Public Domain; Public Interest; Public Funding: focussing on the 'three Ps' in scientific research

Dr Charlotte Waelde and Mags McGinley•

1. Introduction

As a result of legislative changes, IPRs have become broader, stronger and longer. Copyright and the sui generis database right, the IPRs of particular relevance to this paper, are no exception. Over the last ten years the contours of existing rights have altered and new rights have been created. One effect of these changes has been a diminution of the public domain. In other words, the power exercisable by right holders has expanded, while the room for manoeuvre by those seeking to use 'free' elements has diminished. In this way drawing on what has gone before, becomes both more difficult and more costly: more difficult because more of what is currently created is owned, and therefore may not be re-used without permission; more costly because seeking permission to re-use works carries with it transaction costs, and may also entail a payment for use.

These developments have been greeted with alarm in some quarters, most notably by academic commentators working in the legal field, and with some understanding of the problematic issues involved. But the murmurings are spreading to other interest groups including those engaged in scientific research who are becoming more aware of the considerable implications of the expansion of these intellectual property rights for their work. As a result, appeals to 'the public interest' are growing, with questions being asked as to whether the placing of scientific data into the realm of private property, whether through copyright fences or (over) extensive claims made in relation to the contents of databases, is the antithesis of the needs of those who would ensure progress through scientific research. This is most particularly so where it is the public purse that has funded the research results and enabled the 'discovery' of the data.

The purpose of this paper is to discuss those 'three Ps': it is (i) to examine some of the difficulties faced by scientists engaged in scientific research who may have problems working within the constraints of current copyright and database legislation, where property claims can place obstacles in the way of research, in other words, the public domain; (ii) to look at perceptions of the public interest and ask whether copyright and the database right reflect understandings of how this concept should operate; (iii) to consider the relevance of public funding for scientific research in the context of both the public domain and of the public interest. Finally, some recent initiatives seeking to change the contours of the legal framework will be examined; initiatives based on ideals of the public domain and the public interest in the scientific sector, and which have public funding as a common factor.

2. The public domain – a lawyer's view

'All authorship is fertilised by the work of prior authors and the echoes of old work in new work extend beyond ideas and concepts to a wealth of expressive details."

Any process of creation depends not only on the existence of a wide variety of sources on which a creator can draw, but also upon the accessibility and re-usability of those sources. In this process, works are often cumulative. Innovations build on each other

and research activity is directed towards improvement or application of previous discoveries or works. Sometimes works are sufficiently original to represent a true breakthrough but usually 'new' works are based upon what has gone before.[2] To facilitate this process of creation, the law sets limits on the property right granted in respect of both copyright and the database right. So, for example, the term of both rights is limited[3]; neither ideas nor data are protected, but only the expression of those ideas[4] or collation of that data;[5] a work protected by copyright must be original[6] before it will be protected, leaving certain elements beyond the scope of the law; a database must be the subject of the relevant investment before the sui generis rights will arise;[7] for both, takings of insubstantial parts are permitted,[8] as is fair dealing for the purposes of non-commercial research.[9] In this way, the law mediates between absolute property rights and the commons. Having a mixture of works available means that more works can be produced, drawing upon the sum of existing scientific knowledge. In developing works, scientists do not have to start from scratch every time; rather they can enlarge, extend, sometimes retract and retrench on what has gone before. It is the limits on copyright and database rights that lawyers generally collectively refer to as the public domain:[10] *'[t]he public domain is the law's primary safeguard of the raw material that makes authorship possible.'*[11]

3. Copyright fencing

For some, the size, accessibility and re-usability of the public domain may be more critical than for others. Take the example of a self-employed creative author beavering away in her study, writing original literary novels to great acclaim. The need to extract from the public domain may not actually be very important to this individual. Often, the greater the originality of the work, the greater the approval for the author. Ideas certainly will be important, gleaned perhaps from having access to a wide variety of artistic and literary output. But actual takings from others are likely to be minimal. Indeed, such an author might seek to diminish the existing public domain. What interest does she have in her work falling into the general public domain 70 years after her death?[12] Why shouldn't some payment be made for takings after this period perhaps to be contributed to the general welfare of up and coming authors?[13] Why should third parties be able to quote from her work without payment? She might think it enough that her work is in the public domain once published in the sense that her ideas are available to all who care to read her work. Is that not enough? Certainly she will want attribution.[14] This will be essential to furtherance of her career while alive and to mark 'her' work out from the generality in life and after death. Copyright will be of relevance to enable her and her successors to command some return from the consuming public and to protect her property.

But look across the spectrum; to the research world where full time employees in publicly funded institutions use public money in pursuit of research agendas the outcomes of which ultimately should further the public interest. Within this environment academics from different disciplines tend to work in different ways. Lawyers, for instance are rather solitary creatures when it comes to research. Certainly they might come together in events, such as conferences and colloquia, to discuss the fruits of research, or to garner new ideas; or they might discuss and write articles jointly with colleagues. In so doing the lawyer needs access to a wide variety of literature: whether journal articles, books, statutes, case law or other sources of inspiration. These works are required to buttress and support legal argument. Access to the works is essential, but does (or should) the academic lawyer really *take* more than what might be considered to be in the legal public domain? Attribution is of course essential for

furtherance of an academic career, but much less so for carving out rights. These will be more important to the publishing industry which, in the academic field, has historically operated by taking assignations of copyright from the author which it has, in turn, exploited in furtherance of their own commercial operations.

Science, as an academic discipline is different again. Scientific research is a highly collaborative endeavour. From spatial organisation to actual input into a publication; from many scientists gathered around one computer to names on research papers; the research efforts are incremental, joint[15] and combine the efforts of many individuals pursuing a common goal. Further, it is not only scientists from single institutions who engage in the scientific process, but they do so across institutions, and indeed across borders. Scientists are excellent at organising and attending informal and formal gatherings at which ideas are discussed and tested, as indeed are the practical applications. Further, the work of the scientists is unlikely ever to be 'finished' in the way that a self-employed author's book or an academic lawyer's article is finished. One only has to consider the different versions of software that are produced and disseminated – each one (supposedly) better than the last. Each builds on what went before and that building will no doubt go on and on.

But there are ironies in the legal framework as it applies to science, in particular what some call 'Big Science'.[16] Scientific authorship means seeking to buttress claims by taking as much as possible from the existing body of scientific literature. Quite unlike the work of the self-employed creative author, the progress of science depends upon being grounded in what has gone before. And although a scientist will need to be recognised on any research paper, that appears to be much more about responsibilities than it is about rights.[17] Nonetheless the academic process and assertion of authorship entail claims being made over the results:[18]

> *'Fencing off the public domain and taking it to the realm of private property through authorship might be smart if you want to commercialise the results but it is plainly a self defeating tactic if the claim you are putting forward is not about property, and if it can bring you credit only by being endorsed, used and cited (but not bought as property) by your peers."*[19]

Further, the extent to which copyright should protect the fruits of the scientific process is questionable. The value, both for the scientist and for science, lies in the scientific claims, ideas or facts which are themselves not capable of being protected by copyright: *'consequently it cannot be the scientists' property.*'[20] Yet this is precisely the way in which scientific publications are treated. They are wrapped the cloak of copyright protection, rights are assigned to publishers, who will in turn publish in journals available to those who can afford access.

Accordingly, in science the application of copyright and the resulting public domain seem out of line with the requirements of its users. A scientific researcher would appear to need to take more from existing works than the legal public domain might allow, but in return would need to, and it would appear from the scientific process be happy to, donate more to it. Science is a culture that invites, relishes and thrives on collaborative working environments. Using copyright to maintain proprietary claims over one's own work and, in so doing, diminish the public domain seems the antithesis of this process.

