵¹ In order to provide sufficient incentives to investors in the database industry, the European Commission (EC) has put into place a two-pronged protection for databases.

Firstly the EC with its Directive ironed out the differing standards of copyright protection available in the Member countries. The threshold for copyright protection has been uniformly set at the “author’s own intellectual creation.” However it has been made clear that copyright protection for databases provided for by this Directive shall not extend to their contents and shall be without prejudice to any rights subsisting in those contents themselves.

The more far-reaching protection, which the Directive grants to databases, was the creation of a *sui generis right* of the database owner in the contents of the database. Thus the database owner was effectively granted an exclusive property right in the contents of the database to prevent unauthorized extraction and reutilization of the whole or substantial part of the contents of the database. Interestingly, the accrual of a *sui generis* right to the database maker does not require the showing of a creative achievement or a novel contribution to the prior art; rather it accrues on merely showing that there has been qualitatively and/or quantitatively a substantial investment in the obtaining, verification or presentation of the contents.

The rationale behind this right is set out in the Recitals to the Directive, which states that the Directive seeks to safeguard the position of the database makers against misappropriation of the results of the financial and professional investment made in obtaining and collection of the contents. On this issue the Directive has come in for a lot of criticism and commentators have gone to the extent of saying that the fear of market failure and of chronic under protection that initially motivated the quest for a *sui generis* regime to protect electronic databases has thus given way to the creation of “mini-monopolies over information.”⁵²

EC Directive has also been criticized for a number of other reasons. It is argued that the Directive in its zeal to stimulate the database industry has failed to uphold the fair dealing and public good exceptions. Though the Directive strives to establish rights and obligations for the lawful users, however they are couched in such terms that the obligations certainly outweigh the rights that might accrue to the lawful user. Thus it is contended that under the EC Directive, the most borderline and suspect of all the objects of protection ever to enter the universe of intellectual property discourse—raw data, scientific or otherwise—paradoxically

⁵¹ Directive 96/9/EC of the European Parliament and of the Council of March 11, 1996, on the legal protection of databases

⁵² Debra B. Rosler, The European Union’s Proposed Directive for the Legal Protection of Databases: A New Threat to the Free Flow of Information, 10 High Tech. L. J. 105 at 138

obtains the strongest scope of protection available from any intellectual property regime except, perhaps, for the classical patent paradigm.⁵³

It is also contended that the Directive has effectively vested with the database maker an exclusive property right for “a virtually unlimited duration.” Though the term of the *sui generis* right in a database has been fixed at fifteen years, any substantial change in the database would qualify it for its own term of protection. Such substantial change has to be evaluated qualitatively or quantitatively in terms of the successive additions, deletions or alterations, which would result in the database being considered to be a substantial new investment. Since there are no limits to the number of quantitative or qualitative changes that will qualify for such extensions, and any publisher who continues to make a substantial investment in updating, improving, or expanding an existing database can look forward to perpetual protection thus effectively abolishing the public domain.

In spite of all the criticisms, the Directive with its reciprocity provisions has triggered off strong demands from the database industry across the world for similar protective regimes. The Directive denies protection to databases created outside the EU unless the country of origin offers “comparable protection” to databases created by the EU nationals or residents and the Council agrees to extend protection.⁵⁴ The non-EU database industry fears that the failure to reciprocate the protection afforded by the EU would render their databases open for wholesale piracy in Europe. Responding to the demand of the database industry, there have already been legislative attempts in the United States of America to extend protection to database owners.

In fact the need to enact a law in order to provide protection to compilations of information was felt much before the EC Directive, when the United States’ Supreme Court in *Feist v. Rural Telephone*, overruled the ‘*sweat of the brow*’ doctrine for copyright protection of factual compilations. The Court held that the criterion of originality is a Constitutional mandate and though not stringent, has to be satisfied.

Thus it was decided that there is no copyright in Rural’s white pages holding that the end product is a garden-variety white pages directory, devoid of even the slightest trace of creativity. The Court raising the threshold of originality to more than a *de minimis quantum* of creativity held that though Rural’s selection of listings is selection of a sort, but it lacks the modicum of creativity necessary to transform mere selection into a copyrightable expression.

