Monograph: 2010 - Emerging Information Technologies (II) (published jointly with Novática*)
Guest Editors: Alonso Álvarez-García, Heinz Brüggemann, Victor-Amadeo Bañuls-Silva, and Gregorio Martín-Quetglas

2 Presentation: A Cloudy Services Future — Alonso Álvarez-García, Heinz Brüggemann, Victor-Amadeo Bañuls-Silva, and Gregorio Martín-Quetglas
6 Towards Enterprise 3.0 — Nicolas Bertet, Agustin Chacón-Espinu, and Francisco-Javier Torres-Noguero
12 Adopting Semantic Technologies for the Enterprise for Knowledge Management — Maria-Eugenia Beltrán-Jaunsaras and Javier Carbonell-Pérez

18 Trends in Distributed Curatorial Technology to Manage Data Deluge in a Networked World — Sophia B. Liu
25 SOA and Cloud Technologies, Two Pieces of the Same Puzzle — Luis Rodero-Merino, Luis-Miguel Vaquero-González, Juan-Antonio Cáceres-Expósito, and Juan-José Hierro-Sureda
30 Virtualization as Support for SOA and Cloud Computing — Carlos Pérez-Conde and Víctor Díaz-Villanueva
36 Competitiveness and Technological Development in Europe – The Next 10 Years: Interview with Neelie Kroes, Vice President of the European Commission and European Digital Agenda Commissioner — UPGRADE Editorial Team
40 Cloud Computing and Web Research: Interview with Ricardo Baeza-Yates, VP of Yahoo! Research for Europe, Middle East and Latin America — Gregorio Martin-Quetglas
42 The Present and Future of the Web: Interview with Bert Bos, co-inventor of CSS (Cascading Style Sheets) — Encarna Quesada-Ruiz
47 Cloud Computing and Software Services: Interview with Simon Davies, Solution Architect specialising on Microsoft’s Windows Azure Cloud Platform — Gregorio Martin-Quetglas

UPNET (UPGRADE European NETwork)

50 From Novática (ATI, Spain)
Open Access
Intellectual Property in Publishing and Research: Open Access in Biotechnology, Life Sciences, and Software — Fionn Murtagh

CEPIS NEWS

65 Selected CEPIS News — Fiona Fanning

* This monograph will be also published in Spanish (full version printed: summary, abstracts, and some articles online) by Novática, journal of the Spanish CEPIS society ATI (Asociación de Técnicos de Informática) at <http://www.ati.es/novatica/>.
Open Access

Intellectual Property in Publishing and Research: Open Access in Biotechnology, Life Sciences, and Software

Fionn Murtagh

© Novática, 2010

This paper will be published, in Spanish, by Novática, Novática, <http://www.ati.es/novatica>, a founding member of UPENET, is a bimonthly journal published by the Spanish CEPIS society ATI (Asociación de Técnicos de Informática – Association of Computer Professionals).

We show some of the parallels between three sectors: (i) research, in particular research and scholarly publishing; (ii) software, and the diversity of its rights management ecosystem; and (iii) biotechnology, with its restricted intellectual property ecosystem and declining levels of innovation. A core aspect of the research process is to be found in scholarly publishing. Some of the most advanced forms of scholarly, research publishing, relating to publishing practices including citation, are evident in biotechnology and the life sciences. Motivation for Open Access, for example, is far and away the most pronounced in the life sciences. We look at how this ties in with the evolution of the management, generally, of intellectual property. Computing, with its basis in computational reasoning, can and should play a central role in this evolution. In fact we can already discern a future view of pharmaceuticals as a new form of software.

Keywords: Computational Reasoning, ICT, Innovation, Intellectual Property Rights, Open Access, Publishing, Research Process, Software.

1. Introduction

1.1 Convergence of Computing and the Life Sciences

The term ICT (Information and Communications Technologies) will be used here to denote computing-in-the-large, including telecoms, networking, hardware including photonics, design, applications, and so on. We will use the term computing to be somewhat more restrictive, involving software, software engineering, applications – all areas built on computational thinking [1].

ICT has a crucial role to play in drug creation; in early warning and rapid response systems when faced with pandemics; and in developing new, responsive technologies to face the threats of epidemics and pandemics. But ICT goes way further than just an infrastructure for the life sciences or for any other area. In fact, as noted in [2], "Software and networks can express ideas in the conventional written sense as well as create (express) infrastructures that allow ideas to circulate in novel and unexpected ways." Software and networks are therefore simultaneously product, process and environment.
1.2 Open Access Background

Open Access relates to postprints. The diverse Open Access policies of varied research funding organisations internationally are logged by the UK coordination group, SHERPA [3]. Open Access policies in regard to “peer-reviewed research outputs” cover self-archiving, including recommended or mandated; what and when; publishing; and data.

From the publication process, consider for the moment preprints. Consider license choices for a new submission to the widely used arXiv preprint server, [http://arxiv.org/help/license]: (i) granting arXiv.org a non-exclusive and irrevocable license to distribute; (ii) Creative Commons Attribution; (iii) Creative Commons Attribution-Noncommercial-ShareAlike; (iv) Creative Commons Public Domain Declaration. In all there are six license choices available from Creative Commons, [http://creativecommons.org/licenses].

The modern researcher has become knowledgeable (at least in default) of a lot of the legalities of research publication outputs. "The Creative Commons licenses allow authors to grant the use of their work in about a dozen different ways – that is, the license itself comes in versions. One can, for instance, require attribution, prohibit commercial exploitation, allow derivative or modified works to be made and circulated, or some combination of all of these. These different combinations actually create different licenses each of which grants intellectual-property rights under slightly different conditions.” [2]. This is hardly new of course. As Kelty [2] notes: "the success of the Internet as a market infrastructure and as a [unique] entity comes in part because of the recognition of the limitations of the intellectual-property system – and Free Software in the 1990s was the main experimental arena for trying out alternatives".

One issue we address in this article is to trace out how free software is linked indirectly to Open Access. That nexus has everything to do with intellectual property.

Consider the licenses pertaining to software, a range of which are discussed in [2]: copyright, open source, free software, GNU General Public License (“often referred to as a beautiful, clever, powerful ‘back’ of intellectual property law”, [2]), freeware, shareware, and of course patents.

Having touched on software with its many aspects that give rise to its intellectual property let us now switch to analogous concerns in the area of publishing. The system of publication can include "citation, use, reuse, borrowing, building on, plagiarizing, copying, emulating", "reframing text, problems, arguments", and "artistic transformation" [2].

1.3 About This Article

While largely focused on Open Access in this article, we find it curious that journal publications have so much of the limelight, to the exclusion of data and software, and also other forms of publication like books. It is rather reminiscent of the Sherlock Holmes and Watson barking dog story (i.e., the curious incident in the night was that the dog did not bark). Aspects such as this are grist for the mill in this article.

We will locate Open Access within the broader context of, to begin with, digital rights, and then intellectual property generally.

We begin with scholarly research as an industrial or business process. A somewhat far-fetched assertion, the reader will perhaps claim? Not at all. Clearly scholarly research rubs up against the publishing sector immediately and directly. Education too can be viewed as a business. We seek to go further here, and to show how research processes – and not just the published or protected outputs of research – can influence or even be converted into business processes.