4. Databases and Science

It is hard perhaps for the lawyer to appreciate the extent to which science depends on the collection of data and other information, and the subsequent re-use of that material.[21] Lawyers might be familiar with databases. Lexis and Westlaw are well-known examples in the legal field. The academic lawyer will, of course, wish to access and re-utilise contents of such databases in their own 'new' creations. But the extent to which they do so might be limited. Academic lawyers after all, like to consider themselves to be original at least in some senses of the word. For science it is quite different. Data and other information is generated on an exponential basis, and held within vast databases. The progress of science depends on the re-use of that data for a variety of purposes. It can also be hard for the non-scientist to appreciate the size and importance of these databases to the scientific community. Useful examples have been given by Hey and Trefethen in their paper *The Data Deluge: An e-Science Perspective*:[22]

> *'in the field of engineering, consider the problem of health monitoring of industrial equipment. The UK e-Science programme has funded the DAME project - a consortium analyzing sensor data generated by Rolls Royce aero-engines. It is estimated that there are around 100,000 Rolls Royce engines currently in service. Each trans-Atlantic flight made by each engine, for example, generates about a Gigabyte of data per engine – from pressure, temperature and vibration sensors. The goal of the project is to transmit a small subset of this primary data for analysis and comparison with engine data stored in three data centres around the world. By identifying the early onset of problems, Rolls Royce hope to be able to lengthen the period between scheduled maintenance periods thus increasing profitability. The engine sensors will generate many Petabytes of data per year and decisions need to be taken in real-time as to how much data to analyse, how much to transmit for further analysis and how much to archive. Similar (or larger) data volumes will be generated by other high-throughput sensor experiments in fields such as environmental and earth observation, and of course human health-care monitoring.*
>
> *A second example from the field of bioinformatics …. It is estimated that human genome DNA contains around 3.2 Gbases which translates to only about a Gigabyte of information. However, when we add to this gene sequence data, data on the 100,000 or so translated proteins and the 32,000,000 amino acids, the relevant data volume expands to the order of 200 Gigabytes. If, in addition, we include X-ray structure measurements of these proteins, the data volume required expands dramatically to several Petabytes, assuming only one structure per protein. This volume expands yet again when we include data about the possible drug targets for each protein – to possibly as many as 1000 data sets per protein. And there is still another dimension of data required when genetic variations of the human genome are explored. To illustrate this bioinformatic data problem in another way, let us look at just one of the technologies involved in generating such data generation. Consider the production of X-ray data by the present generation of electron synchroton accelerators. At 3 seconds per image and 1,200 images per hour, each experimental station generates about 1 Terabyte of X-ray data per day.'*

For those unfamiliar with the terminology, a useful comparison is given:

'A Large Novel	*1 Mbyte*
The Bible	*5 Mbytes*
A Mozart Symphony (compressed)	*10 Mbytes*
OED on CD	*500 Mbytes*
Digital Movie (compressed)	*10 Gbytes*
Annual production of refereed journal literature (~20k journals; ~2M articles)	*1 Tbyte*
Library of Congress	*20 Tbytes*
The Internet Archive (10B pages) (From 1996 to 2002)	*100 Tbytes*
Annual production of information (print, film, optical & magnetic media)	*1500 Pbytes'*

The conclusion from the examples given is that *'e-Science data generated from sensors, satellites, high-performance computer simulations, high-throughput devices, scientific images and so on will soon dwarf all of the scientific data collected in the whole history of scientific exploration.'*

While these examples illustrate the volume of data being generated and collected within databases, they also, by implication, carry the message that the scientist depends on access to and re-utilisation of the data so collected to advance scientific knowledge.[23] It is here that the database right has the potential to cause problems, as has the practice of 'locking' data within private databases to which the key may be available only at a price.

4.1 Database Directive

In 1996 in the press release promulgated at the time that the Database Directive was agreed in Europe, the then Single Market Commissioner Mario Monti said, '*[t]his innovative and comprehensive measure will ensure an appropriate level of protection for database makers and investors throughout the EU.*' It was further stated in that document that *'the Directive strikes a balance between the interests of the manufacturers of databases and the legitimate interests of their users. Particular account has been taken of situations in which the extraction of contents of databases is required for … scientific research.*' [24] Two years later, Jans Gaster, one of its 'godfathers' (his own word) said of the right: *'evaluation of court decisions in database cases has shown that against all odds no particular problems have arisen in the Member States in applying the Directive in practice.*'[25] In 2003, by contrast, the Royal Society said of the database right in a report *Keeping Science Open: the effects of intellectual property policy on the conduct of science* (hereafter *Keeping Science Open*): *'the fair dealing exception under UK law, in line with the EC Directive permits only extraction and not re-utilisation. Re-utilisation is an essential part of scientific endeavour, and so this limitation does not address the scientific community's needs. The effects of these limitations are difficult to assess quantitatively but in our view they will in the longer term if vigorously enforced become a serious impediment to scientific research and hence to the national interest.*'[26]

So what is the reality? A key concern in relation to the Database Directive has been a perception that the sui generis right seems close to the grant of an intellectual property right in data and information per se, allowing only limited extractions for the purposes of non-commercial research. The result is that scientists may suffer restrictions on access to, and ability to re-use the raw data necessary for scientific progress.[27] In other words, it may represent a further, and significant, restriction on the public domain. But what then

might the exceptions under the Database Directive permit? The rights of the maker of the database are to prevent the extraction[28] and re-utilisation[29] of the whole or a substantial part[30] of the contents of the database. A scientist can extract and re-utilise an insubstantial part of the contents of the database without permission and substantial parts without permission for the purposes of non-commercial research[31] (assuming always that access can be gained). What then amounts to extraction and re-utilisation and what is insubstantial? Although a number of cases have been heard before courts of Member States concerning various aspects of the Database Directive, it was only in November 2004 that the European Court of Justice (ECJ) was given the opportunity to pass judgement on questions of the interpretation of aspects of the Directive referred to it in four cases.[32] Of these cases, one concerns details of horseracing fixtures,[33] the three others details of football league matches.[34]

It should be said at the outset that the judgements handed down by the ECJ have come as a surprise to many commentators. The worries that the rights might go so far as to attach to data themselves were not allayed by the line taken by the Advocate General Stix-Hackl (AG) in her opinions, published in June 2004. The ECJ often follows the opinion of the AG, but not in these cases. Instead the ECJ may have been driven by the concern that an over-broad interpretation of the Directive could result in the data themselves being the subject of property rights. In rejecting such a view, the ECJ, while taking a wide view of what falls under the definition of a database in the Directive, has both substantially narrowed which of those databases will qualify for the sui generis rights of extraction and re-utilisation, and set a high threshold on what will qualify as a substantial part for the purpose of infringement of the rights.

4.2 Database: definition

The definition of a database in the Directive refers to *'a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.*'[35] This covers both on and off-line databases. In one of the cases referred to the ECJ, *Fixtures Marketing v Organismoa Prognostikon Agnon Podosfairou (hereafter OPAP),*[36] the court, while confirming the breadth of the definition, emphasised that the independent materials must be systematically or methodically arranged and individually accessible and include technical means *'such as electronic, electromagnetic or electro optical processes'* to enable the retrieval of independent materials, or in the case of a non-electronic base, an index or table of contents.[37] Thus the fundamental requirements for qualification would appear to be the inclusion of a means of retrieving its constituent materials, and the works and data must be separable without the value of the contents being affected.[38]

But although a collection of data might fall within this definition, it does not follow that the maker will qualify for the sui generis right. Set against the purpose of the Directive which is to protect the investment in the obtaining, verification or presentation of data, the ECJ has sought to draw a distinction between the investment needed in the *creation* of the data per se, from the investment needed in the *creation of the database.* Only the maker of those databases or parts of databases in which substantial investment (in the form of financial resources and/or time, effort and energy[39]) has been expended in the obtaining, verification or presentation of the information will qualify for the sui generis right.

4.2.1 Obtaining

In referring to the investment needed in *obtaining* the contents of the database the ECJ said that this *'must be understood to refer to the resources used to seek out existing independent materials and collect them in the database and* **not** *[emphasis added] to resources used for the creation as such of independent materials.'*[40] So the materials to be placed in the database must already exist as independent materials. Only when those materials exist, and thereafter investment is expended in the collection of those materials, will this criterion be met. Of course it is far from easy to distinguish investment in the creation of material from investment in obtaining or collation in particular where the same body or person is responsible for the two activities. In *British Horseracing Board v William Hill* (hereafter BHB)[41] the database in question comprised inter alia information on over one million horses, and in particular pre-race information on races held in the UK. The latter information includes the name, place and date of the race concerned, the distance over which the race is to be run, the criteria for eligibility to enter the race, the date by which entries must be received, the entry fee payable and the amount of money the racecourse is to contribute to the prize money. When trying to find the line between the creation and the collation of the data, the ECJ said that the investment in the selection, for the purpose of organising horse racing, of the horses admitted to run in a race related to the *creation* of the data which make up the lists for those races which appear in the BHB database.[42] Thus the investment by BHB in that activity was not relevant when considering the criterion of obtaining.

