14 Open Source Software Projects as User Innovation Networks

Eric von Hippel

Free and open source software projects are exciting examples of user innovation networks that can be run by and for users—no manufacturer required.[1] Such networks have a great advantage over the manufacturer-centered innovation development systems that have been the mainstay of commerce for hundreds of years: they enable each user, whether an individual or a corporation, to develop exactly what it wants rather than relying on a manufacturer to act as its (often very imperfect) agent. Moreover, individual users do not have to develop everything they need on their own: they can benefit from innovations developed by others and freely shared within the user community.

User innovation networks existed long before and extend far beyond free and open source software projects. Such networks can be found developing physical products as well. Consider and compare the following examples of early-stage user innovation networks—one in software, the other in sports.

Apache Web Server Software

The Apache Web Server (which is free and open source software) is used on server computers that host Web pages and provide appropriate content as requested by Web browsers. Such computers are the backbone of the Internet-based World Wide Web infrastructure.

The server software that evolved into Apache was developed by University of Illinois undergraduate Rob McCool for, and while working at, the National Center for Supercomputing Applications (NCSA). The source code as developed and periodically modified by McCool was posted on the Web so that users at other sites could download, use, modify, and further develop it.

When McCool departed NCSA in mid-1994, a small group of webmasters who had adopted his server software for their own sites decided to take on the task of continued development. A core group of eight users gathered all documentation and bug fixes and issued a consolidated patch. This "patchy" Web server software evolved over time into Apache. Extensive user feedback and modification yielded Apache 1.0, released on December 1, 1995.

In the space of four years and after many modifications and improvements contributed by many users, Apache has become the most popular Web server software on the Internet, garnering many industry awards for excellence. Despite strong competition from commercial software developers such as Microsoft and Netscape, it is currently in use by more than 62 percent of the millions of Web sites worldwide.[2]

High-Performance Windsurfing

High-performance windsurfing, the evolution of which was documented by Shah (2000), involves acrobatics such as midair jumps and turns. Previously, the sport tended to focus on traditional sailing techniques, using windsurfing boards essentially as small, agile sailboats.

The fundamentals of high-performance windsurfing were developed in 1978 in Hawaii by a group of like-minded users. The development of a major innovation in technique and equipment was described to Shah by high-performance windsurfing pioneer Larry Stanley.

In 1978, Jurgen Honscheid came over from West Germany for the first Hawaiian World Cup and discovered jumping, which was new to him, although Mike Horgan and I were jumping in 1974 and 1975. There was a new enthusiasm for jumping, and we were all trying to outdo each other by jumping higher and higher. The problem was that the riders flew off in midair because there was no way to keep the board with you—and as a result you hurt your feet, your legs, and the board.

Then I remembered the "Chip," a small experimental board we had built with footstraps, and thought "It's dumb not to use this for jumping." That's when I first started jumping with footstraps and discovering controlled flight. I could go so much faster than I ever thought, and when you hit a wave it was like a motorcycle rider hitting a ramp; you just flew into the air. All of a sudden, not only could you fly into the air, but you could land the thing—and not only that, but you could change direction in the air!

The whole sport of high-performance windsurfing really started from that. As soon as I did it, there were about 10 of us who sailed all the time

together and within one or two days there were various boards out there that had footstraps of various kinds on them and we were all going fast and jumping waves and stuff. It just kind of snowballed from there.

By 1998, more than a million people were engaged in windsurfing, and a large fraction of the boards sold incorporated the user-developed innovations for the high-performance sport.

Over time, both of these user innovation networks have evolved and become more complex. Today, although they look different on the surface, they are in fact very similar in fundamental ways. Both have evolved to include many thousands of volunteer participants. Participants in free and open source software projects interact primarily via the Internet using various specialized Web sites volunteer users have set up for their use. Participants in innovation sports networks tend to interact by physically traveling to favorite sports sites and to types of contests that innovative users have designed for their sport. Most users of free and open source software simply "use the code," relying on interested volunteers to write new code, debug others' code, answer requests for help posted on Internet help sites, and help coordinate the project. Similarly, most participants in an evolving sport simply "play the game," relying on those so inclined to develop new techniques and equipment, try out and improve innovations developed by others, voluntarily provide coaching, and help to coordinate group activities such as leagues, and meets (Franke and Shah 2003).

