21 | Free Science

Christopher Kelty

What is the *value* of science? In speculating about the success of open source/free software (OS/FS), users and advocates often suggest that it is "like science." It has characteristics of peer review, open data subject to validation and replication, and a culture of academic freedom, credit, civility, and reputation. The point of this comparison is that these characteristics directly contribute to producing (morally or technologically) better software, just as science is improved by them. This begs the question: what exactly is the value of either endeavor—financial, personal, aesthetic, moral, or all of these? How can we specify it?

This chapter investigates the value of science from the perspective of its social constitution; in particular, the importance of law, informal norms, and technology. It poses two related questions: "Is science like open source/free software?" and "Can you do science without open source/free software?"

Two Economies of Science

In studies of OS/FS, the question of motivation inevitably arises, and is usually answered in terms of reputation. Reputation, it is asserted, is like money, and governs how people make choices about what software they use or to which projects they contribute. A similar issue infuses the social and historical study of science: here the question of motivation concerns what might be called the "remunerative structure of science"; that is, the difference between cash payment for knowledge and ideas, and the distribution of reputation, trust, or credit for knowledge and ideas. On the one hand, many people (Merton 1973; Polanyi 1969; Mirowski 2001; Mirowski and Sent 2002) suggest that it is the latter that keeps science on the right track towards truth and objectivity. Much like the claim in OS/FS that openness and freedom lead to better software, the structure of

remuneration through credit and public acknowledgment in science is said to ensure that the truest truths, and neither the cheapest or the most expensive ones emerge from the cauldron of scientific investigation.

On the other hand, the political economy of science is also deeply embedded in the health and progress of nations and societies. Science (like software) simply must be paid for somehow, and most scientists know this, even if they like to ignore it. What's more, if it is to be paid for—by governments, rich people, or corporations—it is probably required to contribute to their agenda somehow. In a representative democratic society, this means that the funding of science is done on condition that it contributes to "progress." It is only through science and technology (or so many economists have concluded) that growth, progress, and increasing prosperity are even possible. Scarce resources must be effectively distributed or the value of the whole enterprise collapses. Markets and money are one very effective way of achieving such allocation, and science, perhaps, should not be an exception.

The tension between these two demands can be summed up in two different questions concerning "value": (1) What is the best way to achieve efficient allocation of scarce resources? and (2) What is the proper way to organize a secular scientific and technological society so that it can contribute to the improvement of question 1? Needless to say, these questions must be kept separate to be meaningful. Where OS/FS appears, it is often in response to the subordination of question 2 to question 1. Free software licenses, the open collaborative ethic of OS/FS hackers, and the advocacy of lawyers and economists are all ways of reminding people that question 1 is not the only one on the table. This issue strikes science and technology at its heart—especially in its European and American forms in the universities and research labs. It is left to scientists, engineers, and managers in these places to insist on a continual separation of these two questions. In a practical sense, this separation means maintaining and improving systems of remuneration based on the principles of peer review, open access, experimental verification and the reduction of conflicts of interest. Without these, science is bought and sold by the highest bidder.

Doing Science

Is science like open source/free software? Yes, but not necessarily so. There are far too many examples in science of secrecy, meanness, Machiavellian

plotting, and downright thievery for us to believe the prettied-up claim that science is inherently characterized by openness and freedom. Curiously, this claim is becoming increasingly accurate. From the sixteenth century on, norms and forms of openness have improved and evolved alongside the material successes of science and technology. The creation of institutions that safeguard openness, peer review, trust, and reputation is coincident with the rise and dominance of scientific and technical expertise today. The myth of a scientific genius toiling away in an isolated lab, discovering the truths of nature, bears little resemblance to the historically situated and fundamentally social scene of Robert Boyle demonstrating his air pump before the assembled Royal Society. Though it is easy to show how political, how contextual, or how "socially constructed" science is, this is not the point I am making (for some canonical references in this field, see Bloor 1976; Barnes 1977; Collins 1985; Pickering 1984; Latour 1986; Haraway 1997). Rather, the point is that the creation and maintenance of the institutions of science over the last 400 years has been a long, tortured, and occasionally successful attempt to give science the character of truth, openness, and objectivity that it promises. However, we are not there yet, and no scientists are free from the obligation of continuing this pursuit.