4.2.2 Verification

As part of this 'creation' of the data, BHB (or at least the company who carried out this work) had set up a call centre manned by about 30 operators who record telephone calls entering horses in each race. This was followed by checks to ensure the identity and status of the person entering the horse. Thereafter, and to take part in the race, the trainer must confirm the horse's participation by telephone by declaring it the day before the race. The operators need to ascertain the horse can be authorised to run the race. A central computer allocates a saddle cloth number to each horse and determines the stall from which it will start with the final list of runners being published the day before the race.[43] This activity is aimed at verifying the accuracy of the data, but according to the ECJ such checks were made at the stage of creating the list for the race in question and thus constituted investment in the creation of the data and not in the verification of the contents of the database when in existence.[44]

4.2.2 Presentation

Finally on the investment required for the presentation of the contents of the database, this was raised in *OPAP*. In that case it was said that the expression *'investment in the presentation of the contents of the database'* concerns *'the resources used for the purpose of giving the database its function of processing information, that is to say those used for the systematic or methodical arrangement of the materials contained in that database and the organisation of their individual accessibility.*' The ECJ went on to say (rather unhelpfully) that the *'presentation of a football fixture list … is closely linked to the creation as such of the data which make up the list. It cannot therefore be considered to require investment independent of the investment in the creation of the data.*'[45]

4.3 Application of the ECJ's decision to scientific databases

So what then of the investment in scientific databases? It would seem that, at least initially, much investment will be directed towards the creation of data rather than the obtaining or collation of that which exists. Does this then mean that the criterion of

'obtaining' will not be met, at least by data creators? The overlap between creation and collation was acknowledged in the national court in BHB where it was recognised that there may be difficulty in separating the investment in creation from investment in collation particularly where one body is responsible for the two activities: *'As one would expect, effort put into creating the actual data which is subsequently collected together in the database is irrelevant. … On the other hand, the efforts which go into gathering all the data together, including the dates of fixtures, is relevant. … In practice, where one person both creates the underlying data and gathers it together, as BHB does, it may be difficult to draw a sharp dividing line between the two activities.'*[46] This was echoed by the AG in her opinion given on 4 June.[47] By comparison, the ECJ seems to have come down strongly in favour of the view that the two activities can be separated, and it is only by virtue of the relevant investment in the second, that the criterion for subsistence of the database right will be met.

But this raises further anomalies and complications. If the scientist who 'discovers' data for inclusion in a database is funded externally, for instance through one of the research councils in the UK, might that enable the funding council to qualify for the right through meeting the requirement of obtaining? Here it would seem that the link between the investment required for obtaining that data, and the actual creation of the data, might be broken. Even if it was thought the investment in those circumstances was directed towards the creation of the data, might it be possible to develop a model whereby the scientists create the data, but then a third party would invest in purchasing or, in the words of the Directive, obtaining that data once discovered? Such an approach would lead to the development of complex contractual matrices designed to ensure that someone can qualify for the right.

And what of verification and presentation of the contents? The sui generis right does not extend to computer programs used in the making or operation of a database,[48] so beyond that what might amount to verification and presentation in scientific databases? Much verification of data must go on prior to its inclusion in a database; how much then occurs once encompassed within it? Certainly issues of provenance are of great concern to scientists, especially when data moves between databases: where has data come from, what operations have been performed on the data, when and by whom? Could such investment where relevant to any particular database be considered as directed towards those types of verification activities that qualify the maker for the sui generis right? Presentation is equally as tricky. Clearly data within databases need to be presented in a form that is usable for other scientists and in so doing huge amounts of metadata can be generated. But to what extent might such presentation of the contents result from the operation of a computer program and thus be inapplicable in determining the subsistence of the right?

Suffice it to say many questions remain over the extent to which scientific databases might qualify for the sui generis right. Whereas at first blush it might have appeared that many might fall outwith the necessary criteria, most particularly because of the definition the ECJ has given to 'obtaining', it is far too early to argue that the contents of scientific databases fall into the public domain as a result of the ruling, however much that might benefit scientists and the progress of science.

4.4 Extraction and re-utilisation

Assuming that the investment in at least some scientific databases will meet the necessary criteria, the maker will qualify for the sui generis right to prevent unauthorised extraction

or re-utilisation of the whole or of a substantial part of the contents of the database.[49] The Database Directive defines 'extraction' as the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium *by any means or in any form,*[50] and 're-utilisation' to mean any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, by on-line or other forms of transmission.[51]

So what is the scope of these rights? Suppose a scientist received the required data from a third party, rather than extracting it directly from the database. Will she infringe? Is the re-utilisation right exhausted once the contents of a database have first been made available to the public, so any subsequent use would not infringe? A narrow interpretation would mean that science could continue unfettered by proprietary claims by database makers: so only direct extraction would infringe and the scientist could re-utilise the contents of the database freely after the first time those contents had been made available to the public. Conversely a broad definition might create problems: the extraction right would be infringed if a substantial part originally derived from the database, no matter the source for the end user; and takings from a copy of a database made available to the public would continue to infringe – in other words the right would not be exhausted.

In BHB the ECJ said that, as acts of unauthorised extraction and re-utilisation from a source other than the database concerned *'are liable … to prejudice the investment of the maker of the database, …direct access to the database was not a prerequisite'* for infringement of the rights.[52] Further, while the sui generis right does not extend to cover consultation of a database,[53] nonetheless the consent of the maker of the database to consultation does not entail exhaustion of the right. Thus, it does not matter whether the data are extracted or re-utilised directly from the database, or through the medium of a third party. If a substantial part of the contents of a protected base are in issue, then the rights of extraction and re-utilisation will be infringed no matter the source of the data. At first blush, and despite the seeming concern of the ECJ that the rights should not be over extensive, this might seem that the rights could extend to data per se. However, an important proviso is that the rights of extraction and re-utilisation only attach to a substantial part of the contents of a database. Extraction and re-utilisation of insubstantial parts do not infringe. What then is a substantial part? Or, to put it another way, what is an insubstantial part of a database?

4.5 Insubstantial/Substantial part

A typical question from a scientist might be what percentage of the contents of a database amounts to being insubstantial? One third of the contents? One half? In other words, how much falls into the public domain? But any answer is not as simple as a fixed figure. The test for determining what is substantial is both quantitative and qualitative.[54]

The ECJ has said that a substantial part evaluated *quantitatively* refers to the volume of data extracted from the database and must be assessed in relation to the volume of the contents of the whole of that database. If a user extracts and/or re-utilises a quantitatively significant part of the contents of a database whose creation required the deployment of substantial resources, the investment in the extracted or re-utilised part is proportionately equally substantial.[55] At first blush this might seem as if the part extracted must be judged by the size of the database as a whole. Indeed, in BHB despite

having suggested that the investment was in the creation of the data per se, the ECJ went on to comment that the extraction by William Hill of the names of the horses running in a particular race, the date, the time and/or name of the race and the name of the racecourse did not constitute a substantial part evaluated quantitatively – being only *'a very small proportion'* of the whole of the database. So how much is substantial? The ECJ did not quote a figure or percentage in this part of the judgement. However, when considering the test for when the '*repeated and systematic extraction and/or re-utilisation of insubstantial parts of the contents of the database*'[56] would infringe the sui generis right, the ECJ said that this measure *'prohibits acts of extraction … which could lead to the reconstitution of the database as a whole, or at the very least a substantial part of it … whether those acts were carried out with a view to the creation of another database or in the exercise of an activity other than the creation of a database.'*[57] If a similar test is used in relation to determining a quantitatively substantial part of a database, and when considering the size of scientific databases, it would appear a quantitative threshold would seldom be reached. Substantial surely must relate to something significantly over 50% even if it did not lead to reconstitution of the database. But it must be noted that the right only applies to that part of the database which has resulted from the relevant investment. So it may be that a database is very large – as the BHB base – but that part only qualifies for the right. The ECJ noted that the BHB database contained lists of horses – as well as the information on races. It could be that the list of horses, as opposed to the pre-race information, does qualify for the sui generis right as having been the subject of the correct investment in obtaining, verifying or presenting that data. If this were the case, then a quantitatively substantial part of that part of the database need not be a quantitatively substantial part of the database as a whole.

