

Ya Shon (Christine) Huang

The Law and Economics of Databases: A Balancing Act

Ya Shon (Christine) Huang

Abstract

In this paper, I demonstrate that the existing legal frameworks for database protection are inadequate – the American framework under-protects databases, while the European framework over-protects. This paper presents an economic analysis of the current scope of legal protections for databases versus the ideal, with an especial emphasis on the role of intellectual property rights in providing these protections, and concludes with proposals for an ideal system. After an overview of the current systems of legal protections for databases in the United States (US) and the European Union (EU), there will be an explanation of how different types of laws (competition, contract, and most importantly, intellectual property rights) impact the production and innovation of databases. The analysis will show that intellectual property rights are the most comprehensive and efficient form of legal protection due to its ability to limit transaction costs, provide adequate incentives for production, and maintain a reasonable barrier to entry. The scope and specificity of legal protections affects both the static and dynamic efficiency of markets, impacting not just the health of market functions, but also the growth and development of innovation. The proper scope of protection and whether intellectual property rights are the optimal source of protection depends on the economic nature of databases. If databases were to be treated as a commodity by the law, it is necessary for the law to accurately reflect the type of good that it is regulating, as different types of goods require different incentive structures. The paper will conclude with recommendations for an ideal legal system for the protection and regulation of databases, starting with accurately defining databases and ending with reasonable terms for copyrights. The solution is to find the balance between the US and the EU systems in terms of issue definition, economic incentives, and legal theory.

Introduction

With the rise of “data driven decision making” and the exponential growth of information goods (of which databases are a subset), databases are becoming increasingly valuable resources that will enable innovations across markets and borders. Because databases are becoming key components of increasingly complex decision-making processes, the policies and regulations of data also will rise in importance due to their ability to enable or limit the usage of data on individual or institutional levels. The legal protection of databases comes from various sources, including contracts, licensure, and antitrust, but the area of most significant interest is intellectual property (IP), specifically that of copyright.

This paper will be examining the efficiency of the current US and EU legal models for IP rights, and the extent to which databases can or should be protected. The robustness of copyright systems will be evaluated by weighing how well the system maximizes the utility individuals derive from owning IP rights against how well the system maximizes

the utility society derives from allowing others to own IP rights. The study of this topic will be broken down into three parts. The first part will consider how databases are currently protected, the legal purposes for protecting databases, and how those protections shape the distribution of rights and powers. The second part will discuss the economic purposes of protecting databases, and how those protections shape the distribution of wealth. The third part will explore an alternative to the existing legal systems, and explain why a more middle of the ground approach is preferable. This will be done by comparing legal and economics foundations of the US and EU legal systems, and discussing the effects of each.

The analysis in this paper must necessarily combine legal and economic theory, as it is the interaction between the two that should guide the regulation of databases. Databases are a “good” (in the economic sense) whose properties have changed dynamically in recent times. Therefore, the legal remedies for their protection should reflect these changes. This study of databases begins with clarifying how current legal systems address the legal protection of databases, discussing the purposes of such laws, and how they differ between the two of the most prominent legal systems: that of the United States, and of the European Union. Following the contrasts between the legal two systems will be a discussion of the economic dynamics that these systems create. The paper will seek to show that these two systems are at extreme ends of the spectrum (under- and over-protection), and that a modified system in the middle ground would be more economically efficient.

Legal Protection of Databases Today: Two Extremes

The general definition of databases for both the US and the EU usually fall under the category of compilations, with compilations defined as a literary production composed of the works of others, arranged according to a certain method. In other words, collectives or collections of materials, including but not limited to facts, data, or works of art/ literature. The nuance within this definition is that there is an important distinction between the copyright of the actual data, and also the copyright of how the database is assembled, or ‘compiled’. Additionally, there is a distinction between the US and EU standards with regards to what kinds of works merit copyright protection. Where the US primarily recognizes the creativity (but not all intellectual activity) that goes into the creation of the database as the threshold, the EU recognizes both the creativity and investment that goes into creating either the data or the databases, hence the additional ‘sweat of the brow’ investment doctrine that the EU implements in its assessment of database copyrights. Under the EU’s investment doctrine, copyrights can be awarded on the basis of substantial investment alone.

In order to compare the condition of having database protection versus the condition of not having database protection, it is important to highlight the differences between how copyright has evolved under the United States model as opposed to the European Union model. This is because the United States has no formal, specific copyright protections for databases beyond conforming to the Berne Convention of 1886, which only grants ‘thin’ copyright protection to databases due to the difficulty of classifying databases within the current US intellectual property regime. The landmark case for the United States is the *Feist vs. Rural* Supreme Court case, where the court found no infringement of

copyright on the basis that the underlying information that made up the database were not, in themselves, protected by copyright.¹ Following the requirement of minimal creativity by authors, the method of compilation utilized by the plaintiff (Rural) was not sufficiently original to warrant copyright protection, which resulted in the court ruling non-infringement for the defendant. Minimal creativity simply means any sort of intellectual effort to somehow transform or, in this case, re-arrange the materials. The plaintiff did not use an unusual or specific arrangement in their compilation of uncopyrightable facts (people's names, addresses, and phone numbers), and so did not receive copyright protection for their phonebook. The European Union avoids this type of distinction by having a *sui generis* (of its own category) right for database protection after passing the Database Directive in 1996.

