

# The Challenge of Heterogeneously Licensed Systems in Open Architecture Software Ecosystems

Thomas A. Alspaugh, Hazeline U. Asuncion, and Walt Scacchi

Institute for Software Research

University of California, Irvine

Irvine, CA 92697-3455 USA

{alspaugh,hasuncion,wscacchi}@ics.uci.edu

## Abstract

*The role of software ecosystems in the development and evolution of open architecture systems has received insufficient consideration. Such systems are composed of heterogeneously-licensed components, open source or proprietary or both, in an architecture in which evolution can occur by evolving existing components or by replacing them. But this may result in possible license conflicts and organizational liability for failure to fulfill license obligations. We have developed an approach for understanding and modeling software licenses, as well as for analyzing conflicts among groups of licenses in realistic system contexts and for guiding the acquisition, integration, or development of systems with open source components in such an environment. This work is based on empirical analysis of representative software licenses and heterogeneously-licensed systems, and collaboration with researchers in the legal world. Our approach provides guidance for achieving a “best-of-breed” component strategy while obtaining desired license rights in exchange for acceptable obligations.*

## Biographies

**Thomas Alspaugh** is a project scientist at the Institute for Software Research, University of California, Irvine, and an adjunct professor of Computer Science at Georgetown University. His research interests are in software engineering, requirements, and licensing. Before completing his Ph.D., he worked as a software developer, team lead, and manager in industry, and as a computer scientist at the Naval Research Laboratory on the Software Cost Reduction or A-7 project.

**Hazel Asuncion** is an assistant professor in the Department of Computing and Software Systems at the University of Washington, Bothell. Her research interests focus on architecture, workflows, and software acquisition.

**Walt Scacchi** is a senior research scientist and research faculty member at the Institute for Software Research, University of California, Irvine. He received a Ph.D. in Information and Computer Science from UC Irvine in 1981. From 1981-1998, he was on the faculty at the University of Southern California. In 1999, he joined the Institute for Software Research at UC Irvine. He has published more than 150 research papers, and has directed 60 externally funded research projects. In 2007, he served as General Chair of the 3rd. IFIP International Conference on Open Source Systems (OSS2007), Limerick, IE. In 2010, he chaired the Workshop on the Future of Research in Free and Open Source Software, Newport Beach, CA, for the Computing Community Consortium and the National Science Foundation. He also served as Co-Chair of the Doctoral Consortium at the 5th. and 6th. IFIP International Conference on Open Source Systems (OSS 2009, OSS2010). Last, in 2012, he served as General Co-Chair of the 8th. IFIP International Conference on Open Source Systems (OSS2012).

# 1 Introduction

A substantial number of development organizations are adopting a strategy in which a software-intensive system is developed with an *open architecture* (OA) [27], whose components may be open source software (OSS) or proprietary with open application programming interfaces (APIs). Such systems evolve not only through the evolution of their individual components, but also through replacement of one component by another, possibly from a different producer or under a different license. With this approach, the organization becomes an integrator of components largely produced elsewhere that are interconnected through open APIs as necessary to achieve the desired result. An OA development process results in an ecosystem in which the integrator is influenced from one direction by the goals, interfaces, license choices, and release cycles of the component producers, and in another direction by the needs of its consumers. As a result the software components are reused more widely, and the resulting OA systems can achieve reuse benefits such as reduced costs, increased reliability, and potentially increased agility in evolving to meet changing needs. An emerging challenge is to realize the benefits of this approach when the individual components are *heterogeneously licensed*, each potentially with a different license, rather than a single OSS license as in uniformly-licensed OSS projects, or a single proprietary license when acquired from a vendor employing a proprietary development scheme.

This challenge is inevitably entwined with the software ecosystems that arise for OA systems. We find that an OA software ecosystem involves organizations and individuals producing and consuming components, and supply paths from producer to consumer; but also

- the OA of the system(s) in question,
- the open interfaces met by the components,
- the degree of coupling in the evolution of related components, and
- the rights and obligations resulting from the software licenses under which various components are released, that propagate from producers to consumers.

An example software ecosystem is portrayed in Figure 1.

In order to most effectively use an OA approach in developing and evolving a system, it is essential to consider this OA ecosystem. An OA system draws on components from proprietary vendors and open source projects. Its architecture is made possible by the existing general ecosystem of producers, from which the initial components are chosen. The choice of a specific OA begins a specialized software ecosystem involving components that meet (or can be shimmed to meet) the open interfaces used in the architecture. We do not claim this is the best or the only way to reuse components or produce systems, but it is an ever more widespread way. In this paper we build on previous work on heterogeneously-licensed systems [14, 30, 2] by examining how OA development affects and is affected by software ecosystems, and the role of component licenses in OA software ecosystems.

A motivating example of this approach is the Unity game development tool, produced by Unity Technologies [37]. Its license agreement, from which we quote below, lists eleven distinct licenses and indicates the tool is produced, apparently using an OA approach, using at least 18 externally produced components or groups of components:

1. The Mono Class Library, Copyright 2005-2008 Novell, Inc.

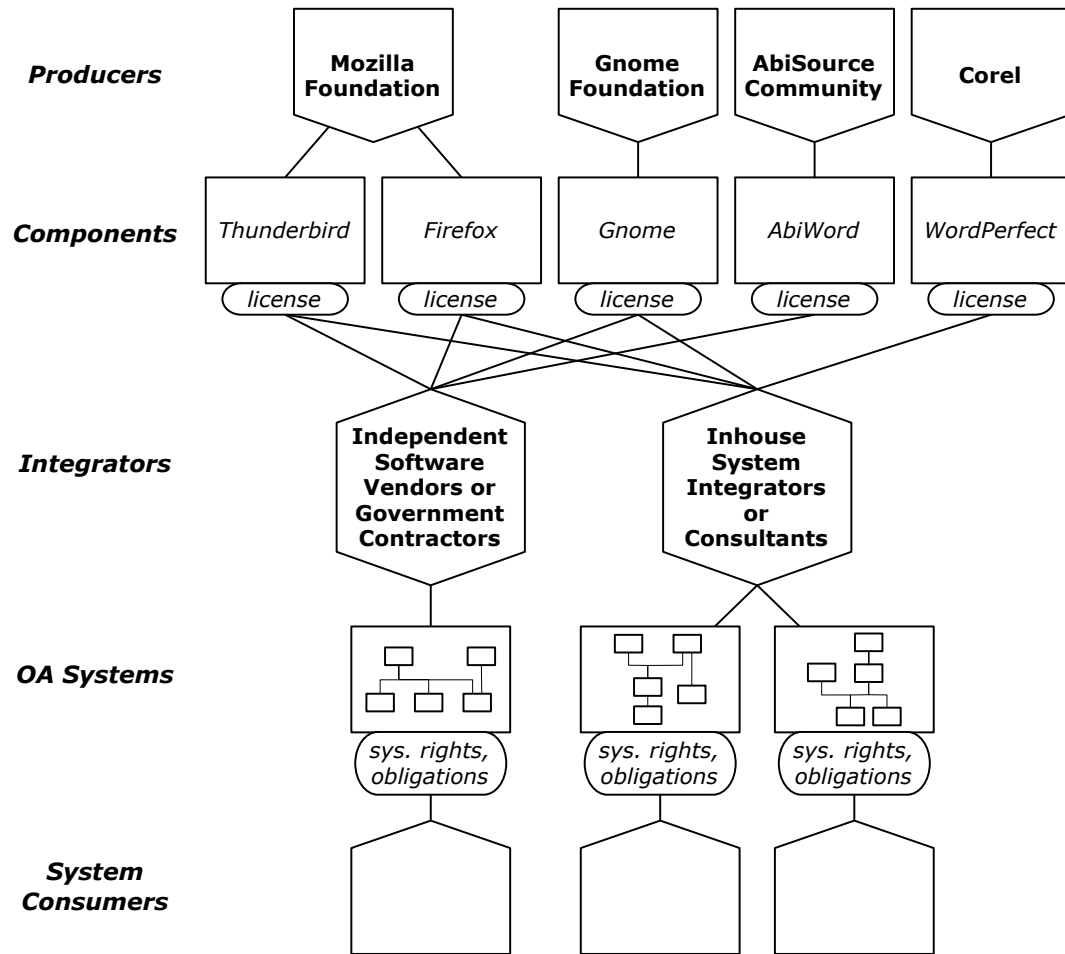


Figure 1: An example of a software supply network within a software ecosystem in which OA systems are developed

2. The Mono Runtime Libraries, Copyright 2005-2008 Novell, Inc.
3. Boo, Copyright 2003-2008 Rodrigo B. Oliveira
4. UnityScript, Copyright 2005-2008 Rodrigo B. Oliveira
5. OpenAL cross platform audio library, Copyright 1999-2006 by authors.
6. PhysX physics library. Copyright 2003-2008 by Ageia Technologies, Inc.
7. libvorbis. Copyright (c) 2002-2007 Xiph.org Foundation
8. libtheora. Copyright (c) 2002-2007 Xiph.org Foundation
9. zlib general purpose compression library. Copyright (c) 1995-2005 Jean-loup Gailly and Mark Adler
10. libpng PNG reference library
11. jpeglib JPEG library. Copyright (C) 1991-1998, Thomas G. Lane.
12. Twilight Prophecy SDK, a multi-platform development system for virtual reality and multimedia. Copyright 1997-2003 Twilight 3D Finland Oy Ltd
13. dynamic.bitset, Copyright Chuck Allison and Jeremy Siek 2001-2002.
14. The Mono C# Compiler and Tools, Copyright 2005-2008 Novell, Inc.
15. libcurl. Copyright (c) 1996-2008, Daniel Stenberg <daniel@haxx.se>.
16. PostgreSQL Database Management System
17. FreeType. Copyright (c) 2007 The FreeType Project (www.freetype.org).
18. NVIDIA Cg. Copyright (c) 2002-2008 NVIDIA Corp.

An OA system can evolve by a number of distinct mechanisms, some of which are common to all systems but others of which are a result of heterogeneous component licenses in a single system.

***Component evolution*** — One or more components can evolve, altering the overall system's characteristics (for example, upgrading and replacing the Firefox Web browser from version 3.5 to 3.6).

***Component replacement*** — One or more components may be replaced by others with different behaviors but the same interface, or with a different interface and the addition of shim code to make it match (for example, replacing the AbiWord word processor with either Open Office or MS Word. ).

***Architecture evolution*** — The OA can evolve, using the same components but in a different configuration, altering the system's characteristics. For example, as discussed in Section 4, changing the configuration in which a component is connected can change how its license affects the rights and obligations for the overall system. This could arise when replacing email and word processing applications with web services like Google Mail and Google Docs.

***Component license evolution*** — The license under which a component is available may change, as for example when the license for the Mozilla core components was changed from the

Mozilla Public License (MPL) to the current Mozilla Disjunctive Tri-License; or the component may be made available under a new version of the same license, as for example when the GNU General Public License (GPL) version 3 was released.

***Changes to the desired rights or acceptable obligations*** — The OA system’s integrator or consumers may desire additional license rights (for example the right to sublicense in addition to the right to distribute), or no longer desire specific rights; or the set of license obligations they find acceptable may change. In either case the OA system evolves, whether by changing components, evolving the architecture, or other means, to provide the desired rights within the scope of the acceptable obligations. For example, they may no longer be willing or able to provide the source code for components within the reciprocity scope of a GPL-licensed module.

The interdependence of integrators and producers results in a co-evolution of software within an OA ecosystem. Closely-coupled components from different producers must evolve in parallel in order for each to provide its services, as evolution in one will typically require a matching evolution in the other. Producers may manage their evolution with a loose coordination among releases, for example as between the Gnome and Mozilla organizations. Each release of a producer component create a tension through the ecosystem relationships with consumers and their releases of OA systems using those components, as integrators accommodate the choices of available, supported components with their own goals and needs. As discussed in our previous work [2], license rights and obligations are manifested at each component’s interface, then mediated through the system’s OA to entail the rights and corresponding obligations for the system as a whole. As a result, integrators must frequently re-evaluate an OA system’s rights and obligations. In contrast to homogeneously-licensed systems, *license change across versions is a characteristic of OA ecosystems*, and architects of OA systems require tool support for managing the ongoing licensing changes.

We propose that such support must have several characteristics.

- It must rest on a license structure of rights and obligations (Section 5), focusing on obligations that are enactable and testable. For example, many OSS licenses include an obligation to make a component’s modified code public, and whether a specific version of the code is public at a specified Web address is both enactable (it can be put into practice) and testable. In contrast, the GPL v.3 provision “No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty” is not enactable in any obvious way, nor is it testable — how can one verify what others deem?
- It must take account of the distinctions between the design-time, build-time, and distribution-time architectures (Sections 4 and 5) and the rights and obligations that come into play for each of them.
- It must distinguish the architectural constructs significant for software licenses, and embody their effects on rights and obligations (Section 4).
- It must define license architectures (Section 5).
- It must provide an automated environment for creating and managing license architectures. We have developed a prototype that manages a license architecture as a view of its system architecture [3, 4, 5].

- Finally, it must automate calculations on system rights and obligations so that they may be done easily and frequently, whenever any of the factors affecting rights and obligations may have changed (Section 6).

In the remainder of this paper, we survey some related work (Section 2), provide an overview of OSS licenses and projects (Section 3), define and examine characteristics of open architectures (Section 4), introduce a structure for licenses and outline license architectures (Section 5), and sketch our approach for license analysis (Section 6). We then close with our conclusions (Section 7).