One compelling study of how science has become analogous to the OS/FS movements is the work of Robert K. Merton, the American sociologist who first attempted to think through what he called the "normative structure of science"—a sociological account of scientific action that focused on the reward system and the ethos of science (Merton 1973). The ethos of science (not unlike the famous "Hacker Ethic," Himanen 2001) is that set of norms and forms of life that structure the activity of scientists across nations, disciplines, organizations, or cultures. Merton identified four norms: universalism, communism (Merton's word), disinterestedness, and organized skepticism.

These norms are informal, which is to say that they are only communicated to you by your becoming part of the scientific establishment—they are not written down, and are neither legally nor technically binding (along the same lines as "You are a hacker when another hacker calls you a hacker"). However, despite the informal character of these norms, the institutions of science as we know them are formally structured around them. For example, communism requires a communication structure that allows the communally owned property—ideas, formulae, data, or results —to be disseminated: journals, letters, libraries, university postal systems,

standards, protocols, and some more or less explicit notion of a public domain.

Or, another example. The norm, disinterestedness, is not an issue of egoism or altruism, but an institutional design issue. For disinterestedness to function at all, science must be closed off and separate from other parts of society, so that accountability is first and primarily to peers, not to managers, funders, or the public—even if this norm is continually under assault both from within and without. Similarly, organized skepticism is not simply methodological (whether Cartesian doubt or acceptable "p" values), but institutional as well—meaning that the norms of the institution of science must be such that they explicitly, if not exactly legally, promote the ability to maintain dissent even in the face of political power. Otherwise, truth is quickly compromised.

To take a historical example, consider Robert Boyle, as told by Steven Shapin and Simon Schaffer (1985) in *Leviathan and the Air Pump*. Boyle's genius lay not only in his formulation of laws concerning the relation of temperature, pressure and volume (a significant achievement in itself), according to Shapin and Schaffer, Boyle's activities also transformed the rules of modern experimentalism, of "witnessing" and of the means for establishing modern facts. Boyle's experimental air pump was seventeenth-century "big science." It required Boyle's significant fortune (he was, after all, the son of the Earl of Cork), access to master glass blowers and craftsmen, a network of aristocratic gentlemen interested in questions of natural philosophy, metaphysics, and physics. Perhaps most importantly, it required the Royal Society—a place where members gathered to observe, test, and "debug" (if you will) the claims of its members. It was a space by no means open to everyone (not truly public—and this is part of the famous dispute with Thomas Hobbes, which Shapin and Schaffer address in this book), because only certain people could be assumed to share the same language of understanding and conventions of assessment. This is a shortcoming that OS/FS shares with Boyle's age, especially regarding the relative absence of women; the importance of gender in Boyle's case is documented in (Potter 2001); there is much speculation, but little scholarship to explain it in the case of OS/FS.

To draw a parallel with OS/FS here, the Royal Society is in some ways the analog of the CVS repository: demonstrations (software builds), regular meetings of members (participation in mailing list discussion), and independent testing and verification are important structural characteristics they have in common. They both require a common language (or several), both natural and artificial.

Hackers often like to insist that the best software is obvious, simply because "it works." While it is true that incorrectly written software simply will not compile, such an insistence inevitably glosses over the negotiation, disputation, and rhetorical maneuvering that go into convincing people, for instance, that there is only one true editor (emacs).

A similar claim exists that scientific truth is "obvious" and requires no discussion (that is, it is independent of "our" criteria); however, this claim is both sociologically and scientifically simplistic. It ignores the obvious material fact that scientists, like programmers, organize themselves in collectivities, dispute with each other, silence each other, and engage in both grand and petty politics. Boyle is seen to have "won" his dispute with Hobbes, because Hobbes science was "wrong." This is convenient shorthand for a necessary collective process of evaluation without which no one would be right. It is only after the fact (literally, after the experiment becomes "a fact") that Boyle's laws come to belong to Boyle: what Merton called "intellectual property." A science without this process would reduce simply to authority and power. He with the most money pronounces the Law of the Gases. The absurdity of this possibility is not that the law of the gases is independent of human affairs (it is) but that human affairs go on deliberately misunderstanding them, until the pressure, so to speak, is too great.