As explained in previous sections, intellectual property law is at the frontier of ensuring the health of the database industry, and also an effective way of ensuring a balance between innovation and access. Within the domain of intellectual property, copyright is the main source of legal protection for databases. Logistically, while databases are products that exhibit traits of both expressive and technical works, they are not rigorously technical enough to claim patent protection, and thus have a better claim to copyright protection. However, the manners in which databases are compiled, curated, or organized are expressive of original thought and creativity, while the creation of databases is indicative of a technically innovative nature, especially in modern methods of creating and maintaining databases. Additionally, the curation, compilation, and organization of databases are increasingly done through codes in a variety of programming languages, which further complicates the question of how wide the scope of legal protection for databases should extend.

The discrepancies in copyright law across national boundaries are also a challenge, as there are many different systems but no internationally unifying system. While there is not currently a common 'international copyright', there are some loose agreements between nations to govern copyrights across borders. As such, though copyright standards still deviate across countries and regions, there does exist some minimum standards that most countries have agreed to adhere by, such as not granting copyright indefinitely, or reciprocating protection across borders. Among the most prominent of these agreements are the Universal Copyright Convention, the Berne Convention, and membership in the United Nations-affiliated but functionally independent global forum, the World Intellectual Property Organization (WIPO).

In terms of such international agreements, the Universal Copyright Convention was adopted in 1952, while the Berne Convention was adopted before it, in 1886. The Berne convention requires the application of each country's copyright protections to works published within said country.² This was the cause for disputes, ranging from particulars about the duration of protection to more serious, philosophical disagreements about having limited, non-renewable terms at all (as opposed to allowing perpetual renewal). These philosophical disagreements would go on to shape the ways that copyright develops in the US, the EU, and Commonwealth systems. The Universal Copyright Convention was designed

1 Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340 (1991).

2 Berne Convention Implementation Act of 1988, Pub. L. No. 100-568, 1988.

to overcome some of these disagreements, while one of the most prominent dissenting members was the United States.³ Among these conflicts, the United States and other systems to it granted protection for fixed, renewable terms (as opposed to the non-renewable term under the Berne Convention), and required registration with the Copyright Office (forbidden under the Berne Convention). Though these two different types of international agreements have since started converging towards a middle ground, there are still too many gaps between these systems of copyright laws for implementation on a truly global scale. The WIPO plays a role in the process of reconciling these differences. Though it does not create international legislation, it serves as an international platform through which minimum standards for international intellectual property can be debated, and different systems of intellectual property legislation researched.

For the purposes of this paper, the copyright laws of the US will be contrasted to those of the EU, since they represent not only the two most dominant systems of copyright laws, but also two systems derived from different philosophical and theoretical foundations.

Although other legal protections for databases like contracts or competition (anti-trust in U.S. law) do exist, copyright is still the most dominant form of protection because it confers property rights to the owners of databases. This is because a property right like copyright gives the owner the powers to exclude a priori, instead of the power to seek damages a posteriori. As copyright law currently stands, the United States and the European Union have two different models of copyright by which databases are influenced. The most salient point of differentiation between these two models is the underlying philosophical justifications upon which the two models are built. The United States model has primarily Utilitarian foundations (greatest good for the greatest number), while the European Union model⁴ combines the Utilitarian and Personal Rights justifications (personal rights exceed overall societal gain in importance). These differences in the scope of protection for intellectual properties (and subsequently, databases), directly impact the economic efficiency of these models. Under-protection can lead to market failure by discouraging newcomer entry into the market, while over-protection can lead to market failure by creating artificially high barriers that block newcomers completely.

In the United States, copyright protection is only provided for compilations (like databases) if there is significant creativity in the selection or arrangement, since copyright in the United States is meant to protect the expression of ideas, rather than ideas or information themselves. The model makes sense if it only regulates the universe of literary and expressive works since they do behave like public goods once released in the market. To extend the exact same regulations to information commodities does not make sense. Current copyright protections for databases are very thin and unreliable, since databases usually fail the ‘modicum of originality’ requirement or the original authorship requirement. Databases are usually protected through alternative measures such as licensure, contracts, DRM programs, etc.

3 Universal Copyright Convention, 1952.

4 The European Union is also sometimes referred to as the European Economic Community in discussions of copyright. However, for the sake of simplicity this paper will use the terms “US model” and “EU model” to denote the different legal systems.

Under this model, databases are treated as compilations, which means that there is no protection for the actual data or information that make up the database, unless the data or information itself is also copyrightable. The landmark case for this distinction is *Feist Publications, Inc. v. Rural Telephone Service Co*, where the court ruled that the manner in which Rural's phonebook was compiled did not satisfy the 'modicum of creativity' threshold to qualify as a work meriting intellectual property protection.⁵ This ruling essentially limits the scope of intellectual property protection for databases, and prevents only the taking of selection or arrangement of data. After the *Feist v. Rural* decision, compilations (of which databases are an example) do not qualify for copyright protection. The most significant impact of this ruling is that compilations are copyrightable only with respect to the selection, coordination, or arrangement of material. Therefore, in order for a work to qualify, it would need to demonstrate sufficient creativity in their "selection, coordination or arrangement" to render them "original works of authorship" entitled to copyright protection. This stems from the interesting characteristic of US law where facts are not copyrightable. As databases are usually compilations of facts, it is very rare for databases to qualify for copyright.

