The Legal Protection of Databases

Mark J. Davison
Monash University

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge CB2 1RP, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge, CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface Plantin 10/12 pt *System* LaTeX 2_ε [TB]

A catalogue record for this book is available from the British Library

ISBN 0 521 80257 1 hardback

Contents

1 Introduction

We live in the Age of Information. Information is money. So is time. The economies of the First World are dominated by the creation, manipulation and use of information and the time it takes to do so. These economies do not suffer from a shortage of information; they suffer from the difficulties associated with collecting, organising, accessing, maintaining and presenting it. Databases are designed to help deal with these difficulties. They are collections of information arranged in such a way that one or more items of information within them may be retrieved by any person with access to the collection containing those items.[1] Therefore, databases are big business because they contain important and copious amounts of information and they reduce the time taken to access that information.[2] And where there is big business, the law and lawyers inevitably follow.

But information is more than money and databases are more than big business. Information and databases are critical to science, the legal system itself, education and all those aspects of life that are improved by them. Consequently, there are important issues of social and political policy to be considered in the regulation of access to, and use of, databases. Again, where there are such critical issues at stake, the law has a role to play.

There is an inevitable tension between the commercial and the socio-political role of databases that leads to complexities in developing an appropriate model for their legal protection. In fact, given the diverse range of areas in which databases can be used, any one of a variety of legal models may be appropriate in any given context. One of the criticisms of general references to the importance of information is that they fail

[1] This is a very rough working definition of a database. The various issues concerning the definition of a database are discussed in later chapters, especially Chapters 3 and 4.

[2] 'In 1989, the world-wide turnover for online database and real time information services accounted for around 8.5 billion ECU.' In 1996, the estimated size of the European Market electronic information supply market was £5.138 billion. A Consultative Paper on United Kingdom Implementation: Directive 96/9/EC of 11 March 1996 on the legal protection of databases copyright directorate, The Patent Office, DTI, August, 1997 at para. 2.1.5 and Annex 2.

to differentiate between different categories of information.[3] The same criticism could be levelled at any legal system that applied a 'one size fits all' approach to the regulation of databases. It is no surprise then that a number of different legal models for protection of, and access to, data and databases have arisen.

Why have databases become an important issue

The transition of many First World economies from industrially based economies to information-based economies is a relatively recent phenomenon. It is a consequence of an explosion in information and the means by which it can be disseminated that results in turn from far-reaching technological and scientific developments.[4] In particular, advances in digital technology have facilitated the creation of databases. Large amounts of data can be created in, or converted into, digital form, and scanners and other devices permit the digital conversion of data. Alternatively, data can be originally produced and stored in digital forms that are perceived by humans as text, pictures, tables, spreadsheets and other easily recognisable formats. The digitisation of data in turn reduces storage costs. For example, if the DNA structure of the human genome were compiled in hardcopy it would occupy 200,000 pages.[5] The physical storage of such documentation in digital form can be achieved with a few CDs.

This expanded capacity to store data is complemented by an increased capacity to access and use it. It is facilitated by computer programs that enable quick and reliable searching and retrieval of data. Computer networks also allow on-line use of databases, thus increasing ease of access and marketability. These increased abilities to store and disseminate information, in turn, have increased the production of information. This is due to the relationship between the production of information and the availability of existing information. Existing information and access to it are critical to the creation of new data and information.[6] This creative process is like a spiral in which the users of existing data actually add

[3] See Chapter 6 for a discussion of this point.

[4] 'It has been estimated that the volume of the increase annually in information generated today equals the total information in circulation in the world fifty years ago.' Explanatory Memorandum to the Proposal for a Council Directive on the legal protection of databases COM(92) 24 final – SYN 393, Brussels, 13 May 1992.

[5] Human Genome Project Information at http://www.ornl.gov/hgmis/publicat/primer/fig14.html.

[6] At this point, the terms 'data' and 'information' are being used interchangeably. Possible distinctions between the two and the relevance of those distinctions are discussed in Chapter 6.

value to that data in the process of using it, thus generating more new data and information.