User Innovation Networks "Shouldn't Exist," But They Do

Manufacturers, not users, have traditionally been considered the most logical developers of the innovative products they sell. There are two major reasons for this. First, financial incentives to innovate seem on the face of it to be higher for manufacturers than for individual or corporate users of a product or service. After all, a manufacturer has the opportunity to sell what it develops to an entire marketplace of users. Individual user-innovators, on the other hand, are seen as typically benefiting primarily from their own internal use of their innovations. Benefiting from diffusion of an innovation to the other users in a marketplace has been assumed to require some form of intellectual property protection followed by licensing. Both are costly to attempt, with very uncertain outcomes.

The second reason is that for an innovation to achieve widespread diffusion invention and development must be followed by production, distribution, and field support. Because these tasks involve large economies of scale for physical products, it has generally been assumed that

manufacturers have major cost advantages over individual users and networks of users. How could users possibly accomplish these tasks as cost-effectively as manufacturers?

Yet, implausible or not, user innovation development and consumption networks clearly do exist. Moreover, when products they develop compete head-to-head against products developed by manufacturers—Apache against Microsoft's and Netscape's server software, for example—the former seem capable of beating the latter handily in the marketplace. Not only do these networks exist, they even triumph! As Galileo is said to have murmured after officially recanting his statement that the earth moves around the sun: "And yet it moves!" What is going on here?

Conditions that Favor User Innovation Networks

We argue that complete fully functional innovation networks can be built up horizontally—with actors consisting only of innovation users (more precisely, "user/self-manufacturers"). Of course, nonuser enterprises can also attach to or assume valuable roles in user innovation networks. Red Hat and IBM provide well-known examples of nonuser involvement in the free and open source software context; professional sports leagues and commercial producers of sports equipment are examples in the case of user sports networks. It is only our contention that nonusers are not essential, and that "horizontal" innovation networks consisting only of users can develop, diffuse, maintain, and consume innovations.

Horizontal user innovation networks can flourish when (1) at least some users have sufficient incentive to innovate and do so, (2) at least some users have an incentive to voluntarily reveal their innovations and the means to do so, and (3) diffusion of innovations by users can compete with commercial production and distribution. When only the first two conditions hold, a pattern of user innovation and trial will occur, followed by commercial manufacture and distribution of innovations that prove to be of general interest.

Innovation by Users

Users have sufficient incentive to innovate when they expect their benefits to exceed their costs. Clearly, many innovators have a use-incentive for innovating in free and open source software projects. Thus, Niedner, Hertel, and Hermann (2000) report that contributors of code to open source projects asked to agree or disagree with statements regarding their

possible motivations for this ranked gain from "facilitating my work due to better software" as the highest-ranked benefit (average level of respondent agreement with that statement was 4.7 on a scale of 5). Similarly, 59 percent of contributors to open source projects sampled by Lakhani and Wolf (chap. 1, this volume) report that use of the output they create is one of the three most important incentives inducing them to innovate. Empirical research also documents the presence of user innovation in many additional fields. Thus, Enos (1962); Knight (1963); Freeman (1968); Rosenberg (1976a); von Hippel (1988); Shaw (1985); and Shah (2000) are among those finding that users, rather than manufacturers, are often the initial developers of what later become commercially significant new products and processes.

Innovation also has been found to be a relatively frequent activity among users that have a strong interest in a product or process area, and it tends to be concentrated in the "lead user" segment of user populations (see Table 14.1).[3]

Research on innovation-related incentives and capabilities provides a theoretical basis for all of these findings. Conditions under which users will—and will not—have incentives to innovate have been explored (von Hippel 1988). In addition, low-cost access to "sticky"[4]—costly to transfer—information has been found to be an important enabling factor for user innovation (von Hippel 1994; Ogawa 1997). Thus, information important to successful innovation, such as need and context of use information is *generated* at user sites and is naturally accessible there, but it can be very costly to move from users' sites to outside developers. For example, the conditions that cause software—and jumping windsurfers—to fail are available "for free" at the site of a user with the problem, but can be very costly to reproduce elsewhere. Also, information about user needs and the context of use is not static. Rather, it evolves at the user site through "learning by doing" as the user experiments with prototype innovations. (Recall from the windsurfing example that users *discovered* that they could and wanted to control the direction of a board when it was in the air only *after* they began experimenting with the prototype footstraps they had developed.)