The problems with the US copyright regime lies with the discrepancy between the actual nature of databases as a good (how it behaves as a good in markets), and the type of good that this regime is meant to address. In fact, the current US copyright regime is poised to disincentivize the production of databases by treating databases as normal public goods (non-excludable, non-rivalrous goods with positive income elasticity of demand, i.e. national security or education). The discrepancy largely stems from the way copyright is centered on creativity and method of expression, providing very thin protections to works like databases. This thin copyright of the US is sufficient in protecting intellectual products such as literary or creative works, but insufficient in protecting those of more technical natures like databases. The language of copyright laws emphasize the creativity in "arrangement, coordination, or selection of data in databases", but the threshold of creativity does not adequately capture the increasingly finer distinguishing characteristics of databases that merit protection and databases that do not merit protection. The complete abandonment of the "sweat of the brow doctrine" is problematic for an intellectual good that is in truth a mix between public and private goods of the common pool resource variety. This type of good is non-excludable but rival in nature, which naturally requires a different set of legislation than a pure public good, which the current system is designed for.

Under the European Union Database Directive, copyright protection is granted if there is some creativity in selection or arrangement much like in the United States, but with a *sui generis* (separate) right for database materials. The copyright component prevents others from taking the selection or the arrangement of data. The *sui generis* component protects the investments that went into obtaining, verifying, and eventually, presenting the actual data in the form of a database. In short, the *sui generis* right offers protection on the basis of "sweat of the brow", rewarding investments of labor and capital separately from the investments of creation. The scope of this regime's protections is much wider than that of the US, which only rewards the investments of creativity. Basically, the Database Direc-

5 Merges, Robert P., Peter Seth Menell, and Mark A. Lemley. *Intellectual Property in the New Technological Age*. 6th ed. New York: Wolters Kluwer, 2012. 440-449.

tive and the sui generis right combined prohibits unauthorized extraction or re-utilization of substantial parts of data, giving owners of data, not just the database, exclusive property rights. It is important to note that while the sui generis right is not explicitly an intellectual property right according to Articles 7.1 and 7.3 of the Directive, it effectively acts as one.⁶ As Derclaye's assessment of the sui generis right's nature states, the sui generis right can be conceptualized as the codification of unfair competition law (specifically the prevention of parasitism) into an intellectual property right.⁷

At first glance, the EU's sui generis right might seem to solve the problem of the U.S.'s under-protection. However, the introduction of the new sui generis database right makes it unclear whether it maintains the balance between the rights of creators with the rights of users. Like the U.S. legal system, the EU system also lacks definition with respect to the identification of databases and the typology of databases, and actually has a greater gap in protection of fact-centric works, since it does not clearly distinguish between types or sources of data, as the US does. Furthermore, while the Directive protects substantial investment in the creation of databases, it does not specify what constitutes 'substantial' and does not provide a de minimis standard. Lastly, while there are some fair use exceptions to the sui generis, these are not many in number and can therefore create too many loopholes for creators to unnecessarily monopolize information to the point where society loses utility. Therefore, it can be observed that the EU system is also problematic, in the sense that it widens the scope of protection by too wide a margin.

Where the US under-protects the copyright of databases, the EU over-protects them. By issuing both copyright and the sui generis protection as found in the Database Directive, databases essentially enjoy a two-fold protection. Since the passing of the Database Directive, databases enjoy separate protection under copyright laws and sui generis rights. This is a problem not only because of the two-fold protection, but also because it creates a great discrepancy in international copyright relations. To put the differences between the US model and the EU model succinctly, the differences are those between means and ends. The system of the US is Utilitarian in its most basic justification, because the ultimate end goal of its copyright laws is not to protect the right of the author, but rather to maximize the overall social utility. This is evident in the Court's rhetoric, which emphasizes the granting of intellectual property rights as a way to incentivize the creation of such works as protected by copyright, rather than a way to preserve any fundamental rights that an author might deserve. Thus in the US treatment of copyrightable works, the originality of the work is of utmost importance: underlying data does not usually qualify for protection in terms of originality, only those parts of the database as created through intellectual activity (manner of arrangement, selection of data, etc.) are copyrightable. The purpose of this is to reward only the investments of intellectual activity, not those investments typical of what is called 'sweat of the brow' in legal scholarship – the investments of time, labor, or finances.

6 *European Union (EU): Directive No. 96/9/EC of the European Parliament and of the Council, of 11 March 1996 on the legal protection of databases*, Art. 7.1-7.3, March, 1996. <http://www.wipo.int>

7 Derclaye, Estelle. *The Legal Protection of Databases a Comparative Analysis*. Cheltenham, UK: Edward Elgar, 2008. 253-257.

The granting of intellectual property rights in the US is therefore a means to the end of societal utility maximization, whereas the system of the EU is a means by which to protect the ends of personal rights. The laws of the EU are based more on the Personal Rights of copyright owners, or the philosophy of the ‘droit d’auteur’ (literally: right of author). This philosophical justification manifests itself most clearly in the EU’s creation of the Database Director, or more specifically, Article 4. This Article states that “the author of a database shall be the natural person or group of natural persons who created the base or, where the legislation of the Member States so permits, the legal person designated as the right holder by that legislation”⁸

The stress on authorship highlights the philosophical priority, especially since (as per Article 5) the author has the exclusive right to carry out or to authorize reproductions, translations, derivative works, distributions, communications to the public, and reproductions of the database, with rather loosely defined ‘fair use’ exceptions.