Merton and others who study science and technology like to point out just how widespread and extensive this system of disputation, credit, and reward is: it includes eponomy (the naming of constants, laws, and planets), paternity (X, father of Y), honors, festschrifts, and other forms of social recognition, prizes like the Fields medal or the Nobel, induction into royal societies, and ultimately being written into the history books. These mechanisms are functional only in hindsight; it is perhaps possible to say that science would still proceed without all these supports, but it would have neither collective existence in nor discernible effect on the historical consciousness and vocational identity of practicing scientists. That is to say, the question of motivation is meaningless when considered in isolation. It is only when considered as a question of institutional evolution and collective interaction that motivation seems to have a role to play. In the end, it is equally meaningless to imagine that people have a "natural" desire to pursue science as it is to suggest that we are somehow programmed to desire money. Curiosity and greed may be inevitabilities (this hangs on your view of human nature), but the particular forms they take are not self-determining.

Funding Science

Of course, such informal norms are all well and good, but science costs money. On this point, there is no dispute. In Boyle's day, air pumps were like linear accelerators: expensive and temperamental. Even books could be quite dear, costing as much as the air pump itself (Johns 1998). For Boyle, money was no object; he had it, and other people around him did too. If they didn't, then a rich friend, a nobleman, a patron could be found. The patronage structures of science permeated its institutional, and perhaps even its cognitive, characteristics (Biagioli 1993). By contrast, twentieth-century science looks very different—first, because of massive philanthropy (Carnegie, Rockefeller, and others); second, because of massive military and government investment (Mirowski 2002; Mirowski and Sent 2002); and third, because of massive "soft money," research and development and contract investment (this most recent and rapid form of the commercialization of science differs from field to field, but can be said, in general, to have begun around 1980). The machines and spaces of science were never cheap, and have gotten only less so. The problem that this raises is essentially one of the dispensation of credit and return on investment.

Sociologists of science have attempted to finesse this difficulty in many of the same ways as observers of OS/FS: through notions of "reputation." Gift economies, in particular, were the study of a short article by Warren Hagstrom. He attempted to explain how the contributions to scientific research—such as giving a paper or crediting others—made the circulation of value an issue of reciprocity that approximated the gift-exchange systems explored by Marcel Mauss and Bronislaw Malinowski (Hagstrom 1982). Bruno Latour and Steve Woolgar also explore the metaphors of non-monetary exchange in science, in the course of their work on the construction of facts in laboratories. They suggested that there is a "cycle of credit" that includes both real money from granting agencies, governments, and firms and the recognition (in the form of published articles) that leads full circle to the garnering of grant money, and so on ad infinitum. In this cycle, both real money and the currency of reputation or credit work together to allow the scientist to continue to do research (Latour and Woolgar 1979). Here, the scientist wears two masks: one as the expert witness of nature, the other as the fund-seeking politician who promises what needs to be promised. Most scientists see the latter as a necessary evil in order to continue the former (see Latour 1986 for an alternate account). There is a similarity here with OS/FS programmers, most of whom, it is

said, keep their day jobs, but spend their evenings and weekends working on OS/FS projects (Raymond 2001).

In these studies to date, the focus has been on the remuneration of the scientists, not the return on investment for the funders. In the cases of early modern patronage systems, the return to the patron was not strictly financial (though it could be), but was often also laden with credit in a more circumscribed and political sense (for example, the status of a monarchy or of a nation's science; see Biagioli 1993). In a similar sense, philanthropists build for themselves a reputation and a place in history. Government and military funding expects returns in specific areas: principally war, but also health, eradication of disease, and economic growth. Soft money, venture capital, and research and development, on the other hand, are primarily interested in a strictly calculated return on investment (though here too, it would be disingenuous to suggest that this were the only reward—venture capitalists and corporations seek also to be associated with important advances in science or technology and often gain more in intangible benefits than real money). The problem of funding science is never so clean as to simply be an allocation of scarce resources. It includes also the allocation of intangible and often indescribable social goods. Strangely, Robert Merton called these goods "intellectual property" (Garfield 1979).

