2002

Comments on Open Source Genomics

Michael J. Meurer
Boston University School of Law

Raymond's writings,[13] one of the key things that he believes makes open source work is this immediate applicability: that users, in his phrase, can "scratch their own itch." As a user, you can take the open source code and use it to solve problems right away and have immediate applicability. It is not clear, when we look at a realm like biotechnology with extremely long lags and large costs to commercialize products, that this essential element is present.

Just to wrap up, the discussion here is really provocative and suggests a lot of interesting issues. I certainly concur that there are all these parallels between academia and open-source. At the same time, however, the broader extensions, especially the claim that biotechnology should be adopting some GPL-like licensing, are ones I am less comfortable with.

Thank you. (*applause*)

PROFESSOR MICHAEL MEURER:

Well, we've made it to the end of the day. I am very appreciative of the performance that everyone has given today. Thank you.

I am going to make a few comments in the spirit of those made by Josh. I think the topic of Dan's paper is very interesting. It covers a lot of ground, and I do not have time to talk about more than a single issue. Josh's began discussing the similarity between academic research and open source, and then analyzed GPL. I want to go back to the first issue Josh raised and argue that, unlike Josh, I see some fairly significant differences between open source software and genomics. I want to talk about the cooperative ethic in the open source movement and compare it to the cooperative ethic in bioinformatics and argue that I do not see the ethic being as sustainable in the genomics world as it is in the open source world.

The first point I want to make is that history matters. I do not know if we should look at Dan's talk as a lament or a prescription. It might have been possible if we went back to 1990 and started from that point, but given our history over the past 10 years, it would be a tough row to hoe. It is tough to reverse things when we have a lot of people that are committed to an exclusive and proprietary approach. Secondly, I want to talk about incentives that bring people to cooperate in a situation in which they might free ride. My comments are based on a paper that Josh Lerner and Jean Tirole have written about the open source movement, so I am borrowing a lot from Josh's work, and he therefore cannot tell me I have it wrong.

Josh talks about how the positive incentives to cooperate include, first of all, a career concern. You can build a reputation for being a star and then be an attractive employee. Secondly, there is a norm of sharing. Working against that norm in the open-source world, and more strongly in the biotechnology

[13] *E.g.*, ERIC RAYMOND, THE CATHEDRAL AND THE BAZAAR (O'Reilley & Associates 1999).

world, though, is the possibility of cashing in on property rights, and for an academic, getting tenure. One factor that is considered in tenure decisions is how many patents have you filed. That is also a factor determining what kind of lab support you get, how many grants you get. Thus, the incentives might be a little bit stronger in the academic life sciences to grab the prize rather than continue to contribute to the public good.

The way I think about it is in terms of game-theoretic models of cooperation. Economists long ago puzzled over why OPEC was so successful. A lot of work was done in the 1980s to explain how we can achieve cooperation in situations in which people would be naturally inclined to cheat – cheat, for example, on the rules of OPEC to undercut prices that prevail in the cartel. The answer is that you could cheat and profit handsomely for a short period of time, but you will lose cooperation in the future. You might see punishment directed specifically at you, so that you balance off the short-term gains from cheating against the long-run loss. That is what sustains OPEC. It is that kind of trade-off that we want to think about when we look at situations like open-source software or academic science and consider whether we can sustain in a non-cooperative environment free of strong property rights, free of integration into a single firm, a rich level of cooperation.

I want to talk about two contrasts. First, I think there is a bigger defection incentive – and Josh was touching on this in his closing comments – in the bioinformatics situation. Secondly, a technical concern to economists when they look at repeated games is whether there is a capital T, whether the "world" is going to end or whether it will continue on for a long period of time. The OPEC model of cooperation supposes that Saudi Arabia, Iran, Iraq, all the countries that are members, think they are going to be cooperating for an indefinite period of time. They do not see the end of the game coming at any particular point. On the other hand, if you think the end of the game is coming at any particular point, then there is a kind of backwards induction that leads to unraveling and undermines the possibility of cooperation. Nobody wants to cooperate in the last period, and they start moving back in time and nobody cooperates at the beginning either.

The reason I think a lot of bioinformatic projects are different is that you are moving along to a target, to a drug, and the game is over at that point. With software, it is continually mutating. New features are added; it is moved to new platforms; there are more efficient routines, etc. I do not see permanent, perpetual use of software products, but I also do not see, as a member of the open-source movement, a particular date when the development of Linux is coming to an end either.