Following *Feist*, the database industry in the United States of America (USA) was already lobbying for an effective database protection regime, which received a further impetus with the passage of the EC Database Directive.⁵⁵ Though there is a consensus amongst the concerned parties that there is a need for adequate protection for the database industry, the

⁵³ J H Reichman, Charting the Collapse of the Patent-Copyright Dichotomy: Premises for a Restructured International Intellectual Property System, 13 *Cardozo Arts & Enter. L. J.* 480 at 488-489

⁵⁴ Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases—Recital 56: third country protection must be “comparable;” Art 11(3): the determination on comparable protection and agreement to extend protection is made by the European Council (on proposal from the Commission)

⁵⁵ Amanda Perkins, “*United States Still No Closer to Database Legislation*” (2000) *EIPR* 366

form of protection has been a contentious issue. The congressional bill H.R. 3531⁵⁶ with an intent to “promote investment and prevent intellectual property piracy with respect to databases” proposed a new hybrid *sui generis* right distinct from and additional to copyright. However, following stiff opposition from the academic and scientific community against the protectionist stance of the bill, the same was not even put to vote.

Currently there are two versions for the proposed database legislations in the USA. The H.R. 354 titled “Collections of Information Anti-Piracy Act” and H.R. 1858 titled “The Consumer and Investor Access to Information Act of 1999.” A perusal of the proposed legislations clearly reflects the divergent approaches prevailing in the USA vis-à-vis database protection. Since it is certain that any grant of *sui generis* right to the database owners would be unconstitutional, the USA models for database protection are based on the law of unfair competition.⁵⁷

The scope of liability under the Antipiracy Bill is undoubtedly narrower than the EC Directive as it renders making available of all or a substantial part of a ‘Collection of information,’ unauthorized only if it causes material harm to the original database maker’s primary or relative market. Similarly, extraction is also prohibited if it causes material harm to the primary market. However the Antipiracy Act has been criticized for not being a true unfair competition model as the intent is to protect investments through the creation of an indirect property right and not to forbid unfair conduct.

The Access Bill on the other hand, is more of a true unfair competition model as it focuses on unfair conduct, that is, it forbids the duplication another’s database and selling and distributing the same in commerce in competition with the original database. This prohibition is no doubt narrower than the Antipiracy bill much to the delight of the research and academic formalities.

Furthermore, the Antipiracy bill contains a general fair use permission which permits making available or extracting in both commercial and non-profit contexts, if it is reasonable and for delineated purposes such as comment, criticism, teaching, research, etc. The bill also contains a specific exemption for non-profit educational, scientific or research purposes only if their activity does not harm the database maker’s primary market. The Access Bill on the other hand creates full immunity for scientific, educational or research purposes whether for profit or non-profit. Thus even substantial market harm to the database maker would be permitted so long as the conduct is not for the purpose of competition with the database maker.

⁵⁶ Database Investment and Intellectual Property Antipiracy Act of 1996 was introduced by Representative Carlos Moorhead on May 23, 1996.

⁵⁷ Supra note 5

III. INDIA: EXISTING LEGAL REMEDIES

The forms of protection available in India to the database providers are enumerated hereunder:

- Copyright protection
- Law of contracts
- Information Technology Act, 2000
- Technological measures for electronic databases

(a) Copyright Protection

The Indian Copyright Act 1957 provides an inclusive definition of “literary works” so as “to include computer programmes, tables and compilations including computer databases.”⁵⁸ The Indian copyright law thus specifically recognizes copyright protection for compilations and computer databases as literary works. Copyright has been defined in the Indian law to include several exclusive rights vested with the owner of the work, namely:

- right of reproduction in any material form including the storage on any medium by electronic means;
- right of publication;
- right of communication to the public or public performance;
- right to make a film or sound recording;
- right to make translations or adaptations.