The pressure to provide specific legislative protection for databases has arisen from the increase in the mass of raw data available in almost every area of commerce and science, the increased technological ability to create databases containing those data and to provide easy access to them. These are coupled with the increased technological ability of others to reproduce those databases and a perceived lack of adequate protection from existing legal regimes, such as copyright. The same technology that has expanded the role and usefulness of databases permits quick and easy reproduction of those databases or large parts of the data contained within them. 'Robots' and other computer technology can be used to download data from databases with little effort or human intervention. This reproduction can take place anywhere on the planet, provided the person arranging for the reproduction has access to the necessary computer infrastructure. Consequently, database owners have claimed that they require additional legislative protection to protect their investment in the creation and marketing of databases from free-riders who can quickly and easily reproduce the databases created and maintained by them.

The structure of this book

This book examines various models of legal protection for databases. A brief explanation of those models is given at the beginning of Chapter 2, where the various basic legal principles relevant to nearly all jurisdictions are covered. In particular, Chapter 2 deals with some basic principles of copyright, unfair competition law, contract and competition or anti-trust law as they apply to databases. These principles are referred to throughout the book.

Chapter 3 examines the European Union (EU) Directive on the Legal Protection of Databases 1996 (the Directive),[7] including both the copyright protection and the *sui generis* protection that has been conferred by the Directive. This examination includes the history of the Directive, the justifications provided for it and its important features. In addition, Chapter 3 examines the impact on database protection of the provisions of the EU Copyright Directive on the harmonisation of certain aspects of copyright and related rights in the information Society 2001 (the Copyright Directive).[8] The provisions of the Copyright Directive

[7] Directive 96/9/EC of 11 March 1996 on the Legal Protection of Databases, OJ No. L77, 27 March 1996, pp. 20–8.
[8] Directive 2001/29/EC, OJ No. L167, 22 June 2001, pp. 10–19.

concerning the circumvention of effective technological measures that are designed to protect copyright material also apply to the *sui generis* right conferred by the Directive. Consequently, those provisions are an important aspect of the protection provided for databases.

The examination in Chapter 3 of the history of the Directive reveals that the initial EU moves for *sui generis* protection proposed a very limited protection clearly separated from the copyright protection of databases. However, Chapter 3 also shows that the final form of *sui generis* protection under the Directive is, in fact, a hybrid of the generous scope of protection under former UK copyright law and the restrictive exceptions provided in the copyright law of many continental countries. The latter are probably quite justified in a copyright scheme that requires high levels of originality before conferring any copyright protection at all. However, they are inadequate in a legislative scheme that confers protection on unoriginal databases. The effect of this hybrid approach has been to confer an extraordinary degree of *sui generis* protection. The argument is also made that the *sui generis* protection provided by the Directive is inappropriately and inextricably entwined with copyright law and that, in a number of technical respects, the Directive is worded in such a way that it provides protection, even beyond its intended scope. An example of this latter point is the broad definition of a database.

Chapter 4 examines the legislation transposing the Directive in a number of the Member States and some of the emerging case law relating to that legislation. This examination further illuminates some of the ambiguities in the wording of the Directive and different approaches that have been taken to its transposition.

Chapter 5 examines the protection provided by copyright and the tort of misappropriation in the United States. Copyright and misappropriation principles have underpinned the different proposals that have been made in the United States for *sui generis* protection. Yet the tort of misappropriation has itself had a chequered history. Considerable judicial attention has been given to the theoretical basis of the tort and its consequent scope with resulting differences in the operation of the tort. Consequently, it is not surprising that different pieces of proposed legislation that have all been (allegedly) based on misappropriation have proposed quite different degrees of protection. The lesson to be learned from this is that if the concept of misappropriation is to be incorporated into *sui generis* legislation, it needs to be defined with some precision. Chapter 5 also examines the different pieces of proposed legislation and compares them with the tort of misappropriation and the Directive. This examination reveals a move away from the approach taken in the Directive towards one with wider exceptions to protection and a less

restrictive approach to the use of information for transformative or wealth-producing uses of information. There are also provisions that are designed to ensure public access to information produced by government or with government funds.