I want to push further with that point, and my crude understanding of open-source software and genomics might be a liability, but at any rate . . . It seems as if there is no pot of gold for anyone to grab for at any particular point in time in the software creation process. There is no point in time where you would say this is a great time to defect, because the gains are going to be very large today. I know I will lose something in the future, but I just cannot resist the temptation. I see that as being more of a problem, though, in the

bioinformatics world or the genomics world, where there will be these pots of gold that people see are ripe to be seized. That is more likely to create an irresistible temptation to defect. The basic difference in the genomics process, the bioinformatics process, is that step by step, we are adding depth to research tools, and we are producing more information that is going to be helpful in finding an end product, but it is that end product where we cash in on the market. That is what is going to create the risk of bailing out at a particular point in time.

That is only a slice of the issues that Dan is talking about in a very interesting paper, but that is where I guess I get off the bus at the first stop with regard to the similarity between open-source and academic science.

Thank you very much. (*applause*)

PROFESSOR BURK:

I found most of these comments very helpful. Let me start by just replying briefly to Mike's comments. They surprised me a little bit. I think you are right that history does matter, and as people like Rebecca Eisenberg have talked about, we have this very long history in the academic research community of supposedly adhering to these norms of communality and sharing and so on. It is not immediately clear to me, however, that you are going to know when the last round is in that situation either. These people are, for example, in the genome project, engaged in incrementally adding to this large database, and when we are done with the genomics, then we move on to the proteomics. Thus, it seems to me that there might be much larger temptation to defect in the open-source community where the product times, as we said, are very short. We know what Linux is worth right now, and as Josh pointed out, you can use it immediately. We are not going to have to wait for drug development, and there seems to be a very big payoff available immediately. Where the community is actually one that is relatively young, it does not have much of a long or entrenched history with the kinds of mechanisms that we see in the scientific community like journals and peer review and so on that have grown up. I think your point is well taken, but I might actually come out the other direction, taking the signals to indicate there is more change of defection in open source.

PROFESSOR MEURER:

I would like to respond. You characterize the cooperative problem differently than I did, and you are looking at lots of people who are just in the business of doing life science research, and they will stick with it for a long period of time. Instead I was looking at particular projects with definite end points. When you look at open-source, when you look at Sendmail, you have a definite project and community people are working on that – Linux is probably the best example – so I am not sure which is the better way to look at it. I appreciate your point. One difficulty with your point, though, is that it

becomes harder to identify the community when you have lots of projects with different players. The identity of the players can be different on different projects, and that creates a new problem for sustaining cooperation when people are coming and going. If the project is specific and the players are the same, then cooperation is easier to sustain.

PROFESSOR BURK:

I think that is right, and one of the thoughts that occurred to me during your presentation is that, a little bit like antitrust analysis, it depends on how we define the market. It is going to depend on how we define the project and what we think the likely defection points might be. I am not entirely certain what those are, but I am not sure that I would intuitively think that they are what you immediately represented. I found Josh's comments to be equally useful. In part, I had thought about this problem of immediate applicability. At least some of the bioinformatics information we are talking about, some of the data, is going to be immediately applicable. As Incyte and Celera have proven, I can take this information and immediately license it out to people who are doing the end product research. Now, that is not as good as getting the end product itself, but there is some applicability there.

Again, I am not sure; it depends on which product we are talking about. Is the information itself the product, or is it an input into a downstream product that is going to happen later on? Let me just mention one problem that I thought Josh was going to raise but did not, which bears some more thinking. Josh talked about enforcement. In the open-source area there is at least the threat of enforcement by the Free Software Foundation or by Evan Moglen or other folks who are wandering around supposedly anxious to sue people if they violate the GPL, although I agree about the strategic ambiguity point. It is not clear who would perform that function if the scientific community had gone the other direction in genomics, who would be the heavy that would enforce these rights. Say that Reed Adler had had his way and NIH had in fact gotten these patents, would NIH really go out and sue people who were trying to privatize or tie up these genome sequences? It is not clear yet that that would really have happened, so one of the things that would have to be thought about would be, is there really anybody who would enforce that? I have said enough, and Professor Eisenberg wants to chime in.

PROFESSOR EISENBERG:

Yes, I am very happy that you are looking into this. This something I have been thinking about for awhile. I think that the open-source and open genome systems may have more in common ideologically and rhetorically than economically. A few things I wanted to sort of highlight for you, that I think you need to talk about that did not make it into your time-limited presentation. First is the role of the government in all of this, the sponsor of the Human Genome Project, which is just huge. It was very expensive. At least in the

early 90s, this was a lot of money we were talking about. This was really different from open source, where a bunch of guys were working in their garages or whatever, around their workstations. You really need a lot of money in order to crank up these things.