On the matter of a *qualitative* part of the database, this refers to '*the scale of the investment in the obtaining, verification or presentation of the contents … regardless of whether that represents a quantitatively substantial part of the general contents of the protected database.'* A quantitatively negligible part of the contents of a database may in fact represent, in terms of obtaining, verification or presentation, significant human, technical or financial investment.[58] This test would appear to require analysis of the investment that has been made in that part of the database that has been extracted. Conceptually, this throws up some difficulties, as it is not clear whether only the specific areas of a database in which there has been sufficient investment are protected by the sui generis right or whether adequate investment overall in a database or part of database results in the whole being protected. If the former is the case, then to what degree of granularity should this be taken?

If the qualitative test is approached from the first angle (i.e. only the specific areas in which there have been investment are protected and a crude granularity test applies) an oddity arises. What is strange is that the test for determining whether the sui generis right exists in the first place requires there to have been substantial investment in the obtaining, verification or presentation of the contents. If there has been, and the right exists, then it would appear to mean that any part is qualitatively a substantial part. Reverting to the BHB example: if it were found that those parts of the database containing information on the horses, but not concerned with the pre-race information, satisfy the tests for the subsistence of the right, then arguably any part of that section must be substantial. Certainly such a test may require going back to look at the relevant investment in that part, but nonetheless, it would appear that any segment or part of a protected database must be qualitatively a substantial part. That said, an instinctive reaction to such an argument is that it cannot be correct. If it were it might, in effect, result in data themselves being protected. As has been pointed out, that is not the

purpose of the sui generis right in the Database Directive and is a conclusion that the ECJ seemed keen to avoid.

Looking at the same situation from the perspective that the whole of a database or part of database is protected even where the relevant investment has been concentrated in certain areas and not in others, and applying a more refined test of granularity, an alternative conclusion may be reached. Despite the fact that a database or part of database has qualified for protection as a result of sufficient investment overall, the quantitatively insubstantial part of that database or part of database that has been extracted may not be a segment in which there has been significant investment. Take the example of a protected database or part of database which contains eight segments within the area that is protected. Perhaps large amounts of investment were required to obtain, verify and present the contents of segments one to three and a comparatively negligible amount was required for segments four to eight. When the eight segments are treated as a whole and the investment in each of them is averaged out over segments one to eight, the total investment is still sufficiently substantial to provide that the whole of that database or part of database is protected. If someone were to extract only segment five from that database or part of database (which is quantitatively an insubstantial part) then, despite the fact that the database or part of database as a whole qualifies for protection, an analysis of the areas of investment would show that the scale of investment in segment five was not significant and therefore insufficient for that segment to be treated as qualitatively substantial.

Which of these two approaches should be taken depends on the courts interpretation of the investment requirement and how it applies to different areas of a database. For a conclusive answer to be given in this area, further elaboration by the ECJ is required.

4.6 Database fencing

At first blush it would appear that the decision of the ECJ in the database cases is good news for scientists keen to re-utilise data and so progress science. However, as hinted above, it may be that complex contractual arrangements between interested parties as to sources of funding aimed at satisfying the criterion of 'obtaining' may soon be a feature of even the smallest venture into research which results in the generation of data to be incorporated into a database. Further difficulties may also arise, most notably in relation to databases the contents of which may, or may not have been generated in whole or in part through public funding, but which are held in private hands. Merely because the contents of the database do not attract the database right it does not thereby follow that scientists engaged in scientific research will be able to gain access to the contents of the database. Technological protection measures, contract and high tariffs are all likely to be common features surrounding those databases held in the private sector. That charges may be made for access to data contained within a database whether or not the database right subsists was acknowledged by the ECJ in BHB: *'The fact that a database can be consulted by third parties through someone who has authorisation for re-utilisation from the maker of the database does not… prevent the maker from recovering the costs of his investment. It is legitimate for the maker to charge a fee for the re-utilisation of the whole or a part of his database which reflects, inter alia, the prospect of subsequent consultation and thus guarantees him a sufficient return on his investment.*[59] Such access is likely to be at a high price, particularly for first comers, to reflect the lack of control thereafter. Of course, should a third party gain access to a substantial amount of data from a non-protected database, then the maker of the original database would have no recourse against that party. But the maker might as against the original taker on

whom contractual terms could be imposed in the event of a third party gaining access to the data with attendant (high) liability provisions. It also remains to be seen the extent to which the original creator of data might attempt to impose contractual conditions on a third party concerning use and on-going control of the data whether by that third party or by another with whom the third party might contract: in other words, use contract to exert control over downstream innovations. This is a question to which we will return below.

So what advice may be given to the scientist seeking to carry out research within the boundaries of the legal public domain concerned to avoid proprietary and contractual claims by the database maker? As can be seen, any answer is far from simple and may often lead to the comment 'it depends', which is hardly useful for a scientist whose concern is to progress science unfettered by legal niceties.

5. The Public Interest

This rather unsatisfactory state of affairs in relation to scientific research and the public domain has led to growing calls for a reassessment and rebalancing of the rationalisations for the grant of private property rights in intellectual products in the scientific sphere, and in particular for a greater weight to be given to the public interest. But that, in turn, begs the question as to what is meant by the public interest.

The public interest is not a unitary concept: different public interests will be relevant in different scenarios and need to be weighted differently depending on the circumstances. The task for any commentator, regulator or adjudicator is to find the appropriate *balance* of the numerous public interests that may exist in any given situation. Finding this balance will not only involve comparing the relative importance of one public interest to another but also involve the contemplation of the interconnection of public interests. By this is meant that merely identifying *a* public interest or a set of interests may not be sufficient to support a claim without considering the broader impact that these may have in turn on other public interest networks.[60]

But it is not only different public interests that require to be carefully calibrated. The distinction between public and private interests is another factor that requires consideration. The relationship between public and private interests will be relevant in ascertaining the appropriate balance or weight to be given to a selection of interests. In this the difference between public and private interests can be a complicated one. If private interests are dressed up as public and not recognised for what they truly are, a danger may arise that the state granted-monopoly becomes too heavily weighted in favour of the private interest. Thus when undertaking any balancing exercise of various interests it is important to recognise an interest for what it is (public or private) and to ensure that the interests that are being balanced against one another are like and like (i.e. private interest v private interest or public interest v public interest).

At a level of abstraction the intricacies have been described thus:

> *'Given the political and cultural framework of a particular society and the economic resources at its disposal, the public interest is the aggregate of the fundamental goals that the society seeks to achieve for all of its members – not for a majority of its members or for any large and powerful group, but for all of the people within the society. Considered separately, a society's goals are often in conflict with one another, and in that case there must be a balancing. The art of government consists of achieving a harmonious rather than a destructive balance among conflicting goals.'*[61]

In K*eeping Science Open*,[62] descending into a little more detail, it was acknowledged that IP rights can stimulate investment and *'aid the conversion of good science to tangible benefits'* noting at the same time that they could *'hinder the free exchange of ideas and information on which science thrives.*' So finding the correct balance, whilst essential, is no easy task. Some protection has to be given to afford the appropriate incentive. But giving too much can work against just the public interest goals in dissemination that policy seeks to further. In an era when increasing emphasis is placed on technology and economic imperatives as key drivers in reform of the law this may be of particular concern. *'Creation can be encouraged or discouraged, depending on the status assigned creators by society. Copyright, whose position has been complicated by the development of new technologies, is a decisive factor. The production policies of commercial distribution of works of the mind are determined primarily, and much more strictly than before by market principles. Accordingly, legal standards are being drafted or revised in order to adjust classical copyright laws to the new economic imperatives.'*[63]

There is no question that in recent years there has been a greater focus on the economic importance of copyright within the creative industries. They grew by an average of 6% per annum between 1997 and 2002 and accounted for 8% of Gross Value Added (GVA) in 2002.[64] Recognition of the economic importance of this sector of the economy has no doubt had some influence on the expansion of rights over recent years and, in the UK, led to the creation of an initiative supported by the Government designed to bring interested parties together to discuss ways in which activities such as music downloading and copying of films might be tackled.[65] The Database Directive grew from pressure exerted by a strong lobby of organisations concerned with the collection of data, who wished their investment to be protected. Certainly it has been argued that an immediate result of the introduction of protection was a growth in the database industry.[66] However, whether that has been maintained is questionable, with some arguing that *'European database production returned to predirective levels almost immediately'*[67] suggesting that legal protection, at least in this area, has not provided the appropriate incentive to greater investment.