Under the Indian law, the author in relation to a database would either be the creator or if the same is computer-generated then under section 2(d)(vi), the person who causes the work to be created. This would mean the person who takes the initiative and responsibility for creating the database.

It is pertinent to mention that India itself being a commonwealth country follows the “sweat of the brow” doctrine. The Indian courts have therefore protected compilations involving minimal originality stating that “no man is entitled to steal or appropriate for himself the result of another’s brain, skill or labor even in such works.”⁵⁹ This rationale was followed in several cases and it was held that a compilation of addresses developed by any one by devoting time, money, labor and skill though the source may be commonly situated amounts to a ‘literary work’ wherein the author has a copyright.⁶⁰

In the first ever instance of copying of electronic databases, the Delhi High Court has been instrumental in injuncting an Italian infringer who had copied the plaintiffs online herbal database onto its Website, hosted by an American ISP. Pursuant to the Delhi High court restraining order, which was brought to the notice of the concerned ISP located in the USA,

⁵⁸ Section 2(o) of the Copyright Act, 1957.

⁵⁹ Govindan v Gopalakrishna AIR 1955 Madras 391

⁶⁰ Burlington Home Shopping Pvt. Ltd. v Rajnish Chibber & Anr. 1995 PTC (15) 278

the ISP removed the infringing content on its own accord and furnished the complete details of the infringer, who had rented space on the ISP's Website.⁶¹

Despite the strong protection that had hitherto been offered to databases by the Indian judiciary, in a recent case involving law reports,⁶² the Delhi High Court preferred to follow the Feist standard of 'a modicum of creativity' and denied protection to copy-edited judgments. The court held that changes consisting of changes of spelling, addition of quotations and corrections of typographical mistakes are trivial and no copyright exists therein.

The plaintiffs had also pleaded that the entire law report is a compilation or a database and entitled to copyright protection because it involved selection, collation, arrangement and coordination of various judgments which requires tremendous input of skill, labor, discretion, judgment and expenditure. Though the court did not give a specific finding on this point, but the eventual refusal to grant an injunction against the defendants suggests that the court disapproved the Plaintiffs' assertions.

Therefore it seems that the Delhi High Court has raised the threshold of the requirement of originality from that of skill, labor and judgment to a minimum level of creativity, which most databases would not be able to satisfy. Thus, there is bound to be a strong argument that since the law of copyrights has failed to protect the collections, there should be an alternate stronger protection. In fact it has been acknowledged by the IT Taskforce in its proposed National IT Policy⁶³ that effective copyright protection system is a pre-requisite for development of creative works in the electronic medium and thus there is a need to strengthen the Indian copyright law in this direction. It has been further suggested that there is a need for global harmonization of copyright laws through the adoption of the TRIPS Agreement and the two WIPO Treaties.

(b) Contractual Protection

The law of contracts complements copyright law in so far as database protection is concerned. With the advent of online databases, it is possible to make the end-user agree to the contractual obligations as deemed fit by the database provider. Thus, if the end-user infringes any of the obligations stipulated in the contract, it could amount to breach of contract under The Indian Contract Act, 1872.

(c) Information Technology Act

"Electronic databases" are also protected by the newly enacted Information Technology Act, 2000. Section 43 of the Act reads as:

⁶¹ Himalaya Drug Company Vs. Sumit, Suit No.1719 of 2000.

⁶² Eastern Book Company & Ors. v Navin J. Desai & Anr.

⁶³ IT Action Plan (Part III): Long Term National IT Policy proposed by the National Taskforce on Information technology and Software Development available at <http://it-taskforce.nic.in/vsit-taskforce/actplan3/chap5.htm>

“43. *Penalty for damage to computer, computer system, etc.*—If any person without the permission of the owner or any other person who is in charge of a computer, computer system or computer network,—

(a) ...

(b) downloads, copies or extracts any data, computer database or information from such computer, computer system or computer network including information or data held or stored in any removable storage medium;

he shall be liable to pay damages by way of compensation not exceeding one crore rupees (approximately US\$20,000) to the person so affected.”