The themes we will reflect on in the following sections of this paper are as follows:

1. Research processes and models are needed now by business and industry. The heart and soul of research processes and models: rights management.
2. A fundamental outcome and product of research: the article in a journal. A basic manifestation of research: citation.
3. The general intellectual property picture in the life sciences and how it is changing.
4. Open Access, in particular in the life sciences.
5. Pharmaceuticals as the new software.

2 The Quiet Revolution in Research and its Profound Impact on Society

In this section we advance the view that research process is itself a most important output of research. Furthermore a research process has potential business and industrial impact.

2.1 Research, Impact and the Changing University

The standard view nowadays is to see (blue-skies, fundamental, curiosity-driven, academic, ...) research as threatened by undue demands arising from business and industry. Research funding agencies use a "strategic review" and focus on "impact" as ways in which business and industry can exert greater influence on research. This curious view has, in our opinion, come about by seeing a research camp, and a research exploitation camp, as clearly demarcated. We find a clear separation between research and its exploitation to be quite unrealistic.

In a radical change of perspective we claim that the increasing threat to research by business and industry is quite an incorrect view of what is happening. Instead, we have business and industry, mostly with research funding and related government agencies as mouthpiece, saying, as it were: please express the processes, methodologies and models of research in such as way that we can take them on board, and replace or invigorate our crusted and dated work methods.

The university as an institution is changing fast. Kelty [2] has the following neat characterization of this: "New questions about competition and profitability have emerged from the massive proliferation of hybrid commercial and academic forms, forms that bring with them different traditions of sharing, credit, reputation, control, creation, and dissemination of knowledge and products that require it. The new economic demands on the univer-
sity – all too easily labeled neoliberalization or corporatization – mirror changing demands on industry that it come to look more like universities, that is, that is give away more, circulate more, and cooperate more. ...

The proliferation of hybrid commercial-academic forms in an era of modifiability and reusability, among the debris of standards, standards processes, and new experiments in intellectual property, results in a playing field with a thousand different games, all of which revolve around renewed experimentation with coordination, collaboration, adaptability, design, evolution, gaming, playing, worlds and worlding.

2.2 Accountability and Deliverables of Funded Research

It has been predicted that the American Reinvestment and Recovery Act (ARRA, the Obama stimulus package) will create about 402,000 jobs for one year, through an additional investment of $20 billion in research. Accountability though is critical: "Within 2 years, the public will want to be informed about the impact of the stimulus on the economic recovery. Were the estimates accurate? How can they be validated? And, in the longer term, what were the impacts of the reinvestment strategy on scientific knowledge, economic growth, and job creation?" [4]. Research nowadays is no longer blue-skies, if it ever was.

Accountability of research and measurement of impact is high on the agenda everywhere now. Is this the end of altruistic, socially-beneficial research as it once was – or as it was once perceived? No, – research has become very different in the way it is structured and organized but, as we will show, hugely more relevant and more beneficial.

A first important point to make is that appearances are not always what they are held up to be. For example the 1968 social upheaval in France and in many other countries is often viewed in terms of youth versus authority. But as we exemplified in a recent article [5], in fact the May events in France served, at least in our example, to bring the academy closer to industry. Internships became popular as a result of 1968, with university students seeing internships now as an important indication of practical relevance of their work.

Research was once seen as a source of collective knowledge. Here the early Royal Society in 17th century London comes to mind, with so much care and attention given to collecting and describing practical life- and work-based experience of varied occupations. Indeed Kelty [2] sees the Royal Society then as an equivalent of free software now, or the emergence of the publishing industry in earlier centuries. Apache and Linux, for instance, were and remain "collective social experiments", and involve "experimental expansion of the Internet".

Look too at how [6] the modern scientific period was begun in the century before then, on the back of voyaging and trading. The 16th Dutch, in particular, gave us "a new global economy", in fact, an "information economy", that "marked the beginnings of science as we know it" (in the words of a review by Lisa Jardine). Cook [6] describes very well how science, the information economy, and globalization, were related in their beginnings in the early mercantile period.

2.3 Performance Indicators of Funded Research

Research now has moved beyond these goals. Research has become the production of intellectual goods. Based on intellectual goods, research has become the production of intellectual properties.

We should understand intellectual property (IP) quite generally because there are areas where there is considerable evolution right now relative to traditional contexts. Consider e.g. research publishing itself. The research journal is usually traced to the Philosophical Transactions and the Journal des Savants (later renamed Journal des Savants), both started in 1665. Nowadays research journals dominate the scene for research and scholarly output. There are other outlets, including conference proceedings, books (monographs are at issue here), and in our more enlightened and also evolving times, data sets, software, and various forms of multimedia object. The trend towards reproducible research [7] is only starting and has implications for data and software. We have therefore a changing landscape of research product and process, with attendant rights and responsibilities.

It would be nice to think that intellectual goods and IP in the research arena are showing the way forward for digital rights management. Maybe in the future the general roll-out to society at large of rights management in research will appear to us as reminiscent of how the web all started because physicists wanted to exchange documents.

So research has become the production of intellectual goods, some of which result in intellectual properties.

We have yet to arrive at a way to measure in the absolute what these intellectual goods bring to society. But it is clear that research has become the mainstay of any modern economy.

It is interesting to note though that powerful ways of measuring relative impact are now core to the way that research is carried out. Moreover the model of how research is comparatively assessed has become a mirror for many other areas of social activity. Examples include government processes; and business practice. Let us draw out just how the impact of research is assessed in a relative and not absolute way.

2.3 Research Benchmarking, Rankings, Scorecards, Audits

Benchmarking has come to the fore, as a new managerial approach. This is not a command and control, centralized, hierarchical way of organizing research or other social activities. Instead we have league tables, rankings, scorecards – a way to organize activities that is distributed and autonomous. This is not so much government any more as governance.

Universities everywhere look towards the Shanghai Jiao Tong University rankings; student satisfaction surveys in universities and institutes of learning are reported on in the newspapers; countries have their scorecards for telecoms infrastructure, for Business Expenditure on Research and Development (BERD), and so on; the UK had the Research Assessment Ex-
ercise and will have the Research Evaluation Framework, Australia will have its Excellence in Research for Australia; and so on and on. Within companies there are examples, and within governments lots of examples too, of rankings and scorecards (cf. [8]). The rise of governance fits well with competitiveness through benchmarking. It fits really well with modern development models of the knowledge economy and of the nurturing of human capital.

There is an insightful article by Bruno on this [9] (see also her book, [10]), that links the new benchmarking-based, and competitiveness-oriented governance to the Lisbon strategy in Europe, – growing the European knowledge economy and associated human capital through research. Furthermore, governance has a clear managerial principle underpinning it, where government could be said to have had an underpinning legal principle.

Since the 1990s, there have been enormous changes – behind the scenes in research and the way it is structured on the global level. The mix of competition and cooperation is now at the core of research. Already significant aspects of government and of business are in tow. Government and business have taken on board how research is carried out and copied the benchmarking models established in research.

It is true that there is still evolution in how impact of research is measured in absolute terms. But in relative terms, what we see is that the impact of research is captured through competitiveness and innovation. These in turn have been taken on board as essential organizational drivers by government and business.

Just like the web’s take-off through Marc Andreesen’s Mosaic browser in early 1993, once again research innovation – this time through organisation and a new dynamic – is changing society completely.