Furthermore, by treating data itself as an intellectual product, the EU legislation is emphasizing the role of authorship in determining whether something deserves to be protected by the granting of intellectual property rights. And rather than just creativity, the standard for meriting protection also includes considerations of investment.

Now that the differences between the two extremes are clear, it is necessary to first clarify how these different legal systems affect the economics and incentive structures of database production, regulation, and innovation before discussing how the ideal system should be structured

Economics of Intellectual Property, Competition and Contract Law

This section will give a succinct account of the roles of each of the three neighboring areas of the law, explain the extent to which each provides legal protection, and conclude that intellectual property rights is currently the area that best fits the system requirements for protecting databases. In the next section, I will explain why databases are not purely private or public goods, and how that changes the way we should think about the intellectual property rights framework.

Databases need regulation for the same reasons that other products of intellectual activity need them: to protect the financial interests of the owners of such products such that they will continue to create them, but also limiting ownership rights such that there remains a reasonably low barrier to entry. Legal protection is primarily achieved through three separate areas of the law: intellectual property, competition, and contracts. Contract law, while important, does not impact the market structure or incentives on the same level as intellectual property and competition, and so will not be discussed in this paper since the topic is concerned with how legislation shapes market behavior, rather than the behavior of individual entities.

In order to determine the optimal scope of legal protection for databases, it is necessary to first understand the two main areas of legislation that govern databases: intellectual property and competition law. These two areas of law shape the economics of

⁸ *European Union (EU): Directive No. 96/9/EC of the European Parliament and of the Council*

databases by defining them as goods, limiting monopolies thereof, as well as creating incentive structures for innovation and competition. In this section, the purposes of intellectual property and competition laws in the regulation of markets will be explained first. Then, there will be a discussion on the relationship between these two areas of the law, and how they work together. This is to build an understanding of where the boundaries of intellectual property end, and where those of competition law begin, so that a regime capable of regulating databases with more economic efficiency than the currently dominant ones can be identified.

Both intellectual property and competition law are designed to maximize the efficiency of markets. However, it can be argued that intellectual property aims to maximize the efficiency of production in the long term (dynamic efficiency), as opposed to maximizing the efficiency of short-term production (static efficiency). Since dynamic efficiency aims to maximize the efficiency of production in the long run, this is the efficiency most relevant to the production of databases, as the database industry requires the development of newer, better technologies in order to continue its upward trajectory.

The purpose of intellectual property is extremely similar to that of traditional, tangible property: affording protection to products in order to prevent and remedy market failures. Intellectual property rights allow the markets for products of intellectual activity (such as databases) to function properly by increasing efficiency in several ways.

First of all, intellectual property rights allow for the existence of markets by way of creating a system through which allocations of resources can take place efficiently through peaceful transactions (as opposed to through violent means as in a Hobbesian state of nature). The existence of ownership enables trade by facilitating trades, where the ownership of those products of intellectual activity can be bought, sold, transferred, or rented.

Secondly, intellectual property rights increase dynamic efficiency by creating an incentive structure that rewards investment in intellectual properties, thus also enabling the dissemination of information. This is probably the most common justification for the existence of intellectual property rights. Basically, the creators of intellectual products must have some way of recouping their investments as incentive to undertake the investment in the first place. Intellectual property rights can supply this incentive very efficiently (especially for products with both private and public good characteristics), as it gives the owners ways to trade and make profits off of the intellectual products by way of engaging in trade. On the other hand, there is the conflicting question of creating social good through the dissemination of information and knowledge. This conflict between private and public interests is one of the biggest challenges in the study of intellectual property rights: finding the balance between giving a wide enough scope of protection to incentivize investment and creation, but not giving so much protection that the creative process is inhibited by the intellectual property rights. The question is even more complex (and will be discussed more in-depth later) for intellectual products like databases, which are relegated to the gray area between literary and technical works.⁹

Lastly, property rights allow for the exclusion of usage, which also plays a role in

9 Landes, William M., and Richard A. Posner. "Antitrust and Intellectual Property." In *The Economic Structure of Intellectual Property Law*, 372-380. Cambridge: Harvard University Press, 2003.

the efficient allocation of resources. This is useful in both the cases of private and public goods. For example, private goods are rivalrous, and property rights would allow for the most economically efficient scenario of having a single individual or a select few individuals decide on the rules for exclusion. On the other hand, public goods are non-rivalrous (i.e. usage by multiple individuals does not diminish its value), and the most economically efficient scenario in this case would be to disseminate the use of the good as widely as possible. By enforcing intellectual property rights, exclusion from consumption will be possible, but the goods can remain non-rivalrous. This way, the producers can make profits, but the public can also maximize benefits from dissemination.

On the other hand, while competition law (also known as anti-trust law in the US) is also meant to regulate the market, it does so by promoting competition so that the consumer can benefit from an absence of unfair monopolies, rather than granting bounded monopolistic rights to owners as intellectual property law does. The objectives of competition law are achieved by imposing a posteriori restrictions on undesirable behavior, unlike the a priori restrictions that intellectual property imposes.¹⁰ For example, competition law punishes those who gain monopolies or other types of market dominance unfairly (i.e. bribing, collusion), or those who abuse their dominant positions to extract unfair benefits from other firms or from consumers, thus unfairly restraining trade and deviating from optimal utility (i.e. monopolistic price-setting). In other words, the purpose of competition law includes preventing abuses of the legal powers that come from intellectual property rights, particularly the potential for abusing the monopoly power that intellectual property rights grant, temporarily or otherwise.