“Computer databases” have been given a wide definition⁶⁴ and means “representations of information, knowledge, facts, concepts or instructions in text, image, audio, video that are being prepared or have been prepared in a formalized manner or have been produced by a computer, computer system or computer network and are intended for use in a computer, computer system or computer network.” On similar lines, data has also been defined⁶⁵ to mean “a representation of information, knowledge, facts, concepts or instructions which are being prepared or have been prepared in a formalized manner, and is intended to be processed or has been processed in a computer system or computer network, and may be in any form (including computer printouts magnetic or optical storage media, punched cards, punched tapes) or stored internally in the memory of the computer.” Though the provision has not been subjected to judicial interpretation, it can be observed that the protection extended to electronic databases is a broad one.

It is pertinent to mention that not only is there no requirement for the databases to be original but also the IT Act does not contain any ‘fair dealing’ exceptions to infringement as provided in the Copyright Act 1957. Furthermore there is no requirement that at least a substantial amount of the work should be downloaded, copied or extracted. Thus, even insubstantial amounts of information is protected by the IT Act. In fact “mere data” has also been extended protection against unauthorized downloading, copying or extraction.⁶⁶

(d) Protection of Technological Measures

Lastly, the digital revolution has provided in the hands of the database providers various tools to control the access and free-riding appropriations from their databases, such as encryption devices, etc. In this context it is pertinent to mention that the WIPO Copyright Treaty 1996 obligates the members to provide adequate legal protection and effective legal remedies against the circumvention of effective technological measures that are used by authors in connection with the exercise of their rights under that Treaty or under the Berne Convention and that restrict acts, in respect of their works, which are not authorised by the authors concerned or permitted by law.

⁶⁴ Explanation (ii) to Section 43 of the Information Technology Act 2000.

⁶⁵ Section 2 (1) (o) of the Information Technology Act 2000.

⁶⁶ It is pertinent to note that the definition of data is wide enough to include any data that is in electronic form.

Although the Copyright Act 1957 in its present form does not provide legal remedies against the circumvention of technological measures, the impending amendment in the law is expected to incorporate such a provision. In fact the Drafting Committee set up by the Core Group on Amendments to the Copyright Act, 1957 has in the draft amendments incorporated a section⁶⁷ for the “Protection of Technological Measures.” Therefore, if and when the Copyright Act is amended to include the said provision, the protection of databases through technological measures would also be enhanced. In fact, database owners are already deploying various technological measures including encryption watermarks, etc. For example NISCOM (National Institute of Science and Commerce) plan to deploy these measures while making its TKDL Website accessible.

(e) Unfair Competition

The law of unfair competition or misappropriation is also proposed as an effective remedy for database protection. This is one method of the protection of databases, which has the backing and support of the scientific and academic community. In fact there are pending legislations in the United States of America for the protection of databases based on the law of unfair competition.

However, the Indian courts have not yet recognized the residuary right to unfair competition. However the Monopolies and Restrictive Trade Practices Act, the Consumer Protection Act and the various Intellectual Property statutes have inherent statutory provisions based on the law of unfair competition.

From the above, it can be seen that the Indian legal regime is not fully adequate to cater to the emerging database industry; we therefore need to think in terms of a *sui generis* regime. However the exact scope and modalities of this regime will need to be worked out extensively and is not within the scope of this paper.

⁶⁷ Section 65A of the Draft Amendments recommended by the Drafting Committee in its meeting held on July 29, 2001, which reads as “(1) Any person who procures or uses any device with the intention of circumventing an effective technological measure applied for the purpose of protecting any of the rights conferred by this Act, or possesses such a device for the said purpose shall be punishable with imprisonment which may extend to two years and shall also be liable to fine.”

CONCLUSION

The economic justification for the Intellectual and Industrial Property Rights (IPRs) have from their very inception revolved around the “incentive” argument. It has been contended by the proponents of IPRs that the creation of inventions and various works, which are protected by these rights calls for protection against infringement; otherwise the incentive to invest time and other resources in the creation of such products is lost. The “incentive” argument has resulted in the evolution of various new rights to help protect various new works.