It is important to distinguish, however, the metaphorical from the literal use of intellectual property: in the case of the scientist, reputation is inalienable. No one can usurp a reputation earned; it cannot be sold; it cannot be given away. It may perhaps be shared by association; it may also be unjustly acquired—but it is not an alienable possession. Intellectual property granted by a national government, on the other hand, exists precisely to generate wealth from its alienability: inventors, artists, writers, composers, and yes, scientists, can sell the products of their intellectual labor and transfer the rights to commercialize it, in part or in whole, by signing a contract. The reputation of the creator is assumed to be separate from the legal right to profit from that creativity. This legal right—intellectual property as a limited monopoly on an invention or writing—is often confused with the protection of reputation as an inalienable right to one's name; this is not guaranteed by intellectual property law (on this confusion throughout history, see Johns 1998).

This confusion of the metaphorical and the literal uses of intellectual property goes both ways. Today it is virtually impossible to step foot in a lab without signing a licensing agreement for something, be it a machine, a tool, a process, a reagent, a genetic sequence, or a mouse. Many of the

things biologists or engineers once traded with each other (cell lines, testing data, images, charts, and graphs) are now equally expected to generate revenue as well as results. The race to publish is now also a race to patent. It might even be fair to say that many scientists now associate success in science with return on investment, or see the "free" exchange of ideas as more suspicious than a quid pro quo based on monetary exchange (see Campbell et al. 2002). This metaphorical confusion necessitates a closer look at these practices.

Valuing Science

The metaphors of currency and property in science meet in a peculiar place: the Science Citation Index. Citation indices give one a very prominent, if not always precise, indicator of value. It is a funny kind of value, though. Even though citation is quantifiable, not all reputation depends on citations (though some tenure committees and granting agencies beg to differ on this point). Qualitative evaluation of citing practices is an essential part of their usefulness. Even though work that isn't included in such databases is at a rather serious disadvantage, reputationally speaking, science citation indices do not simply measure something objective (called reputation). Rather, they give people a tool for comparative measure of success in achieving recognition.

Robert Merton clearly understood the power of citation indexing—both as currency and as a kind of registration of intellectual property for the purposes of establishing priority. In the preface to Eugene Garfield's 1979 book *Citation Indexing*, Merton says, "[Citations in their moral aspect] are designed to repay intellectual debts in the only form in which this can be done: through open acknowledgment of them" (Garfield 1979, viii). He thus makes citations the currency of repayment. But he goes even further, explaining scientific intellectual property in a manner that directly parallels the claims made for OS/FS's success as a reputation economy:

We can begin with one aspect of the latent social and cultural structure of science presupposed by the historically evolving systematic use of references and citations in the scientific paper and book. That aspect is the seemingly paradoxical character of property in the scientific enterprise: the circumstance that the more widely scientists make their intellectual property available to others, the more securely it becomes identified as their property. For science is public, not private knowledge. Only by publishing their work can scientists make their contribution (as the telling word has it) and only when it thus becomes part of the public domain of science can they truly lay claim to it as theirs. For the claim resides only in the recognition of the source of the contribution by peers. (Garfield 1979, vii–viii)

This claim is remarkable, but not dissimilar to that remarkable claim of OS/FS (particularly open source) advocates—that openness results in the creation of better software. Merton here claims as much for science. The incentive to produce science depends on the public recognition of priority. The systems involved in making this property stick to its owner are reliable publishing, evaluation, transmission, dissemination, and ultimately, the archiving of scientific papers, equations, technologies, and data. As stated previously, this priority is inalienable: when it enters this system of registration, it is there for good, dislodged only in the case of undiscovered priority or hidden fraud. It is not alienable intellectual property, but constant; irretrievably and forever after granted. Only long after the fact can diligent historians dislodge it.

Who grants this property? The key is in Merton's paradox: "the more widely scientists make their intellectual property available to others, the more securely it becomes identified as their property" (Garfield 1979, vii). That is, no one (or everyone) grants it. The wider the network of people who know that Boyle is responsible for demonstrating that under a constant temperature gas will compress as pressure is increased, the more impossible it becomes to usurp. Only by having a public science in this sense is that kind of lasting property possible. A privatized science, on the other hand, must eternally defend its property with the threat of force, or worse, of litigation. While a public science tends toward ever greater circulation of information in order to assure compensation in reputation, a private science must develop ever more elaborate rules and technologies for defining information and circumscribing its use. Not only is a private science inefficient; it also sacrifices the one thing that a public science promises: progress.