But even against this backdrop, it is never easy to develop specific rules and there are many examples of where it might be questioned whether the balance of public interests has been correctly set, or where private interests may have masqueraded as public. Much concern has been expressed for instance in relation to recent changes made to the law of copyright with the introduction and implementation of the so-called Infosoc Directive,[68] an instrument drafted primarily in response to entertainment industry concerns about the ease with which digital music files and films could be copied over the Internet on a one-to-one basis.[69] This measure requires Member States to make it unlawful to circumvent technical protection measures designed to protect works protected by copyright,[70] has narrowed the research exemption from commercial to non-commercial research,[71] and would have appeared to validate the use of contract to shape the copyright monopoly.[72]

Thus right holders are apparently given the power to determine the contours of the public domain. While the legislation is in its infancy in the UK, having only been implemented in late 2003,[73] concern has already been expressed as to the potential impact in the fields of research and education.[74] It has been suggested that the power of the entertainment lobby in influencing the regulators to their own ends has entailed a balance that reflects private commercial concerns rather than the public interest in advancement.

Similarly, in relation to the database legislation, it has been questioned whether a proper public interest balancing exercise was undertaken by the EU regulators[75] and the charge laid that there was not *'an explicit analysis of the social or public interest costs for the proposed protection of investment.'*[76] As a result, database laws *'set a new milestone for mischief by virtually abolishing even the concept of a public domain and by abrogating the public interest components of intellectual property policymaking.'*[77] This may have come about, not only because there was a failure to balance public interests, but also because private interests may have been dressed up as public ones, merely being clothed as public to aid passage through the legislative process. Certainly the interests of the database maker are served by the grant of the sui generis right because the investment is protected.[78] This in turn has been argued to serve a greater public interest because of the resulting increase in production of databases which can in turn facilitate scientific development.[79] But when the very limited fair dealing exceptions to the database right[80] are considered, where extraction of a substantial part of the contents of a database is permitted only for the purposes of non-commercial research,[81] but a scientist is not then able to *re-utilise* the contents in the furtherance of scientific research,[82] questions arise as to what balance of interests were considered by the regulators minds when agreeing to such measures.

It is questionable whether the economic imperatives that drive the entertainment and commercial database industries and on which many recent reforms have been based are best suited to the progress of science. Even reverting to the most basic justifications, scientists, particularly where publicly funded and working within publicly funded institutions, do not need and generally have no interest in the 'incentive' copyright is designed to engender; proprietary claims made to the contents of database can only inhibit the free flow of information and thus seem the antithesis of the sharing, collaborative ethos on which science thrives. The power of both copyright and the database right as instruments of public policy should not be underestimated. The danger is where regulators fail to take into account the matrix of public and private interests that are affected by changes to the law.

6. Contract

Connecting both copyright and the database right, two issues arise in which the strength of the public interest in the advancement of science may be severely tested.[83] The first relates to the potential of contract to constrain the already limited public domain at the behest of the copyright owner or database maker. A question that has bedevilled the copyright community is the extent to which contract might be used to alter the contours of copyright, for instance by 'preventing' a third party from using a work for the purposes of non-commercial research or private study.[84] The current weight of academic opinion seems to suggest that, in the public interest, such contractual provisions should not be tolerated.[85] It can only be a matter of time before such clauses appear in 'database' contracts and no doubt subject to the same critical analysis. A second issue relates to the accessibility of works protected by copyright and data held within databases, where

technological protection measures are used to deny access except at a price, or subject to conditions. It may be that, beyond the limited control provided by competition law and licensing regimes,[86] those who hold the keys to the technological locks will, de facto, be the guardians of the public interest. Where those same guardians are driven by private commercial concerns, the public interest in the advancement of science, along with the public domain, could be sorely constricted in favour of shorter term economic imperatives.

7. Public Funding

The underlying theme of the discussion until now has been in relation to the progress of science as advanced within publicly funded institutions through the expertise of publicly funded research scientists. The copyright and database frameworks make only small concessions as to who is using protected information and for what purposes, and no distinction when considering sources of funding used to advance science. The diminishing public domain combined with the focus on economic aspects of the public interest may be essential to sustain a thriving commercial scientific sector and entertainment industry, but where public funding makes crucial advances possible, then questions arise as to the appropriateness of the current framework.

At its most basic, large numbers of academics who work in scientific research are not 'in it for the money' in the sense that they either look to, or depend upon a commercial return from their innovative work. They are, after all, paid by the State for their research prowess through their contracts with their employing institutions. Thus, in devising a legal framework that protects the commercial value of created works, prompts scientists to place fences around intellectual endeavour, and rewards those who gather what is created, the many and varied funding routes though which such creations may come into being are not being taken into account. Equally the traditional academic publishing model in the UK, geared as it is towards the rating of Universities and their departments in the Research Assessment Exercise,[87] and which encourages the assignation of copyright to publishers of research results, must be questioned as to its sense within the scientific research process.[88] It is a practice which only further endorses copyright fencing as discussed above.[89]

An extra layer of complication is added through the present pressure on academics and Universities to commercialise their results. Herein lie murky waters providing a good illustration of the difficulties in balancing both public interests and public and private interests. The funding crisis in Higher Education has been well documented in recent years.[90] There are many calls on public funding, of which research in Higher Education Institutions is only one. But what is in effect a decrease in public funding directed towards research has led to an increase in emphasis on the business aspect of universities. Where historically there has been an academic tradition which has valued purity of research, nowadays stress is laid on the commercialisation of the research that takes place: research that translates into IP rights in their many and varied forms.

Where public money has funded the creation of data and other research, there is an argument that the results should remain in the public domain, freely accessible to all for the maximum public benefit. But this brings its own problems. Particular concern has been expressed by a number of bodies, including The Royal Society,[91] in relation to results of research funded by the public purse being drawn into and exploited by the private sector to the detriment of both the advancement of science and of the public

interest.[92] How then, on the one hand, to ensure the results of publicly funded research are freely available to build upon, whilst on the other hand ensuring that those results in turn remain free? The Royal Society has suggested that where the fruits of publicly funded research are exploited by private commercial concerns, *'scientists ensure that any publicly funded data that are made available to private databases are done so non-exclusively, and that at least one repository of the information is liberal regarding access to and use and manipulation of the data.'*[93] Where there is research collaboration between business and universities the Lambert Review[94] (published in December 2003) recommended that '*the common starting point for negotiations on research collaboration terms should be that universities own any resulting IP, with industry free to negotiate licence terms to exploit it*'[95] a recommendation which seems to have found little favour with those same businesses.

It has been stated that IP regimes should be designed so that the subject matter of each one has relatively homogenous needs for protection.[96] There are strong arguments to suggest that where research and creation of data is publicly funded then different factors should drive protection, and a regime developed which takes account of public interest considerations that will ensure the accessibility and re-use of the results. It is to the responses to these pressures that we now turn.

8. Public and Private Initiatives in response to tensions within and amongst the 'three Ps'.

The limited public domain, the public interest that underpins the scientific research process within publicly funded institutions, together with substantial public funding combine to suggest that the current copyright and database frameworks largely geared towards economic imperatives may not be best suited to scientific progress. That is not to dispute the fact that scientific progress is important for the economic contribution that it can make, however indirectly. It is however to argue that when these three elements come together, the framework, as it currently stands, may be placing barriers in the way of what *could* be achieved in the way of progress. But, there is a paradox. A diminishing public domain, a balance of interests which might be argued to favour the private rather than the public sphere, and a policy which requires research institutions to commercialise results does not appear to have entailed a reduction in scientific outpouring.[97]

But two pressing issues beyond the 'three Ps' add weight to the argument that now is the time to reassess these factors in scientific research. The first is that the expansion of property rights is not without its adverse consequences. If property rights can be claimed in upstream material, then not only can re-use be prevented, but those same property rights can also be exerted in downstream or derivative works leading to many and varied ownership claims in scientific advances. With property rights, so conditions can be exerted whether they concern exploitation or further re-use. Property rights may also be used to extract ever-higher tolls to access works and data, ultimately being affordable only by the few. As discussed above, in many cases it might be public money that is expended on the initial creation or scientific discovery, which is then exploited in the private commercial arena. And there is always the unanswerable question: how much more research and scientific progress would be made if these domains had been at the forefront of regulators minds when developing a proprietary framework?[98]

The second issue relates directly to advances in technology. Whilst the power of technology to underpin scientific progress has long been acknowledged, the

comparatively recent development of computing capabilities has reinforced this relationship and consequently highlighted the potentially huge impact that technology may have on science and scientific research. From open standards in the development of the programs needed to house and to manipulate data in ways useful to science (open source); through the ease by which the latest scientific research results can be accessed and re-used when made freely available on the Internet (open access), to the freedom that can be exercised by cohorts of scientists seeking to explain scientific phenomena and thus develop understanding (open science), so technological advances have made and continue to promise opportunities never before imagined.