That brings into focus the Bayh-Dole Act.[14] The Bayh-Dole Act quite deliberately constrains the NIH in its ability to stop grantees from pursuing patents. Nonetheless, within that constraint, NIH has tried to play a role as enforcer of open genome. They have tried to do that through their adherence to these international "Bermuda Rules," which you must look into or discuss. These rules, agreed to by the international participants in the Human Genome Project in the mid-90s, in 1995 or 1996, require the deposit of DNA sequence information into public databases within twenty-four hours of getting sequence information. That looked, at the time, to some people, like it was going to short-circuit patent applications because it is tough to get a patent application on file, much less decide what you want to patent, within twenty-four hours. In fact, however, it does not prevent anybody from filing patent applications. It does not prevent others who mine a genetic bank for interesting tidbits of information from filing patent applications. The NIH has not constrained, and I do not think within the law can constrain, grantees from filing patent applications. I am not sure they even want to do so because they want to fulfill the mandate under the Bayh-Dole Act of promoting commercial development of products that come into view as a result of this research. In fact, they have been filing their own patent applications. One final thing I want to say is that your spin on what Reed Adler was up to is one spin –

PROFESSOR BURK:

It is Reed's spin. (*laughter*)

PROFESSOR EISENBERG:

No, it is not Reed's spin; it is Reed's occasional spin. It is not what Reed would say in writing at the time. Check out Reed's *Science* article on topic, check out Bernadine Healy's *New England Journal of Medicine* article at the time. They were telling a story that it was much more consistent with the Bayh-Dole Act's mandate that they patent the results of their research for the purposes of transferring to the private sector for commercial development. It was not anything like copyleft.

PROFESSOR BURK:

Yes, I have actually been very interested in the role of the government also, partly for the reasons that you mentioned, but even more so because of the

[14] 35 U.S.C. §§ 200-212 (2000).

problem that I mentioned of coordination. You have got these two communities that look uncoordinated, yet they are producing these very sophisticated types of outputs. Part of the research that Josh and my colleague David McGowan have been doing in the open-source community indicates that it is not nearly as uncoordinated as it looks. Linus Torvalds keeps a pretty heavy thumb on the development of the Linux kernel using copyleft to make sure that it does not fork and it goes the way he wants it to, and the government, I think, has played a huge role in coordinating, through meetings and transfers and all kinds of other things. It is not nearly as Hayekian as it first appears, and I think you are right that that role is very important.

PROFESSOR LERNER:

It is interesting to think about the ideological similarities of these two movements. One of the core issues in both is the hostility toward intellectual property. In the last couple of years, we have seen on the one hand advocates arguing against life forms patenting, against natural products patenting, and against patenting of traditional knowledge. We have also had advocates arguing against business method and software patents. In both, there is an underlying hostility toward intellectual property

But it is not clear that the open source movement must be – or even should be – hostile toward intellectual property. Does not the open-source movement depend at its core on copyright protection, which is then leveraged through licenses to control the manner in which the software is disseminated? Thus, it is using intellectual property to force public disclosure. The approach seems perfectly consistent with patent protection as well. I see no reason why the software underlying the open-source movement should not be unpatentable, so long as the patent rights are licensed parallel to the copyright. Likewise, in the genomic field, the existence of patents on bioinformatic methods, databases, and so forth would not necessarily preclude the sharing of knowledge.

PROFESSOR BURK:

You are absolutely right that in the open source area, the copyleft or GNU Public License depends, in the end, on the threat of copyright infringement lawsuit. If you want to work on this piece of code, that is fine. The terms are that you are going to have to make what you do available, and if you do not want to agree to that license and you do the work, then you are the infringer and theoretically we will sue you. I think that is part of the thing that appeals to these people who created the name GNU, Gnu's Not UNIX. This sort of self-referential and recursive use of property rights to try to prevent people from privatizing the material. As I have suggested in the presentation, it seems that that is actually a use of intellectual property very much in line with what Kitch suggested and what Coase suggested to coordinate the development of a resource in a certain way. Now, as we just talked about with Professor Eisenberg, they have not been doing that in the genomic area. They have been

coordinating in other ways through governmental intervention, through grants, through meetings and so on, so it is a very different type of model of how to coordinate that development.

PROFESSOR O'ROURKE:

Well, thank you very much, and thank you for coming today.