Nonetheless, a public science is only possible through publication. The publication and circulation of results is a sine qua non that up until the advent of the Internet was possible only through academic publishers, university presses, and informal networks of colleagues and peers. Reputation was effectively registered through the small size and manifest fragility of the publication network. It has been successful enough and widespread enough that most people now associate the quality of a result with the publication it appears in. We have a well-functioning system, however imperfect, that allows a widely distributed network of scientists to coevaluate the work of their peers.

From an institutional standpoint, this is a very good thing. As I said earlier, science has not always been open or free, but it has become more and more so over the last 400 years. The functions of openness that have

developed in science exist only because publishers, universities, and academic presses—along with scientists—believe in them and continue to propagate them.

To some extent, this system of reputational remuneration has lived in strained but peaceful coexistence with the monetary structure of funding. It is only of late, with the expansion of intellectual property law and the decreasing vigilance of anti–trust policing, that the legal and institutional framework of the U.S. economic system has actually become hostile to science.

Consider the situation scientists face today. Most scientists are forced to explicitly consider the trade-off between the ownership of data, information, or results and the legal availability of them. In designing an experiment, it is no longer simply a process of finding and using the relevant data, but of either licensing or purchasing it, and of hiring a lawyer to make sure its uses are properly circumscribed. In economic terms, the transaction costs of experiment have skyrocketed, specifically as a result of the increased scope of intellectual property and more generally due to the ever increasing dangers of attendant litigation. In scientific terms, it means that lawyers, consultants, and public relations agents are increasingly stationed in the lab itself, and they increasingly contribute to the very design of experiments.

The result is a transformation of science, in which the activity of using an idea and giving credit is confused with the activity of buying a tool and using it. Science becomes no longer public knowledge, but publically visible, privately owned knowledge.

The skeptic might ask: why not let intellectual property law govern all aspects of knowledge? What exactly is the difference between using an idea and buying one? If the copyright system is an effective way of governing who owns what, why can't it also be an effective way of giving credit where credit is due? Such a proposition is possible in the context of U.S. law, and less so in European intellectual property law, which makes an attempt (however feeble) to differentiate the two activities. In German law, for instance, the "moral right of the author" is presumed to be inalienable, and therefore a separate right from that of commercial exploitation. While U.S. law doesn't make this distinction, most U.S. citizens do. Even the firm believer in copyright law wants to protect his reputation; no one, it seems, wants to give up (metaphorical) ownership of their ideas. Unfortunately, these are issues of fraud, plagiarism, and misappropriation—not of commercial exploitation. And these are fears that have grown enormously in an era of easy online publication. What it points to is not a need for

stronger intellectual property law, but the need for an alternative system of protecting reputation from abuse.

The need to somehow register priority and (metaphorical) ownership of ideas is a problem that cannot be solved through the simple expansion of existing intellectual property law. It will require alternative solutions. These solutions might be technical (such as the evolution of the science citation index—for example, cite-seer and LANL) and they might also be legal (devices like the Creative Commons licenses, which require attribution, but permit circulation). In either case, science and technology are both at a point similar to the one Robert Boyle faced in the seventeenth century. A new way of "witnessing" experimental knowledge is necessary. A new debate over the "public" nature of science is necessary.

Many people in OS/FS circles are aware of this relationship between informal reputation and calculable monetary value. Even Eric Raymond's highly fantastic metaphorical treatment of reputation reports an important fact: the list of contributors to a project should never be modified by subsequent users (that is, contribution is inalienable). To do so is tantamount to stealing. Similarly, Rishab Ayer Ghosh and Vipul Ved Prakash (2000) also recognize this nonlegal convention; they combined it with the formal availability of free software packages and created a tool much like the Science Citation Index: it adds up all contributions of individuals by grepping (using a Unix text search command) packages for e-mail addresses and copyrights. We might call what they find "greputation," since it bears the same relation to reputation that money supposedly does to value. That is, it is the material and comparable marker of something presumed to be more complex—the reputation of a scientist—just as money is an arbitrary technology for representing value.

In order to understand why reputation is at stake in science, we might take this analogy a bit further and ask what exactly is the relationship between money and value. Economic dogma has it that money is a standard of value. It is a numerical measure that is used to compare two or more items via a third, objectively fixed measure. This is an unobjectionable view, unless one wants to ask what it is that people are doing when they are valuing something—especially when that something is an idea.