The concentration of innovation activity among the "lead users" in a user population can also be understood from an economic perspective. Given that innovation is an economically motivated activity, users expecting significantly higher economic or personal benefit from developing an innovation—one of the two characteristics of lead users—will have a

Table 14.1
User innovation tends to be frequent and concentrated among "lead users"

Innovation area	Number of users sampled	Percentage developing and building innovation for own use	Were the innovating users "lead users"?
Industrial products			
Printed circuit CAD software (a)	136 user/firm attendees at PC-CAD conference	24.3%	Yes
Pipe hanger hardware (b)	74 pipe hanger installation firms	36%	NA
Library information systems (c)	102 Australian libraries using computerized library information systems	26%	Yes
Apache OS server software security features (d)	131 Apache users	19.1%	Yes
Consumer products			
Outdoor consumer products (e)	153 outdoor-specialty mail-order catalog recipients	9.8%	Yes
"Extreme" sporting equipment (f)	197 expert users	37.8%	Yes
Mountain biking equipment (g)	291 expert users	19.2%	Yes

Sources: (a) Urban and von Hippel 1988; (b) Herstatt and von Hippel 1992; (c) Morrison, Roberts, and von Hippel 2000; (d) Franke and von Hippel 2002; (e) Luthje 2003; (f) Franke and Shah 2003; (g) Luthje, Herstatt, and von Hippel 2002.

higher incentive to innovate and are therefore more likely to do so. Also, given that lead users experience needs in advance of the bulk of a target market, the nature, risks, and eventual size of that target market are often not clear to manufacturers. This lack of clarity can reduce manufacturers' incentives to innovate, and increase the likelihood that lead users will be the first to develop their own innovative solutions for needs that later prove to represent mainstream market demand.

User Incentives to Freely Reveal Their Innovations

Progress and success in user innovation networks is contingent on at least some users "freely revealing" their innovations.[5] Without free revealing, each user would have to redevelop the same innovation in order to use it, resulting in a huge system-level cost, or resort to protecting and licensing their innovations and collecting revenues from other users, which would burden the networks with tremendous overhead.

Research has shown that users in a number of fields do freely reveal details of their innovations to other users and even to manufacturers (von Hippel and Finkelstein 1979; Allen 1983; Lim 2000; Morrison, Roberts, and von Hippel 2000; Franke and Shah 2003). Of course, free revealing is clearly visible in free and open source software networks, and is also clearly present in the sports innovation example; innovating users gather on the beach, inspect one another's creations, and imitate or develop additional modifications that they, in turn, freely reveal.

To economists, free revealing is surprising, because it violates a central tenant of the economic theory of innovation. In this classical view, appropriating returns to innovation requires innovators to keep the knowledge underlying an innovation secret or to protect it by patents or other means. After all, noncompensated spillovers of innovation-related information should represent a loss that innovators would seek to avoid if at all possible, even at some cost. Why then do we observe that some innovation-related information is voluntarily freely revealed?

The answer to this puzzle has several components. First, note that software code (and other public goods) have aspects that remain private to the innovator even after the code has been freely revealed as a public good. This thinking has been codified in a "private-collective" model of innovation incentives (von Hippel and von Krogh 2003). As illustration, consider some of the private benefits retained by users who write and then freely reveal their code. Code may be written precisely to suit the private needs of the code writer—and may serve the needs of free riders less well (Harhoff et al. 2003). Also, the learning and enjoyment gained from actually writing

the code—benefits that have been shown to be highly valued by contributors to open source software projects (Lakhani and Wolf, chap. 1, this volume)—cannot be shared by free riders who only adopt the completed product. Nor can the private reputation of an innovator be shared by a free-riding adopter of that innovation (Lerner and Tirole 2002). Finally, when free riders do adopt and use code that has been freely revealed, that action in itself leads to significant private benefits for the code creator: others will help debug the code; it may be integrated into the authorized OS code, leading others to help update and maintain it; higher use (greater "market share") will yield "network effect" advantages; and so on.