Chapter 6 examines moves to provide additional protection for databases outside of the EU. In 1996, a draft treaty based on the Directive and legislation that had been proposed in the United States[9] was briefly considered at a diplomatic conference hosted by the World Intellectual Property Organization (WIPO). The draft treaty was not adopted but the issue has continued to receive consideration by WIPO since that time. The failure to pass any of the proposed pieces of legislation in the United States has no doubt hampered that process but once such legislation is in place, moves for a treaty are likely to intensify. To date, the EU has suggested its Directive as a template for a treaty on the topic but this has encountered considerable resistance from developing countries. Resistance has also come from international science organisations that are concerned about the potential impact of any *sui generis* legislation on the exchange of scientific information. The relevance of their views to *sui generis* protection is considered in Chapter 7.

In response to this resistance at WIPO, the EU has shifted its focus to its bilateral arrangements with other countries such as those seeking membership of the EU. Consequently, over fifty countries, including the fifteen Member States, either have *sui generis* protection for databases or will acquire it within the next few years.

There are other significant international aspects to the protection of databases associated with these moves. For example, the Directive provides that *sui generis* protection for overseas databases will only be conferred if the nations from which those databases originate also provide materially the same protection for EU databases.[10] This use of reciprocity provisions in intellectual property regimes is a relatively rare departure from the usual international practice of according national treatment to nationals from other nations. One of the reasons for this approach is to place pressure on countries such as the United States to provide reciprocal protection and to create a de facto international model for protection. The implications of this are discussed. In particular, Chapter 6 argues that the EU may be obliged by international agreements to provide national treatment to overseas databases and, consequently, the pressure to provide reciprocal protection is not as great as it may seem. Part of the

[9] The Database Investment and Intellectual Property Antipiracy Act of 1996, HR 3531 of 1996.

[10] Article 11 of the Directive.

basis of this argument relates back to the point made in Chapter 3 that *sui generis* protection is inextricably entwined with copyright. While the Directive describes it as being separate from copyright, a close inspection of the subject matter of protection, the rights conferred and exceptions to those rights suggests that *sui generis* protection is, in reality, a form of copyright.

Chapter 7 analyses the arguments for and against the different models for *sui generis* protection from a theoretical perspective. As with every intellectual property regime, the law in relation to the protection of databases needs to achieve an appropriate balance between the rights of users and the rights of producers or owners of intellectual property. The ultimate objective of this balancing act is to achieve an optimal production and dissemination of the material that is, or could be, contained within databases. Hence, database owners have argued that greater protection is required for databases in order to protect their investment in production. This emphasis on protection of the database maker's investment undoubtedly underpinned the Directive, as making a substantial investment is the litmus test for whether the Directive's *sui generis* protection extends to a particular database.[11] However, this emphasis represents a significant shift in the general approach to the recognition and protection of intellectual property. At least in common law countries, the emphasis in other intellectual property regimes has been on the creation and maintenance of a social contract between creators and users. While encouraging investment is a desirable goal of this social contract, the real question is whether the investment in question is an optimal investment for public purposes. This in itself is a controversial issue, as what constitutes 'optimal' investment is debatable.

In the context of databases, this relationship between producers and users is complicated by the fact that in a number of contexts, the users themselves make significant contributions to the production of the information that is contained within those databases; and this information production is often subsidised by public funds. A particular concern is the relationship between protection of databases and the impact of that protection on research and education, activities essential to the continued production of the very information that finds its way into many databases. Consequently, the book examines the impact of the models for *sui generis* protection on research and education.

As the justification for *sui generis* protection of databases is primarily an economic one, an analysis of that justification inevitably requires some examination of economic arguments for protection; hence, some

[11] Article 7(1) of the Directive.

of Chapter 7 is taken up with this. However, the validity of such theories is ultimately dependent on empirical evidence.[12] At the present time, there is no clear empirical evidence justifying a strong form of *sui generis* protection.[13] Consequently, while those theories are important, they should be treated with some caution, particularly when they suggest the creation of strong intellectual property rights which, if created, will be effectively impossible to rescind. In addition, there are important non-economic aspects of the debate concerning protection of, and access to, databases that receive attention in Chapter 7. One example concerns the availability of information for news reporting and political debate.

The book concludes with a list of basic principles that need to be considered and incorporated into any *sui generis* protection of databases. This list is explained by reference to the preceding analysis in Chapter 7 of the arguments for and against different forms of *sui generis* protection, and is compared with particular aspects of the Directive and the various American bills on the topic that are examined in Chapters 3, 4 and 5.