## 2 Related Work

It has been typical until recently that each software or information system is designed, built, and distributed under the terms of a single proprietary or OSS license, with all its components homogeneously covered by that same license. The system is distributed, with sources or executables bearing copyright and license notices, and the license gives specific rights while imposing corresponding obligations that system consumers (whether external developers or users) must honor, subject to the provisions of contract and commercial law. Consequently, there has been some very interesting study of the choice of OSS license for use in an OSS development project, and its consequences in determining the likely success of such a project.

Brown and Booch [10] discuss issues that arise in the reuse of OSS components, such as that interdependence (via component interconnection at design-time, or linkage at build-time or run-time) causes functional changes to propagate, and versions of the components evolve asynchronously, giving rise to co-evolution of interrelated code in the OSS-based systems. If the components evolve, the OA system itself is evolving. The evolution can also include changes to the licenses, and the licenses can change from component version to version (cf. Footnote 1).

Legal scholars have examined OSS licenses and how they interact in the legal domain, but not in the context of HLSs [13, 28, 33]. For example, Rosen [28] surveys eight OSS licenses and creates two new ones written to professional legal standards. He examines interactions primarily in terms of the general categories of reciprocal and non-reciprocal licenses, rather than in terms of specific licenses. However, common to this legal scholarship is an approach that analyzes the interaction among licenses on a pairwise or interlinked components basis. This analysis scheme means that if system A has OSS license of type X, system B has a licenses of type Y, and system C has license of type Z, then license interaction (matching, subsumption, or conflicting constraints) is determined by how A interacts with B, B with C, and A with C. This follows from related legal scholarship (e.g. Burk [11]) that brought attention to problems of whether or not intellectual property rights apply depending on how the systems (or components) are interlinked (cf. German and Hassan [14]). We similarly adopt this approach in our analysis efforts.

Stewart et al. [34] conducted an empirical study to examine whether license choice is related to OSS project success, finding a positive association following from the selection of business-friendly licenses. Sen, Subramaniam, and Nelson in a series of studies [31, 32, 35] similarly find positive relationships between the choice of a OSS license and the likelihood of both successful OSS development and adoption of corresponding OSS systems within enterprises. These studies direct attention to OSS projects that adopt and identify their development efforts through use of a single OSS license. However, there has been little explicit guidance on how best to develop, deploy, and sustain complex software systems when heterogeneously-licensed components are involved, and thus multiple OSS and proprietary licenses may be involved. Ven and Mannaert [38], Tuunanen et

al. [36], and German and Hassan [14] are recent exceptions.

Jansen and colleagues [18, 19] draw attention to their observation that software ecosystems (a) embed software supply networks that span multiple organizations, and (b) are embedded within a network of intersecting or overlapping software ecosystems that span the world of software engineering practice. Scacchi [29] for example, identifies that the world of open source software (OSS) development is a software ecosystem different from those of commercial software producers, and its supply networks are articulated within a network of FOSS development projects. Networks of OSS ecosystems have also begun to appear around very large OSS projects for Web browsers, Web servers, word processors, and others, as well as related application development environments like NetBeans and Eclipse, and these networks have become part of global information infrastructures [20].

OSS ecosystems also exhibit strong relationships between the ongoing evolution of OSS systems and their developer/user communities, such that the success of one co-depends on the success of the other [29]. Ven and Mannaert discuss the challenges independent software vendors face in combining OSS and proprietary components, with emphasis on how OSS components evolve and are maintained in this context [38].

Boucharas and colleagues [9] then draw attention to the need to more systematically and formally model the contours of software supply networks, ecosystems, and networks of ecosystems. Such a formal modeling base may then help in systematically reasoning about what kinds of relationships or strategies may arise within a software ecosystem. For example, Kuehnel [21] examines how Microsoft’s software ecosystem developed around in operating systems (MS Windows) and key applications (e.g., MS Office) may be transforming from “predator” to “prey” in its effort to control the expansion of its markets to accommodate OSS (as the extant prey) that eschew closed source software with proprietary software licenses.

Other previous work examined how best to align acquisition, system requirements, architectures, and OSS components across different software license regimes to achieve the goal of combining OSS with proprietary software that provide open APIs when developing a composite “system of systems” [30]. This is particularly an issue for the U.S. Federal Government in its acquisition of complex software systems subject to Federal Acquisition Regulations (FARs) and military service-specific regulations. HLSs give rise to new functional and non-functional requirements that further constrain what kinds of systems can be built and deployed, as well as recognizing that acquisition policies can effectively exclude certain OA configurations, while accommodating others, based on how different licensed components may be interconnected.

### 3 Open-Source Software

Traditional proprietary licenses allow a company to retain control of software it produces, and restrict the access and rights that outsiders can have. OSS licenses, on the other hand, are designed to encourage sharing and reuse of software, and grant access and as many rights as possible. OSS licenses are classified as *permissive* or *reciprocal*. Permissive OSS licenses such as the Berkeley Software Distribution (BSD) license, the Massachusetts Institute of Technology license, the Apache Software License, and the Artistic License, grant nearly all rights to components and their source code, and impose few obligations. Anyone can use the software, create derivative works from it, or include it in proprietary projects. Typical permissive obligations are simply to not remove the copyright and license notices.

