

Synthesis

Open Source and Open Content: a Framework for Global Collaboration in Social-Ecological Research

Charles Schweik¹, Tom Evans², and J. Morgan Grove³

ABSTRACT. This paper discusses opportunities for alternative collaborative approaches for social-ecological research in general and, in this context, for modeling land-use/land-cover change. In this field, the rate of progress in academic research is steady but perhaps not as rapid or efficient as might be possible with alternative organizational frameworks. The convergence of four phenomena provides new opportunities for cross-organizational collaboration: (1) collaborative principles related to "open source" (OS) software development, (2) the emerging area of "open content" (OC) licensing, (3) the World Wide Web as a platform for scientific communication, and (4) the traditional concept of peer review. Although private individuals, government organizations, and even companies have shown interest in the OS paradigm as an alternative model for software development, it is less commonly recognized that this collaborative framework is a potential innovation of much greater proportions. In fact, it can guide the collective development of any intellectual content, not just software. This paper has two purposes. First, we describe OS and OC licensing, dispense with some myths about OS, and relate these structures to traditional scientific process. Second, we outline how these ideas can be applied in an area of collaborative research relevant to the study of social-ecological systems. It is important to recognize that the concept of OS is not new, but the idea of borrowing OS principles and using OC licensing for broader scientific collaboration is new. Over the last year, we have been trying to initiate such an OS/OC collaboration in the context of modeling land use and land cover. In doing so, we have identified some key issues that need to be considered, including project initiation, incentives of project participants, collaborative infrastructure, institutional design and governance, and project finance. OS/OC licensing is not a universal solution suitable for all projects, but the framework presented here does present tangible advantages over traditional methods of academic research.

Key Words: *land use; land cover; modeling; open content; free/libre software; open source software;*

INTRODUCTION

Traditional approaches to the communication and validation of scientific research, e.g., peer review, and to the communication of findings, e.g., refereed publication, have been in place in some form since shortly after the development of the printing press in the 16th century (Ziman 1969, Johns 2001). This process of peer review as a mechanism to check for credible information (Burnham 1990, Kronick 1990) and journal publication has resulted in great progress in scientific knowledge over the last four centuries. The process also provides an example of how advances in technology, such as the printing

press coupled with systems for the delivery of mail, can change the speed at which scientific knowledge can grow.

However, the evolution of the traditional process of scientific discourse over the course of four centuries is not without two age-old, and one more recent, problems. The first relates to the costs of scientific publishing, which force the kind of brevity in publication notable in word-count restrictions. A second problem is the cost of the global distribution of paper copies of the journal. The third more recent problem involves the difficulty in making widely available scientific knowledge based on complicated,

¹University of Massachusetts, Amherst, ²Indiana University, ³U.S. Forest Service, Northeastern Research Station

technical expertise often grounded in computing. For example, in the context of modeling land-use change, a topic discussed later in more depth, significant and complex models exist, but they are rarely further developed or applied by anyone outside the original founding group because of the transaction costs required to learn and apply them elsewhere. Moreover, Aber (1997:232) notes that, in the field of ecology, there is a general distrust of such models because it is assumed that these projects and papers are not being held to the same standards of full disclosure and peer review as in other areas of scientific research. We would argue that Aber's conclusion is, at least in part, driven by the first problem of brevity in publishing.

An emerging opportunity?

Drucker (1999) observed that, when new technologies emerge, it often takes years before real innovation using this technology occurs. A similar lag exists in the area of scholarly communication. For example, with the exception of one or two emerging innovators, most "e-journals" follow the design of the old paper model of volumes and issues, with hyperlinks to papers or abstracts. Although this is an important innovation in itself, e.g., making scholarly material more accessible worldwide, Web technologies have the capacity to adapt traditional methods of scientific communication and collaboration to systems that are much more powerful, immediate, and interactive. Internet-based technologies and collaborative approaches that have been around for nearly a decade can be applied in new ways to enhance methods of communicating and build on scientific research. This paper describes some of the features that might be present in a new framework for collaborative scientific research that includes:

1. the Web as an interactive platform for communicating throughout the research process,
2. the traditional concept of peer review as a mechanism for maintaining quality control and encouraging participation,
3. practices for licensing and developing free/libre open source software (FOSS), and
4. recent developments in open content (OC) licensing.

Taken together, these four components create the foundation for a new approach to global scientific collaboration, in the form of what we refer to as a "next-generation peer-reviewed e-journal." Undoubtedly, readers of *Ecology and Society* are familiar with the Web (the first item above) and the traditional peer-review process for evaluating scientific research (the second). Therefore, this paper will elaborate on the third (FOSS licensing and collaborative practices) and fourth (OC) to describe more fully these components and to raise important issues we see looming on the horizon.

To do this, we first describe the concept of FOSS software licensing and related aspects of FOSS collaborative development practices. Second, we describe the emerging OC licensing phenomenon. Third, we address the question some may ask as to whether the FOSS approach is really a special case of the traditional scientific process, because we see several important differences. Fourth, we provide a vision of the future, in particular, how we might apply an OS/OC approach to domains beyond software development in the broader area of scientific research. To do this, we use an example that is directly relevant to *Ecology and Society* readers: the use of FOSS and OC collaborative principles in modeling changes in land use. Using this example, we raise several important issues that will need to be examined in more depth for the proposed collaborative paradigm to become operational.

FREE/LIBRE AND OPEN SOURCE SOFTWARE LICENSING AND COLLABORATIVE DEVELOPMENT PRACTICES

Most software developed by companies or individuals is treated as proprietary with full copyright protection and distributed as compiled, i. e., unreadable, executable programs. In other words, such software is treated as private or toll goods (Ostrom and Ostrom 1977), and the proprietary license agreement typically prohibits further copying of the software for use on other computers.

However, beginning in 1984, Richard Stallman and others who are now part of the [Free Software Foundation](http://www.fsf.org/) (FSF), initiated a social movement in the software industry, arguing that because of the digital properties of software, users should be provided the freedom to use, modify, and

distribute the software as they deem necessary (Stallman 1999, 2002). Consequently, this type of software is often referred to as "free" or "libre" (FL) software (Ghosh et al. 2002), which is different from the commonly used term "freeware." A major innovation in the effort by Stallman was the principle of "copyleft," an approach that uses copyright law in a way that maintains the following freedoms for software users: (1) the right to run the software, (2) the right to read the software source code and modify it, (3) the right to redistribute the original version of the software, and (4) the right to redistribute modified versions of the program (Stallman 1999). The principle of copyleft also stipulates that all future versions of the software be assigned these same principles. The FSF summarizes the approach this way:

"To copyleft a program, we first state that it is copyrighted; then we add distribution terms, which are a legal instrument that gives everyone the rights to use, modify, and redistribute the program's code or any program derived from it but only if the distribution terms are unchanged. Thus, the code and the freedoms become legally inseparable." (For full details, see the [FSF definition of copyleft](#)).

Stallman implemented these principles through a software license referred to as the GNU (GNU's Not Unix) General Public License or GPL.

Open source (OS) software licensing follows these same general principles and is a derivative of Stallman's free software work (Perens 1999). As Weber (2004:16) notes, "Property in open source is configured fundamentally around the right to distribute, not the right to exclude." However, OS licenses do not always follow the complete "freedom standards" set out by the FSF, and therefore are seen as a different "political camp." The difference between FL and OS software has to do with other restrictions provided in these OS licenses that may limit the freedom of users in what they can do with the software (Perens 1999). For example, some OS licenses may allow their software to be used in proprietary software packages, a practice that the FSF rejects. More than 50 OS licenses now exist (see [opensource.org](#)), demonstrating that authors of software have an array

of rights they can keep or relinquish rather than the simple dichotomous choice of either full copyright or "all rights reserved," the default condition, or no copyright or "public domain." For comparisons of various FL and OS software licenses, see Perens 1999 or www.gnu.org.

The permission to copy, modify, and distribute readable software source code found in free/libre or open source software (FOSS) licenses provides at least two potential advantages over the traditional proprietary, full-copyright approach to software development. All FL and most OS software packages are provided at no monetized cost to the end-user. This creates a major incentive for it to be used, especially by people and organizations that are working under limited budgets (Hahn 2002). Second, by providing readable source code and allowing new derivative works to be made, FOSS licenses create the potential to generate a large community of users and developers, larger than any one proprietary organization could create, to develop, test, and debug future versions of the software (Raymond 1998).

This innovation in software licensing, coupled with Internet-based collaborative tools, resulted in a new form of collaboration. FOSS projects represent a form of "commons" (Ostrom 1990, Dietz et al. 2003), but one that differs slightly from the environmental commons that most readers are familiar with (Hardin 1968, Ostrom 1990). In FOSS commons, groups of people act collectively to produce a public good, i.e., the software, rather than over-appropriate the resource (e.g., Hardin 1968). Of particular importance is the fact that FOSS software projects produce this public good through a common property regime (Benkler 2002, Boyle 2003a,b, Schweik 2005). These regimes typically have some type of governance structure, with one or more individuals having the authority to prohibit others from contributing if they so desire (Schweik and Semenov 2003, Weber 2004).

The distributional and readable source requirements of most FOSS serve additional social purposes beyond the issue of property rights. Deep social and cultural roles and norms have been the driving force for these licenses. They promote the ideals of sharing, collaboration, learning from reading the work of others, learning from peer review of one's own software contributions, and feeling part of a "virtual community" or a social movement (Bollier 1999, Stallman 2002). These kinds of social norms

drive people's willingness to participate in the production of many FOSS projects and are important for driving similar collaborations in the context of scientific research. Schweik and Semenov (2003) show that there are many similarities between some of these norms and incentives of FOSS programmers and those faced by scientists or academics. This has important implications for the proposed extension of FOSS approaches into the scientific research domains that we describe later in this paper.

Thousands of FOSS products are currently under development. At SourceForge.net, a prominent Web site for hosting FOSS projects, there are nearly 100,000 hosted FOSS projects and more than 10⁶ registered users. Many of these projects are surely obsolete. However, even if 10% of them are active, this is substantial. The Linux operating system and the Apache Web Server software are perhaps two of the oldest and most prominent examples of FOSS success stories. They can be considered "enterprise" software in that they have achieved a significant global market share and amassed a global user and development community.