A greater awareness of what *might* be achieved appears to underpin a number of high profile moves to construct 'free' spaces more suited to the needs of current and future scientific users. The common thought behind these initiatives is that the legal construct forcing players to work within the confines of frameworks more suited to the private commercial sector has at least the potential to hinder scientific advances in the public sector. Many and varied attempts are now being made to construct spaces in which research may thrive in the interests of progress. Each of the initiatives has in common, to a greater or lesser extent, the 'three Ps'. They are attempts to expand the legal public domain; they rest on the belief that there is a weighty public interest in the dissemination and re-use of scientific research and data; each has an element of public funding, whether for the research per se, or though the funding of those who advocate the strategy.

8.1 Open Source

One well-known example of the use of contract to construct a public domain is 'opensource' software.[99] This refers to computer software whose source code is either one in which no person or organisation has any proprietary interest, or, more commonly, one which is protected by copyright but then distributed under an open-source licence such as the GNU General Public License (GPL).[100] This licence seeks to ensure that the source code will always be available to future developers to build upon. It came about apparently in response to Netscape making its browser freely accessible.[101] The founders wanted to prove to other big businesses that software development could continue apace without the assertion of property rights. Another movement, seeking to achieve similar ends is the Free Software Foundation.[102] Led by Richard Stallman, the motivation behind this initiative appears to be ethical rather than practical; a reaction to the claims (sometimes over extensive) made by software companies in relation to proprietary rights in software. The underlying philosophy appears to revolve around perceptions of the needs of society; in other words, the public interest: *'What does society need? It needs information that is truly available to its citizens---for example, programs that people can read, fix, adapt, and improve, not just operate. But what software owners typically deliver is a black box that we can't study or change. Society also needs freedom. When a program has an owner, the users lose freedom to control part of their own lives. And above all society needs to encourage the spirit of voluntary cooperation in its citizens. When software owners tell us that helping our neighbors in a natural way is ``piracy'', they pollute our society's civic spirit.'* [103] Much of the developmental work is carried out for 'free' by volunteers, many of whom are employed in publicly funded research Institutions.

8.2 Open Access

In response to the concerns over the academic publishing model described above, there has been in recent years a quiet revolution occurring in the academic library and

information sectors. Working with ever-tighter publicly funded budgets, these sectors have been setting up initiatives, generally called 'open access', aimed at making the results of scholarly work freely accessible to all. Key drivers appear to be the belief that it is not in the public interest that public funding which supports the researchers who write the journal articles should be expended a second time in buying back those same fruits of research from publishers, and that the weight of the public interest should be in favour of the results being freely available to other researchers to build upon in their work. Schemes include those led by Sparc Europe and DOAJ. Sparc Europe is an alliance of European research libraries, library organizations, and research institutions which '*advocate change in the scholarly communications market, support competition, and encourage new publishing models (in particular, open access models) that better serve the international researcher community.*' [104] DOAJ is the Directory of Open Access Journals whose aim '*is to increase the visibility and ease of use of open access scientific and scholarly journals thereby promoting their increased usage and impact.*'[105]

For some, open access means no more than making the results of research, in the form of published articles, available in a manner that can be accessed freely by those who need or wish to do so. There would seem to be no attempt to widen the legal public domain by limiting or confining any of the exclusive rights belonging to the owners. But open access is also used to denote different, broader, domains.[106] For example, the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities,[107] an initiative headed by a number of German academic institutions, defines open access as: '*a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.'* Any contributions made to the initiative (which can include original scientific research results, raw data, digitised pictures and scholarly multimedia material) must satisfy the following conditions:

'The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.'

Significant weight has been added to the movement as a result of an investigation into the scientific publishing industry[108] commissioned by the Wellcome Trust, an influential funding body in the UK which distributes public funding to biomedical sciences.[109] What the investigation found was a clash of priorities among those involved in the industry: commercial publishers who want to better their business position; libraries who struggle to buy journals on tight budgets; and researchers who want their research published in reputable journals. The conclusion drawn in the report was that *'[t]he current market structure does not operate in the long-term interests of the research community.'*[110]

Perhaps prompted by the reports produced for the Wellcome Trust, the House of Commons Committee on Science and Technology announced in 2003 that it would carry out an enquiry into '*access to journals within the scientific community, with particular reference to price and availability.'* The particular focus of the committee was on *'what measures are being taken in government, the publishing industry and academic institutions to ensure that researchers, teachers and students have access to the publications they need in order to carry out their work effectively.'*[111] The Committee published their report, '*Scientific Publications: Free for All?*' in July 2004. Within it, they made a number of recommendations among which were that:

'*all UK higher education institutions establish institutional repositories on which their published output can be stored and from which it can be read, free of charge, online'* and that '*Research Councils and other Government funders mandate their funded researchers to deposit a copy of all of their articles in this way.*'[112] The Government's response[113] in November 2004 largely rejected the advice of the Committee, saying that, while institutional and thematic repositories can play a significant role in the dissemination of research outputs, *'the Government has no present intention to mandate Research Council funded researchers to deposit a copy of their published material in institutional repositories.*'[114]

8.3 Open Science

Another example of the reaction to burgeoning IP rights is that of the human genome project.[115] Funded by an international consortium, the group responsible for the project agreed in 1996 that publicly funded laboratories involved in the project would release all data immediately as it was produced. In other words, this data would be placed 'in the public domain'. Now while there are ironies in this strategy (placing the information in the public domain allowed a rival company, Celera, to use this information for its own commercial purposes), it was decided that contract should not be used as a method of keeping the data free. The rationale was explained by John Sulston, co-founder of the project: *'[W]e all came to the conclusion that more would be lost than gained by our pursuing this line. Our role was to provide the data publicly for anyone to use…Had we claimed ownership, which is what any kind of restriction would amount to, we would be abrogating the very thing that we stood for.'*[116]

So publicly funded information was placed in the public domain (in its widest sense) by scientists who, following their sense of an ethical strategy, believed that the information they discovered was: *'of fundamental importance…which the great majority of people believe should be beyond private ownership.'* [117]

The Organisation for Economic Cooperation and Development has been drawn into the debate. The Committee for Scientific and Technological Policy held a meeting in 2004 during which they focussed on three issues high on the science and innovation policy agendas of their member countries.[118] These were: *'1) promotion of stronger relationships between science and innovation systems, including the changing role of intellectual property rights in stimulating knowledge creation and diffusion; 2) ensuring sustained development of human resources in science and technology; and 3) global-scale issues that call for enhanced international co operation in science and technology.'*[119] Following their deliberations the committee produced a Declaration on *Access to Research Data from Public Funding.*[120] This declaration, as the title suggests, focussed on access to scientific data. Couched in general language, it encourages the respective Governments to work towards the establishment of access regimes to data produced by publicly funded bodies. While the declaration does not go so far as to suggest that publicly funded data should be freely available to all, it does encourage giving thought to how interests in open access for the purposes of research will enhance the quality and productivity of science systems worldwide.

What is noticeable is that in most cases the initiatives emanate from one or a group of individuals. Only latterly have regulators, policy makers and funding bodies seen fit to suggest approaches that might support these efforts, and not all of these have been unqualified.

9. Conclusion

The current copyright and database right regimes are clearly not meeting the needs of scientific researchers. There is obvious concern at the extent to which inaccessibility of the fruits of research might hamper the next generation of researchers; over the inability to engage in 'new' research without meeting expensive property claims; and over the extent to which publicly funded research may end up being exploited for private ends. That there is dissatisfaction can be seen both at the grass roots level and, more recently, in the policy initiatives that have been taken to try and liberalise the fruits of existing scientific research for the newcomers. A different and to some extent more personal 'public interest' has emerged. Emanating from individuals engaged in the research process and now supported by public bodies, the agenda supporting the scientific research process is becoming increasingly liberal and is designed to support those who work within this sector. It is however questionable as to whether these pressures will result in changes to the legislative framework. The trend over recent years has been to increase intellectual property rights, and it is likely that this will continue with regulators focussed on the entertainment and database fields and their clamorous concerns.[121] But the quiet revolution in scientific research has gained a significant volume and seems likely to grow in strength and influence. The long term question will be as to whether it can survive despite the absence of a supportive legislative framework. We have confidence that it will.