A second point important to explaining the practice of free revealing is that profitably creating and serving a market for software you may develop is often not a trivial undertaking. And when benefits from free revealing such as those just described exceed the benefits that are *practically* obtainable from other courses of action such as licensing or selling, then free revealing should be the preferred course of action for a profit-seeking firm.

Finally, we note that the costs associated with free revealing may be low—or in any case, unavoidable—because others who know the same thing will reveal even if you do not. And when the costs of freely revealing an innovation are low, even a low level of benefit can be adequate reward. Competitive losses from free revealing of intellectual property depend upon the degree of rivalry between the software developer and those who may adopt that software as free riders. Thus, users who write and freely reveal software code will expect low losses if they have only low or no rivalry with potential adopters. (For example, there is low rivalry among town libraries: they serve different populations and do not seek to gain market share from each other.) Also, if more than one person or firm has developed a particular piece of software, everyone's decision to freely reveal can be determined by the action of the innovator with the *least* to lose. That is, even those who would prefer to hide their software to keep it from rivals may nonetheless freely reveal if they expect that others will do this if they do not (Lakhani and von Hippel 2003).

Innovation Diffusion by Users

"Full-function" user innovation and production networks—no manufacturer required—are possible only when self-manufacturing and/or distribution of innovative products directly by users can compete with commercial production and distribution. In the case of free and open source software, this is possible because innovations can be "produced"

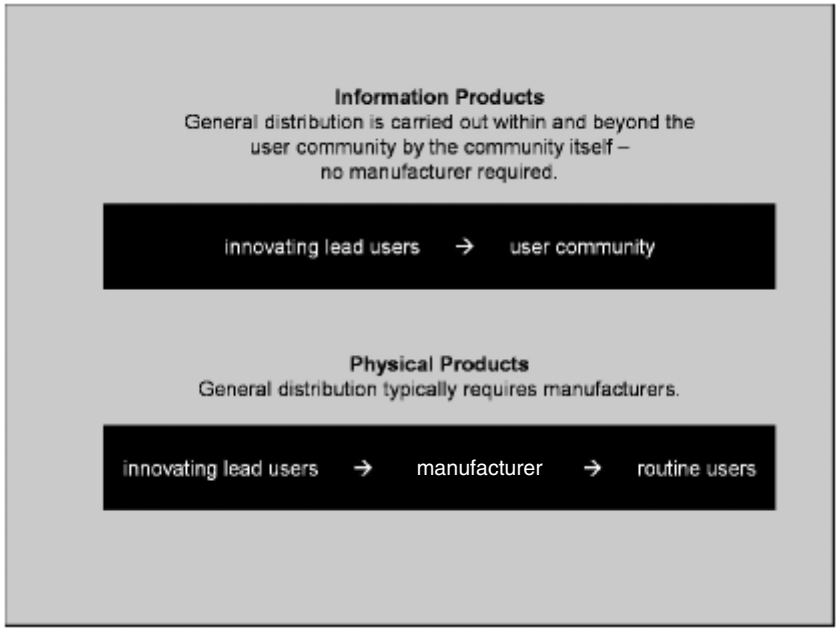

Figure 14.1
How lead user innovations are distributed

and distributed essentially for free on the Web, software being an information rather than a physical product (Kollock 1999). In the case of the sports innovation example, though, equipment (but not technique) innovations are embodied in a physical product that, to achieve general diffusion, must be produced and physically distributed; activities that, as mentioned earlier, involve significant economies of scale. The result, in the case of the windsurfing example and for physical products generally, is that while innovation can be carried out by users and within user innovation networks, production and diffusion of products incorporating those innovations is usually carried out by manufacturing firms (figure 14.1).

Ongoing Exploration of User Innovation Networks

The advent of the Web and consequent public proliferation of free and open source software development projects has focused intense academic attention on the phenomenon of user innovation networks in general, and free and open source software in particular. The thousands of extant free

and open source software projects represent natural experiments that academics and others can study to better understand this phenomenon. Among the issues being explored now are conditions under which free and open source software projects can be expected to succeed, how they can be most successfully managed, and what attracts the interest of volunteers. We can expect rapid progress on these fronts.