Nevertheless, although the copyleft licensing principle is a major innovation that has led to the emergence of FOSS, there are other institutional and infrastructural components that drive and support FOSS collaborations. Given the complexity of this topic and the brevity needed for this paper, we cannot address all of the collaborative components here. However, we will mention three important ones: (1) a deeply committed and/or funded participant base, (2) an Internet-based collaborative infrastructure, and (3) an established system of project governance. Let us elaborate on each, in sequence.

First, in the early days of FOSS development and in the literature about this phenomenon, the emphasis was on the volunteer nature of the developer communities who contributed to "flagship" FOSS software like the Linux operating system or the Apache Web Server. Indeed, this volunteer nature still drives many FOSS projects. An argument often made in support of the FOSS software development model is that it has the potential to create teams with participant numbers rivaling or exceeding anything commercial firms can establish, as expressed in the now famous "with more eyes, all bugs are shallow" quote by Raymond (1998). However, more recent studies of FOSS

software report that most projects have only a handful of participants (Dempsey et al. 2002, Ghosh et al. 2002, Krishnamurthy 2002, Healy and Schussman 2003), and, even in some projects with very large numbers of people associated with them, only a small percentage of the participants appear to be performing the work (Warsta and Abrahamsson 2003). However, what is clearly demonstrated in some of the real success stories in FOSS, such as the case of Linux and the Apache Web Server, is that the early developers were passionate in what they were doing and often driven by a feeling of a social movement such as Stallman's "freedom" ideas or the challenge of "taking on" the dominance of Microsoft products. These social movements motivated some developers to contribute their time and intellectual property without monetary compensation. Today, we are witnessing what appears to be a shift by some companies in the software industry to place human and organizational capital behind the development of FOSS products that are considered of strategic interest, e.g., Hewlett Packard, IBM, etc. In addition, some governments are starting to place more emphasis on the use and possible support for FOSS (Hahn 2002). In short, either a passionate developer community, some financial support, or both are required for a FOSS project to succeed.

Second, regardless of the size of the participant community, a set of Internet-based tools for communication and coordination is required (Dube and Pare 2001, Kelly and Jones 2001, Halloran et al. 2003, Shaikh and Cornford 2003). Components include version and concurrency control, build management, workplace management, and change request and release management (Asklund and Bendix 2002). Many of these projects use systems like Concurrent Versioning System or CVS, project management platforms like SourceForge.net, and various communication tools such as chat/instant messaging, discussion forums, Wikis, email, FOSS project Web pages, etc.

Third, as projects become larger in scope and in number of participants, it is likely that some form of governance structure will be required to make decisions about important project issues (Schweik and Semenov 2003). Such a governance structure might include: (1) assigning priorities to features to be included in new versions of the software, (2) defining rules and procedures on how production will proceed, and (3) assisting in the resolution of disputes between team members. For example,

Sharma et al. (2002:13) note that FOSS communities create and abide by sets of rules that are modified over time as the project matures. Studies such as Divitni et al. (2003) and Shaikh and Cornford (2003) provide examples of conflict in FOSS settings. This is one area of FOSS projects that is less understood but could be important for understanding why some projects succeed and others fail.

To summarize, a licensing innovation and a motivated or financially supported developer community using Internet-based collaborative infrastructure tools and working within some set of governance principles are the core elements of FOSS software collaboration projects. This configuration provides the foundation for a new paradigm in collaborative scientific research. However, making that transition requires attention to the next innovation in FOSS licensing: what we refer to as open content licensing.

For readers interested in more information and studies on elements of FOSS, there are at least three relatively large research repositories on the FOSS phenomenon. See, for example, the [MIT Free/Open Source Research Community Web site](#), the [Open Source Resources for Researchers](#), and the [Open Source Software Research program](#) at the Institute for Software Research at the University of California at Irvine.

THE EMERGENCE OF OPEN CONTENT LICENSING

Although the principle of copyleft and FOSS software licensing have been around for two decades, it is not widely recognized that this licensing innovation is applicable to other intellectual property content in addition to computer code (Bollier 1999, Schweik and Grove 2000, Stallman 2001, Weber 2004).

This idea was first advanced by Stallman's group, who felt the "freedom philosophy" should apply not only to software, but also to user guides, technical documentation, etc. In this context, they developed the [GNU \(GNU's Not Unix\) Free Documentation License](#) or GFDL to govern the use, modification, and distribution of software documentation. They also set requirements on what parts of the document must remain unmodified from version to version,

such as the original author's copyright notice, the terms of distribution, and a list of previous authors (Stallman 1999).

Around the year 2001, a new set of licenses following similar principles was developed by people associated with [Creative Commons.org](#) that can be applied to work created by authors or artists. These developers have created a set of Creative Commons (CC) licenses that allow authors to select, in a modular way, which rights they want to reserve and which to relinquish related to the content they have produced. Key issues and questions related to the software licensing above drive various CC license "module" choices, including:

1. distribution. Can readers freely copy and distribute this intellectual property? If yes, the user of the content can download and freely distribute the associated content.
2. derivative works.
Can derivative work be made using the content as a base? If yes, the user of the software or content can create new works based on the available and readable digital files.
3. If derivative works are permitted (yes to question 2A), must these derivations fall under the same license as the parent work? In other words, must they form part of a "viral" licensing scheme (Pavlicek 2000), or can they be distributed under a different licensing scheme?

author attribution. Is author attribution required? Answering yes to this question requires the user of software or content to give credit to the original author(s) who developed the work.

One distinction between Stallman's GFDL and the CC licenses is that the CC licenses are more modular. This means that authors can select, through the answers to the above questions, which rights to maintain and which rights to give away. With such licensing, authors of other content, e.g., music, artwork, academic or scientific papers, now have the opportunity to hold "some rights reserved" (Stix 2003), in a fashion similar to what has occurred in the domain of OS software development. In this paper, work other than software or software documentation will be referred to as open content

(OC).

Good examples of the use of Stallman's GFDL can be found at various "wiki" sites such as [Wikipedia](#) and [Wikibooks](#). Interested readers may want to review the entry for "open content" in Wikipedia, which provides a list of some emerging open content projects.

IS THE FREE OR OPEN SOURCE SOFTWARE APPROACH SIMPLY A SPECIAL CASE OF THE TRADITIONAL SCIENTIFIC PROCESS?

Bezroukov (1999a,b) and others have suggested that the traditional method of submitting papers to refereed journals and publishing one's findings is similar to the process of developing free or open source software (FOSS). However, in our view, there are some important differences (Schweik and Semenov 2003). First, in FOSS collaboration settings, the entire research product (software) is shared with the community, including the "research" (software development) process, rather than just the equivalent of a final paper that summarizes methods and provides end results. This differs from traditional publishing, in which space restrictions in journals limit what can be provided to the community. Second, the open access nature of FOSS collaboration over the Internet increases the size of the community that might contribute to the project. Third, the right to freely copy and distribute copies of the intellectual property (software) is very different from the policies of many scientific journals that hold full copyright and require the reader to obtain permission before copies can be made. Fourth, FOSS collaboration over the Internet potentially increases the speed at which innovations can be published. Systems of peer review in FOSS contexts might take similar amounts of time compared to traditional peer-review processes in scientific journals, but the results, e.g., improvements to a program module, can be published much more quickly, sometimes almost immediately after the peer-review process is completed.

In short, the convergence of (1) collaborative principles related to FOSS development, (2) the emerging area of open content (OC) licensing, (3) the Web as a platform for scientific communication, and (4) the traditional concept of peer review provides the foundation for a new paradigm in the

production of collaborative scientific research. Weber (2004:17) articulates this nicely: "The open source process has generalizable characteristics, it is a generic production process and it can and will spread to other kinds of production." We refer to this emerging paradigm in science as open source- or content-based collaborative research, or OS/OC research for short. The remainder of this paper focuses on important issues that will need further thought to make this transition to OS/OC research. To place these issues in some context, we focus our discussion around one area of research relevant to readers of *Ecology and Society*: modeling land-use change.

APPLYING AN OPEN SOURCE/OPEN CONTENT COLLABORATION APPROACH IN SCIENTIFIC RESEARCH: THE EXAMPLE OF MODELING LAND-USE CHANGE

Modeling land-use change is an important area in the domain of ecology and society and one that is ripe for an open source/open content (OS/OC) collaborative endeavor for several reasons. First, there is substantial global interest in such tools. Scientists and government officials desperately need models of this type to help them better understand how land-use systems function, to explore the effects of various policy or planning initiatives, and to assist in the building of consensus (Dutton and Kraemer 1985, King and Kraemer 1993). Second, in the current system of research and publication, cross-organizational participation in model development is happening at a snail's pace, if at all.

We base this statement on recent reviews of available models of land-use change (EPA 2000, Agarwal et al. 2002, Grove et al. 2002). These reports identify more than 30 different models currently available, each of which uses different approaches and technologies. Some of these models are in the public domain, some are licensed for free or open source software (FOSS), and others are proprietary packages. Although having a relatively large number of models of land-use change is beneficial in terms of innovation, building upon this prior work is difficult because these models often (1) require special disciplinary expertise that may not be available, (2) are technologically and theoretically complex, (3) require extensive data, and (4) are not readily available, in part because of

publishing limitations. Consequently, a researcher or policy maker who wants to try to apply a model to another geographic location faces substantial transaction costs. In short, this model-review literature makes it apparent that the scientific community needs to find better ways to build upon each other's work, because it appears that most of the evolution of individual models over time comes from the same groups that initially developed them.

Parker and colleagues (2002:212) raise similar issues in their discussions of the challenges of integrated assessment and modeling. They note that some of the barriers to advancing models, including models of land-use change, are fragmentation in education and the separations between scientists in different disciplines, between scientists and modelers, between application modelers and software interface designers, and among scientists, decision makers, and the community. For all of these reasons, we believe that building a land-use modeling effort based around an OS/OC collaborative paradigm might lead to some real productivity enhancements in this area of scientific research and provide a mechanism to close some of these gaps between stakeholders. Just how might this be accomplished?

In August 2003, we held an interdisciplinary workshop to initiate such an effort (Schweik et al. 2003). Several models of land-use change were represented, and participants came from a variety of academic institutions and organizations active or with an interest in modeling land-use change. UrbanSim (Waddell 1998, 2000), a simulation model for the integrated planning and analysis of urban development, was one such model that garnered particular interest from workshop members, and it is the one we use to better describe the issues of OS/OC modeling we see ahead. Interested readers can find a short summary of UrbanSim in Appendix 1. Technically, it has a "model coordinator" program and five submodules: Accessibility, Economic and Demographic Transition, Household and Employment Mobility, Household and Employment Location, and Real Estate Development.