Going on from this, it will be very interesting to see how rights management and the myriad IP processes that have evolved and developed in research are generalized to all of society in the coming years. We have no doubt they will be, if only because research has established itself as the major source of innovation, change and renewal in our time.

Biotech illustrates this very well: "science creates novel business requirements that cannot be fulfilled with ... existing approaches" [11]. Pisano [11] discusses why biotechnology came about in the way it did. Biotechnology, based on biotech start-ups, was based in biology, creating large protein molecule drugs using genetic engineering. Venture capital did not go far enough. So funding partnerships were established with major pharmaceutical companies "in exchange for future product rights". Thus "a market for know-how emerged as a response to a gap in the market for capital" [11]. The R&D (research and development) arms of pharmaceutical companies were "largely dominated by chemists", creating drugs through "small chemically synthesized molecules". So a natural question was why did the major pharmaceutical companies not do R&D in-house, or why did they not deal directly with universities? They did in fact continue with both. There were other reasons for their partnerships with only partially venture capital funded biotech start-ups. This included the fact that large protein drugs cannot be broken down in the human gut, so they need to be administered other than orally such as by injection. "Collaborations with specialized [biotech] companies [therefore] provided a vehicle for exploring the terrain with a fairly limited and reversible commitment" [11]. Other explanations for the raison d’être of biotech companies include their work being too applied for universities and too early-stage for many pharmaceutical companies.

3 Rights Management Underlying Journal Publishing

3.1 The Journal Form of Publication

The move towards Open Access involves many actors. In this we adopt an author perspective rather than the perspectives of many other stakeholders including librarian, printer, publisher, organisation or institute, consortium (which increasingly funds purchase of research publications), funding agency, state agencies, councils or other bodies (responsible for funding and support), and so on and on.

The form of research publishing of interest to us here is sometimes called scholarly publishing. A curious aspect of interest in this area is that journals alone are mostly under discussion in Open Access, but far less so conference proceedings, or indeed books. One could add that other forms of research output are off the radar screen including data, software, media products such as film (movie), economic reports, etc. While data (cf. [5]), conference proceedings, and other research outputs postdated the journal in their origins, the same cannot be said for the book. Be that as it may, we consider now what lessons journal publishing have for us in regard to intellectual property.

3.2 Origins of the Research Journal and Associated Property Rights

For Guédon [12], scholarly journal publishing has become "big business" and this is seen as a problem. However his insights into early intellectual property in this area are valuable. As is well known, what are commonly regarded as the first scientific journals came about in 1665. These were the Philosophical Transactions (Phil. Trans.) of the Royal Society of London in early 1665, and the Journal des Scavants in the same year. For Guédon, the latter was more journalistic. Phil. Trans. The however "aimed at creating a public record of original contributions to knowledge ... the Parisian publication followed novelty while the London journal was helping to validate originality". They "introduced clarity and transparency in the process of establishing innovative claims in natural philosophy, and, as a result, it began to play a role not unlike that of a patent office for scientific ideas. The purpose was to tame and police 'scientific paternity' and priority controversies and intellectual polemics ... [to] help create internal rules of behavior" in the scientific community. Phil. Trans. was established and edited by Henry Oldenburg (c. 1619 to 1677), Secretary of the Royal Society.

The primary function of the Phil.
Trans. was not (as such) general communication between peers, nor dissemination to non-scientists, but instead “a public registry of discoveries”. In a way reminiscent of how a comet is named after its first (in modern times, of course) observer (see also [13]), “Republic of Science claimed the right to grant intellectual property to scientific ‘authors’ and Phil. Trans. was its instrument of choice”.

Through the argument of Guédon [12], this first scientific journal was a means for creating intellectual property. Johns [14] points to how suspicion and mistrust accompanied early publishing so that Oldenburg was also targeting an “innovative use of print technology” [12]. A third related objective of Oldenburg was to have London “do for science what Paris was striving to do for taste: ... become the universal ... arbiter of natural knowledge”. These aims together still effect us, “Oldenburg’s long shadow”. (Another reason again, which we will not further pursue, was for Oldenburg to increase his earnings.)

So journals originating in Oldenburg’s prototypical scientific – indeed scholarly – journals are to be seen “as registers of intellectual property whose functions are close to that of a land register”. Guédon also locates time limited intellectual property in terms of parallels of king versus parliament: time-limiting the rights that one has on one’s own scholarly output is due in his view to a democratizing influence.

Further: “The design of a scientific periodical, far from primarily aiming at disseminating knowledge, really seeks to reinforce property rights over ideas; intellectual property and authors were not legal concepts designed to protect writers – they were invented for the printers’ or stationers’ benefits.”

So, interestingly, Guédon locates the motivation for intellectual property not with the author but with those producing the printed, published form of product. We have seen Oldenburg’s motives but now the publishers enter into the picture. Note though that the creation of intellectual property is just being carried out in various ways that reinforce each other to give us today’s rules of the game in intellectual property as it relates to scholarly publishing. “Oldenburg’s shadow stretches far indeed; from the original intent to simplify the management of scientific intellectual property to the subsequent possibility of evaluating scientists’ performance”.

The context as much as the author-scientist led to this particular form of intellectual property. What we find convincing enough in the role of the printer or stationer (cf. Section 3.4 below) is that the article, or collection of articles in a journal, or other forms of printed product (pamphlet, treatise), became the most important paradigm. Other possible forms did not. Examples could include: the experiment; or the table of experimental data; or catalogs or inventories. Note that the latter have become extremely important in, e.g., observational data based sciences such as astronomy, or the processed or derived data based life sciences. What is interesting is that the publication – and primarily the article – remains the really dominant form of research output.

### 3.3 From Rights and Intellectual Property to Quantitative Performance Evaluation

Guédon [12] continues: “Probably more clearly than ever before, the digital, networked, world reveals that refereed science publishing is closely intertwined with the evaluation of scientists; in fact, it stands closer to evaluation than to communication. As the digital world grows in importance, the evaluation question will loom ever larger.”

This is for the following reasons. Firstly, as already noted: “Journals are rather inadequate when it comes to communicating quickly and efficiently; they are much better at validating and evaluating the relative worth of scientific authors. They are adequate to preserve the memory of science over the long haul (several centuries in the best of circumstances).”

Alternatively: “Journals can help to evaluate; they are not evaluation per se.”

A mechanism pushing this trend even further is “the distinction between article and journal”.

This leads to the real nub of the current developments that see Open Access increasingly on the agenda, together with citation-based impact assessment: “In fact, in the digital world, selection through usage becomes the dominant question to be addressed and solved. In other words, the peer review process tends to extend to the whole community almost immediately.”

At this point, however, we disagree with the anti-elitist, anti-publisher, and anti-commercial lines of reasoning of [12]. The seemingly radical digital democratization (part and parcel of Open Access and citation-based impact assessment of research) is nothing other than a reduction to pure numerical properties. Any single metric is not appropriate for multivariate reality. In fact the superficiality of this view leads to the necessity of evaluation being once again invented in the new, 100% vox pop world. How is not indicated by Guédon – that is left to the “learned societies, university administrators, and, of course librarians”. This is not very satisfactory since longer term sustainability cannot be guaranteed.