What is very exciting, it seems to us, is that innovation networks exclusively by and for users, networks that by any yardstick of traditional economics shouldn't exist, can create, diffuse and maintain complex innovation products without *any* manufacturer involvement. This means that in at least some, and probably many, important fields users can build, consume, and support innovations on their own independent of manufacturer incentives to participate and manufacturer-related "agency costs."[6]

Direct development and diffusion of innovations by and for users via horizontal user innovation networks can improve individual user's abilities to get what they really want—because they have an increasingly practical and economical pathway to "do it themselves." As we learn to understand these networks better, we will be in a position to improve such networks where they now exist and may be able to extend their reach and attendant advantages as well.[7]

Notes

1. In the "functional" sources of innovation lexicon, economic actors are defined in terms of the way in which they expect to derive benefit from a given innovation. Thus, firms or individuals that expect to profit from an innovation by in-house use are innovation "users." Innovation "manufacturers," in contrast, are firms or individuals that expect to profit from an innovation by selling it in the marketplace (von Hippel 1988). By user "network," I mean user nodes interconnected by information transfer links that may involve face-to-face, electronic, or any other form of communication. User networks can exist within the boundaries of a membership group but need not. User innovation networks also may, but need not, incorporate the qualities of user "communities" for participants, where these are defined as "networks of interpersonal ties that provide sociability, support, information, a sense of belonging, and social identity" (Wellman, Boase, and Chen 2002, 4).

2. Netcraft April 2003 Web Server Survey, http://news.netcraft.com/archives/2003/04/13/april_2003_web_server_survey.html.

3. Lead users are defined as users of a given product or service type that combine two characteristics: (1) lead users expect attractive innovation-related benefits from a solution to their needs and are therefore motivated to innovate, and (2) lead users

experience needs that will become general in a marketplace, but experience them months or years earlier than the majority of the target market (von Hippel 1986). Note that lead users are not the same as early adopters of an innovation. They are typically ahead of the entire adoption curve in that they experience needs before *any* responsive commercial products exist—and therefore often develop their own solutions.

4. The stickiness of a given unit of information in a given instance is defined as the incremental expenditure required to transfer that unit of information to a specified locus in a form useable by a given information seeker. When this cost is low, information stickiness is low; when it is high, stickiness is high. A number of researchers have both argued and shown that information required by technical problem-solvers is indeed often costly to transfer for a range of reasons (von Hippel 1994). The requirement to transfer information from its point of origin to a specified problem-solving site will not affect the locus of problem-solving activity when that information can be shifted at no or little cost. However, when it is costly to transfer from one site to another in useable form—in my term, sticky—the distribution of problem-solving activities can be significantly affected.

5. When we say that an innovator "freely reveals" proprietary information, we mean that all existing and potential intellectual property rights to that information are voluntarily given up by that innovator and all interested parties are given access to it—the information becomes a public good. Thus, free revealing of information by a possessor is defined as the granting of access to all interested agents without imposition of any direct payment. For example, placement of nonpatented information in a publicly accessible site such as a journal or public Web site would be free revealing under this definition. Note that free revealing as so defined does not mean that recipients necessarily acquire and utilize the revealed information at no cost to themselves. Recipients might, for example, have to pay for a journal subscription or an Internet connection or a field trip to acquire the information being freely revealed. Also, some may have to obtain complementary information or other assets in order to fully understand that information or put it to use. However, if the information possessor does not profit from any such expenditures made by information adopters, the information itself is still freely revealed, according to our definition (Harhoff et al. 2003).

6. *Manufacturers* are the agents of users with respect to new products and services. It is their job to develop and build what users want and need; they do not want the products for themselves. The trouble is that, when manufacturers' incentives don't match those of users—and they often do not—users end up paying an agency cost when they delegate design to manufacturers. A major part of this agency cost takes the form of being offered products that are not the best possible fit with users' needs, even assuming that manufacturers know precisely what those needs are. Manufacturers want to spread their development costs over as many users as possible, which

leads them to want to design products that are a close enough fit to induce purchase from many users rather than to design precisely what any particular user really wants.

7. Recent working papers on free and open source software and user innovation by many researchers can be downloaded from the Web sites http://opensource.mit.edu and http://userinnovation.mit.edu. These sites are intended for those interested in keeping up-to-date on, and perhaps contributing to, our understanding of these phenomena.