As an example, UrbanSim demonstrates the multidisciplinary nature and the levels of technical, theoretical, and empirical sophistication involved in current-day models of land-use change (see EPA 2000 and Agarwal et al. 2002 for descriptions of other models of land-use change). Applying the

model involves data gathering, formatting, and a variety of other steps; some of these must be taken manually by the analyst, whereas others can be carried out automatically by programs such as the model coordinator (see Appendix 1). As in FOSS development situations, to make improvements or alternations to a submodule, some programming skills may be required. However, from the standpoint of making scientific advances in modeling land-use change, moving to an OS/OC approach extends the paradigm to new areas. Intellectual contributions are required not only from modeler/programmers but also from theoreticians and scientists working in disciplines that range from economics, geography, and political science to sociology and ecology, to name just a few, who are interested in the processes of land-use change, as well as practitioners such as policy analysts and planners who are interested in what information or scenario projections these models may reveal for a geographic area of interest. Moreover, different driving variables of change and processes will become important as these models are applied in various regions of the Earth. When applications of models like UrbanSim expand and are applied to other geographic areas, which is already happening, it is likely that the logic within various submodules will need to be changed to capture the effect of different relevant processes. For example, we would expect some variation in the causes or drivers of change (Lambin et al. 2003) when comparing the contexts of Phoenix, Arizona, USA, and Baltimore, Maryland, USA.

There are many ways that individuals could contribute to an OS/OC project in general and more specifically to the further development of models of land-use change like UrbanSim. People with the technical skills to "work under the hood" of a model, e.g., researchers with programming skills, could write or modify existing code or contribute entirely new modules, thereby extending the operations, manipulations, or analyses that data sets might be subjected to. Others with some programming knowledge but with less interest in writing new code or functions could serve a critical role in code review, which is a form of peer review in a software context, or in code documentation.

Scientists who lack programming skills could help identify the bugs in a model module by using and testing it. With complex modeling frameworks, model users outside the core model development team could also make critical contributions by

developing user manuals or other modeling documentation or conceptualizing new functions such as model validation approaches and tools (Pontius et al. 2004) for someone else to program. In addition, new model users and developers can facilitate the "many eyes" approach (Raymond 1998), in which new ideas and insights may be contributed to the original model that its developers would not have considered. A critical aspect of this approach would be to provide enough distance-learning material to allow the model to be explored and applied by researchers from disciplines other than those represented in the model's development. Similar collaborations could involve users in the development of "how to" documentation, or theorists in providing key insights through papers placed under some form of OC licensing. Last, but perhaps most important, theoreticians who are not modelers can still add crucial elements to the collaborative endeavor by contributing the theoretical base through papers that provide the foundation for the logic of particular model components.

From an empirical standpoint, an OS/OC licensing approach would encourage the application of a model like UrbanSim in different geographic areas. Different users or interested stakeholders could contribute data, perhaps also under an OC license, that could be used to test a single model in two different locations that have different social-ecological situations. These comparative applications would likely result in new specifications for model refinements. Models licensed as "new derivative works permitted" could then be tailored accordingly by the same researcher or another in the collaborative community, leading to insights that would not have been apparent by running the model in only a single location. Under OS/OC, researchers could collaborate in the parameterization of models for different locations. For example, Clarke (K. Clarke, *personal communication*) recently reported calibration results of the SLEUTH model for land-use change applied to 11 different cities. Here, five different parameters that drive model growth, namely dispersion, breed, spread, slope, and road gravity, were reported for each study, building what Clarke referred to as the first inventory of the "DNA" of urban land-use change. The development of these kinds of parameter libraries for various land-use change settings, e.g., urban, rural, forest, agriculture, etc., at different spatial and temporal scales would be particularly useful mechanisms for cross-site analysis, model testing, and the

identification of proximate and underlying drivers of change (Geist and Lambin 2002).

This describes the vision we are working toward, which is applicable to many areas of scientific research, not just to land-use modeling. We think that this is appropriate to all kinds of scientific research, from basic research in which papers and analytic approaches are licensed as OC and shared, to more applied research such as in this land-use modeling context. Perhaps a reviewer of a previous version of this manuscript said it best: "... OS and OC techniques when brought together within a project infrastructure [discussed below] may potentially stimulate the advance of science and scientific knowledge through a more open experimentation, innovation, evaluation, versioning, and e-journal publishing regime." We now turn to a discussion of some critical issues that will need to be addressed to make such a vision viable.

Initiating open source/open content collaboration (Issue 1)

A central first issue is how to initiate an OS/OC collaborative endeavor in scientific research, such as the one we are trying to set up in the context of modeling land-use change. The question here is: just how do FOSS collaborations start? And what does this mean for OS/OC-based collaborative science?

As seen in Fig. 1, FOSS projects tend to, but do not always, follow a three-stage trajectory (Schweik and Semenov 2003, see Schweik 2005 for a more detailed discussion of this figure). They are often initiated by a small group of people who possess programming skills and a common software need (Nakakoji et al. 2002). These small groups work to develop a core piece of software that may not even at this point be licensed FOSS (Fig. 1, Stage 1). At some point, the team members decide that the software is ready to "go open," assign the software an FL or OS license, and make it available over the Internet (Fig. 1, Stage 2). Many software developers hope that soon after that their projects will achieve "high growth" (Fig. 1, Stage 3) by gaining a larger development community and a sizable user community; this is one measure of project success (Crowston et al. 2003, 2004, Stewart 2004). However, projects that continue to be vibrant in terms of development even though they never attract large groups of participants should also be considered successful (Fig. 1, Stage 3, Small Group,

Stable Use). In fact, many studies now report that the majority of OS projects never reach high growth and involve only small numbers of individuals (Ghosh and Prakash 2000, Dempsey et al. 2002, Ghosh et al. 2002, Krishnamurthy 2002, Healy and Schussman 2003, Capiluppi et al. 2003).

We should note that, in recent years, large groups or organizations that have already developed closed source software have decided to relicense it and make it open; IBM is perhaps the best example of this. Consequently, not all FOSS projects follow the trajectory in Fig. 1. Nevertheless, regardless of how they get started, in each stage shown in Fig. 1, the project could experience a loss of momentum or interest and an eventual loss of participation terminating in project death. In sum, the success of a FOSS project is critically dependent on some community of practitioners, including developers and users, whatever the size, to ensure that the project continues and evolves over time (Nakakoji et al. 2002).

What does this mean for the idea of OS/OC scientific collaboration projects? It is likely that these kinds of projects will follow the same trajectories shown in Fig. 1. Many will begin with a small virtual team of interested collaborators and perhaps remain that way throughout the project's lifetime. Alternatively, they could generate a larger following or participation base over time. One emerging example that is closer to an OS software example but exhibits elements of OC and scientific collaboration is the [R-Project for Statistical Computing](#), which appears to be growing in popularity.

We expect that an OS/OC scientific collaboration project in an area such as modeling land-use change will have a greater chance of achieving a high-growth type of success if it is able to attract an online community of practitioners, including scientists, model developers, policy makers, concerned citizens, etc. Consequently, we would hypothesize that projects that are sponsored or sanctioned by a well-established scientific association will have an edge over projects that are not. Projects sponsored by a well-known and respected association will be seen by other potential participants as more legitimate and worthy of their time.

In our specific case of trying to initiate such collaboration around the land-use change model UrbanSim, we recently proposed these ideas at the

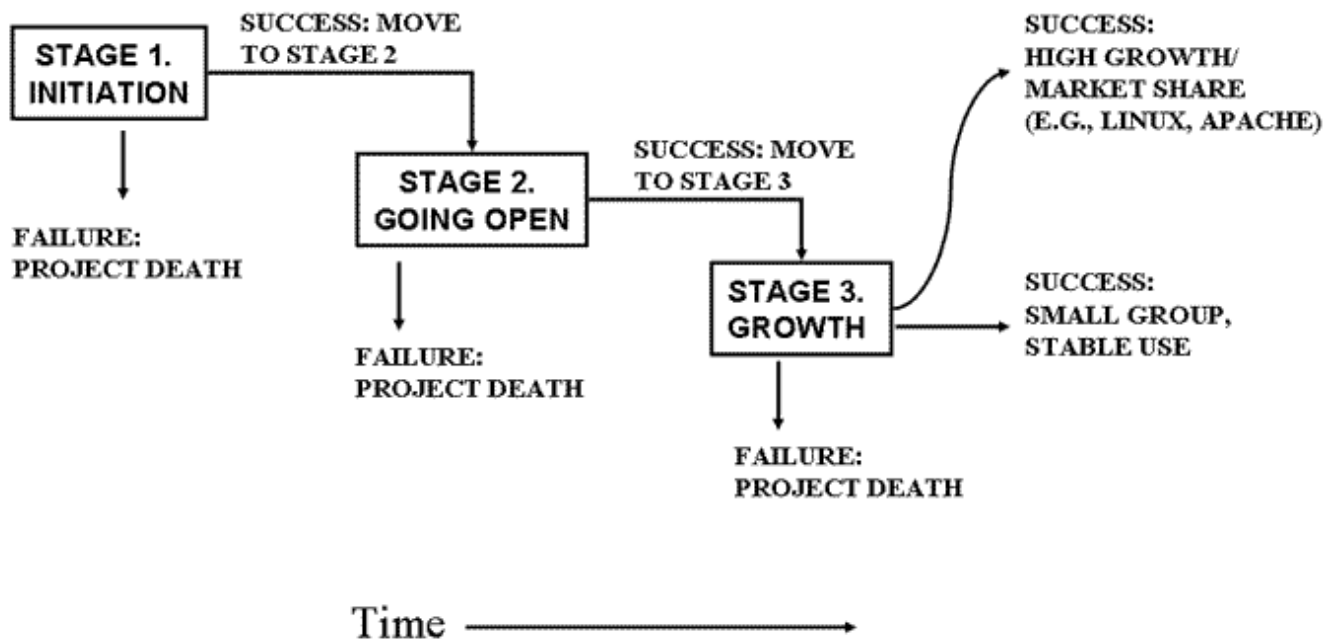
first workshop ever held for users of this model. In addition to the UrbanSim model development team, in attendance at this workshop were policy makers and modeling technicians from a variety of metropolitan planning organizations, all of whom were trying to apply the model to help them make regional planning decisions related to transportation infrastructure, smart growth, etc. Other participants included interested parties from several U.S. universities. When asked, most attendees were receptive to the idea of OS/OC collaboration. Clearly, attention to incentives for participation are key, but we left the meeting optimistic that people would be willing to share their expertise through the act of contributing some documentation, as long as someone else would write up their experiences in an area of the modeling effort that they have yet to face.