• Dr Charlotte Waelde, Co-Director, AHRB Research Centre for Studies in Intellectual Property and Technology Law, School of Law, University of Edinburgh, and Mags McGinley, Legal Research Associate, Digital Curation Centre, University of Edinburgh.

[1] J Litman, "The Public Domain" (1990) 39 *Emory Law Journal* 965 at 968.

[2] Perhaps the much quoted statement made by the scientist Sir Isaac Newton over three hundred years ago concerning a leap forward in understanding still best sums up the derivative nature of creativity: *'What Descartes did was a good step. You have added much in several ways and especially in taking the colours of thin plates into philosophical consideration. If I have seen further it is by standing on the shoulders of Giants'* Newton to Hooke 5th February 1676; 1,416 *The Correspondence of Isaac Newton Turnbull and others* eds. Cambridge University Press Vol. II 1676-1687. Note that some argue that there is a hint of irony in this statement. Hooke was physically very small, and one line of thought is that Newton was acknowledging using work of others who came before him, but not that of Hooke.

[3] Copyright Designs and Patents Act 1988 (hereafter CDPA) s 12. Directive 96/9 of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases (hereafter the Database Directive) Article 10.1

[4] Agreement on Trade-related Aspects of Intellectual Property Rights 1994 (TRIPS) Article 9(2).

[5] It is the investment in the obtaining, verification and/or presentation of data that is protected. Database Directive Article 7.1.

[6] The originality threshold in the UK is low: the author must have exerted the requisite labour, skill or effort in producing the work *Ladbroke* v *William Hill* [1964] 1 All ER 465, 469.

[7] Database Directive Article 7.1.

[8] CDPA s 16(3)(a). Database Directive Article 7.1.

[9] CDPA s 29. Database Directive Article 9.b. However, re-utilisation of the contents of the database is not so permitted

[10] P Samuelson, "Preserving the Positive Functions of the Public Domain in Science" (2003) 2 *Data Science Journal* 192

[11] J Litman, "The Public Domain" (1990) 39 *Emory Law Journal* 965 at 967. For more on the public domain see "The Public Domain" (2003) 66 No 1 & 2 *Law & Contemp. Probs.* (the whole issue) available @: <http://www.law.duke.edu/journals/lcp/lcptoc66winterspring2003.htm > and in particular J Boyle, "The Second Enclosure Movement and the Construction of the Public Domain" (2003) 66 *Law & Contemp. Probs.* 33 available @: <http://www.law.duke.edu/shell/cite.pl?66+Law+&+Contemp.+Probs.+33+(WinterSpring+2003)>

[12] CDPA s 12. The term of protection has always been keenly debated and proposals to increase the term hotly contested. For example in a speech delivered in the House of Commons on 5th February 1841, Thomas Macaulay took exception to the proposal to extend the copyright in books to lifetime of the author plus 60 years: *'The principle of copyright is this. It is a tax on readers for the purpose of giving a bounty to writers. The tax is an exceedingly bad one; it is a tax on one of the most innocent and most salutary of human pleasures; and never let us forget, that a tax on innocent pleasures is a premium on vicious pleasures'.*

[13] Termed by some '*domain public payante'*.

[14] CDPA s 77.

[15] In the joint copyright sense. CDPA s 10(1).

[16] Said by some to be '*science that requires massive capital investment but which is also expected to yield very significant results'.* http://www.wordiq.com/definition/Big_Science

[17] Credit and responsibility seem closely intertwined in the scientific field mainly as a result of fraudulent scientific claims. For instance the International Committee of Medical Journal Editors requires that authorship credit for articles in that journal must be based on substantial contributions to 1.conception and design, or analysis and interpretation of data; and 2.drafting the article or revising it critically for important intellectual content; and 3.final approval to be published. Further, when a large, multi-centre group has conducted the work, the group should identify the individuals who accept direct responsibility for the manuscript. http://www.icmje.org/#author

[18] See the claims made by C McSherry in *Who Owns Academic Work? Battling for Control of Intellectual Property* (2001) at 219.

[19] M Biagioli. "Rights or Rewards? Changing Frameworks of Scientific Authorship", in M Biagioli and P Galison (eds), *Scientific Authorship. Credit and Intellectual Property in Science* (2003) at 257.

[20] Ibid at 254.

[21] International Council for Science, *Scientific Data and Information – A Report of the CSPR Assessment Panel,* December 2004. Available @: <http://www.icsu.org/Gestion/img/ICSU_DOC_DOWNLOAD/551_DD_FILE_PAA_Data_and_Information.pdf>

[22] Available @: < http://www.ecs.soton.ac.uk/~ajgh/DataDeluge(final).pdf >

[23] '*The next generation of research breakthroughs will rely upon new ways of handling the immense amounts of data that are being produced by modern research methods and equipment, such as telescopes, particle accelerators, genome sequencers and biological imagers….Similar developments are having an impact in the arts and humanities, and in the social sciences.'* A Vision for Research, Research Councils UK, December 2003.

[24] @ < http://europa.eu.int/rapid/pressReleasesAction.do?reference=IP/96/171&format=HTML &aged=1&language=EN&guiLanguage=en >

[25] J Gaster, *Communications Law* Vol 5 No. 3, 200 at 97.

[26] Available @ < http://www.royalsoc.ac.uk/displaypagedoc.asp?id=11403> at 23.

[27] International Council for Science, *Scientific Data and Information – A Report of the CSPR Assessment Panel,* December 2004. Available @:
<http://www.icsu.org/Gestion/img/ICSU_DOC_DOWNLOAD/551_DD_FILE_PAA_Data_and_Information.pdf>

[28] Database Directive Article 7.2(a).

[29] Database Directive Article 7.2(b).

[30] Database Directive Articles 7.2 (a) and (b).

[31] Database Directive Article 9(b).

[32] *British Horseracing Board* v *William Hill* C-203/02 (from the Court of Appeal, England and Wales); *Fixtures Marketing Ltd.* v *Svenska Spel* AB C-338/02 (from the Hogsta Domstol, Sweden); *Fixtures Marketing Ltd* v *OY Veikkaus Ab C-46/02* (from the Vantaan Darajaoikeus, Finland); and *Fixtures Marketing Ltd.* v *Organismoa Prognostikon Agnon Podosfairou (OPAP)* C-444/02) (from the Monomeles Protodikio Athinion, Greece). The opinions were delivered on 8 June 2004 and the judgements by the ECJ on 4 November 2004.

[33] *British Horseracing Board* v *William Hill* C-203/02 (hereafter BHB).

[34] *Fixtures Marketing Ltd.* v *Svenska Spel AB* C-338/02; *Fixtures Marketing Ltd* v *OY Veikkaus Ab* C-46/02; and *Fixtures Marketing Ltd.* v *Organismoa Prognostikon Agnon Podosfairou (OPAP)* C-444/02.

[35] Database Directive Article 1(2).

[36] OPAP Case C-444/02.

[37] Ibid para 30.

[38] Ibid para 32.

[39] Database Directive Recital 40.

[40] *Fixtures Marketing Ltd.* v *Organismoa Prognostikon Agnon Podosfairou (OPAP)* C-444/02

[41] BHB Case C-203/02.

[42] Ibid para 38.

[43] Ibid para 14.

[44] Ibid para 40.

[45] OPAP Case C-444/02 Para 51. Also *Fixtures Marketing Ltd.* v *Svenska Spel AB* C-338/02 para 35.

[46] *British Horseracing Board Limited and Others* v *William Hill Organisation Limited* [2001] 2 C.M.L.R. 12. Laddie. J.

[47] *British Horseracing Board Ltd* v *William Hill Organisation Limited*, Case 203/02 Opinion of Advocate General Stix-Hackl delivered on 8 June 2004 paras 37-49.

[48] Database Directive Recital 23.

[49] There is also the right the right to prevent repeated and systematic extraction and or re-utilisation of insubstantial parts of the database contents. Database Directive Article 7.5.

[50] Database Directive Article 7.2.(a).

[51] Database Directive Article 7.2.(b).

[52] BHB Case C-203/02 para 53.

[53] Ibid para 54.

[54] Database Directive Article 8(1).

[55] BHB Case C-203/02 para 70.