### 3.4 The Much Maligned Publisher

Sometimes in Open Access debates, including in Guédon [12], the publishers, driven by the profit motive, are portrayed in self-aggrandizing terms. (Chapter titles in [12] include: “Scholarly Journals as a New Publishers’ Eldorado”, and “How Commercial Publishers Managed to Harness the Digital Revolution into a Counterrevolution.”) Historically [15], and still now legally, the publisher is the one most likely to take the rap, over and above the author. Before the French revolution the trades of bookseller, publisher and printer were very much overlapping. Napoleon played a big role in the development of the modern bookshop. Through a regulatory decree in early 1810 a clear distinction was introduced between bookseller and printer. The bookseller was therefore distinguished from the printer from 1810 onwards in France. Both bookseller and printer were distinguished from author. This allowed the printer to come under the watchful eye of the state, in order to monitor non-conformism or sedition. The number of permitted printers was fixed.
whereas the number of booksellers was open to free competition. The 1810 decree held sway until 1870. By then the bookstore was fully established.

In this article, we are engaged in a very different discussion compared to the (by now) more traditional librarian versus publisher one, or Open Access versus commercial publishing. We are far more interested in current trends. We are also keenly interested in the diverse cultural aspects of research and their related business and industrial sectors.

3.5 Origins of Citing in Scholarly Publishing
The self-promoting of one’s research wares was already usual enough in the 17th century. The French royal librarian, Naudé "perhaps harmlessly noted that the noise (bruit) and the vogue of the books of one time or place may not be the same in others" (p. 374).

"As early as 1611, a certain jurist named Stephanus held that a university ought to accord superior precedence to an academic in view of good publications over mere seniority... The 1640/48 statutes of Frankfurt a.d.O. made an exception concerning precedence, normally set by seniority, for, among other reasons, a man with a big name (magni alicuius nominis vir). ... The [Prussian] ministry urged the professors at Halle, the Prussian flagship university at the time, to publish 'more in accord with the taste of the time' (sich mehr nach dem Geschmack der Zeit zu richten). That was nothing less than a ministerial call to make fashionable noise." (p. 375).

"... Göttingen historian of universities, Michaelis [in the 1760s and 1770s], noted at the time that the worst way to make an academic appointment was 'according to the praise of the learned periodicals'. Who was this reviewer after all? "If this Anonymus should be a beginner, an uninformed about the matter, a student, a degenerated master; a friend, an auditor of the author or, indeed, upon removing the critic's mask, the author himself" – what then? "

3.6 Examples of Change in Research Process: Teams, Movement away from the Monograph

Coming now to authorship, there has been an overall shift towards teamwork in authorship, clearly enough led by the life sciences and by "big science". In the highly cited journal, Nature, it has been noted [17] that "almost all original research papers have multiple authors". Furthermore (in 2008), "So far this year... Nature has published only six single-author papers, out of a total of some 700". What is however very clear is that mathematics or statistics or related methodology work rarely ever appears in Nature. While social networks of scientists have become very important, notes Whitfield [20], nonetheless there is room still for a counter-current in scholarly activity: "... however finely honed scientists' team-building strategies become, there will always be room for the solo effort. In 1963, Derek de Solla Price, the father of authorship-network studies, noted that if the trends of that time persisted, single-author papers in chemistry would be extinct by 1980. In fact, many branches of science seem destined to get ever closer to that point but never reach it."

With online availability now of the scholarly literature it appears, counter-intuitively, that relatively fewer, rather than more, papers are being cited and, by implication, read. Evans [17] finds that: "as more journal issues came online, the articles referenced tended to be more recent, fewer journals and articles were cited, and more of those citations were to fewer journals and articles". Evans continues: "The forced browsing of print archives may have stretched scientists and scholars to anchor findings deeply into past and present scholarship. Searching online is more efficient and following hyperlinks quickly puts researchers in touch with prevailing opinion, but this may accelerate consensus and narrow the range of findings and ideas built upon." Again notwithstanding the 34 million articles used in this study [17], it is clear that there are major divides between, say, mathematical methodology and large teams and consortia in the life and physical sciences.

In a commentary on the Evans article [17], Couzin [21] refers to "herd behavior among authors" in scholarly publishing. Couzin concludes by pointing to how this trend "may lead to easier consensus and less active debate in academia". We would draw the conclusion that mathematical thinking – if only because it lends itself poorly to the particular way that "prevailing opinion" and acceleration of "consensus" are forced by how we now carry out research – is of great importance for innovation and new thinking.

The change in research and scholarly publishing has implications for book publishing. Evans [17] notes this: "The move to online science appears to represent one more step on the path initiated by the much earlier shift from the contextualized monograph, like Newton’s Principia or Darwin’s Origin of Species, to the modern research article. The Principia and Origin, each produced over the course of more than a decade, not only were engaged in current debates, but wove their propositions into conversation with astronomers, geometers, and naturalists from centuries past. As 21st-century scientists and scholars use online searching and hyperlinking to frame and publish their arguments more efficiently, they weave them into a more focused – and more narrow – past and present."

Undue focus and narrowness, and "herd behavior", are at odds with the Hayashi (Chikio Hayashi, 1918–2002) and Benzécri (Jean-Paul Benzécri) vision of science, as we have argued in [5]. Fortunately, this vision of science has not lost its sharp edge and its innovative potential for our times.

3.7 Citations in an Online Information Context
Evans [17] notes how "... the shift from browsing in print to searching online facilitates avoidance of older and less relevant literature." He continues: "Moreover, hyperlinking through an online archive puts experts in touch with consensus about what is the most important prior work – what work is broadly discussed and referenced. With both strategies, experts online bypass many of the marginally related articles that print researchers skim. If online researchers can more easily find prevailing opinion, they are more likely to follow it, leading to more citations referencing fewer articles."
Consensus: "Agents view others’ choices as relevant information – a signal of quality – and factor them into their own reading and citation selections. By enabling scientists to quickly reach and converge with prevailing opinion, electronic journals hasten scientific consensus."

And, through "haste" or the speeded up process: "Findings and ideas that do not become consensus quickly will be forgotten quickly."

The net effect of the dominance of recency, consensus and relevance in the sense that now usually holds is a research culture that is very foreign to, e.g., the mathematician for whom all findings and results are constitutive of an overall intellectual edifice.

3.8 Divergence between Disciplines

Citation cultures differ greatly across fields and sub-fields. Citation rates for a typical article are as follows (mid-2008, [22]):
- in mathematics and computer science < 1
- in chemistry and physics: 3
- in clinical medicine: little higher
- in life sciences: > 6

In discussing the evolution and direction of the biosciences, Pisano [11] notes that the journals Science and Nature are among the "institutions of basic science", and that "it would be hard to imagine what the life sciences would look like today without" a range of universities, organizations, projects – and these journals.

The following critical view of the legacy of US Bayh-Dole legislation, the University Small Business Patent Procedures Act introduced in 1980 and changing the universities’ role from public good to individual benefit, comes from [23].

"In fact, it was the life sciences – in particular, biotechnology – that started universities down the slippery commercial slope in the first place. Even before the Bayh-Dole Act, pharmaceutical companies were eagerly trolling campuses, looking for projects to finance. After the law was passed, they stepped up their efforts, but now with renewed zeal for keeping potential trade secrets from competitors.

While patients have benefited from the growing supply of new medications, the universities have obtained patents not only for the actual substances but also for the processes and methods used to make them, potentially hampering discovery of even more beneficial treatments."