Incentives for participation and e-journal publishing (Issue 2)

Perhaps the biggest puzzle in recent years concerning FOSS collaboration is the question of why volunteer programmers donate their time and intellectual property (Glass 1999, Ghosh 2003, Lee et al. 2003). Recent studies note some motivations, which include some personal software need, a software crisis, intellectual stimulation, the desire to belong to some community, a wish to contribute to the free software movement, or simple altruism (Feller and Fitzgerald 2002, Lakhani et al. 2002, Hertel et al. 2003). Human capital and signaling theory provide an economic explanation: participation in OS projects allows actors to learn from others' source code and receive feedback from others on their own contributions (Hann et al. 2002, Voightmann and Coleman 2003). In this regard, contributions are seen as future-oriented investments in building one's career (Johnson 2002, Lerner and Tirole 2002, Lee et al. 2003). Moreover, high-quality contributions act as a signaling device for a participant's programming abilities (Lee et al. 2003) and help to establish a reputation (Sharma et al. 2002), which could lead to future work opportunities (Lee et al. 2003).

However, in some cases of FOSS software, teams are not composed entirely of volunteers. In one study, nearly one-third of the FOSS developers surveyed were directly paid by employers to participate (Ghosh et al. 2002). Wichmann's (2002) study of 25 companies active in Linux development found that most were largely motivated by self-

Fig. 1. The trajectory of free/libre or open source software projects (adapted from Schweik and Semenov 2003). Open content projects may follow a similar trajectory.



interest, e.g., product standardization, cost savings, strategies to weaken competition, and efforts to make their own products compatible with FOSS products.

What does this mean for applying the OS/OC collaborative paradigm to other scientific domains such as modeling land-use change? The motivations of FOSS programmers and scientific researchers are not all that different (Schweik and Semenov 2003). Many in academia, such as graduate students and junior or even senior faculty, would be willing to hone new skills through the distance-learning components of reading "source," e.g., models, papers, etc., and peer review with feedback. The motivation to "signal" one's abilities is also important, particularly for junior scientists looking for jobs or to gain prestige. As one scholar put it during our workshop (Schweik et al. 2003): "Had I known about it, I would have gladly licensed my model OS in graduate school. That way others might have used it and it would have gotten my name more

widely known."

Nevertheless, to signal one's abilities by posting intellectual property, one's name needs to be associated, over time, with that submission. We predict that a key motivation in OS/OC projects will be the assurance that the intellectual contributions to the broader project by individual participants will be tracked and archived over time. In the context of FOSS code, this type of tracking is mandated as part of the licensing agreement. In the General Public License (GPL), for example, the authors of a completely new program place a copyright notification comment as original authors of a work at the start of the source code with a pointer to the location in which the full notice can be found. The GPL requires this copyright statement to stay with all future derivative works based on that original code. It also requires that authors of new derivative work update the software comments with a prominent notice that changes were made, who made them, and the date of the change. In the OS/

OC context, scientists will be more likely to contribute new research content if they are able to maintain copyright over their original submissions in the same way a programmer does using the GPL. Mechanisms will need to be developed to attach similar copyright information to any research product, e.g., a paper, a data set, an analytic module of some sort. This is easily accomplished in content in the form of documents and could be done by specifying the copyright information or update histories in metadata documentation for other components like data sets. In short, the infrastructure built to support an OS/OC collaboration in science will need to include a good historical record of how someone contributed over time to a new model module, to a new derivative paper on land-use change theory, to empirical findings, or to other project content.

Another aspect related to the incentives of scientists and academics to participate in OS/OC projects is a consideration of the demands and expectations of their employers. For many scientists, having their work published in high-quality, refereed journals is a key measure of success in their field and an important measure used for job promotion. This is particularly true in situations in which a junior academic is working toward tenure and could provide a significant disincentive to contribute to an OS endeavor. Consequently, we believe that, for any viable OS/OC scientific collaboration to succeed, it must be placed in the context of current evaluation systems at universities and scientific research organizations. This means that publishing in this context would require not just a collaborative Web site, but also a "next-generation e-journal" component that follows the traditional peer-review process.

We refer to this as a "next-generation e-journal" because it would publish not only traditional peer-reviewed "final" content, such as papers on theory related to land-use change or the results of empirical studies applying a particular model, but also work in progress in the form of working papers, new or revised versions of complete models or subcomponent modules, and new distance-learning material related to the model. In some instances, even data sets used by models might be "published" after some level of peer review, such as a data set on economic projections for a country that might be used in another application of the model. In other words, all the components of the development or application of a model could be published in this e-

journal, broadly defined. This is vastly different than the maximum 30 or so pages that most journals, even most e-journals, currently accept.

Admittedly, this idea is radical and one that is difficult to impose on the current scientific culture. It means a change in what we consider publishable. However, recently both computer scientists and librarians have recognized that this change is needed; there is subtle evidence toward a shift from publication as product to publication as process (Lougee 2004). For example, Berghel (2001:18) sees a future in which the phrase "digital library" will refer to "articulated processes and procedures" rather than a "dull and lifeless era of publishing ... of 'things.'"

We should note that, traditionally, researchers tend not to share work until it is fairly complete, and then first as a working paper to a more limited audience, such as in a conference presentation, for the purpose of soliciting feedback before submission to a journal for publication. How might this process work in the next-generation e-journal we propose? And how could such a system help to protect the intellectual property of an author given that it is being offered publicly but not yet in a final, more traditional published form? This question involves both incentives to participate, i.e., the current issue being discussed, and aspects related to project infrastructure, the issue that follows this section.

Scientists associated with Boston College's [Urban Ecology Institute](#) and the U.S. Forest Service are discussing one possible approach as they work to develop the Urban Ecology Collaborative (UEC). This collaborative would involve people from a variety of governmental agencies and universities studying urban ecology in a number of U.S. cities. One can think of the UEC idea as a kind of professional research association focused on a common research area or agenda. In this case it is around urban ecology, but a similar idea could be developed around land-use modeling, ecology and society, or any research program, however broadly defined. In ideal cases, the research program would be associated with a formal professional organization that already publishes an associated scientific journal, and this idea would extend that journal.

The idea proposed at UEC meetings is that the Internet collaboration infrastructure could be designed to serve three levels or "tiers" of

publishing. The first tier would be a "gallery" that would contain high-quality, peer-reviewed information about the broad research program, in this case urban ecology. The gallery would summarize the program, make available educational materials, and provide information on breaking news and events in the field. In some instances, this could be analogous to the Web site of an established professional association. The content on these pages would likely fall under full copyright protection.

A second parallel tier within the UEC would be the formal or official e-journal, such as *The Journal of Urban Ecology*, in which only refereed components would be published after going through a traditional review process. This would be accessible from the main gallery page. However, if UEC decided to take an OS/OC approach as proposed here, what is published could be more broadly defined and include not only final drafts of theoretical or empirical papers, for example, but, in the case of land-use modeling, new and peer-reviewed versions of model modules, new data sets, distance-learning materials, or other analytical products that are not, in today's environment, typically considered peer-reviewed publications. It would be up to the UEC governance body to determine the appropriate OS/OC licensing options to use for each type of published material.

Also connected to the main gallery and e-journal pages would be the third tier of the collaborative system, probably an intranet system, that UEC associates refer to as the "virtual lab." This intranet would be accessible only to researchers or organizations more formally connected to the professional association or research collaborative who are actively working on program-related projects. Mechanisms could be included within the virtual laboratory for posting draft research products and would allow the solicitation of feedback from peers within a smaller, but still public, community. Like the e-journal tier, the governance body would need to consider appropriate OC licensing options, but it may be at this level that the licensing promotes the most freedom, e.g., new derivative works are permitted. Models of such a facility, although not intranets, are the [ArXiv.org](http://arxiv.org) preprint server in physics or the opensource.mit.edu preprint server for papers studying FOSS issues, or the working paper series that are commonplace in many research organizations. However, like the e-journal tier, this system would allow for preprints of other "working content," e.g., data, model modules, as well, and

would have other Internet-based communication mechanisms such as threaded discussion lists. The whole point is this: in some respects, this virtual lab tier would be an online system that would mimic some of the early information sharing and dialog that occurs at regular in-person conferences. Technically, it would be fairly trivial to provide an automatic email to subscribers based on interest keywords about new working papers, and at the e-journal tier as well, that have been recently posted, much as formal publishers have already done. Such a notification system and a preprint or working content-like archive would provide an incentive for the early prepublishing of work in progress, thus speeding up the research process. At the same time, it would protect the author's personal interests in being known as the generator of the ideas posted to this working content server.

Project infrastructure (Issue 3)

A third important issue is the development of project infrastructure to support OS/OC group collaboration, which includes the tiers idea. However, there are other critical Internet-based components that will comprise a next-generation e-journal to support OS/OC collaboration.

In FOSS projects, group collaboration is supported through Web-based communication and version control systems. For example, the OS project management Web site www.sourceforge.net provides group communication functions and software version control systems based on the "Concurrent Versioning System" or CVS (Fogel 1999). CVS and other version control systems:

- provide parallel read/write access with password protection; archive versions of software; allow for the retrieval of modules; allow for new submissions and protect against the problem of overwriting and errantly eliminating others' work; document changes and change comments over time, i.e., author tracking; provide analysis functions to identify differences between module versions; and email subscribers when project components are moved, updated, or deleted. A similar set of functions, along with at least two other functions, would be required to support the land-use change modeling effort, and these functions would need to support a broader set of submission types.

For example, consider a hypothetical case using the UrbanSim model as an example (Appendix 1).

Researcher A might be interested in applying UrbanSim to a location at the fringe of the urban/agriculture interface. To begin, she may first decide to read an existing OC-licensed paper on the theoretical drivers used in UrbanSim for a western U.S. application. She realizes that some component of the logic does not entirely apply to an east coast situation. Consequently, she might decide to download the text of this paper and, assuming that the license attached to the paper permits new derivative works to be made, revise it to fit her case's context. The OC license could be viral, meaning that the new derivative work is required to follow the same license of its parent document. It could require her to submit this new derivative paper back to the project for peer review. Submission of the paper would invoke a function similar to the posting of enhancements of software (number 4 above), but for papers or documentation. The paper would then undergo a peer review and eventually be rejected or accepted and added to the "production" project library or e-journal.