Additionally, competition law generates more economic welfare by creating incentives to innovate. Because price systems are restricted to prevent excessive monopolistic or oligopolistic behavior, producers of databases must find other ways outside of high prices to generate value. Innovation would be the answer to such a necessity, as it is the most cost effective way for producers to add value. In turn, innovation will drive healthy competition, leading to more societal benefits for consumers.

The common underlying assumption of both intellectual property and competition laws is that the free market competition will yield the most economically efficient balance of innovation and wealth generation for the greatest number. However, free markets are still shaped by the decision-makers, and neither firms nor individuals can or should be assumed to act rationally at all times. Though the ways in which they regulate markets may seem to be in conflict, it is more accurate to characterize them as complementary bodies of law – there is a division of labor between these two areas in the overall goal of maintaining beneficial free market competition. Intellectual property rights create the grounds for innovation and evolution by protecting those who introduce innovative products and services. Competition law seeks to keep those activities and powers that intellectual property grants under control, such that the success of the few are not achieved by unjustifiable expenses of the others. Both seek to improve dynamic efficiency and long run economic growth. Both ultimately serve to further the interests and welfare of the consumers.

Despite shared goals, intellectual property rights should be prioritized simply because it is the first layer in the foundation for dynamic economic efficiency in the database markets. Without a properly delineated set of intellectual property rights, the

10 Ibid.

very basics of trade will be difficult, and there will be little competition to observe, much less regulate. Due to this hierarchy of priority, the second part the study will focus on the economics of intellectual property rights (especially copyright), and how they apply to databases, products that exist in the gray area between literary and technical works.

Databases: Public or Private Goods?

The second part of determining the optimal scope of legal protection is to take an even closer look at the economics at play under the different legal systems. In particular, the definition of databases affects the way the system is structured, and ultimately how databases are treated. By producing different incentive structures, systems can either regulate a market efficiently, fail to put boundaries on economic activities, or stifle a market with too many rules. To begin, it is necessary to define databases in an economic context. Defining the type of goods that databases are will inform whether or not the legal system is balancing incentives to create with incentives for innovation. The definition of databases as goods and the balance between incentives are the two challenges that this section seeks to discuss, and explain how controlled monopolies and competition can actually foster innovation.

The difficulty of trying to fit databases into the current copyright models lies with the changing nature of both copyright and databases, especially since databases can take a variety of forms. The most prevalent analysis of databases is to treat databases as public goods (non-excludable, non-rivalrous goods) and to discuss copyright law as a means to reduce transaction costs and to further the dissemination of knowledge by giving those who produce intellectual public goods a way to recoup investments. Copyright started out as a simple mechanism by which the law can ensure each man can claim the profits of any intellectual product he may create, when it was first conceived as an alternative to the old patent laws inspired by the Gutenberg press.¹¹ Authors were compensated for the printing of their works with copyrights, and printers were compensated for their printing technology with patents. In short, copyright is meant to give protection to the rights of those who produce intellectual goods that become public goods if released to the public – goods that become non-excludable and do not diminish in value even when consumed by many.

Databases, on the other hand, are now increasingly produced by the private sector, rather than by the government, which was traditionally the case until recently. Databases have also become more complex in their structure and designs, corresponding to the demands created by the proliferation of big data and exponential application of the Internet of Things (IOT). This is due to the rising prevalence of computers and super-computers, which can store and manipulate large amounts of big data, greatly speeding up the capture, analysis, curation, transfer, and visualization that goes into the processing of databases. The prevalence of databases in decision-making is such that they are now often a necessity, especially for those industries that deal with predictive analyses. The applications of databases are now more varied and different from traditional economic considerations. Depending on the needs of the owner, a database can consist simply of some information

11 Merges, Robert P., Peter Seth Menell, and Mark A. Lemley. "Brief History of Copyright Protection." 430-434.

stored on excel spreadsheets, or a quantity of information so big that they need to be stored in relational database management systems (RDBMS) or cloud servers, and need to be manipulated with programs or languages such as SQL. Additionally, while databases can be produced solely for the purposes of trading and profit, many companies now also create databases in-house. This makes the markets of databases difficult to define, as one database can arguably be products within multiple markets, depending on how it is utilized. Most importantly, new applications mean new business models, which in turn require a reconsideration of copyright law's adequacy in providing a balanced amount of protection.

Another difficulty arises in distinguishing between code and database, especially due to the distinctions drawn between databases and computer software, whereas the proliferation of big data has often blurred the line of where 'public information' ends and 'authored creation' begin. There has not been a definite distinction between the two for databases, and data that has been manipulated by code or by calculations lie in the gray areas of legal definition. However, as the current US system specifies that code enjoys its own copyright and the EU system does not specify whether code is protected with or as a part of databases, this paper will not discuss this issue extensively. It is sufficient to note the existence of this controversy, because it indicates an incomplete understanding of what type or types of goods databases should be classified as.