Indeed, from a system that had evolved to protect the creative works of authors, inventors and artists, the Intellectual property system has now come to embody a system that is geared towards protecting investments, irrespective of their literary or inventive merit. In today’s context, economic imperatives mandate that “information goods” be protected adequately.

The only drawback of extensive protection would be the problem of access to these works by the academic/scientific community.⁶⁸ This problem of potential non-access is heightened in the case of developing countries (such as India), where the psyche of the social and economic thinkers has always been against the grant of intellectual property rights.

The “incentive argument” in these countries does not augur well either with the scientific and academic communities or with the general populace. One of the reasons for this dislike of the IPR system is that these are considered as alien concepts imposed by the capitalist economies. The problem is aggravated in the context of “database” rights, since the traditional argument that stronger Intellectual Property Regimes would facilitate the transfer of technology to the developing countries does not hold good in the case of the “database” industry, where the key subject matter is “data.”

Like all other developing nations, India has also had a history of continuously opposing any increase in the scope of intellectual property protection. Although the advent of the TRIPS Agreement had brought about many changes in the Indian psyche towards Intellectual Property Rights, the recognition for new forms of intellectual property rights still remains a very contentious issue.

Over the years, India has felt the need for new forms of protection. Excellent examples where India is proactively working towards increased protection would be areas such as traditional knowledge⁶⁹ and geographical indications.⁷⁰ Thus, where it is felt that the interests of the domestic industry need to be protected new forms of protections are generally granted.

There is no gainsaying the fact that the database industry in India is a rapidly growing one. In fact, specific projects recognizing the need for database creation are now emerging in

⁶⁸ This argument has been substantially elaborated upon in the chapter titled “Issues”

⁶⁹ In fact, a first draft of a proposed legislation in this area has already been completed. In fact, in the Draft Science Policy Document (Science and Technology Policy–October 29, 2001), it was stated that IPR systems that protect innovations raising out of traditional knowledge generation will be evolved.

⁷⁰ The Geographical Indications Act has already been passed and is yet to be notified.

India. An excellent example of this would be a project initiated by SENET.⁷¹ Launched in 1997, this project would aim at pioneering and creating and promoting databases for the small enterprise sector. SENET⁷² intends to use computer technology to create/help create a series of electronic databases.

In order to attract more investments to this field, it is critical that India provide for an adequate protection regime. As has been seen in the preceding chapter dealing with the Indian database industry, India's database industry could be considerably leveraged through its IT capabilities and this perhaps would be Indian key strength in the years of come. In the words of the renowned scientist, Dr. Mashelkar, India is "ancient, diverse and mathematical."⁷³ The "ancient" quality of India translates into an immense wealth of "traditional knowledge." The "diverse" quality translates into a varied gene/blood pool and consequently a wealth of genomic data. The "mathematical" quality translates into India's IT capability," which would then be used to leverage the database industry in general.

The problem of potential non-access could be solved by carving out specific exceptions in favor of the academic and research community. The argument that government data cannot be commercialized does not hold much ground in the light of the fact that the Indian government has now started commercializing its data, especially in areas where it is technologically advanced, such as remote sensing. More importantly, with the digital era churning out tons of data today, one should structure the regime in such a way that mere data is not protected; rather only databases that involve the expenditure of substantial investment/resources should be protected. One also needs to clearly define what constitutes a "database." Of course, whilst providing for adequate database protection and thereby providing incentives to the industry, one should at the same time ensure that data collection is not done at the cost of "privacy" and other rights of individuals. This aspect is however outside the ambit of this study-suffice it to say that India lacks an adequate privacy legislation and therefore it is now imperative that we provide for such legislation.

[End of document]

⁷¹ This project is in specific recognition of the fact that the database industry in India requires both promotion and investment.

⁷² See <http://www.smallindustryindia.com/schemes/senet.htm>

⁷³ Per statement of Dr. Mashelkar, Director General, CSIR. India