Similarly, this or another researcher with sufficient skills may then decide to take the logic used in the theoretical paper and implement a new version of UrbanSim's Household Location module (Appendix 1, Fig. A1.1). Assuming that the module falls under an OS license such as the GPL, once her modifications are complete, she would be required to submit the new derivative module back to the project for peer review. After receiving a positive review from the project's "editorial board," this new module could then be added to the project to replace the older module or as a parallel "sibling" module. Future users of the UrbanSim model might then select between the "west coast household location module" or the "east coast household location module."

Critical to making the collaborative ideas proposed here work will be special attention to some kind of mechanisms for tracking all types of author contributions, including theoretical and empirical papers, modules, and other project documentation such as distance-learning materials. We mentioned earlier that, in FOSS projects, the licensing requires some historical log of author contributions in the comment area of the main header of the computer program or in some kind of credits file (Tuomi 2004). However, in the vision of the next-generation e-journal we are describing, even more sophisticated methods of tracking author contributions will be needed because of the vital

role such tracking plays in encouraging scientist participation. Potentially new ways for authors to "cite" their contributions will need to be devised, and, given the speed with which innovations could be published, citations will be more like the way Web pages are cited currently, i.e., journal, year, issue, number, date, and link.

A reader might raise the important concern that some, particularly junior, scientists may be hesitant to publish ideas that are in more of a "work in progress" stage on a more public site for fear of being "scooped" by someone else who would then publish a more formal, peer-reviewed paper on the same ideas. However, new systems for tracking "idea submissions" into the preprint or virtual lab section of the e-journal will be required to help document early, i.e., in the publication process, contributions made by project participants to give them some level of protection of the ownership of these ideas. In this context, contributions in the form of smaller "idea modules" could be attributed to programmers or authors of software documentation that could be referenced by them in, for example, a curriculum vitae.

Another concern with an OS/OC approach to scientific research might be a blurring or clouding of authorship, which could be an especially important issue for junior scientists. Of course, in many instances, peer-reviewed publication of a paper, or a new version of a module, will involve a process similar to the current one. Some negotiation between contributors will perhaps be necessary to decide how their names will be listed on the submission. In instances in which the responsibility for a contribution is even less clear, there are other ways for contributors to receive credit that enhances their personal job situations. For example, in many academic contexts, external letters of support are also vital for promotion, and this type of activity in "clouded" collaboration could bring an active and important contributor some valuable exposure in the scientific community. In addition, what we are suggesting is that, by documenting intellectual contributions through some kind of "history log" of all work products, scientists could document in their curricula vitae not only papers that they have published in the formal peer-reviewed e-journal but also contributions in other collaborative parts of the project as part of their service to the discipline. This suggests a potential change to the culture of what we document as an intellectual contribution and how we document it, and we expect these changes

will be gradual and not easily accomplished before they become as acceptable as more traditional forms of publishing. However, in the meantime, junior academic scientists can treat these contributions as a form of academic service and cite them as such.

Existing systems of scientific research tend to rely on bibliographic measures such as the number of times a paper is cited, and where it is cited, as surrogates for the importance of research contributions. The infrastructure we are proposing also provides opportunities to develop new measures of the significance of someone's contribution. For example, consider two separate submodule submissions (Modules A and B) by two different authors for UrbanSim. Both modules provide different functions for the land-use change model in which they are embedded. Over time, Module A gains wide interest from other modelers or stakeholders who use UrbanSim, and several new theoretical papers on the logic or derivatives of this submodule are submitted. Module B, on the other hand, spurs little new interest and no new derivatives. It could be argued that the original author of Module A provided a more important contribution because of the activity, interest, and derivative works that followed. Quantitative measures could be generated by the system that would measure how many derivative works were spawned by a submission, and these numbers could be added to the author's curriculum vitae as another measure of intellectual contribution.

Still another method of measuring the importance of the contribution that could be applied to all types of research content would be the addition of a voting or recommendation function attached to the content. This follows the lines of the ranking system used in Amazon.com, in which researchers who have read a paper or reviewed a model's logic could then submit their comments and give it an importance ranking of, e.g., one to five stars following the Amazon.com model. Authors of the contribution could then contact those who have offered comments or criticism to hold further private discussion or initiate future collaboration, or they might prefer to engage in public debate and discourse through some communication function the e-journal infrastructure could provide.

In addition to the functions listed in the earlier discussion of version control systems, e.g., CVS, there is a need for at least two other major components related to project infrastructure that are

not prominent in FOSS development projects: (1) metadata and data management functions and (2) distance-learning components. Regarding the former, in the proposed application of modeling land-use change, data will be a primary input to the modeling effort, and data availability or access is often a major hurdle to developing a model for a particular geographic area. However, in some countries, standards for documenting data, particularly geographic data, now exist (see, for example, the [U.S. National Spatial Data Infrastructure guidelines](#)), and there are efforts to develop data clearinghouse servers (see www.fgdc.gov/clearinghouse/clearinghouse.html) to share such data across organizations. Other groups are working on standards for developing "Web services" that would allow for real-time access and sharing of geographic data over the Internet (see opengeospatial.org). In the context of projects like land-use modeling, some kinds of data, such as raster land-cover data, may be too large to provide over the Internet efficiently, but at least their metadata could be made available, including information on where these data reside, how they can be acquired, and their OC licenses.

Another important component of the next-generation e-journal infrastructure requiring special attention and prominent display is distance-learning. This feature can help recruit and build new membership. In the context of modeling land-use change, learning material would include general documentation on the modeling approach, more detailed literature on the steps involved in running the model, details on model subcomponents such as modules, frequently asked questions, etc. All material could be licensed OC permitting future improvements, subject to peer review, by anyone wishing to contribute. The e-journal might also include, as another incentive for participation, a prominently displayed list of "key enhancements needed" that might serve as a place in which, for example, graduate students could find potential thesis topics.

To conclude this section on infrastructure, we should note that, along with software version control systems like CVS, there are other more recent and FOSS-licensed tools that provide some of the collaborative functions required for this next-generation e-journal we are proposing. For example, the "bugzilla" software provides functions for the reporting and the tracking of software bugs (bugzilla.org). Other software exists that provides

document management capabilities. The [DSpace software](#) (Smith et al. 2003) is a Web-based repository system to support and archive digital research and educational material produced by faculty and staff at research universities or other research organizations. There are also Web-based portal and content management systems such as [Plone](#), [Zope](#), and [Fedora](#), as well as other support programs for Internet collaboration such as [wikis](#). Moreover, other online systems for e-journal publication and editorial management, like the systems used by this journal, [Ecology and Society](#), are helping to define the infrastructure needed to support the next-generation e-journal we describe. Although none of these have all the functions required, especially in terms of relating metadata such as OC licensing to various research products such as data, model modules, papers, etc., there are enough tools available that could possibly be configured with modest support to establish the core infrastructure to support this next-generation e-journal. There is the need, and an opportunity, for some organization or group of inspired individuals with programming skills to take on the development of this next-generation e-journal infrastructure and then perhaps be willing to provide their solution to the world via a FOSS license.

Project institutional design and governance (Issue 4)

Because FOSS projects involve common property regimes (Schweik 2005), project governance is an important issue but one that has not been deeply researched (Schweik and Semenov 2003, Weber 2004). Ghosh et al. (2002) report that most of the OS software projects they studied were directed by a single "lead developer" who maintained a centralized decision-making structure. For instance, studies on the OS Linux operating system describe the lead developer as a "benevolent dictator" working with a team of "trusted lieutenants" who are experts in a particular domain (Moody 2001, Sharma et al. 2002, Shaikh and Cornford 2003). In some OS projects, would-be developers earn the status of trusted lieutenants by first working at the periphery of the project, e.g., offering bug reports, and then progressively contributing more in terms of time, effort, and source code. If their contributions are deemed to be high-quality or significant, these developers may move toward the core of the development team, thus acquiring more say in project governance decisions. Other FOSS

projects have a different management structure. For example, Jorgensen's (2001) study of the FreeBSD OS project identifies a nine-member team heading the project. Members are from and elected by the pool of developers with authority to post changes. However, as Bezroukov (1999b:17) notes: "... in each [OS] project in particular, there are political systems with corresponding and sometimes fuzzy hierarchical structures."

Some early studies suggest that decisions related to project direction are reached by consensus (Fielding 1999, Markus et al. 2000, Mockus et al. 2000). Established systems of rules, shared norms of behavior, voting systems, and monitoring and sanctioning systems appear to be important in some FOSS projects (Sharma et al. 2002). Most projects have some sort of established conventions or norms of behavior that all must follow, and some of these enhance the effectiveness of Internet-based coordination (Bonaccorsi and Rossi 2003).

Consequently, an important issue in extending the FOSS collaboration paradigm to collaboration in scientific research will be designing the decision-making structure and the system of operational rules and dispute-resolution procedures for the virtual team. It is conceivable that the governance structure may ultimately become a combination of how journals today are run and organized, e.g., an editor and editorial board, and how FOSS software projects are organized.

It is interesting to note that the trajectory toward membership on some governing board in a scientific research context will mirror in many ways what was described earlier in the context of FOSS projects. Researchers will start at the periphery, as undergraduates, then continue working on the project as graduate students and student researchers, then post-docs, junior faculty, and ultimately senior faculty or researchers surrounded by the preceding others. This is similar to the trajectory described above in the context of FOSS developers. However, what may be different is the amount of time that elapses before an appointment to a governance position occurs. In a FOSS setting, it may take only a year or two for a gifted programmer to demonstrate his or her skills and knowledge to the broader group, whereas in a scientific context it might take an equally outstanding person a decade or more to achieve a similar status within the research community.

A related institutional design issue that will need to be addressed early by the initial governance team is the kind of OS/OC licensing to be used. Each type of project output, e.g., model modules, model usage documentation, empirical papers, theoretical papers, and data sets, will need to be assigned some OS or OC license. For example, we are advocating that whatever form or technology the land-use model uses, e.g., a computer program, a statistical script, etc., it should be placed under some form of OS/OC license that allows the free copying of the model, requires the model source to be readable, and permits the development of new derivative works based on the model structure. In some instances, the model developers might decide to apply these conditions to all related products, including the model modules, their documentation, data, and even theoretical papers. However, there will be situations in which more restrictive licensing is warranted. For example, we expect that most empirical papers on a particular model application should probably be licensed with a "no derivative work" component, because these types of papers report findings from a particular study at a particular time.