All of these different elements mean that databases are neither perfect public goods nor perfect private goods. A public good is one that is both non-excludable and non-rivalrous, or one from which individuals cannot be excluded from use, and where the use of the good by multiple users does not diminish its value. A private good is one that's excludable and scarce, or one from which individuals can be excluded from use (unless they buy the good or obtain usage otherwise), and where the use of the good by multiple users diminishes its value. Although databases are typically treated as public goods, there is reason to believe that if databases are to be treated as commodities, they would be somewhere in between the spectrum of public and private goods, as they exhibit characteristics of both in the marketplace, depending on the nature of the database.

While they are non-excludable if released to the public, the value of databases can, in fact, diminish in value if there are multiple users, especially unauthorized users. This is because the principal value of some commercial databases resides in its exclusivity, either in terms of point of access, access in general, or singularity of source. Take for example databases that consist of real-time or time-sensitive data. A major part of the value proposition of such databases is the advantage of obtaining data before others, or gaining access to the up-to-date information possible. If a release to a third party occurs (against whom the database owner cannot seek legal remedy due to how contract law works), then the database not only becomes non-excludable, but also diminishes in value because it loses the exclusivity of early access, which can be interpreted as a loss in scarcity.

Alternatively, while the obtaining of 'free copies' of databases without time-sensitive data does not theoretically affect its integrity or objective, such a situation does affect its market value, again due to the decrease in exclusivity or scarcity of this product.¹² In these aspects, a database more closely resembles a private good. Thus the actual position

12 Boldrin, Michelle, and David Levine. "Intellectual Property and the Efficient Allocation of Social Surplus from Innovations." In *Property Rights Dynamics: A Law and Economics Perspective*. Milton Park, Abingdon, Oxon: Routledge, 2007. 101.

of databases on the spectrum spanning from public to private good would be determined by the nature of the database in question, and the data that it is comprised of. Singularity of source refers to databases that include information that only exists in that database or in a very limited number of databases, because the information is laborious and difficult to generate or to discover. However, these processes do not necessarily or typically involve creativity and would fail to obtain copyright protection. Although it is technically true that the information existed before being generated or discovered by the database author(s), it would not have been generated or discovered otherwise. It would be unwise to exclude these types of information products from protection due to a simple threshold test, since it would reduce the amount of such information products in the market (due to inability of producers to recoup costs). Again, these scenarios illustrate how it is almost arbitrary to limit copyrights to works of creativity or originality in the absence of a way to protect products that require otherwise intellectually and financially demanding investments.

The numerous ways in which to categorize databases indicate that a simple threshold like that of the U.S. system or the somewhat arbitrary commitment to personal rights over societal gains of the EU system is also not precise enough to effectively determine an appropriate scope of protection for each database.¹³ Accordingly, the economic purpose of protecting databases is not just to preserve value for those who produce databases, but also for those who are on the consumers of databases. Legal protection can increase welfare not only by preventing dead weight loss in the form of inefficient trade, but also by conserving product value.

Defense of The Middle Ground

In this section, I argue that the economic arguments in the earlier sections necessitate a revision of the law on databases. More specifically, I make the case for a middle ground between the current extremes exemplified by the US and EU systems.

Changes in the creation, dissemination, and application of databases fueled by technology's increasingly rapid growth have outstripped the progress of the laws that govern their activities. Although agility is not a word normally associated with the law, it might become a necessary characteristic when applied to industries of dynamic change. The role of the database industry is so central to developments across sectors that it is of utmost importance that it is regulated in an economically efficient manner. It can be argued that the key to success in creating an economically efficient system of regulation for databases lies in the balancing of multiple moving parts, as the problem itself is multi-faceted. The biggest challenge with aligning these moving parts is to find a balance between maintaining the incentive to create and the incentive to innovate.

From the legal perspective of the debate, it is a question of balancing the individual's economic interests with those of society. This is why ownership of intellectual products is an integral part of the ideal legal system, such that the creativity and labor that goes into producing them can be rewarded. In the absence of ownership, there can be no incentive to create, as individuals have nothing to gain from creating databases that will

13 Derclaye, Estelle. *The Legal Protection of Databases a Comparative Analysis*. Cheltenham, UK: Edward Elgar, 2008. 32-33

ultimately lead to exploitation from society, especially if the databases created is somehow unfairly disseminated without consent. From the economic perspective of the debate, it is a question of stabilizing the markets by setting boundaries for different players. Without clear boundaries, it would be too easy to abuse rights without serious repercussions, which would create unfair advantages and thus discourage the creation and innovation that lead to dissemination of knowledge to the public. Therefore, the wants and needs of the public must keep those of the individual in check, and vice versa.

Putting these two analytical perspectives together, the ideal system would lie somewhere between the threshold system of the United States and the two-pronged system of copyright and sui generis right of the European Union. Where the United States system under-protects databases by a lack of specification, the European Union over-protects by establishing a system that could potentially provide two-fold protection. To find a point of balance between these two systems, some degree of agility is necessary in order to satisfy considerations from both the legal and the economic side.

To begin with, there should be a defined set of typologies of databases, such that the appropriate remedies can be applied to the appropriate types of products. Starting with the definition, one needs to recognize the fact that the category of 'database' encompasses a very broad spectrum of products, with varying levels of complexity, creativity, and volume. It would appear that a single threshold, such as the US system relies on, would be insufficient to effectively regulate the market activities of databases – separating the universe of databases into two parts would be a gross over-generalization. Instead, there should be a multi-tiered approach to determine the scope of rights granted, so that different levels of investments in intellectual creativity and labor can be rewarded accordingly. For example, a database that is high in creativity and labor should not be awarded the same rights as a database that is high in creativity but low in effort, because overall, the latter did not require the same investments (intellectual or otherwise) to create as the former.