Wellsing [24] notes the differences and maybe divergence across disciplines: "The nature of the debate over the use of IP varies substantially by industry sector. Some sectors, such as ICT, want a rapid transfer of know-how into products. The Creative Industries want the transfer of skilled people often with the capacity to work in multidisciplinary teams. In contrast the Life Sciences and Pharmaceutical sectors frequently need strong IP protection to justify further staged investment and to secure returns on any products successfully placed in the market."

Universities and higher education institutes have differed in their approaches to intellectual property. A particular case in point is how "professor’s privilege" in regard to intellectual property ownership has been ceded in most countries by now to the professor’s institute (see, e.g., [25]).

Wellsing [24] comes quickly down on the side of benefits accruing, via the university, to the economy, society and environment, and not just financial; and furthermore that this benefit has to be measured. "Strategies which deliver a faster, more flexible interface with business and more co-operation between actors in the innovation system seem likely to return the greatest dividends to the UK economy. At the same time, by taking this approach, universities should be able to generate greater levels of contract funding for translational research from their commercial partners."

4 Intellectual Property in the Life Sciences: Focus on Patents

4.1 Role of Life Sciences in Patenting

From [26] it is clearest if we continue this quotation in full in this section: "Life sciences contribute the lion’s share of patent revenues at leading US universities, outpacing contributions from physical science, information technology, and other fields. Life scientists also supply most of the inventions patented by the 10 technologically strongest US institutions, according to a study provided to The Scientist by research and consulting company, CHI Research. In 2001, the top 10 US universities generated 689 life science patents, compared with 263 in information technology and 245 in all other technology categories.

These life science discoveries contribute 42% to the overall tech strength portfolio at these universities, compared with a 35% contribution by information technology and a 23% contribution from all other technology categories. The technical strength of a university indicates the extent to which its inventions influence others in their fields, as determined by how often its patents are cited by other patents.

Life sciences contribute more than do other fields to institutions’ technological and research strengths because ‘companies in the life sciences have a lot of experience in commercializing ideas from universities,’ says John Fraser, director of technology transfer at Florida State University (FSU). "[Companies in] engineering and other fields do not have this same experience because they tend to develop ideas themselves. The engineering fields also are more process oriented, while the life sciences are more product oriented. That makes a difference."

4.2 Crisis of Innovation in Pharmaceuticals

As noted in [27] dating from early 2009: "The number of new drugs approved by the FDA (Food and Drug Administration) has fallen linearly from 53 in 1996 to 17 in 2007, the same number as in 1983. A slight bump upwards to 21 approvals in 2008 includes 3 drugs eventually approved on reconsideration and 3 radiocontrast agents. In other words, it doesn’t buck the trend. This coincides with downward pressure on drug pricing. The growth in prescription drug sales – 10% of the $286.3 billion US healthcare budget in 2007 – had plummeted even before the present crisis: generics now account for roughly 60% of the market and are rapidly growing in market share.

Pharma has reacted by shedding jobs in the US – more than 100,000
over the past 5 years – and moving research to join drug production in lower cost economies overseas. Anticipated future revenue for the industry has shifted dramatically towards Asia. The current crisis is likely to accelerate these trends.”

4.3 Exclusivity of Exploitation by Pharmaceutical Companies and its Loss

“The pharmaceutical sector is one of the main users of the patent system.” This is quite an interesting assertion, from the Executive Summary of the Final Report on the EC (European Commission) Competition Inquiry [28]. And: “Contrary to what might be assumed, blockbuster medicines’ patent portfolios show a steady rise in patent applications throughout the life cycle of a product. Occasionally they show an even steeper increase at the end of the protection period conferred by the first patent” [29]. The EC Preliminary Report of November 2008 [29] had the appearance of being somewhat less sympathetic to the originating companies (i.e. the companies that initially researched and then protected their invention), compared to the Final Report [28].

The EC initiated an enquiry into competitiveness in the pharmaceutical sector in January 2008, leading to a Preliminary Report in November 2008 [29] and the Final Report in July 2009 [28]. Motivations were “delays in the entry of generic medicines to the market and the apparent decline in innovation as measured by the number of new medicines coming to the market”. Prescription medicines for human use only were at issue.

The European Union market for medication was worth over EUR 138 billion ex factory, and EUR 214 billion retail, in 2007 [30]. The latter value corresponds to 2% of European Union GDP (gross domestic produce). The decline in novel medicines reaching the market was evidenced by the total number of on average 40 between 1995 and 1999, and in the 2000s this has averaged 27 [31]. The effective patent protection period from product launch to the first generic launch is over 14 years in 2009, up from 10 years in 2000 [31]. After about two years, generic companies have the effect of taking the erstwhile price for the drug down by 40%. Generic market share varies a lot between countries: in Poland it is highest at 56%, whereas it is lowest in Ireland (13%), France (15%) and Finland (16%) [29].

Companies in the “highly regulated and R&D driven” pharmaceutical sector are considered as (i) originating, driving research and the regulatory process including clinical trials, and protecting products through time-limited patent protection; and (ii) generic companies, entering markets post-patent protection with equivalent products.

Interestingly, 35% of the compounds taken on board by originator companies are acquired or licensed from third party, often small or medium sized biotech, companies.

<table>
<thead>
<tr>
<th>Publishing: traditional, still current university research; pharma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Need and use:</strong> researcher (including PhD student) patient</td>
</tr>
<tr>
<td><strong>Purchase decision:</strong> librarian (influenced strongly by researcher) prescribing doctor (pharmacist in some cases)</td>
</tr>
<tr>
<td><strong>Costs paid:</strong> university central funds covered by block university funding from the state (contributing to library consortium) insurance contributory by patient, or by employer or by state (possibly national scheme)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Current trends in the Open Access context; open questions for pharma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Need and use:</strong> researcher patient</td>
</tr>
<tr>
<td><strong>Purchase decision:</strong> researcher prescribing doctor (pharmacist in some cases)</td>
</tr>
<tr>
<td><strong>Costs paid:</strong> grant covered by competitive research funding award (ultimately by the state) New economic and pharma models, supported by new business models and financial instruments?</td>
</tr>
</tbody>
</table>

Table 1: Publishing: the traditional and the co-existing newer model in academic, university research, and for comparison, the pharmaceutical situation.
There are close parallels between academic publishing and the pharmaceutical and medical sectors. The users in both cases, researchers, and patients, usually have the products—journals, say, or medication—paid for by others such as librarians or insurance. Table 1 contains a schematic comparison taking into account in the lower panel the need now for cost subsidization for many journals of immediately Open Access availability of content (we raise the question of what the life sciences, and in particular pharmaceuticals, parallels could be). The table shows how there can be by and large price insensitivity in university research from the researchers and maybe even limited price sensitivity by the librarians. On the pharma side, there is great lack of price sensitivity by patient and prescribing doctor.

The pharma and medication industry distinguishes between primary and secondary patenting. The latter include "different dosage forms, the production process or...particular pharmaceutical formulations" [28]. Further, in order to prolong the exploitation of (time limited) protected pharmaceutical products use is made of numerous defensive patents, referred to as "patent clusters" or "patents thickets". The Final Report notes that "individual medicines are protected by up to nearly 100 product-specific patent families, which can lead to up to 1,300 patents and/or pending patent applications...". The Final Report notes some "awareness by patent holders that some of their patents might not be strong", and also that defensive patenting has the clear benefit of slowing down the ability of generic products to come forward and take over the role of the originating company's product. A further mechanism to create legal, intellectual property ownership uncertainty is through (legitimate) "divisional patent" application, which can split an initial patent application. However the original patent parent is not changed as such. In March 2009 the European Patent Office introduced time limitations on filing of divisional patents.