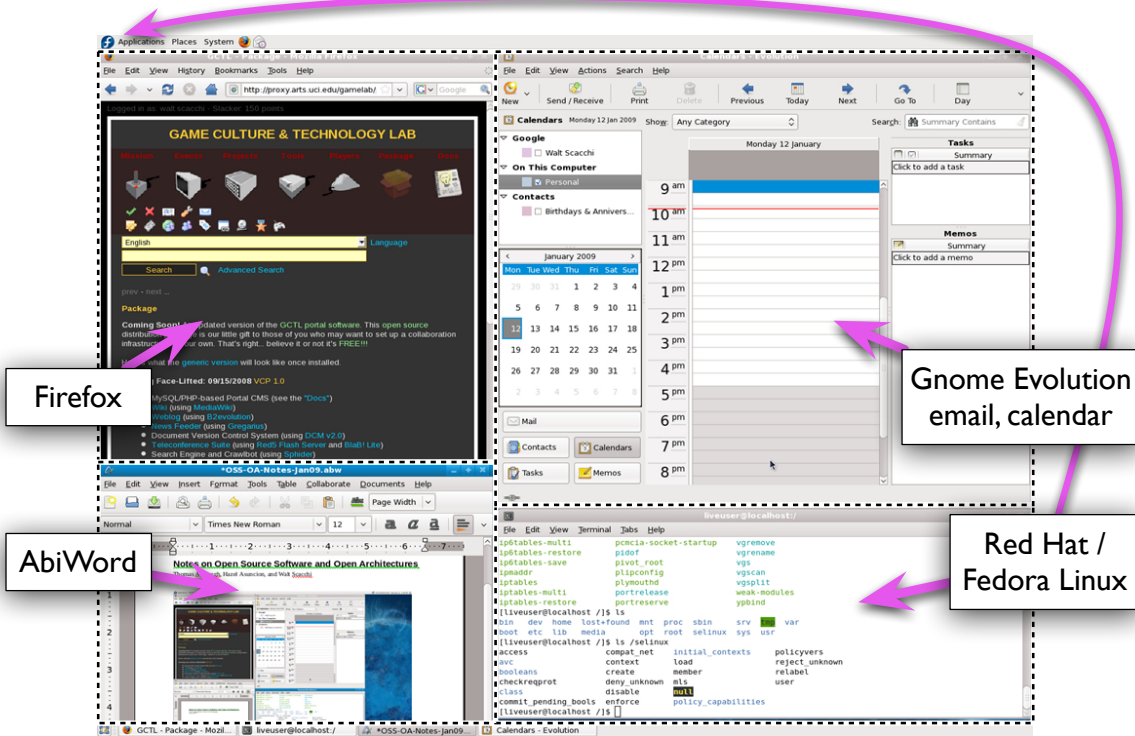


Figure 2: A heterogeneously-licensed composite system

*Reciprocal* OSS licenses take a more active stance towards sharing and reusing software, imposing obligations with respect to the original software in exchange for rights, and also reciprocally on any future derivative versions of it in exchange for the right to create and distribute the derivative versions. The most demanding reciprocal licenses such as GPL impose the obligation that reciprocally-licensed software not be combined (for various definitions of “combined”) with any software that is not in turn also released under the reciprocal license. The goals are to increase the domain of OSS by encouraging developers to bring more components under its aegis, and to prevent improvements to OSS components from vanishing behind proprietary licenses. Example reciprocal licenses are GPL, the Mozilla Public License (MPL), and the Common Public License.

Both proprietary and OSS licenses typically disclaim liability, assert no warranty is implied, and obligate licensees to not use the licensor’s name or trademarks. Newer licenses often cover patent issues as well, either giving a restricted patent license or explicitly excluding patent rights.

The Mozilla Disjunctive Tri-License licenses the core Mozilla components under any one of three licenses (MPL, GPL, or the GNU Lesser General Public License LGPL); OSS developers can choose the one that best suits their needs for a particular project and component.

The Open Source Initiative (OSI) maintains a widely-respected definition of “open source” and gives its approval to licenses that meet it [26]. OSI maintains and publishes a repository of approximately 70 approved OSS licenses.

Common practice has been for an OSS project to choose a single license under which all its products are released, and to require developers to contribute their work only under conditions compatible with that license. For example, the Apache Contributor License Agreement grants



enough of each author’s rights to the Apache Software Foundation for the foundation to license the resulting systems under the Apache Software License. This sort of rights regime, in which the rights to a system’s components are homogeneously granted and the system has a single well-defined OSS license, was the norm in the early days of OSS and continues to be practiced.

More recently it has become increasingly common for OSS (and mixed) systems to be composed of components from several different organizations governed by different licenses. An example of such a heterogeneously-licensed system, this one composed entirely of OSS components, is shown in Figure 2.

## 4 Open Architecture

Open architecture (OA) software is a customization technique introduced by Oreizy [27] that enables third parties to modify a software system through its exposed architecture, evolving the system by replacing its components. Increasingly more software-intensive systems are developed using an OA strategy, not only with OSS components but also proprietary components with open APIs (e.g. [37]). Using this approach can lower development costs and increase reliability and function [30]. Composing a system with heterogeneously-licensed components, however, increases the likelihood of conflicts, liabilities, and no-rights stemming from incompatible licenses. Thus, in our work we define an OA system as a *software system consisting of components that are either open source or proprietary with open API, whose overall system rights at a minimum allow its use and redistribution, in full or in part.*

It may appear that using a system architecture that incorporates OSS components and uses open APIs will result in an OA system. But not all such architectures will produce an OA, since the (possibly empty) set of available license rights for an OA system depends on: (a) how and why OSS and open APIs are located within the system architecture, (b) how OSS and open APIs are implemented, embedded, or interconnected, and (c) the degree to which the licenses of different OSS components encumber all or part of a software system’s architecture into which they are integrated [30, 1].

The following kinds of software elements appearing in common software architectures can affect whether the resulting systems are open or closed [8].

**Software source code components**—These can be either (a) standalone programs, (b) libraries, frameworks, or middleware, (c) inter-application script code such as C shell scripts, or (d) intra-application script code, as for creating Rich Internet Applications using domain-specific languages such as XUL for the Firefox Web browser [12] or “mashups” [25]. Their source code is available and they can be rebuilt. Each may have its own distinct license.

**Executable components**—These components are in binary form, and the source code may not be open for access, review, modification, or possible redistribution [28]. If proprietary, they often cannot be redistributed, and so such components will be present in the design- and run-time architectures but not in the distribution-time architecture.

**Software services**—An appropriate software service can replace a source code or executable component.

**Application programming interfaces/APIs**—Availability of externally visible and accessible APIs is the minimum requirement for an “open system” [24]. Open APIs are not and cannot be licensed, and can limit the propagation of license obligations.

**Software connectors**—Software whose intended purpose is to provide a standard or reusable way of communication through common interfaces, e.g. High Level Architecture [22], CORBA, MS

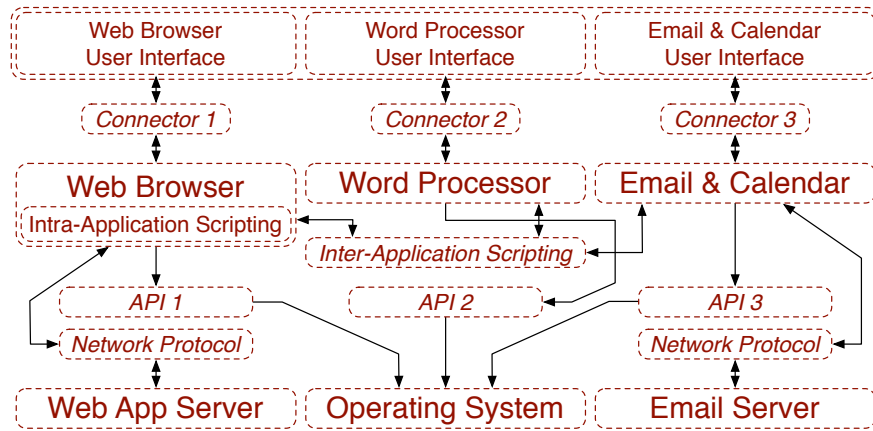
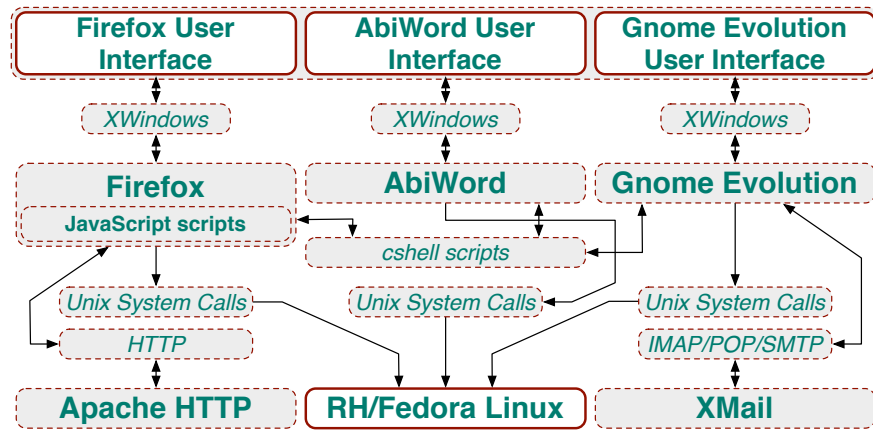


Figure 3: The design-time architecture of the system of Figure 2



Components/connectors not visible in Figure 2 are shown in gray

Figure 4: A build-time architecture describing the version running in Figure 2.