A couple of points need to be made about the issues with which the book does not deal. In particular, it does not cover in any detail the law of confidential information or trade secrets as it applies to databases. This is because the emphasis is upon databases that are available to the public, or at least those members of the public with sufficient resources and interest to acquire access to them. Consequently, the emphasis in the legal analysis is upon proposals for *sui generis* protection for databases that cannot rely upon the protection of the law of confidential information. Legal issues surrounding privacy and databases are also not considered here, although obviously privacy in the context of databases is an important issue in its own right. Nevertheless, the emphasis in this book is on database owners, rights and their appropriate nature and extent, rather than the privacy rights of those whose details may be included in a database.

While it would be superfluous to repeat the details of Chapter 7 here, a couple of general observations about the book's conclusions are worthwhile to assist the reader in the course of the following chapters. The ultimate conclusion of the book is that there is justification for some *sui generis* protection of the investment involved in the creation and presentation of databases. This view is taken by various independent organisations and even those who have expressed concerns about the possibly excessive nature of any *sui generis* protection.[14] In many jurisdictions, the protection provided by copyright is insufficient. However, the justification

[12] P. Drahos, *A Philosophy of Intellectual Property* (Dartmouth, Aldershot, 1996), p. 7.
[13] US Copyright Office Report on Legal Protection for Databases, August 1997, pp. 76–7.
[14] Ibid., p. 78. Statements of Andrew Pincus, General Counsel, US Department of Commerce, Joshua Lederberg (on behalf of the National Academy of Science and Ors),

only extends to quite limited protection over and above that presently conferred by copyright, contract and other means. Any international agreement or legislation on the topic needs to acknowledge and respond to the diverse types of information in databases and the diversity of their potential uses. A simplistic approach which confers strong exclusive property rights in all databases and which applies to all uses of those databases does not meet that need. Such an approach runs the risk of treating all information as a commodity for all purposes.

In particular, there is a need to ensure that public access to information created with government funds or subsidies is not completely lost. This is an important issue. For example, governments, universities and other non-profit organisations supply more than one-third of the funds devoted to research and development[15] and the process of government also generates large amounts of information that are valuable both in a commercial sense and to the democratic process.

The latest American proposals for *sui generis* protection based on misappropriation principles have addressed some of the difficulties, and demonstrate an appreciation of the complexities associated with legislation concerning such a diffuse area. Hence, there are a number of exceptions provided for in the latest proposed legislation and protection is based on misappropriation principles. Nevertheless, it is too simplistic to just accept the view that any *sui generis* protection should be based on misappropriation principles. As argued in Chapter 6, misappropriation is a nebulous concept and it must be given a concrete form that is relevant to the area of its application. The latest American proposals still provide generous protection that approximates exclusive property rights, even though they are ostensibly based on misappropriation principles. In addition, the relationship between any prohibition on misappropriation, copyright and contract law needs to be addressed in some detail. While those proposals have considered these issues, there is some room for improvement.

In contrast to the more sophisticated American response to the issue of *sui generis* protection, the Directive adopts an approach conferring broad exclusive property rights with few, if any, meaningful exceptions.

and Charles Phelps (on behalf of the Association of American Universities and Ors) to the Subcommittee on Courts and Intellectual Property of the Judiciary Committee on the 1999 Bill (Collections of Information Antipiracy Act of 1999) on 18 March 1999, pp. 62–506 (Pincus, pp. 51–100; Lederberg, pp. 189–205; Phelps, pp. 223–53).

[15] E.g. between 1992 and 1997 more than 33 per cent of all research and development in the USA was funded by government, universities or other non-profit organisations. 'Statistical Abstract of the United States' (Bureau of Statistics, Washington DC, 1998). The same was also true for the UK between 1992 and 1996: 'Annual Abstract of Statistics No. 135 of 1999, Table 19.1' (Office for National Statistics, London, 1999).

Consequently the Directive has greatly exceeded, in a number of respects, what is necessary or desirable. These include the manner in which it defines a database, the scope of *sui generis* protection provided, the insufficiency of exceptions to *sui generis* protection and an excessive period of protection for database contents. In addition, it is critical that the distinction between copyright and *sui generis* rights be maintained if separate protection regimes are created. Again, the Directive has failed to make this distinction, resulting in a number of difficulties. For those and other reasons, the Directive should not become a template for the international protection of databases.