OC licensing of theoretical papers presents a particularly interesting and potentially difficult problem. Recall the earlier example of a researcher writing a new derivative theoretical paper related to household location decisions in the eastern United States based on an earlier paper by someone else on the theoretical drivers in a western U.S. context. The result would be two separate theoretical papers, the east coast version and the original west coast version. This differs from the traditional approach to publishing research, because the second version of the paper may have substantial sections of text taken verbatim from the first paper, with new text added. This is similar, of course, to what might occur in a software documentation update situation under the GNU (GNU's Not Unix) Free Documentation License described earlier.

Readers will quickly note that this treads dangerously close to the issue of plagiarism. If the licensing for the research paper permits new derivative works, the situation exists in which someone could download a paper, revise it only slightly and in a trivial manner, and then put his or her name on it as an additional author. With regard to this situation, we paraphrase the words of a reviewer of an earlier draft of this paper: "The potential complexity of sorting out the degree of a

unique contribution could be overwhelming as this moves from this easy case to more substantial changes that require subjective interpretation to assess ... This will require a robust solution if e-publishing is to co-exist with OS/OC licensing."

Although we cannot profess to have all the answers to this question, we do have some ideas on how to address it. A conservative approach might be to use licenses that do not allow derived work for academic papers in general, but that still promote other licensing options such as free copying and distribution. Through simple Internet-based open access, these alternative licensing options will yield more easy access worldwide and will likely lead to a more rapid evolution of the field. Other project output, such as distance-learning documentation, might be licensed with permission for derivative works to promote more rapid improvements to learning materials.

However, in situations in which projects want to take greater advantage of the innovation that could occur in a licensing environment that grants permission for derivative works, there are several additional options for managing and communicating how much of a contribution has been made by the various authors. First, as we stated earlier in the *Infrastructure* section, version control and author tracking systems will need to be devised in next-generation e-journal platforms to ensure that the intellectual contributions of the original and subsequent authors are tracked over time. As we discussed earlier, to a substantial degree these issues already exist and have been dealt with in the software and documentation setting of FOSS projects, e.g., history logs. However, further empirical work is needed to investigate whether and how this issue has been addressed in the software and documentation contexts. Second, version control systems like CVS not only store "change logs" of how OS software has been modified from version to version along with who made these changes, but they also archive the past versions of the software modules themselves. The same could be done in an OC setting in which the project infrastructure keeps previous versions of other project content, such as theoretical papers. If this were done, it would be a fairly trivial exercise for someone to review the changes made between two versions to assess the real intellectual contribution made by a subsequent author. In fact, commands like the CVS "diff" (difference) command allows one to identify where the differences lie when

comparing two different versions of the same content module.

Project finance (Issue 5)

The last issue we will raise relates to the financing of OS/OC-based research collaborations. Skeptics question the viability of the volunteer FOSS model over the long term, given that there is no financial support. For example, this question was raised at a meeting at UNESCO in Paris on issues related to science and the public domain (Esanu and Uhlir 2004), at which a participant argued that the only way FOSS projects would succeed was if a "government" supported the endeavor. Although this is an empirical question yet to be answered, it is highly likely that, at least for major projects, some level of financing will be required. In fact, in the FOSS domain, projects can be found that fall under a variety of different financial support schemes, including (1) the government funding or subsidy model (Hahn 2002); (2) philanthropic funding, such as efforts undertaken by the Andrew W. Mellon foundation; (3) corporate consortia (Hildebrand 2004); (4) corporate investment (Webb 2004); (5) venture capital/investment banking; and (6) a hybrid/mix of these. The argument could be made that the long-term success of some of these enterprise OS software projects, perhaps even Linux, is a consequence of the investments firms like IBM, Sun Microsystems, and others are making in such projects (Ghosh et al. 2002, Wichmann 2002). However, it remains an open question as to when such funding is absolutely needed for a project to survive vs. when a FOSS project can exist solely through a volunteer base.

The same question can be raised when one considers extending the FOSS development paradigm into the area of scientific collaboration. However, in the context of academic contributions, it seems that some "industry support" may already exist. Universities expect their faculty to undertake research that contributes to a larger research program. Service to a broader professional community is seen as a positive thing in most faculty annual evaluations. If a faculty member wanted to participate in a virtual OS/OC collaboration as part of his or her research program, most universities would be supportive of this so long as the individual continued to meet traditional measures of scholarship and productivity, such as publications in refereed outlets. This incentive raises again the

importance of closely linking or merging the OS/OC research collaboration to a refereed e-journal, because it is likely that this would greatly enhance the number of possible researchers willing to contribute.

This leads to the question of how such an e-journal might be financed. Traditionally, scientific journals tend to be published either by academic or professional societies or by commercial publishers. A study recently commissioned by the Wellcome Trust (SQW Ltd. 2004) estimates the cost to be about U.S. \$2750 per article in good- to high-quality subscriber-pays journals and about U.S. \$1425 per article in medium-quality journals. This covers activities such as managing the referee process, editorial support, subscription management, sales and marketing, and final printing costs. Author-pays journals tend to be lower, in the range of U.S. \$1950 for good- to high-quality journals and U.S. \$1025 for medium-quality journals (SQW Ltd. 2004:3). This same study assumes that the cost of electronic publishing is about the same, given that the cost of maintaining the electronic system replaces conventional distribution costs. Our idea of moving to a Web-based e-journal that includes some level of referee process not only for theoretical and empirical papers but also for models and model documentation, data, and distance-learning documents will probably increase the amount of time required by reviewers and various component editors and raise the costs of these activities.

It is an open question whether traditional journal publishers following a traditional user-pays subscription business and taking an all-rights-reserved philosophy would embrace such an idea. Commercial journals tend to provide e-journal access as a companion to the already established printed version or as part of a larger package of journal subscriptions provided in bulk to large institutional subscribers such as research libraries. However, other financing models are being explored to promote more open access to scientific information, such as the author-pays-to-publish model with open access to the journal as well as the model in which academic institutions or research libraries are taking on the publishing of disciplinary e-journals as part of their mission (Shortliffe 2004, Suber 2004). In this regard, an interesting question requiring further research is whether any components of the business models in the area of OS software (e.g., Wichmann 2002) are at all applicable and transferable to the process of

publishing scientific journals. IBM, for example, has managed to turn OSS into a profitable venture. Could some of what drives their profitability be transferred to the publishing domain? For example, some of IBM's profits undoubtedly come from service and support of OS software. Could a possible model based on this idea be for a university to set aside some funds received through a distance-learning program to help support a next-generation e-journal published by its library and edited by some of its faculty?

In sum, we cannot admit to having a solution to the financing issue, other than to say that experiments in alternative e-journal funding to promote open access are emerging, and a sizable number of FOSS packages are available with some of the needed functionality to support such endeavors. Professional societies, foundations, or government agencies such as the U.S. National Science Foundation could potentially invest in supporting the development of a next-generation e-journal infrastructure and provide this infrastructure under a FOSS license for other groups to use. Other financing will surely be needed to support such e-journals, but possibly using the user-pays model or, in the case of specialized research areas, perhaps through support as part of a research library mission at a particular institution. Whatever approach is taken, clearly financing is an issue that will need to be carefully considered for the ideas here to work.

CONCLUDING REMARKS

A primary innovation in free or open source software (FOSS) is the development of new forms of copyright through alternative licensing structures. Importantly, these licenses provide the right to distribute the software, rather than focusing solely on the right to exclude people from using the software (Weber 2004). Another important innovation is how collaborative development groups operate and are governed, which to a large degree is still an area open for study (Schweik and Semenov 2003). Weber's (2004) recent book, *The Success of Open Source*, is a significant step in the right direction, arguing, as it does, that FOSS is a series of experiments in social organization.

Recently, scholars have begun to recognize that the FOSS collaborative paradigm is not limited to software (e.g., Bollier 1999, Schweik and Grove 2000, Stallman 2001, Schweik and Semenov 2003,

Creative Commons 2004, Weber 2004). It has the potential, under the OC licensing approach, to be applied to any intellectual domain that requires a team of thinkers to tackle a problem. For example, OC licensing has already been used to encourage collaboration or "new derivatives" in areas such as music and art (Creative Commons 2004). There are several "wiki" sites that are using the GNU (GNU's Not Unix) Free Development License for the construction of encyclopedias and other online books, e.g., [Wikipedia](#) and [Wikibooks](#). There is an emerging movement to apply OC licensing to educational and scientific initiatives as well (see [Science Commons](#)). There are also emerging open source software projects that exhibit some of the scientific collaboration characteristics we have described. For example, the [R-Project for Statistical Computing](#) includes review processes for publishing software, online publishing of user-generated documents, author attribution, etc. This paper is an attempt to think through some of the issues that such OC licensing-based collaborations might encounter in one scientific area: modeling land-use change (Schweik et al. 2003).

However, it should be stressed that the FOSS collaborative paradigm is not a panacea, and in recent years there has been significant hype over it as a phenomenon. Most of the existing studies of FOSS projects are on the success stories such as Linux, the Apache Web Server, and a few others. Clearly, many FOSS projects fail for a variety of reasons such as the lack of a critical mass of developers, the lack of infrastructure, poor governance, or infighting among developers. Many of the more than 100,000 projects on Sourceforge.net and other sites are probably stalled or dead in terms of active production. On the other hand, the FOSS collaborative paradigm in software has produced a large number of real collaborative experiments, and there are enough real success stories and interest, including serious industry and government interest, to testify to what we believe to be its very great potential.

Like FOSS collaborations, OS/OC-based collaborations will require experimentation, and many will fail. However, the possibility is there that this innovation in collective action, applied to important social-ecological questions and to other fields of scholarly inquiry, has the potential to produce important breakthroughs in areas in which the benefits of enhanced collaborative effort can make a

difference. In the context of modeling land-use change, we believe that the move toward an OS/OC paradigm needs to be an evolutionary movement, working within the existing research-publishing paradigm, rather than a revolutionary movement. We are not advocating the destruction of the existing process of peer review and publication. Rather, we see modifications to this paradigm whereby new ideas can be incorporated into the traditional form of scientific collaboration. Combinations of OS and OC licensing, the new collaborative infrastructure we described, and the advancement of e-journal capabilities provide the possibility for more rapid progress than is possible within the existing structure of scientific research and publication.