.NET, Enterprise Java Beans, and GNU Lesser General Public License (LGPL) libraries. Connectors can also limit the propagation of license obligations.

**Methods of connection**—These include linking as part of a configured subsystem, dynamic linking, and client-server connections. Methods of connection affect license obligation propagation, with different methods affecting different licenses.

**Configured system or subsystem architectures**—These are software systems that are used as atomic components of a larger system, and whose internal architecture may comprise components with different licenses, affecting the overall system license. To minimize license interaction, a configured system or sub-architecture may be surrounded by what we term a *license firewall*, namely a layer of dynamic links, client-server connections, license shims, or other connectors that block the propagation of reciprocal obligations.

Figure 3 shows a high-level view of a reference architecture that includes all the kinds of software elements listed above. This reference architecture has been instantiated in a number of configured systems that combine OSS and closed source components. The configured systems consist of software components such as a Mozilla Firefox Web browser, Gnome Evolution email client, and

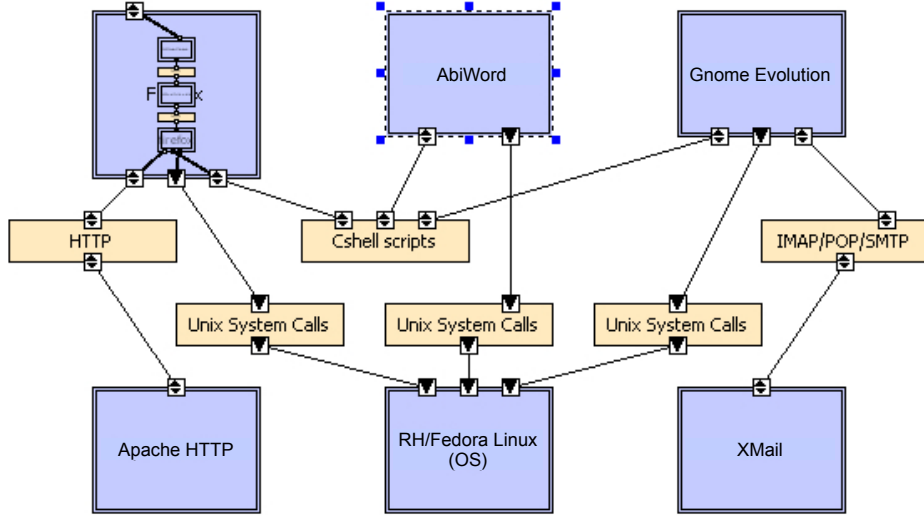


Figure 5: Instantiated build-time architecture (Figure 2) within ArchStudio [16]

AbiWord word processor (similar to MS Word), all running on a RedHat Fedora Linux operating system accessing file, print, and other remote networked servers such as an Apache Web server. Figure 4 shows a build-time architecture instantiated with those choices. Figure 5 is a screenshot of the instantiated architecture in our extension of ArchStudio [16], where it is one view of the architecture data structure whose automated analysis is discussed and shown in later sections. Components are interconnected through a set of software connectors that bridge the interfaces of components and combine the provided functionality into the system’s services.

The topology of the build-time architecture also determines the OA software ecosystem of the system, with its dependencies on suppliers and (implicitly) the evolution paths that ecosystem can take, in the context of design and instantiation choices that involve different suppliers of components of the same sort, or more extensive changes that involve suppliers of components of a different sort. Figure 6 shows the reference architecture of Figure 3, annotated with the supplier organizations implied by the instantiations of the build-time architecture of Figure 4. The choices are a result of desired functional abilities and nonfunctional qualities, and may also be influenced by desired supply-chain characteristics, licensing regimes, and future software ecosystem evolution paths. All these choices, however, are limited by software license constraints and interactions, as described in the next two sections.

## 5 Software Licenses

Copyright law is the common basis for software licenses, and gives the original author of a work certain exclusive rights: the rights to use, copy, modify, merge, publish, distribute, sub-license, and sell copies. The author may license these rights, individually or in groups, to others; the license may give a right either exclusively or non-exclusively. After a period of years, copyright rights enter the public domain. Until then copyright may only be obtained through licensing.

Licenses typically impose obligations that must be met in order for the licensee to realize the assigned rights. Common obligations include the obligation to publish at no cost any source code you modify (MPL) or the reciprocal obligation to publish all source code included at build-time

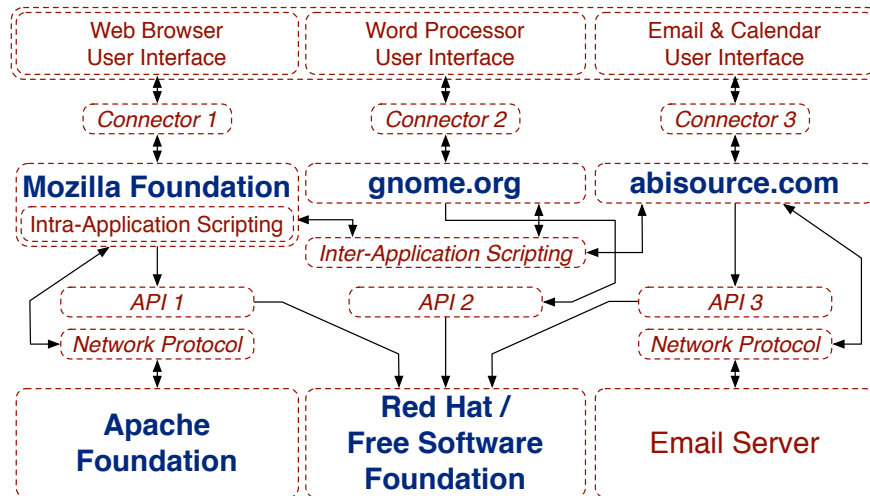


Figure 6: Reference architecture components supplied by an organization, indicating the integrator’s dependencies on suppliers, mediated by interfaces and licenses

or statically linked (GPL). The obligations may conflict, as when a GPL’d component’s reciprocal obligation to publish source code of other components is combined with a proprietary component’s license prohibition of publishing its source code. In this case, no rights may be available for the system as a whole, not even the right of use, because the two obligations cannot simultaneously be met and thus neither component can be used as part of the system.