[56] Database Directive Article 7(5). The purpose of this provision stems from a concern to *'ensure that the lack of protection of the insubstantial parts does not lead to their being repeatedly and systematically extracted and/or re-utilised'.* Common Position (EC) No 20/95 adopted by Council on 10 July 1995 (OJ 1995 C 288 at 14.

[57] BHB Case C-203/02 para 87

[58] Ibid para 71.

[59] BHB para 57.

[60] G Laurie and C Waelde, "Privacy, Property and Personalities: Whatever Happened to the Public Interest?" Forthcoming. Currently on file with the authors. At 22

[61] Ringer, "Authors' Rights in the Electronic Age: Beyond the Copyright Act of 1976" (1981) 1 *Loyola Entertainment Law Journal* 1.

[62] Ibid n. 26

[63] UNESCO, Third Medium Term Plan (1990-1995), adopted in November 1989, para 195.

[64] Statistics taken from Department of Culture, Media and Sport Creative Industries Economic Estimates Statistical Bulletin, August 2004 available @:
<http://www.culture.gov.uk/global/research/statistics_outputs/creative_industries_eco_est.htm >

[65] For example the Westminster Media Forum Seminar on Intellectual Property and Rights Ownership, addressed by the Rt Hon Estelle Morris MP, London, 9 December 2004.

[66] J Reinbothe, untitled paper presented at WIPO's *Protection of Databases Workshop* on 16 September

1999, Geneva.>

[67] S Maurer, P Hugenholtz, H Onsrud, "Europe's Database Experiment" (2001) Vol 294 *Science* 789-790 at 790.

[68] European Parliament and Council Directive 2001/29/EC of 22 May 2001, OJ 2001, L167/10 (hereafter the Infosoc Directive).

[69] C Waelde and H MacQueen "From Entertainment to Education: the Scope of Copyright?" (2004) 3 *Intellectual Property Quarterly* 259 at 270.

[70] Infosoc Directive Articles 6 and 7.

[71] Infosoc Directive Article 5.3 (a).

[72] Infosoc Directive Article 6.4.

[73] Implemented in the UK in the Copyright and Related Rights Regulations 2003, SI 2003/2498.

[74] See generally C Waelde and H MacQueen "From Entertainment to Education: the Scope of Copyright?" n 69 above.

[75] Chapman, "A Human Rights Perspective on Intellectual Property, Scientific Progress and Access to the Benefits of Science". Available on the WIPO website @:
<http://www.wipo.int/tk/en/activities/1998/humanrights/papers/pdf/chapman.pdf >

[76] ibid

[77] J Reichman and P Samuelson, "Intellectual Property Rights in Data?" (1997) 50 *Vanderbilt Law Review* 51 at 164.

[78] Whether it provides an incentive is a different matter as discussed infra.

[79] For a paper questioning whether this has in fact happened see n 67 above.

[80] Database Directive Article 9. Note that these are optional for Member States and not mandatory.

[81] Database Directive Article 9.b.

[82] ibid.

[83] Database makers/vendors are protecting their data with both clever business strategies and technology. N. Gallini and S. Scotchmer, "Intellectual Property: When is it the best incentive system?" (2002) Vol 2 *Innovation Policy and the Economy* 51

[84] CDPA s 29(1).

[85] That does not of course prevent such provisions appearing on a regular basis in copyright licences.

[86] *RTE and ITP* v *Commission* Joined Cases C-241/91 P and C-242/91 P [1995] ECR I-743*; IMS Health* v *NDC Health* Case 418/01; Technology Transfer Block Exemption Regulation Commission Regulation (EC) No 772/2004 of 27 April 2004 on the application of Article 81(3) of the Treaty to categories of technology transfer agreements

[87] < http://www.rae.ac.uk >

[88] This model assumes that the academic owns the copyright in the research output – an assumption that would appear not to be in accordance with the legislative provisions of the CDPA s 11(2), although arguably accepted by custom and practice.

[89] For an 'out of the box' discussion attended by academics, librarians, policy makers, representatives of collecting societies and publishers as to what the contours of a 'Copyright Free World in Higher Education' might look like see < http://www.law.ed.ac.uk/ahrb/publications/online/ipfreeworld.pdf >

[90] See e.g.< http://education.guardian.co.uk/egweekly/story/0,5500,784805,00.html. > Certainly research is to receive £120 million more in funding in 2005, but (as with other areas) this does not keep pace with demand nor the potential of what could be done if more were available.

[91] Ibid n 26.

[92] In addition to the initiatives discussed here see also US Public Access to Science Act H.R. 2613, a proposal designed to make research funded by the American government exempt from copyright protection; an Australian government support of open access to scientific results <http://www.dest.gov.au/Ministers/Media/McGauran/2003/10/mcg002221003.asp. > For a useful repository of both public and private initiatives see:
<http://www.earlham.edu/~peters/fos/2003_10_19_fosblogarchive.html >

[93] Ibid n 26 Ch 5

[94] Lambert Review of Business-University Collaboration. Final Report December 2003. ISBN: 0-947819-76-2.

[95] Recommendation 4.1.

[96] N. Gallini and S. Scotchmer, "Intellectual Property: When is it the best incentive system?" (2002) Vol 2 *Innovation Policy and the Economy* 51 at 53

[97] Ibid n. 22

[98] For arguments that intellectual property rights do inhibit the progress of science see e.g. J Reichman and P Uhlir, "A Contractually Reconstructed Research Commons for Scientific Data in a Highly Protectionist Intellectual Property Environment" (2003) 66 *Law & Contemp. Probs.* 315 available @:
<http://www.law.duke.edu/shell/cite.pl?66+Law+&+Contemp.+Probs.+315+(WinterSpring+2003)>, M Heller and R Eisenberg, "Can Patents Deter Innovation? The Anticommons in Biomedical Research" (1998) Vol 280 *Science* 698 available @: <http://www.sciencemag.org/cgi/content/full/280/5364/698> and *Keeping Science Open.*

[99] See < http://en.wikipedia.org/wiki/Open_source. > Licences must meet ten conditions in order to be considered open source licenses. For further information see <http://www.opensource.org/docs/definition.php. >

[100] < http://www.gnu.org/copyleft/gpl.html >

[101] < http://www.oscommerce.com/about/philosophy >

[102] < http://www.gnu.org >

[103] < http://www.gnu.org/philosophy/why-free.html >

[104] < http://www.sparceurope.org/ >

[105] < http://www.doaj.org/. > On 11 January 2005 there were 1408 journals in the directory and 62788 articles.

[106] Trends towards open access, in a variety of forms, are gaining ground in a number of countries around the world. Many initiatives are led by those directly involved. E.g. DAREnet, a joint initiative by the Dutch universities, which facilitates digital access to the results of their research. Available @: <http://www.darenet.nl/en/toon >; WSIS Declaration of Principles & Plan of Action <http://www.itu.int/wsis/documents/doc_multi-en-1161%7C1160.asp >; call by twenty-five Nobel Prize winning scientists on the U.S. government to make all taxpayer-funded research papers freely available. < http://www.usatoday.com/news/science/2004-08-29-free-research_x.htm >

[107] < http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html >

[108] *An Economic Analysis of Scientific Research Publishing* available @: <http://www.wellcome.ac.uk/assets/wtd003182.pdf >

[109] The Wellcome Trust says that it planned to distribute more than £400 million (approximately $662 million) in funding to biomedical sciences in the year 2003–2004. < http://www.wellcome.ac.uk/ >

[110] Ibid n. 108 at iv. See also the second report produced by the Wellcome Trust looking at costs and business models. Available at < http://www.wellcome.ac.uk/assets/wtd003184.pdf >

[111] < http://www.parliament.uk/parliamentary_committees/science_and_technology_committee/scitech111203a.cfm >

[112] Available @: <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf>

[113] Available @: <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/1200/120006.htm >

[114] Ibid.

[115] J Sulston "Intellectual Property and the Human Genome" in P Drahos and R Mayne (eds) *Global Intellectual Property Rights: Knowledge, Access and Development* (2002) 61-73

[116] Ibid at 66.

[117] Ibid at 72.

[118]Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique

[119] < http://www.oecd.org/document/0,2340,en_2649_34487_25998799_1_1_1_1,00.html >

[120] Declaration on *Access to Research Data from Public Funding* adopted on 30 January 2004 in Paris.

[121] See for example the recent UK Film Council Report *Film theft in the UK*
Anti-Piracy Task Force: an analysis and recommendations for action. December 2004. Available from Communications Department, UK Film Council, 10 Little Portland Street, London W1W 7JG.