Responses to this article can be read online at:
<http://www.ecologyandsociety.org/vol10/iss1/art33/responses/>

Acknowledgments:

We are extremely appreciative of the anonymous reviewers who provided very insightful and helpful comments on a previous version of this manuscript. Support for this study was provided by the U.S. Forest Service's Burlington Laboratory (4454) and Southern and Northern Global Change Programs; the Cooperative State Research Extension, Education Service, U.S. Department of Agriculture, Massachusetts Agriculture Experiment Station, under Project Number MAS00847; the Center for Policy and Administration and the Department of Natural Resources Conservation at the University of Massachusetts, Amherst; the Center for the Study of Institutions, Population and Environmental Change at Indiana University (NSFSBR 9521918); and an NSF CAREER award (NSFIIS 0447623). Finally, this manuscript benefited from discussions with participants at a workshop on open-source development of land-use/land-cover change models (see open source land-use/landcover report) sponsored by the NSF Biocomplexity program (grant #NSFSBR 0083744). The findings, recommendations, and opinions expressed are those of the authors and do not necessarily reflect the views of the funding agencies.

LITERATURE CITED

- Aber, J. D.** 1997. Why don't we believe the models? *Bulletin of the Ecological Society of America* 78 (3):232-233.
- Agarwal, C., G. M. Green, J. M. Grove, T. Evans, and C. M. Schweik.** 2002. *A review and assessment of land-use change models: dynamics of space, time, and human choice*. CIPEC Collaborative Report Number 1. USFS Publication NE-GTR-297. Joint publication of the Center for the Study of Institutions, Population and Environmental Change at Indiana University-Bloomington and the U.S. Forest Service. Burlington, Vermont, USA. [Online] URL: http://www.fs.fed.us/ne/newtown_square/publications/technical_reports/pdfs/2002/gtrne297.pdf.
- Askund, U. and Bendix, L.** 2002. A study of configuration management in open source software projects." *IEE Proceedings of Software Engineering* 149(1):40-46.
- Benkler, Y.** 2002. Coase's penguin, or, Linux and the nature of the firm. *Yale Law Journal* 112 (3):371-446.
- Berghel, H.** 2001. Digital village: a cyberpublishing manifesto. *Communications of the ACM* 44 (3):17-20.
- Bezroukov, N.** 1999a. Open source software as a special type of academic research (a critique of vulgar Raymondism). *First Monday* 4(10). [Online] URL: http://firstmonday.org/issues/issue4_10/bezroukov/.
- Bezroukov, N.** 1999b. A second look at the cathedral and the bazaar. *First Monday* 4(12). [Online] URL: http://firstmonday.org/issues/issue4_12/bezroukov/.
- Bollier, D.** 1999. *The power of openness: why citizens, education, government and business should care about the coming revolution in open source code software*. [Online] URL: <http://h2oproject.law.harvard.edu/opencode/h2o/>.
- Bonaccorsi, A., and C. Rossi.** 2003. Why open source software can succeed. *Research Policy Special Issue on Open Source Software*

Development. [Online] URL:
<http://opensource.mit.edu/papers/rp-bonaccorsirossi.pdf>.

Boyle, J. 2003a. The second enclosure movement and the construction of the public domain. *Law and Contemporary Problems* 66(1/2):33-75.

Boyle, J., editor. 2003b. The public domain. *Law and Contemporary Problems* 66(1/2):1-483.

Burnham, J. C. 1990. The evolution of editorial peer review. *Journal of the American Medical Association* 263(10):1323-1329.

Capiluppi, A., P. Lago, and M. Morisio. 2003. Evidences in the evolution of open source projects through changelog analyses. In J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL:<http://opensource.ucc.ie/icse2003>.

Creative Commons. 2004. *Choose license*. Creative Commons, San Francisco, California, USA. [Online] URL:
<http://creativecommons.org/license/>.

Crowston, K., H. Annabi, and J. Howison. 2003. Defining open source project success. *Proceedings of the 24th International Conference on Information Systems (ICIS 2003)*. (Seattle 2003). [online] URL:
<http://opensource.mit.edu/papers/crowstonannabihowison.dpf>.

Crowston, K., H. Annabi, J. Howison, and C. Masango. 2004. Towards a portfolio of FLOSS project success measures. [Online] URL:
<http://opensource.mit.edu/papers/crowston04towards.pdf>.

Dempsey, B., D. Weiss, P. Jones, and J. Greenberg. 2002. Who is an open source developer? *CACM* 45:67-72.

Dietz, T., E. Ostrom, and P. C. Stern. 2003. The struggle to govern the commons. *Science* 302:1907-1912.

Divitini, M., L. Jaccheri, E. Monteiro, and H. Traetteberg. 2003. Open source processes: no place for politics? In J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open*

Source Software Engineering. [Online] URL:
<http://opensource.ucc.ie/icse2003>.

Drucker, P. F. 1999. Beyond the information revolution. *Atlantic Monthly*. [Online] URL:
<http://www.theatlantic.com/issues/99oct/9910drucker.htm>.

Dube, L., and G. Pare. 2001. Global virtual teams. *Communication of the ACM* 44(12):71-73.

Dutton, W. H., and K. L. Kraemer. 1985. *Modeling as negotiating: the political dynamics of computer models in the policy process*. Ablex, Norwood, New Jersey, USA.

EPA. 2000. *Projecting land-use change: a summary of models for assessing the effects of community growth and change on land-use patterns*. EPA/600/R-00/098. Environmental Protection Agency, Office of Research and Development, Cincinnati, Ohio, USA.

Esanu, J. M., and P. F. Uhlig. 2004. *Open access and the public domain in digital data and information for science; proceedings of an international symposium*. National Academies Press, Washington, D.C., USA.

Feller, J., and B. Fitzgerald. 2002. *Understanding open source software development*. Addison-Wesley, London, UK.

Fielding, R. Y. 1999. Shared leadership in the Apache project. *Communications of the ACM* 42(4):42-43.

Fogel, K. 1999. *Open source development with CVS*. Coriolis, Scottsdale, Arizona, USA.

Geist, H. J., and E. F. Lambin. 2002. Proximate causes and underlying driving forces of tropical deforestation. *BioScience* 52(2):143-150.

Glass, R. 1999. Of open source, Linux and hype. *IEEE Software* 16(1):45-51.

Ghosh, R. A. 2003. Cooking pot markets: an economic model for the trade of free goods and services on the Internet. *First Monday* 3(3). [Online] URL:
http://www.firstmonday.org/issues/issue3_3/ghosh/.

Ghosh, R. A., and V. V. Prakash. 2000. The orbiten free software survey. *First Monday* 5(7). [Online] URL: http://firstmonday.org/issues/issue5_7/ghosh/.

Ghosh, R. A., G. Robles, and R. Glott. 2002. *Free/libre and open source software: survey and study*. Technical report. International Institute of Infonomics, University of Maastricht, Maastricht, The Netherlands. [Online] URL: <http://www.infonomics.nl/FLOSS/report/index.htm>.

Grove, J. M., C. M. Schweik, T. P. Evans, and G. M. Green. 2002. Modeling human-environmental dynamics. Pages 160-188 in K. C. Clarke, B. E. Parks, and M. P. Crane, editors. *Geographic information systems and environmental modeling*. Prentice Hall, Upper Saddle River, New Jersey, USA.

Hahn, R. W., editor. 2002. *Government policy toward open source software*. Brookings Institution Press, Washington, D.C., USA.

Halloran, T. J., W. L. Scherlis, and J. R. Erenkrantz. 2003. Beyond code: content management and the open source development portal. In J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL: <http://opensource.ucc.ie/icse2003>.

Hann, I., J. Roberts, S. Slaughter, and R. Fielding. 2002. Why do developers contribute to open source projects? First evidence of economic incentives. [Online] URL: <http://opensource.ucc.ie/icse2002/HannRobertsSlaughterFielding.pdf>.

Hardin, G. 1968. The tragedy of the commons. *Science* 162:1243-1248.

Healy, K., and A. Schussman. 2003. *The ecology of open-source software development*. [Online] URL: <http://opensource.mit.edu/papers/healyschussman.pdf>.

Hertel, G., S. Niedner, and S. Herrman. 2003. Motivation of software developers in open source projects: an Internet-based survey of contributors to the Linux kernel. In J. Feller, B. Fitzgerald, S.

Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL: <http://opensource.ucc.ie/icse2003>.

Hildebrand, J. D. 2004. Hidden agendas in Linux land. *Software Development Times*. [Online] URL: http://www.sdtimes.com/cols/opensourcewatch_015.htm.

Johns, A. 2001. The birth of scientific reading. *Nature* 409:287-289.

Johnson, J. P. 2002. Economics of open source software. *Journal of Economics and Management Strategy* 11(4):637-662.

Jorgensen, N. 2001. Putting it all in the trunk: incremental software development in the FreeBSD open source project. *Information Systems Journal* 11(4):321-336.

Kelly, S., and M. Jones. 2001. Groupware and the social infrastructure of communication. *Communications of the ACM* 44(12):77-79.

King, J. L., and K. L. Kraemer. 1993. Models, facts, and the policy process: the political ecology of estimated truth. Pages 353-360 in M. F. Goodchild, B. O. Parks, and L. T. Steyaert, editors. *Environmental modeling with GIS*. Oxford University Press, New York, New York, USA.

Krishnamurthy, S. 2002. Cave or community? An empirical examination of 100 mature open source projects. *First Monday* 7(6). [online] URL: http://www.firstmonday.org/issues/issue7_6/krishnamurthy/.

Kronick, D. 1990. Peer review in 18th century scientific journalism. *Journal of the American Medical Association* 263(10):1321-1322.

Lakhani, K. R., B. Wolf, J. Bates, and C. DiBona. 2002. *The Boston Consulting Group Hacker Survey, Release 0.73*. [Online] URL: <http://www.ostg.com/bcg/bcg-0.73/BCGHackerSurveyv0-73.html>.

Lambin, E. F., H. J. Geist, and E. Lepers. 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environmental Resources* 28:205-241.