However, from the perspective of Georg Simmel, the early twentieth-century German sociologist whose magnum opus is devoted to the subject (Simmel 1978), considering money as something that simply facilitates a natural human tendency (to value things according to cardinal ranking) is a sociologically and anthropologically illegitimate assumption. Humans

are not born with such an objective capacity vis-à-vis the world around them. Rather, since money is a living set of institutions that calibrate value and a set of technologies (cash, check, credit, and so on) that allow it to circulate or accumulate, then humans are caught within a net that both allows and teaches them how to reckon with money—how to count with it, as well as on it. Even if staunch neoclassicists agree that the rational actor of economic models does not exist, that by no means suggests he cannot be brought into existence by the institutions of economic life. To borrow David Woodruff's willful anachronism: "Humans are endowed only with an ordinal sense of utility; they attain something like a cardinal sense of utility ("value") only through the habit of making calculations in money" (Woodruff 1999).

If we consider this insight with respect to the currency of reputation, as well as that of money, we can say the following: the standard of value (money, or the citation) serves only to stabilize the network of obligations thus created: in the case of money economies, a single cardinal value; in the case of citations, a widely recognized, though sometimes disputed reputation. The vast interconnected set of legal obligations that money represents can be universally accounted by a single standard—a cardinal value. But if we reckoned the world of obligations using a different standard—a nonnumerical one, for instance—then humans could also learn to express utility and value in that system. Money, it should be very clear, simply isn't natural.

Therefore, a similar approach to scientific citations would have to focus on something other than their cardinality. And in fact, this is exactly what happens. Citations are simply not fungible. Some are good (representing work built upon or extended), some are bad (representing work that is disputed or dismissed), some are indifferent (merely helpful for the reader), and some are explicit repayments (returning a citation, even when it is not necessarily appropriate). Often the things that are most well known are so well known that they are no longer cited (F = ma, or natural selection), but this could hardly diminish the reputation of their progenitors. It requires skill to read the language and subtleties of citations and to express gratitude and repay intellectual debt in similarly standardized, though not simply quantitative ways. There are whole stories in citations.

This description is equally accurate in open source and free software. Although some might like to suggest that good software is obvious because "it works," most programmers have deep, abiding criteria for both efficiency and beauty. Leaf through Donald Knuth's *The Art of Computer Programming* for a brief taste of such criteria and the interpretive complexity

they entail (Knuth 1997). The scientist who does not cite, or acknowledge, incurs irreconcilable debts—debts that cannot be reckoned in the subtle currency of citations. The more legitimate the information infrastructure of scientific publications, databases, and history books becomes, the more essential it is to play by those rules, or find increasingly creative ways to break them. In money, as in science, to refuse the game is to disappear from the account.

Today, we face a novel problem. The institutions we've inherited to manage the economy of citation and reputation (publishing houses, journals, societies and associations, universities and colleges) used to be the only route to publicity, and so they became the most convenient route to verification. Today we are faced with a situation where publication has become trivial, but verification and the management of an economy of citation and reputation has not yet followed. In the next section, I conclude with two cases where it will very soon be necessary to consider these issues as part of the scientific endeavor itself.

A Free (as in Speech) Computational Science

In 1999, at the height of the dot-com boom, there was an insistent question: "But how do you *make money* with free software?" I must admit that at the time this question seemed urgent and the potential answers seductive. In 2003, however, it seems singularly misdirected. From the perspective of science and technology, where software can be as essential a tool as any other on the lab bench, the desire to make free software profitable seems like wanting a linear accelerator to produce crispier french fries. You *could* do that, but it is a rather profound misunderstanding of its function.

I propose a different, arguably more important, question: "Can you do science without free software?"

By way of conclusion, I want to offer two compelling examples of computational science as it is developing today. The promise of this field is evident to everyone in it, and these examples should be viewed as evidence of early success, but also as early warnings. What they share is a very precarious position with respect to traditional scientific experiment, and—in terms Robert Boyle would understand—traditional scientific "witnessing." It is impossible to think about either of these endeavors without considering the importance of software (both free and proprietary), hardware, networks and network protocols, standards for hardware and software, and, perhaps most importantly, software development methodologies. It is in the need to explicitly address the constitution, verification, and

reliability of the knowledge produced by such endeavors that something like OS/FS must be an essential part of the discussion.