Next, the tiers of such rights should be measured according to a weighted 'formula' of sorts: creativity should have higher rewards than effort, but effort should not be completely neglected either. This is because while the primary purpose of intellectual property is to encourage creativity and innovation, it is important to remember that oftentimes, products of creativity are derived from the fruits of plain, un-creative labor. To cut out the value of labor would be to do a great disservice to the goal of innovation for the sake of social welfare. One way to conceptualize this would be to think of the granting of intellectual property rights as a sort of quid pro quo between the innovator and society. Society rewards the innovator for all of his efforts by giving him a temporary monopoly over his creation, and in return, the innovator gives to society his creation after he has recouped a reasonable amount to make the undertaking worth his while in the first place. In the case of copyright-protected works like databases, the nature of innovation usually occurs in small steps, as opposed to the larger strides observable in patentable technologies, for example. This means that along the way, the unrecognized value of labor becomes a deterrent to new creativity, as potential innovators see no reason to risk not regaining their investments if their product is not guaranteed protection.

The duration of copyright is also of great importance. If copyrights can be renewed indefinitely, then it would be, in effect, a perpetual and everlasting copyright that would harmfully deprive society of the utility of information flow. Although Landes and

Posner argue that the size of public domain of works has a positive relation to the extent of copyright protection, and that the actual percentage of copyright renewals are so low that the possibility of perpetual rights has negligible effect, this analysis of works does not apply to the copyright of databases.¹⁴ Rather, the *Amici Curiae* brief by George A. Akerlof et al. with respect to the *Eldred v. Ashcroft* case (a critical case for copyright terms) better captures the question of copyright renewal for databases. In this brief, Akerlof et al. come to the conclusion that there is no discernible benefit in allowing the possibility of perpetual copyright, due to the inability of such a right to generate additional incentives to create, the additional costs that such a system would impose, and the minimal costs and benefits that invalidate the rationale for such an undertaking.¹⁵ A similar view is expressed by Boldrin and Levine, observing that indefinite extensions of term makes consumers worse off, thus actually decreasing economic efficiency rather than increasing it.¹⁶ These latter analyses are even more relevant for the copyrights of databases, since the incentive to create is much more closely related to the ability to recoup the investments than other types of literary works. Additionally, database production is more sensitive to transaction costs, since a perpetual copyright would create a monopoly on information that might be necessary for the generation of other, new information and ideas, rather than a monopoly on the expression of ideas.

All in all, the biggest and most impactful shortcoming in the legal protection of databases lies with intellectual property, as property is the basic infrastructure upon which markets are built. Combining the insight demonstrated by the European Union and the streamlined practicality of the United States can amend these shortcomings in the current legislation. The resulting system would be a more agile, and being much more easily tailored to new technological trends, it can perhaps also be a starting point for greater future consensus across countries.

Conclusion

In this paper, I have explained the economics of competition, contract, and intellectual property, defined databases as a common pool resource, shown the differences between the two extremes in the legal protection of databases, and defended a potential solution for the current extremes by finding a point of balance between their strengths and weaknesses.

From an economic standpoint, intellectual property law is the most effective and efficient way to protect works like databases, since the existence of property rights is the foundation upon which markets can be built to function. Comparatively, contracts have limited enforceability beyond the immediate parties, and the applicability of competition law requires the prior existence of a market. However, pure copyright alone does not

14 Landes and Posner 212

15 Brief of George A. Akerlof et al., as *Amici Curiae* in Support of Petitioners in *Eldred v. Ashcroft*, 123 S. Ct. 769 (2003), filed May 20, 2002.

16 Boldrin, Michelle, and David Levine. "Intellectual Property and the Efficient Allocation of Social Surplus from Innovations." In *Property Rights Dynamics: A Law and Economics Perspective*. Milton Park, Abingdon, Oxon: Routledge, 2007.

guarantee adequate protection, especially since databases can have such a wide range of properties, from pure public goods (i.e. government produced databases) to common pool resources (i.e. proprietary databases that can lose value if leaked).

The two extremes compared are the United States and European Union legal systems with respect to the protection of databases. While the United States' Supreme Court emphasizes the importance of originality as a defining factor of intellectual works worthy of copyright, the European Union's Database Directive includes both originality and investments (of labor or resources) in the consideration of copyright. The problem with the former is an under protection of valuable works that require both intellectual and financial resources to create, thus unnecessarily increasing the likelihood of market failures in the form of undersupply. The latter errs on the side of too much protection, lending itself to abuse and market failures in the form of monopolies.

The solution to these complications is threefold. First, it is necessary to build a solid foundation of definitions, for databases, for characteristics that merit copyright or sui generis protection, for the scope of rights awarded, and for fair use exception. Secondly, it is necessary to determine different levels of protection according to the amount of originality and to the amount of investments in forms outside of intellectual creativity (monetary investments, intellectual labor, etc.). It is important to maintain that though both originality and 'sweat of the brow' are accounted for, this does not necessarily signify that both should be awarded on the same scale. It is perfectly reasonable for courts to find originality a more compelling reason than labor or financial investment to award copyright, but the latter should not be completely disregarded in the interest of avoiding unnecessary transaction costs at best, or market failures at worst. Lastly, it is critical to clearly define limits to the copyright term, so as to avoid situations of the 'perpetual renewal' or everlasting term, which is counterproductive to the purpose of copyright.