Litigation that is started by the originator, patent-owning company is also a widespread practice. Patent litigation between originator and generic companies is also on the increase—by a factor of four between 2000 and 2007, the Final Report notes. During litigation secondary patents (see above for definition) were mainly invoked, rather than primary patents. In regard to oppositions filed prior to the granting of patents, generic companies "almost exclusively opposed secondary patents", and "prevailed in approximately 60% of final decisions rendered by the EPO" (European Patent Office). Mostly such decisions took more than two years.

Apart from patenting, there are other important and related considerations in terms of market barriers to entry, or market negotiating positions. All medicines "need to obtain a marketing authorization and [mostly] pricing and reimbursement status before they can be put on the market". Objections or litigation by originator companies in regard to marketing authorization rarely worked, but slowed down generic company entry to market by about four months.

From a sample of medicines, the EC Final Report found that 40%, which had lost exclusivity, were the basis of originator companies launching second generation or follow-on medicines. A majority—60%—of litigation cases between originator and generic companies were related to medicines that had moved from being primary medicines to secondary ones. As a ploy against generic companies, "in order to successfully launch a second generation medicine, originator companies undertake intensive marketing efforts with the aim of switching a substantial number of the patients to the new medicine prior to the market entry of a generic version of the first generation product. If they succeed, the probability that generic companies will be able to gain a significant share of the market decreases significantly." The EC Final Report notes that this process could well be in the interest of the patient, by offering new therapeutic or other innovation.

What we have been describing here are product "life-cycle instruments" used by originating companies in their own interest over and above generic companies. Competition between originating companies is also looked into in the Final Report, as is competition or the lack of it (including lack of competition between generic companies) in off-patent or non-exclusive market segments. Background assumptions are clearly enough the following.

"Intellectual property rights promote dynamic competition by encouraging undertakings to invest in developing new or improved products and processes. So does competition by putting pressure on undertakings to innovate." The developing of new or improved products and processes can—not always of course—take the form of "increasing concentration among (large) originator companies as well as the acquisition of biotech companies".

Outcomes of the EC Final Report are recommendations for faster and more efficient intellectual property system, including a (European) Community patent and a unified litigation system. As opposed to a unified Community patent, a patent currently obtained through the European Patent Office is "a bundle of national patents that need to be enforced/challenged nationally" [31]. Promised for 2010 is a new consideration "on the use of personalized medicines and 'omics' technologies in pharmaceutical research and development and on the possible need for new...instruments to support them". At issue here are "new technologies like pharmacogenomics and patient-specific modelling and disease simulators" for personalized medicine. The prospects for personalization of health and medical care based on the association of human illness with genetic make-up is not always viewed with optimism though [32]. This is due to the complexity of common diseases like cancer and diabetes, linked to multiple genetic variations.

5 Open Access in Life Sciences, why and wherefore?

5.1 The Changing Culture of Publishing and Open Access

In Section 3 we pointed to trends that are already discernible and that would give rise to highly significant consequences for science and engineering, as well as technology and industry. Now in this section we will draw out one further facet of this.

Science as we know it can be dated
to 1665 and major aspects of the research game have remained pretty close to how things were done those few hundred years ago.

Increasingly now, Open Access is gaining ground. What this usually implies is that journal or conference proceedings articles ought to be put in institutional or discipline repositories six months (or immediately) following publication. A curious aspect here is that there are other outputs of research which do not figure, – at least not usually. Among these are book publishing, published data and openly accessible, perhaps open source, software.

It is interesting to probe further and see who or what is leading the charge towards Open Access. In our view, this movement is led clearly enough by the medical and health sectors. Firstly it is in these areas that there is potentially a clear resonance with the marketplace for commercialization, and with an expressed need for application and deployment. Secondly, medical and healthcare research is one opening – one vantage point – in regard to the large life sciences sector. Thirdly and most of all, it is organisations like the National Institutes of Health that have gone furthest, most quickly too, in introducing Open Access policies. Between April and May 2008, for example, NIH policy of mandated depositing in PubMed Central became finalized [33].

We discuss in Section 6 how Open Access facilitates search, retrieval and fusion of information. Curiously – certainly from the computing and other perspectives – only journals are covered by the NIH policy. Conference proceedings are not mentioned, nor are software nor data, and certainly not books.

5.2 Wide Diversity of Digital Rights in Research Data and Information

5.2.1 Two Case Studies

Let us look at two recent legal cases relating to intellectual property in research. It is possible that their import will be reversed or modified. It goes without saying too that they relate to the given legal jurisdictions. Notwithstanding these caveats they are interesting, and revealing, in the first case described, of some current perspectives on rights and duties of university employees relative to the fruits of their work. In the second case described, the rights and duties are with respect to digital data.

5.2.2 "The Researcher Owns the Invention, Not the Employer"

The Federal Court of Australia case of the University of Western Australia versus Gray (No 20) [2008] FCA 498, see [34], resulted in a judgment by Justice French on 17 April 2008 which caused great difficulty for university support including incubation of research. Bruce Gray, who held a professorial position in surgery at UWA, developed a novel system for delivering liver cancer treatment. He filed patents in his own name and otherwise, but not linked to UWA, and transferred them to a company he had established. He was sued by UWA, as his employer, who claimed rights to the technology and shares in his company. There were other associated cases, including by Prof. Gray’s company against him for not disclosing his invention link with UWA. The judgment was in his favor: Justice French found that there was a contractual duty to carry out research but no contractual "duty to invent". The implications of this judgment are far reaching. For a start if a company has a contractual link with a university, this UWA vs Gray case means that this may well be non-enforceable relative to the researcher. There are also implications for the content of the contract held by a university employee, in order to enforce invention as a stated duty. (UWA at the time was to appeal the judgment. Meanwhile, Justice French became Chief Justice of the High Court of Australia.)

The implications of this judgment are profound. For a start, any agreement of a commercial nature entered into by a university could be completely ignored by those faculty or students who are concerned by it. Most contracts, after all, do not mandate inventing.

5.2.3 Ownership of Database Content

Prof. Dr. Ulrich Knoop, Albert-Ludwigs-Universität, Freiburg, compiled over 2 years and published a list of 1100 selected German poems between 1720 and 1933. Directmedia Publishing GmbH produced a CD of 1000 selected poems between 1720 and 1900. 856 of them were shared with the Knoop list. The university claimed that the company had infringed ownership of the database. The German Bundesgerichtshof (Federal Court) found in favor of the university on 24 May 2008. See e.g. [35].

The implications of this judgment are potentially major. Poetry or other cultural outputs, especially going back in time as here, are public goods, let us assume. So the list, order, ranking, arrangement, alone is marked as owned.

5.3 Gathering Storm Over Cloud Computing

The burgeoning problems of jurisdiction, of rights and of privacy in the web 2.0 world are sketched out in [36]. This discussion of the interface between computing and the law is again taken up by [37] who notes a wide range of issues arising in the context of cloud computing, including the following summary.