Bioelectric Field Mapping

Chris Johnson is the director of the Scientific Computing Institute (SCI) at the University of Utah. SCI is a truly stunning example of the kind of multidisciplinary computational "big science" that relies equally on the best science, the best software programming, and of course, the best hardware money can buy. Dr. Johnson's bioelectric field mapping project (http://www.sci.utah.edu) extends the possibilities of understanding, simulating, and visualizing the brain's electrical field. It makes EEGs look positively prehistoric.

The project makes sophisticated use of established mathematical and computational methods (methods that have been researched, reviewed, and published in standard science and engineering publications); neuro-anatomical theories of the brain (which are similarly reviewed results); mathematical modeling software (for instance, MatLAB, Mathematica); graphics-rendering hardware; and a wonderful array of open and closed, free and non-free software. It is a testament both to the ethos of scientific ethos and to the best in management of large-scale scientific projects.

What makes Dr. Johnson's project most interesting is the combination of traditional scientific peer-review and his plea for more effective large-scale software management methodology in science. Here is a chance for the best of science and the best of business to collaborate in producing truly exceptional results. But it is here that Dr. Johnson's project is also a subject of concern. In a recent talk, amidst a swirl of Poisson equations, Mesh generation schemes, and finite element models, Johnson pointed out the difficulty of *converting file formats*. Dr. Johnson's admirable attempt to create a computational science that weaves mathematical models, computational simulations, and graphical visualization has encountered the same problem every PC user in the world laments daily: incompatible file formats.

Part of this problem is no doubt the uncoordinated and constant reinvention of the wheel that scientists undertake (often in order to garner more credit for their work). The other part, however, concerns the legal and political context where such decisions are made—and they are usually not made in labs or institutes. This second and more serious problem concerns whether the circumvention of particular file formats in scientific research is affected by the institutional changes in intellectual property or antitrust law. Even if it isn't, knowing requires the intervention of lawyers

aplenty to find out—and the unfortunate alternative is to do nothing. These issues must be addressed, whether through licenses and contracts or through the courts, in order to ensure that the ordinary activity of scientists continues—and does not fall afoul of the law.

The best science, in this case, depends on an analogy with the principles and practices of free software and open source: not only are source code and documentation resources that Dr. Johnson would like to see shared, but so are data (digital images and data from high-end medical scanners) as well as geometric and computational models that are used in the computational pipeline. Only by sharing all of these creations will it be possible for science as a distributed peer-reviewed activity to reach even tentative consensus on the bioelectric fields of human brains.

An Internet Telescope

In order to avoid any suggestion that a large Redmond-based corporation is at fault in any of this, take a second example. Jim Gray, research scientist at Microsoft, has been working on an "Internet telescope," which federates astronomical data from telescopes around the world (http://research.microsoft.com/~Gray). His manifest skill in database engineering (he is a Turing Award winner) and his extraordinary creativity have resulted in a set of database tools that can be used to answer astronomical questions no single observatory could answer—simply by querying a database.

Dr. Gray's project is a triumph of organizational and technical skill, another example of excellent project management combined with the best of traditional scientific research. Dr. Gray is assisted considerably by the fact that astronomical data is effectively worthless—meaning that, unlike genetic sequence data, it is neither patented nor sold. It is, though, hoarded and often unusable without considerable effort invested into *converting file formats*. Gray will ultimately be more successful than most university researchers and amateur database builders, because of the resources and the networks at his command, but the problem remains the same for all scientists: the social and normative structure of science needs to be kept institutionally and legally open for anything remotely like peer-reviewed and reliable knowledge to be possible. All of this data, and especially the programming languages, web services, databases, and Internet protocols that are used in the creation of the telescope, need to remain open to inspection, and remain ultimately legally modifiable without the permission of their owners. If they are not, then scientific knowledge is not "witnessed" in the traditional sense, but decided in advance by lawyers and

corporate public relations departments (Dr. Gray gets around this problem because of his skill in achieving informal and formal participation from each participant individually—but this is probably a solution only Microsoft can afford, and one that does not scale).

Can these sciences exist without free software, or something like it? George Santayana famously quipped: "Those who cannot remember the past are condemned to repeat it." Now might be a time to both remember the past, and to *insist* upon repeating it. Science, as an open process of investigation and discovery, validation, and verification, is not a guaranteed inheritance, but something that had to be created and has yet to be perfected. Openness can not be assumed; it must be asserted in order to be assured.