The basic relationship between software license rights and obligations can be summarized as follows: if the specified obligations are met, then the corresponding rights are granted. For example, if you publish your modified source code and sub-licensed derived works under MPL, then you get all the MPL rights for both the original and the modified code. However, license details are complex, subtle, and difficult to comprehend and track—it is easy to become confused or make mistakes. The challenge is multiplied when dealing with configured system architectures that compose a large number of components with heterogeneous licenses, so that the need for legal counsel begins to seem inevitable [28, 13].

We have developed an approach for expressing software licenses that is more formal and less ambiguous than natural language, and that allows us to calculate and identify conflicts arising from the rights and obligations of two or more component’s licenses. Our approach is based on Hohfeld’s classic group of eight fundamental jural relations [15], of which we use *right*, *duty*, *no-right*, and *privilege*. We start with a tuple  $\langle \text{actor, operation, action, object} \rangle$  for expressing a right or obligation. The actor is the “licensee” for all the licenses we have examined. The operation is one of the following: “may”, “must”, “must not”, or “need not”, with “may” and “need not” expressing rights and “must” and “must not” expressing obligations. Because copyright rights are only available to entities who have been granted a sublicense, only the listed rights are available, and the absence of a right means that it is not available. The action is a verb or verb phrase describing what may, must, must not, or need not be done, with the object completing the description. A license may be expressed as a set of rights, with each right associated with zero or more obligations that must be fulfilled in order to enjoy that right. Figure 7 shows the metamodel with which we express licenses, discussed at greater length in our previous work [3, 4, 5].

Our license model forms a basis for effective reasoning about licenses in the context of actual

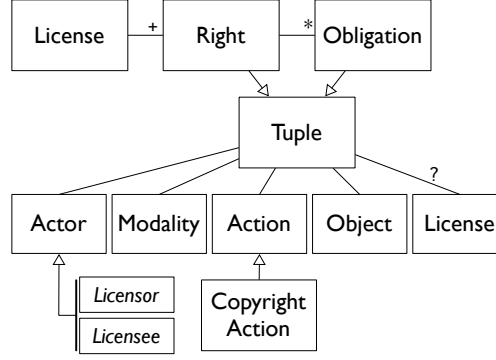


Figure 7: License metamodel

systems, and calculating the resulting rights and obligations. In order to do so, we need a certain amount of information about the system’s configuration at design-, build-, distribution-, and run-time. The needed information comprises the *license architecture*, an abstraction of the system architecture:

1. the set of components of the system;
2. the relation mapping each component to its license (Figure 8);
3. the relation mapping each component to its set of sources; and
4. the relation from each component to the set of components in the same license scope, for each license for which “scope” is defined (e.g. GPL), and from each source to the set of sources of components in the scope of its component.

With this information and definitions of the licenses involved, we can calculate rights and obligations for individual components or for the entire system, and guide heterogeneously-licensed system design.

Heterogeneously-licensed system designers have developed a number heuristics to guide architectural design while avoiding some license conflicts. First, it is possible to use a reciprocally-licensed component through a *license firewall* that limits the scope of reciprocal obligations. Rather than connecting conflicting components directly through static or other build-time links, the connection is made through a dynamic link, client-server protocol, license shim (such as a Limited General Public License connector), or run-time plug-ins. A second approach used by a number of large organizations is simply to avoid using any reciprocally-licensed components. A third approach is to meet the license obligations (if that is possible) by for example retaining copyright and license notices in the source and publishing the source code. However, even using design heuristics such as these (and there are many), keeping track of license rights and obligations across components that are interconnected in complex OAs quickly becomes too cumbersome. Automated support is needed to manage the multi-component, multi-license complexity.

## 6 License Analysis

Given a specification of a software system’s architecture, we can associate software license attributes with the system’s components, connectors, and sub-system architectures, resulting in a license

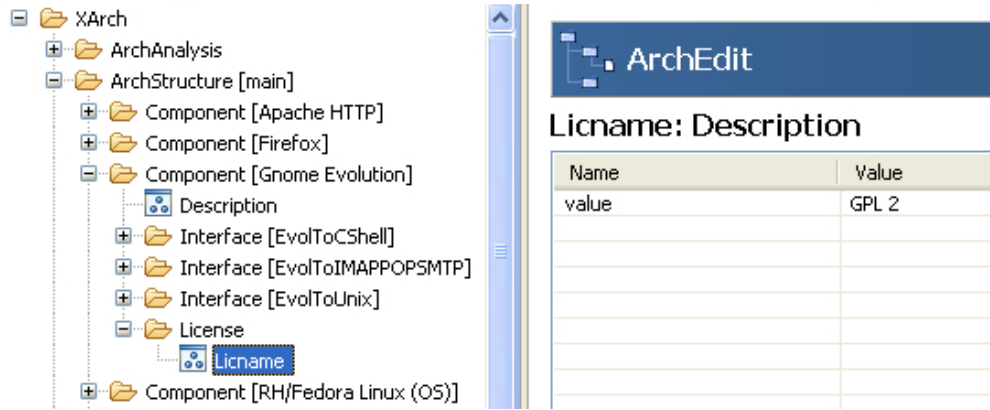


Figure 8: License annotation of Gnome Evolution component seen in Figure 5, as implemented in our extension of ArchStudio4 [3, 4, 5, 7]

architecture for the system, and calculate the copyright rights and obligations for the system’s configuration. Due to the complexity of license architecture analysis, and the need to re-analyze every time a component evolves, a component’s license changes, a component is substituted, or the system architecture changes, OA integrators really need an automated license architecture analysis environment. We have developed a prototype of such an environment [3, 4, 5].

We use an architectural description language specified in xADL [17] to describe OAs that can be designed and analyzed with a software architecture design environment [23], such as ArchStudio4 [16]. We have built the Software Architecture License Analysis module on top of ArchStudio’s Traceability View [6]. This allows for the specification of licenses as a list of attributes (license tuples) using a form-based user interface in ArchStudio4 shown in Figure 8 [16, 23].

We analyze rights and obligations as described below [3, 4, 5] as implemented in our extension of ArchStudio4 [2, 7]. For example in Figure 9 we show the results of a license analysis for the architecture shown in Figure 5, showing no conflicts. If we replace AbiWord with Corel WordPerfect, the license analysis then shows what rights are missing, as seen in Figures 9 and 10. The analysis that determines how this work is described in the remaining subsections.

## 6.1 Propagation of reciprocal obligations

We follow the widely-accepted interpretation that build-time static linkage propagate the reciprocal obligations, but appropriate license firewalls do not. Analysis begins, therefore, by propagating these obligations along all connectors that are not license firewalls.