Future steps to research this topic should include more empirical evidence in order to examine the effects of different legal regimes. Due to the fact that effects of legislations typically take longer to manifest themselves (sometimes periods as long as a generation or two), it is currently unclear whether the European Union's Database Directive and sui generis rights have positive, negative, or neutral impacts, or whether the United States' simple originality threshold is preventing more robust innovation of information products. However, what is clear is that there needs to be a balance between the two, such that the law and economics underlying the database industry can be in harmony.

Works Cited

- Allen, R. G. D. *Mathematical Analysis for Economists*. London: Macmillan and, 1938.
- Becker, Gary S., and Richard A. Posner. *Uncommon Sense*. Chicago: University of Chicago Press, 2009.
- Bitton, Miriam. "A New Outlook On the Economic Dimension of the Database Protection Debate." *IDEA: The Intellectual Property Law Review* 47, no. 2 (2008): 93-165. Accessed July 10, 2015. heinonline.org.
- Blanke, Jordan M. "Vincent Van Gogh, "Sweat Of The Brow," And Database Protection." *American Business Law Journal* 39 (2002): 645-82. Accessed July 10, 2015. heinonline.org.
- Chen, Chuanfu, Gang Wu, Kuei Chiu, Qiong Tang, Juana Tang, and Jie Liu. "Legal Protection for Database: Current States, Problems and Possible Future." *2008 Third International Conference on Convergence and Hybrid Information Technology*, 2008.
- Davison, Mark J. *The Legal Protection of Databases*. Cambridge: Cambridge University Press, 2003.
- Derclaye, Estelle. "An Economic Analysis of the Contractual Protection of Databases." *Journal of Law, Technology, and Policy* 2005, no. 2 (2005): 247-80. Accessed July 10, 2015. heinonline.org.
- Derclaye, Estelle. *The Legal Protection of Databases a Comparative Analysis*. Cheltenham, UK: Edward Elgar, 2008.
- Ficsor, Mihály. "Substantive Provisions of the WIPO Copyright Treaty (WCT)." In *The Law of Copyright and the Internet*. New York, New York: Oxford University Press, 2002.
- Freno, Michael. "Database Protection: Resolving the U.S. Database Dilemma With An Eye Towards International Protection." *Cornell International Law Journal* 34 (2001). Accessed July 10, 2015. heinonline.org.
- Furmanyuk, Andriy, Mykola Karpinskyy, and Bohdan Borowik. "Modern Approaches to the Database Protection." *2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, 2007.
- Hladjk, Jörg. "The Protection of Databases under US and EU Law – Sui Generis Right as an Appropriate Concept? – Part I – US Law." *Computer Law & Security Review*, 2004, 288-92. Accessed July 10, 2015. sciencedirect.com.
- Gallegos, Frederick. "Database Protection: Selected Legal and Technical Issues." *Information Systems Security*, 2000, 1-13.

Greenwald, Douglas. *The McGraw-Hill Dictionary of Modern Economics; a Handbook of Terms and Organizations*. 2d ed. New York: McGraw-Hill, 1973.

Landes, William M., and Richard A. Posner. "Antitrust and Intellectual Property." In *The Economic Structure of Intellectual Property Law*. Cambridge: Harvard University Press, 2003.

Maurer, Stephen M. "Across Two Worlds: Database Protection in the United States and Europe." 2005. Accessed July 10, 2015. www.researchgate.net.

Miceli, Thomas J. *The Economic Approach to Law*. Stanford, Calif.: Stanford Economics and Finance, 2004.

Murray, Andrew. *Information Technology Law: The Law and Society*. Oxford: Oxford University Press, 2010.

Parisi, Francesco. *The Law and Economics of Irrational Behavior*. Stanford, Calif.: Stanford Economics and Finance, Stanford University Press, 2005.

Porrini, Donatella, and Giovanni Ramello. "Intellectual Property and the Efficient Allocation of Social Surplus from Innovations." In *Property Rights Dynamics: A Law and Economics Perspective*. Milton Park, Abingdon, Oxon: Routledge, 2007.

Porrini, Donatella, and Giovanni Ramello. "Intellectual Property and the Efficient Allocation of Social Surplus from Innovations." In *Property Rights Dynamics: A Law and Economics Perspective*. Milton Park, Abingdon, Oxon: Routledge, 2007.

Robin Jeweler. *Protecting Noncreative Databases* (CRS Report No. RS21662). Washington, DC: Congressional Research Service, 2011

Ruse, H. G. "Electronic Agents and the Legal Protection of Non-creative Databases." *International Journal of Law and Information Technology* 9, no. 3 (2001): 295-326. Accessed August 10, 2015.

Stamatoudi, Irini A. *EU Copyright Law*. Cheltenham: Edward Elgar Publishing, 2014.

Zhu, Hongwei (Harry), and Stuart E. Madnick. "Reutilization and Legal Protection of Non-Copyrightable Database Contents." *SSRN Electronic Journal SSRN Journal*, 2006. Accessed August 10, 2015.