- Lee, S., N. Moisa, and M. Weiss.** 2003. Open source as a signaling device: an economic analysis. In J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL: <http://opensource.ucc.ie/icse2003>.
- Lerner, J., and J. Tirole.** 2002. The scope of open source licensing. [Online] URL: http://opensource.mit.edu/online_papers.php.
- Lougee, W. P.** 2004. Scholarly communication and libraries unbound: the opportunity of the commons. Paper presented at the Workshop on Scholarly Communication as a Commons, March 31, Indiana University, Bloomington, Indiana, USA. [online] URL: http://dlc.dlib.indiana.edu/archive/00001250/00/Commons_workshop_Lougee_rev_Mar29b.pdf.
- Markus, M. L., B. Manville, and C. E. Agres.** 2000. What makes a virtual organization work? *Sloan Management Review* 42(1):13-26.
- Mockus, A., R. Fielding, and J. Herbsleb.** 2000. A case study of open source software development: the Apache server. Pages 263-272 in *Proceedings of the 2000 International Conference of Software Engineering*. (Limerick 2000). ACM Press, New York, New York, USA.
- Moody, G.** 2001. *Rebel code: Linux and the open source revolution*. Perseus, Cambridge, Massachusetts, USA.
- Nakakoji, K., Y. Yamamoto, Y. Nishinaka, K. Kishida, and Y. Ye.** 2002. Evolution patterns of open-source software systems and communities. Pages 76-85 in *Proceedings of the International Workshop on Principles of Software Evolution*. ACM Press, New York, New York, USA.
- Noth, M., A. Borning, and P. Waddell.** 2003. An extensible, modular architecture for simulating urban development, transportation, and environmental impacts. *Computers, Environment and Urban Systems* 27(2):181-203.
- O'Reilly, T.** 1999. Lessons from open-source software development. *Communications of the ACM*. 42(4):33-37.
- Ostrom, E.** 1990. *Governing the commons: the evolution of institutions for collective action*. Cambridge University Press, New York, New York, USA.
- Ostrom, V., and E. Ostrom.** 1977. Public goods and public choices. Pages 7-49 in E. S. Savas, editor. *Alternatives for delivering public services: toward improved performance*. Westview, Boulder, Colorado, USA.
- Parker, P., R. Letcher, and A. Jakeman.** 2002. Progress in integrated assessment and modeling. *Environmental Modelling and Software* 17:209-217.
- Pavlicek, R.C.** 2000. *Embracing insanity: open source software development*. SAMS, Indianapolis, Indiana, USA.
- Perens, B.** 1999. The open source definition. Pages 171-188 in C. DiBona, S. Ockman, and M. Stone, editors. *Open sources: voices from the open source revolution*. O'Reilly, Sebastopol, California, USA.
- Pontius Jr., R. G., D. Huffaker, and K. Denman.** 2004. Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling* 179(4):445-461.
- Raymond, E. S.** 1998. The cathedral and the bazaar. *First Monday* 3(3). [Online] URL: http://www.firstmonday.org/issues/issue3_3/raymond/index.html.
- Schweik, C. M.** 2005. An institutional analysis approach to studying libre software "commons." *Upgrade* 6(3), in press. [Online] URL: <http://www.upgrade-cepis.org/pages/pastissues.html#2005>.
- Schweik, C. M., and J. M. Grove.** 2000. Fostering open-source research via a World Wide Web system. *Public Administration and Management* 5 (3). [Online] URL: http://www.pamij.com/5_4/5_4_2_opensource.html.
- Schweik, C. M., T. Evans, and J. M. Grove.** 2003. *Initiating an open source/content landcover change modeling effort; report of a workshop held on August 22-23, 2003*. [Online] URL: http://www.lulc.org/bcworkshop_2003/os_lulc_workshop_report_2003.pdf.
- Schweik, C. M., and A. Semenov.** 2003. The

institutional design of 'open source' programming: implications for addressing complex public policy and management problems. *First Monday* 8(1). [Online] URL:

http://www.firstmonday.org/issues/issue8_1/schweik/

Shaikh, M., and T. Cornford. 2003. Version management tools: CVS to BK in the Linux kernel. In Feller, J., B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL: <http://opensource.ucc.ie/icse2003>.

Sharma, S., V. Sugmaran, and B. Rajgopalan. 2002. A framework for creating hybrid-open source software communities. *Information Systems Journal* 12(1):7-25.

Shortliffe, E. H. 2004. *Electronic scientific, technical and medical journal publishing and its implications: report of a symposium*. The National Academies Press, Washington, D.C., USA.

Smith, M., M. Barton, M. Bass, M. Branschovsky, G. McClellan, D. Stuve, R. Tansley, and J. Harford Walker. 2003. DSpace: an open source dynamic digital repository. *D-Lib Magazine*. 9(1). [Online] URL: <http://www.dlib.org/dlib/january03/smith/01smith.html>.

SQW Ltd. 2004. *Costs and business models in scientific research publishing*. (Online) URL: <http://www.wellcome.ac.uk/assets/wtd003184.pdf>

Stallman, R. M. 1999. The GNU operating system and the free software movement. Pages 53-70 in C. DiBona, S. Ockman, and M. Stone, editors. *Open sources: voices from the open source revolution*. O'Reilly, Sebastopol, California, USA.

Stallman, R. M. 2001. The free universal encyclopedia and learning resource. Pages 257-269 in C. Werry and M. Mobray, editors. *Online communities: commerce, community action, and the virtual university*. Prentice Hall, Upper Saddle River, New Jersey, USA.

Stallman, R. M. 2002. *Free software, free society: selected essays of Richard M. Stallman*. Free Software Foundation, Boston, Massachusetts,

USA.

Stewart, K. 2004. OSS project success: from internal dynamics to external impact. Pages 92-96 in *Proceedings of the 4th Annual Workshop on Open Source Software Engineering*. (Edinburgh 2004). ACM Press, New York, New York, USA.

Suber, P. 2004. *Creating an intellectual commons through open access*. [Online] URL: <http://dlc.dlib.indiana.edu/archive/00001246/>.

Stix, G. 2003. Some rights reserved. *Scientific American* (March). [Online] URL: http://www.slis.indiana.edu/news/story.php?story_id=577

Tuomi, I. 2004. Evolution of the Linux credits file: methodological challenges and reference data for open source research. *First Monday* 9(6). [Online] URL: http://www.firstmonday.dk/issues/issue9_6/tuomi/index.html.

Voightmann, M. P., and C. P. Coleman. 2003. Open source methodologies and mission critical software development. In J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source Software Engineering*. [Online] URL: <http://opensource.ucc.ie/icse2003>.

Waddell, P. 1998. *An urban simulation model for integrated policy analysis and planning: residential location and housing market components of urbanism*. [Online] URL: <http://www.urbansim.org>.

Waddell, P. 2002. UrbanSim: modeling urban development for land use, transportation and environmental planning. *Journal of the American Planning Association* 68(3):297-314.

Waddell, P., and G. F. Ulfarsson. 2005. *Introduction to urban simulation: design and development of operational models*. [Online] URL: [URL:http://www.urbansim.org](http://www.urbansim.org).

Warsta, J., and P. Abrahamsson. 2003. Is open source development essentially an agile method?" Pages 143-147 in J. Feller, B. Fitzgerald, S. Hissam, and K. Lakhani, editors. *Taking stock of the bazaar: proceedings of the 3rd Workshop on Open Source*

Software Engineering. [Online] URL:
<http://opensource.ucc.ie/icse2003/3rd-WS-on-OSS-Engineering.pdf>.

Webb, C. L. 2004. IBM's open source lovefest. Monday, September 13. *The Washington Post* (September 13, 2004).[Online] URL:
http://www.washingtonpost.com/wp-dyn/articles/A17842-2004Sep13.html?nav=rss_technology.

Weber, S. 2004. *The success of open source*. Harvard University Press, Cambridge, Massachusetts, USA.

Wichmann, T. 2002. FLOSS final report. Part 2. Firm's open source activities: motivations and policy implications. Berlecon Research, Berlin, Germany. [Online] URL:
http://www.berlecon.de/studien/downloads/200207-FLOSS_Activities.pdf.

Ziman, J. 1969. Information, communication, knowledge. *Nature* **224**:318-324.

APPENDIX 1. SUMMARY OF THE URBANSIM MODEL

This Appendix provides an overview of the general approach and structure of the UrbanSim model that our open source/open content (OS/OC) is focusing on (Schweik, Evans, and Grove 2003). We include this appendix to give interested readers a little more information on the model to help understand where we are headed with an OS/OC land-use change collaboration. For even more information on UrbanSim, read Waddell (1998, 2000), Waddell and Ulfarsson (2005), and Noth et al. (2003). This summary is based on these articles.

UrbanSim is described as "a software-based simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy" and is designed specifically for use and application in metropolitan regions (<http://www.urbansim.org>). Consequently, UrbanSim models change at fine spatial resolutions, usually with one cell in their spatial grid representing a 150 x 150 m area. Recent empirical applications of UrbanSim include Eugene-Springfield, Oregon; Salt Lake City, Utah; and Honolulu, Hawaii. UrbanSim explicitly models the location choices made by households (housing), businesses (location and jobs), and the location choices of real estate developers. It uses a set of interacting submodels representing these different actor types and corresponding processes in urban environments. Discrete choice modeling (e.g., multinomial logit) is used to predict location choices. Geographic Information Systems are used to integrate data and display model results.

Figure A1.1 provides a graphic of the general structure and processing of UrbanSim (Waddell 2002). Inputs to the UrbanSim model include the initial year data (e.g., current land-use configuration), data from regional economic forecasts provided by an externally developed macroeconomic model, travel access indicators provided by another externally developed travel demand model, and user-specified data that provides input for various public policy scenario investigations. This information gets fed into a "Model Coordinator" computer program that coordinates the processing and dialog between five submodules: Accessibility, Economic and Demographic Transition, Household and Employment Mobility, Household and Employment Location, and Real Estate Development. The Accessibility model predicts the pattern of accessibility by auto ownership level. The Economic and Demographic transition estimates the creation or loss of households and jobs by type. The Household and Employment Mobility submodule organizes movement of households or jobs within the region. The Household and Employment Location submodule models the location choices of households and jobs from the available vacant real estate. The Real Estate Development submodule determines the location, type, and quantity of new construction and redevelopment by developers. Applications of UrbanSim tend to model changes in residential housing and business (employment) locations over 15–20 yr at an annual time step. The UrbanSim computer application is one of the only land-use change models we have identified that is already licensed as OS.

Fig. A1.1. Structure of UrbanSim model (adapted and interpreted from Figs. 3 and 4 in Waddell 2002).