## 6.2 Obligation conflicts

An obligation can conflict with another obligation, or with the set of available rights, by requiring a copyright right that has not been granted. For instance, a proprietary license may require that a licensee must not redistribute source code, but GPL states that a licensee must redistribute source code. Thus, the conflict appears in the modality of the two otherwise identical obligations, “must not” in the proprietary license and “must” in GPL.

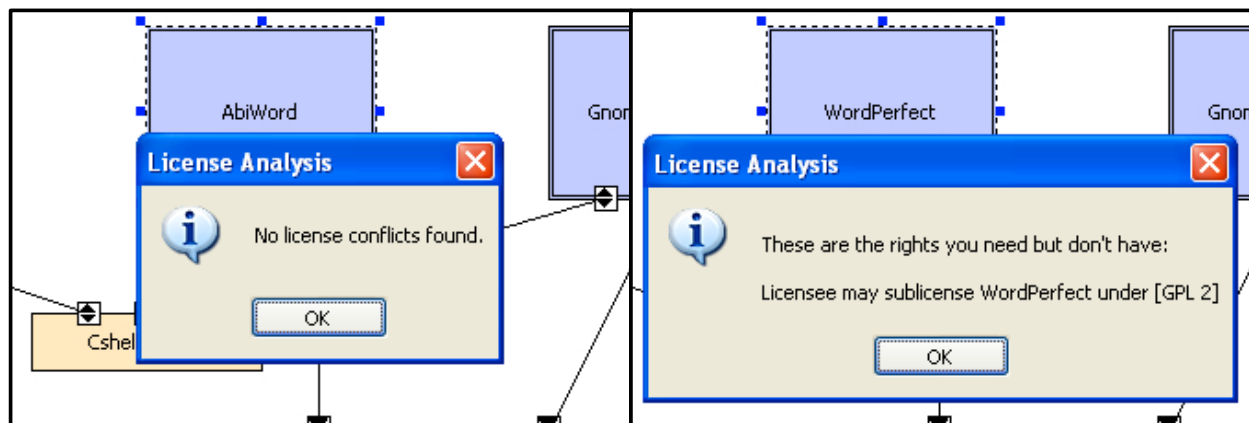


Figure 9: On the left is displayed the result of the automated analysis of the architecture in Figure 6. After replacing the AbiWord word processor with the WordPerfect word processor, and redoing the analysis, the tool shows this alternative design results in license conflicts [3, 4, 5, 7].

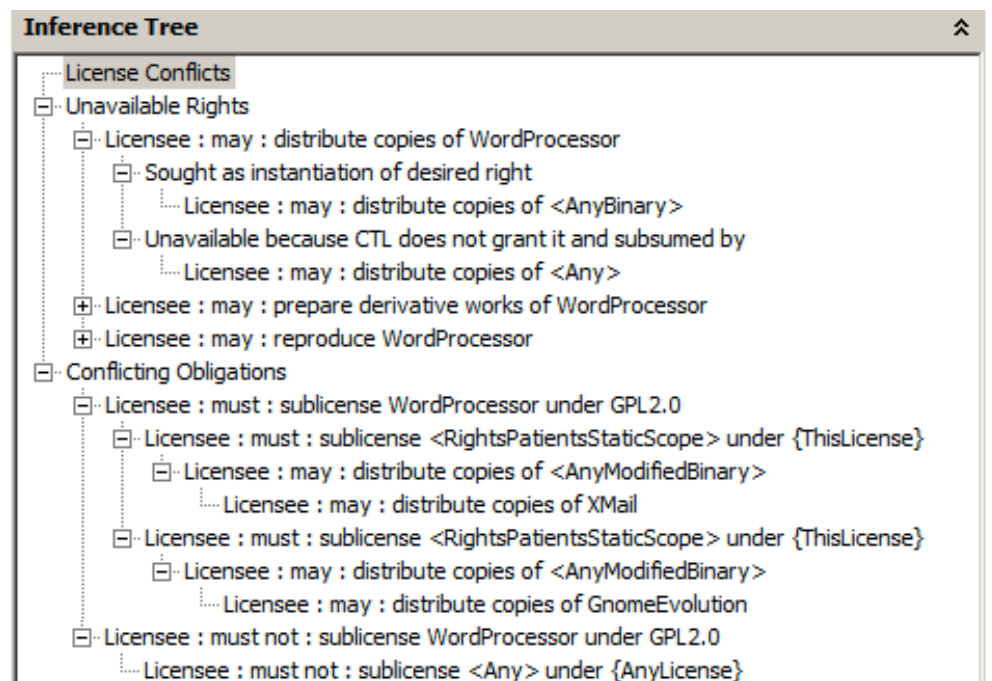


Figure 10: Prototype explanation results for a CTL-GPL2.0 conflict implemented in our extension of ArchStudio4 [3, 4, 5, 7]: (at top) unavailable rights (partially collapsed), (middle) two conflicting obligations.



In order to obtain a particular desired right  $r$  for a specific module or other entity  $e$ , in other words a desired *concrete right*, one of two cases must hold:

- Figure 11 illustrates one step of the application of a license to obtain a desired concrete right  $r$ . In the license of  $r$ 's object  $e$ , we search for an abstract right  $R$  subsuming  $r$ . The figure shows two obligations  $O_1$  and  $O_2$  of  $R$ , which we apply to  $r$ 's object  $e$  in order to obtain  $r$ 's concrete obligations  $o_1$  and  $o_2$ . Depending on what kind of object  $O_1$  has,  $o_1$  could apply to  $e$  itself, in which case  $e = e'_1$ , or to an entity related to  $e$ , or (if  $L$  is a propagating license) to another module linked or otherwise connected to  $e$ . Finally, in order to fulfill  $o_1$  we must have  $o_1$ 's correlative right  $r'_1$ . The same considerations apply for  $O_2$ , of course. The heavy arrow shows the flow of inference from desired concrete right through to required concrete obligations and correlative rights.

## 7 Conclusion

16



how and why an OA system evolves in its ecosystem. We note that license changes across versions of components is a characteristic of OA systems and software ecosystems with heterogeneously licensed components. A structure for modeling software licenses and the license architecture of a system and automated support for calculating its rights and obligations are needed in order to manage a system's evolution in the context of its ecosystem. We have outlined an approach for achieving these and sketched how they further the goal of reusing components in developing software-intensive systems. Much more work remains to be done, but we believe this approach turns a vexing problem into one for which workable solutions can be obtained.

## Acknowledgments

This research is supported by grants #N00244-10-1-0077 and #N00244-12-1-0004 from the Acquisition Research Program at the Naval Postgraduate School, and by grant #0808783 from the U.S. National Science Foundation. No review, approval, or endorsement implied.