"It is likely that in the future cloud computing will go both ways. There will be one market in which services are cheap or free and advertising-supported and in which the customer takes nearly all the risk. This is the typical model for cloud computing agreements at the moment. But there will also be a market for cloud computing in which the service provider takes more of the risk in return for more payment."

Moybray and colleagues return to this overlap area of computing and the law in [38] and propose protocols for privacy management and for data and information obfuscation (i.e. what we might characterize as encryption light protocols).

5.4 Rights Management in Software

There are interesting parallels and analogies between the process of creating software and the process of creation in science.

Kelty [2] discusses how the Creative Commons was launched in December 2002. It was not just aimed at
software but also "created with people like artists, musicians, scholars, and filmmakers in mind". Creative Commons is one answer to how one may "confront the problem of how to stabilize a work in an unstable context: that of shareable source code, an open Internet, copyleft licenses, and new forms of coordination and collaboration". What we are up against is "the changing meaning of the finality of a scholarly or creative work". In a word, ownership and authorship can be split. Furthermore there are lots of different types of user, including in particular those who can take one’s own creative work further. So use and reuse are fundamental – and very tricky in practice.

For software, after all [2], "The distinction between creating software and maintaining it is a commercial fiction driven in large part by the structure of intellectual property." For – "One programmer’s minor improvement is another programmer’s original contribution." What applies to software can apply even more so to digital content, i.e. "mechanics of copyright, the nature of infringement, the definition of software, the meaning of public domain, the difference between patent, copyright, and trade secret, and the mechanics of permission and its granting".

What is important is "availability and modifiability" and Kelty draws out some parallels and analogies with "the realms of education, music, film, science, and writing". In discussion of Linus Torvalds, Kelty counterposes the fixed textbook version of the Minix operating system, due to Tanenbaum and published by Prentice-Hall, as opposed to the modifiable, growing Linux: "the goals of Minix remained those of a researcher and a textbook author: to be useful in classrooms and cheap enough to be widely available and usable on the largest number of cheap computers" And: "By this model, Linux could do neither; it couldn’t be used in the classroom because it would quickly become too complex and feature-laden to teach, and it wasn’t pushing the boundaries of research because it was an out-of-date operating system" [2]. This points to how some aspects of digital rights in software become of necessity quickly intricate.

5.5 Open Innovation Models
The rights management involved in Open Access can be seen as one element of a more general open innovation approach to intellectual property use and dissemination.

Open innovation is one among various collective activities that have been pursued in the ICT area in particular. Examples noted by Broglia [39] include: Linus Torvalds and the development of the Linux operating system; Richard Stallman who established the GNU, "GNU’s not Unix", repository of all levels of software from the early 1980s onwards; in conjunction with GNU, there has been the GPL, Gnu Public License or copyleft, which allows open redistribution and modification of software code so long as others can maintain these rights; the SourceForge repository of open source projects which according to [39] in July 2008 had 180,000 registered projects and 1.8 million users. Broglia proceeds to IBM’s use of Linux; the Red Hat distributor of Linux, set up by Bob Young; and Wikipedia. Extrapolating from these efforts, the point is made as to how collective, free contributed effort and work – the “online volunteer” – is a new, remarkable and by now fully established model. The results are seen too in Second Life group practices, Amazon’s recommender system, and content on FaceBook, YouTube, eBay, and so on.

Vingron [40] sees data as being a side of the life sciences that should be accommodated by open access. An element of change is seen in data: data sets are huge and so must be published (as ancillary to the paper) online. Meanwhile, "There is a strong push for all data to be public (genome sequence, protein structure)".

Richard Jefferson’s BiOS, Biological Innovation for Open Society, <http://www.bios.net>, goes all the way, seeking a fully open innovation environment for data in the biosciences.

Open innovation can go hand in hand with more distributed and less centralized forms of development, and indeed this can even lead to more democratic organizational forms with fewer barriers to entry. While open innovation has much to say in its favor it is nonetheless clear that centralized and proprietary control can also lead to more reliable, more resilient, more robust and more recovery-enabled, forms of development. In the healthcare sector, software systems have been on occasion developed centrally, based on large state contracts, and have failed disastrously. For some, the blame lies in thoroughly inadequate requirements gathering. Others have seen the culprit as the refusal to access an open innovation model, incorporating open source. Johnson [41] presents a wide-ranging overview of these divergent views, based on case studies and experience with the US Veterans Health Administration (VHA). In passing we note that he includes examples of where centralized healthcare ICT systems have been developed and implemented with recognized success. What Johnson finds is that the "political" changes in organisations, for example towards a centralized, standardized, controlled model, or towards an open, distributed, local or regional model, must be understood. The politics and social aspects must also be factored into the equation, as well as the technical (here, the computing) aspects.

6 Pharmaceuticals as the New Software
6.1 Pharmaceutical, Software and Other Goods

While it is clear that the seeds of the future always lie in the present, what we say on pharmaceuticals as a new form of software is one possible forward extrapolation. However an interesting consequence is how this viewpoint fits in well with certain trends in research dissemination and delivery, as discussed by us in regard to Open Access.

Convergence between the computing sciences and the life sciences is ongoing but with a great deal of, as yet, unrealized targets [42]. We will now sketch out a scenario with clear implications for how the underpinning information-based sciences and engineering could relate to the life sciences in a very new way.

What is the greatest product of all time in terms of return on investment and development? It is not Microsoft Windows and it is not any of Boeing’s
programs. It is the drug Lipitor, used for cholesterol treatment, that cost Pfizer half a million dollars to develop, with annual profit in recent years of around $13 billion. The patent on Lipitor runs out in 2010. This is only one drug among many where intellectual property rights are likely to change greatly in the next few years. Instead generics will step into their place, costing an enormous amount less.

Consider software and pharmaceuticals. Both act on the environment. As such they have determinate inputs and they process these inputs in algorithmic ways.

Now consider how both are characterized by, or contrasted by, the following.

- Potentially huge costs to produce, refine.
- Validating shares engineering and art, statistical sciences and social sciences. But in neither domain are the physical or natural sciences really to the fore. (Energy may yet change the picture here, and bring to the fore – for example – the thermodynamics of information: see [43].)
- The cost of reproducing is near zero.

Computer software was non-trivial to demarcate. With reference to the US Copyright Act of 1976, and the USC section 17 of 1980, Kelty [2] notes that: "During the 1980s, a series of court cases helped specify what counted as software, including source code, object code (binaries), screen display and output, look and feel, and microcode and firmware."

Adaptability and modifiability are crucial: "It is a peculiar feature of copyright law that it needs to be updated regularly each time the media change ..." We can consider "gramophones, jukeboxes, cable TV, photocopiers, peer-to-peer file-sharing programs". Further, "new questions arise: how much change constitutes a new work, and thus demands a new copyright license? If a licensee receives one copy of a work, to which versions will he or she retain rights after changes? ... is the XML document equivalent to the viewable document...? Where does the 'content' begin and the 'software' end?". Kelty speaks of "denaturalization" of the software product or the software manual.

Faced with declining innovation in the pharma sector, and following the model of software that reuse is not simply a technical issue, the thought arises as to how to harness collaboration. Some communities (e.g. humanities) would be aghast at liberal creation of derivative works. Other communities (e.g. computing and engineering) revel in reuse. Kelty ([2], p. 291) includes biology in the reveling category. Maybe the life sciences will follow suit, involving for example that the "compulsory licensing of pharmaceuticals [could be] open to analysis" ([2], p. 304).