## References

- [1] T. A. Alspaugh and A. I. Antón. Scenario support for effective requirements. *Information and Software Technology*, 50(3):198–220, Feb. 2008.
- [2] T. A. Alspaugh, H. U. Asuncion, and W. Scacchi. Analyzing software licenses in open architecture software systems. In *2nd International Workshop on Emerging Trends in FLOSS Research and Development (FLOSS)*, May 2009.
- [3] T. A. Alspaugh, H. U. Asuncion, and W. Scacchi. Intellectual property rights requirements for heterogeneously-licensed systems. In *17th IEEE International Requirements Engineering Conference (RE'09)*, pages 24–33, Aug. 31–Sept. 4 2009.
- [4] T. A. Alspaugh, H. U. Asuncion, and W. Scacchi. Presenting software license conflicts through argumentation. In *23rd International Conference on Software Engineering and Knowledge Engineering (SEKE 2011)*, pages 509–514, July 2011.
- [5] T. A. Alspaugh, W. Scacchi, and H. U. Asuncion. Software licenses in context: The challenge of heterogeneously-licensed systems. *Journal of the Association for Information Systems*, 11(11):730–755, Nov. 2010.
- [6] H. Asuncion and R. N. Taylor. Capturing custom link semantics among heterogeneous artifacts and tools. In *5th International Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE)*, May 2009.
- [7] H. U. Asuncion. *Architecture-centric traceability for stakeholders (ACTS)*. PhD thesis, University of California, Irvine, 2009.
- [8] L. Bass, P. Clements, and R. Kazman. *Software Architecture in Practice*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
- [9] V. Boucharas, S. Jansen, and S. Brinkkemper. Formalizing software ecosystem modeling. In *First International Workshop on Open Component Ecosystems (IWOCE'09)*, pages 41–50, 2009.

- [10] A. W. Brown and G. Booch. Reusing open-source software and practices: The impact of open-source on commercial vendors. In *Software Reuse: Methods, Techniques, and Tools (ICSR-7)*, Apr. 2002.
- [11] D. L. Burk. Proprietary rights in hypertext linkages. *Journal of Information, Law and Technology*, 1998(2), 1998.
- [12] K. Feldt. *Programming Firefox: Building Rich Internet Applications with XUL*. O'Reilly Media, Inc., 2007.
- [13] R. Fontana, B. M. Kuhn, E. Moglen, M. Norwood, D. B. Ravicher, K. Sandler, J. Vasile, and A. Williamson. *A Legal Issues Primer for Open Source and Free Software Projects*. Software Freedom Law Center, 2008.
- [14] D. M. German and A. E. Hassan. License integration patterns: Dealing with licenses mismatches in component-based development. In *28th International Conference on Software Engineering (ICSE '09)*, May 2009.
- [15] W. N. Hohfeld. Some fundamental legal conceptions as applied in judicial reasoning. *Yale Law Journal*, 23(1):16–59, Nov. 1913.
- [16] Institute for Software Research. ArchStudio 4. Technical report, University of California, Irvine, 2006. <http://www.isr.uci.edu/projects/archstudio/>.
- [17] Institute for Software Research. xADL 2.0. Technical report, University of California, Irvine, 2009. <http://www.isr.uci.edu/projects/xarchuci/>.
- [18] S. Jansen, S. Brinkkemper, and A. Finkelstein. Business network management as a survival strategy: A tale of two software ecosystems. In *First Workshop on Software Ecosystems*, pages 34–48, 2009.
- [19] S. Jansen, A. Finkelstein, and S. Brinkkemper. A sense of community: A research agenda for software ecosystems. In *ICSE Companion '09: Companion of the 31st International Conference on Software Engineering*, page 187 190, May 2009.
- [20] C. Jensen and W. Scacchi. Process modeling across the web information infrastructure. *Software Process: Improvement and Practice*, 10(3):255–272, July/Sept. 2005.
- [21] A.-K. Kuehnel. Microsoft, open source and the software ecosystem: of predators and prey—the leopard can change its spots. *Information & Communication Technology Law*, 17(2):107–124, June 2008.
- [22] F. Kuhl, R. Weatherly, and J. Dahmann. *Creating computer simulation systems: an introduction to the high level architecture*. Prentice Hall, 1999.
- [23] N. Medvidovic, D. S. Rosenblum, and R. N. Taylor. A language and environment for architecture-based software development and evolution. In *ICSE '99: Proceedings of the 21st international Conference on Software Engineering*, pages 44–53, 1999.
- [24] B. C. Meyers and P. Oberndorf. *Managing Software Acquisition: Open Systems and COTS Products*. Addison-Wesley Professional, 2001.

- [25] L. Nelson and E. F. Churchill. Repurposing: Techniques for reuse and integration of interactive systems. In *International Conference on Information Reuse and Integration (IRI-08)*, page 490, 2006.
- [26] Open Source Initiative. Open Source Definition, 2008. <http://www.opensource.org/>.
- [27] P. Oreizy. *Open Architecture Software: A Flexible Approach to Decentralized Software Evolution*. PhD thesis, University of California, Irvine, 2000.
- [28] L. Rosen. *Open Source Licensing: Software Freedom and Intellectual Property Law*. Prentice Hall, 2005.
- [29] W. Scacchi. Free/open source software development. In *ESEC/FSE 2007: 6th Joint European Software Engineering Conference and ACM SIGSOFT Symposium on the Foundations of Software Engineering*, pages 459–468, Sept. 2007.
- [30] W. Scacchi and T. A. Alspaugh. Emerging issues in the acquisition of open source software within the U.S. Department of Defense. In *5th Annual Acquisition Research Symposium*, May 2008.
- [31] R. Sen. A strategic analysis of competition between open source and proprietary software. *J. Manage. Inf. Syst.*, 24(1):233–257, 2007.
- [32] R. Sen, C. Subramaniam, and M. L. Nelson. Determinants of the choice of open source software license. *Journal of Management Information Systems*, 25(3):207–240, 2009.
- [33] A. M. St. Laurent. *Understanding Open Source and Free Software Licensing*. O’Reilly Media, Inc., 2004.
- [34] K. J. Stewart, A. P. Ammeter, and L. M. Maruping. Impacts of license choice and organizational sponsorship on user interest and development activity in open source software projects. *Info. Sys. Research*, 17(2):126–144, 2006.
- [35] C. Subramaniam, R. Sen, and M. L. Nelson. Determinants of open source software project success: A longitudinal study. *Decis. Support Syst.*, 46(2):576–585, 2009.
- [36] T. Tuunanen, J. Koskinen, and T. Kärkkäinen. Automated software license analysis. *Automated Software Engineering*, 16(3-4):455–490, 2009.
- [37] Unity Technologies. End User License Agreement, Dec. 2008. <http://unity3d.com/unity/unity-end-user-license-2.x.html>.
- [38] K. Ven and H. Mannaert. Challenges and strategies in the use of open source software by independent software vendors. *Information and Software Technology*, 50(9-10):991–1002, 2008.