### 6.2 Rise of Generics and Implications

Walton and Frank [44] writing in 2008 give indicative figures of 80% of all scientists who ever lived being alive today; only 10–20 new drugs approved each year; to get a drug to market it costs now $1 billion and takes 15 years. Now, they state, an estimated $13.1 billion of drugs are going generic in 2008 and $6.7 billion in 2009.

As patents run out, and as competitive pressure prises open current intellectual property holdings, the whole pharma and bio sector will be affected. We are convinced that the early years of this decade (2010 onwards) will see the pharma sector changing greatly.

The metaphor of a supernova would be apt. Note that in saying this we are not only pointing to a massive economic and technological change. Rather, in analogy with, say, a black hole resulting from a supernova, we are suggesting a new source of tremendous dynamism – a melting together of pharma and software. To see how this could be, consider what pharmaceutical products really are: a type of software. They are expensive to produce and essentially have zero cost to reproduce.

The implications of this would be profound. It would lead to a very different health system, including all of insurance, medical and public health domains. To have some feel for the direction of events, one should look very carefully at computer science and the ICT sector. Computer science and engineering are characterized by a wealth of "models" both in research and development and for intellectual property. The dream is one where "Everything that can be delivered digitally will be, at a cost approaching zero, through a bandwidth nearing infinity" [45].

Creating of new drugs would result from ultra high dimensional search and discovery in truly massive data stores. Drug development would need a new pharma-oriented information search and fusion infrastructure at its core.

Our health system would be even more integrally based on the information infrastructure. One could suggest that the ground is already being prepared, through data provision, for the search engines that will power the new pharma informatics and health informatics. Indeed Norvig, Google's Director of Research has claimed that correlation (in massive dataset interrelationships) now is on the verge of replacing (scientific) causation, since all one has to do is let the data speak for themselves [46] (and see also [5]).

To envisage barriers between disciplines coming down is not a bad prospect. Computer science and engineering may be very different in one or two decades from now.

An implication of such change is in research publishing. This is because one thing about the way we carry out our research and scholarly work is that we, across all science disciplines, have become very influenced by the biosciences. Take for example how journal citation rates are based on just two previous years. Extreme recency has come to count greatly in citation practices, and a very small number of highly profiled journals tower over all others. We know how different the scene is in computing: Moed and Visser [47] describe various aspects, not least the role of conferences. The latter – conferences and conference proceedings – represent in a different guise the ever growing importance of recency.

Perhaps the future of scholarly publishing across many disciplines is closer than we think to today's variety of publishing practices in computing, both science and engineering, both theory and application, and embracing both open source and commercial orientations.
6.3 Implications of Pandemics for Patent Protection

To tackle the H1N1 or swine flu pandemic, in May 2009 the World Health Organization (WHO) allowed the leading Indian pharmaceutical corporate Cipla to produce the generic version of the anti-viral medication, Tamiflu.

Hoffmann-La Roche’s oseltamivir, branded as Tamiflu, and GlaxoSmithKline’s Relenza are the only two recognised drugs to treat the pandemic swine flu, H1N1 [48]. Cipla brought a generic drug, zanamivir, substituting for oseltamivir to market in August 2009. The inhalatory drug zanamivir was branded as Virenza in India.

As of 22 August 2009, a withdrawal was expected of zanamivir by regulatory authorities in India. The Hoffmann-La Roche drug, Tamiflu, only is to be used. Meanwhile GSK’s Relenza is a pre-1995 drug and lacks patent protection in India.

What is of note here is the readiness of generic companies to step into the production and distribution breach, and how such a situation unfolds in a possibly socially threatening context.

The pharma majors may also be coming to terms in other ways too with generics, with all that that implies for the big shift in intellectual property and its exploitation: "Until recently, many brand-name drug makers invested the bulk of their research and marketing dollars in the development of blockbuster drugs, only to see their intellectual property and market share to lower-priced generic competitors once patents expired. But now, with an estimated $89 billion in brand-name drug sales in the United States at risk to generic competition over the next five years [...] some drug makers are selling generics to offset revenue declines – as well as wring some post-patent profits from the innovative drugs they developed." [49]

The interesting point to be made is that whether through the force majeure of pandemic, or otherwise, diverse business models and attendant intellectual property frameworks are coming about.

7 Conclusions

In concluding, Pisano [11] reiterates the need for organizational and institutional innovation in biotechnology, in order for it to overcome the slowness of productivity (i.e., rarity of new drug discovery and successfully bringing to market). He noted the role played by the modern corporation in rail and telegraph systems in the nineteenth century. He noted the role played by venture capital in semiconductors, software, computers and communications, in the US in particular, in the latter half of the twentieth. Finally his discussion dealt with the promise of the entrepreneurial biotechnology sector, in its successive generations, roughly 1976–1985; 1986–1992; and 1992–2000. (Note how the first and the third periods are close to the "New Economy" and "Knowledge Society" periods of ICT growth; see [50].)

In order to prevent under-exploited "islands of expertise" greater openness and disclosure are needed. Recall that both research publishing, and patenting and licensing, are forms taken by such expertise and knowledge. Pisano focuses on the latter areas of intellectual property. Among other recommendations, he proposes greater disclosure of clinical trial data much earlier in the development process, and not just after approval. While such disclosure does happen through journal publishing, when this happens is at the discretion of the sponsoring drug company, and will not necessarily include negative as well as positive findings. Pisano’s objective is to overcome the logjam of innovation in the biotechnology sector, and to facilitate investment decisions.

Information and access to it is key. Open Access allows the research literature to be fully internetworked and integrated. Open Access in the life sciences is a not insignificant part of the overall requirements to overcome the historic logjam of innovation in the biotechnology sector.

Two examples follow, intended as no more than existence proofs of the potential that is at issue here.

- An early stage system for open publishing of clinical trials data, with linkages to some other data sources, can be found at [51].
- Interoperability of health records and health data [52]. A "secure, standards-based, [health and medical] data landscape": "a software platform that will allow medical data from at-home devices like glucose meters and blood pressure monitors to be sent automatically to Google Health or other Personal Health Records systems online" [53]. The Continua Health Alliance, <http://www.continuaalliance.org>, with more than 200 company members is pursuing an agenda of building a health systems ecosystem. A range of working groups include technical, wellness solutions, regulatory, test and certification, use case, and policy strategy.

Moving to software, Boyle [54] notes how a creative open "commons" in property rights contributed so much to spurring on the industry. He notes, with worry, how synthetic biology "which shares aspects of both software (programming in genetic code) and genetic engineering" is threatened by the lack of open models of intellectual property.

While the biotechnology industry is evolving slowly but surely to a new software industry the core legacy of software – its tremendously rich ecology of intellectual property rights – has not yet been taken on board by biotechnology. Biotechnology has been stuck, not able to take sufficiently new forms of property rights on board. This is why Open Access in the life sciences generally is so important, to help to make possible the radically new forms of software.

References

[5] F. Murtagh, "Origins of modern data analysis linked to the beginnings and early development of computer science and information


[22] R. Adler, J. Ewing and P. Taylor, Citation Statistics, A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS), Joint Committee on Quantitative Assessment of Research, 11 June 2008.


