

# SPEC KITS

Supporting Effective Library Management for Over 40 Years

Committed to assisting research and academic libraries in the continuous improvement of management systems, ARL has worked since 1970 to gather and disseminate the best practices for library needs. As part of its commitment, ARL maintains an active publications program best known for its SPEC Kits. Through the Collaborative Research/Writing Program, librarians work with ARL staff to design SPEC surveys and write publications. Originally established as an information source for ARL member libraries, the SPEC Kit series has grown to serve the needs of the library community worldwide.

## **What are SPEC Kits?**

Published six times per year, SPEC Kits contain the most valuable, up-to-date information on the latest issues of concern to libraries and librarians today. They are the result of a systematic survey of ARL member libraries on a particular topic related to current practice in the field. Each SPEC Kit contains an executive summary of the survey results; survey questions with tallies and selected comments; the best representative documents from survey participants, such as policies, procedures, handbooks, guidelines, Web sites, records, brochures, and statements; and a selected reading list—both print and online sources—containing the most current literature available on the topic for further study.

## **Subscribe to SPEC Kits**

Subscribers tell us that the information contained in SPEC Kits is valuable to a variety of users, both inside and outside the library. SPEC Kit purchasers use the documentation found in SPEC Kits as a point of departure for research and problem solving because they lend immediate authority to proposals and set standards for designing programs or writing procedure statements. SPEC Kits also function as an important reference tool for library administrators, staff, students, and professionals in allied disciplines who may not have access to this kind of information.

SPEC Kits are available in print and online. The executive summary for each kit after December 1993 can be accessed online free of charge. For more information visit: <http://www.arl.org/publications-resources>.



# SPEC Kit 340

---

Open Source Software  
July 2014

---

**J. Curtis Thacker and Dr. Charles D. Knutson**

---

Brigham Young University

**Mark Dehmlow**

---

University of Notre Dame



ASSOCIATION OF RESEARCH LIBRARIES®

---

Series Editor: Lee Anne George

SPEC Kits are published by the

Association of Research Libraries

21 Dupont Circle, NW, Suite 800

Washington, DC 20036-1118

P (202) 296-2296 F (202) 872-0884

<http://www.arl.org/publications-resources>

[pubs@arl.org](mailto:pubs@arl.org)

ISSN 0160 3582

ISBN 1-59407-921-8 / 978-1-59407-921-4 print

ISBN 1-59407-922-6 / 978-1-59407-922-1 online

Copyright © 2014

This compilation is copyrighted by the Association of Research Libraries. ARL grants blanket permission to reproduce and distribute copies of this work for nonprofit, educational, or library purposes, provided that copies are distributed at or below cost and that ARL, the source, and copyright notice are included on each copy. This permission is in addition to rights of reproduction granted under Sections 107, 108, and other provisions of the US Copyright Act.



The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (R1997) Permanence of Paper for Publications and Documents in Libraries and Archives.

## SURVEY RESULTS

<b>Executive Summary</b> .....	11
<b>Survey Questions and Responses</b> .....	17
<b>Responding Institutions</b> .....	66

## REPRESENTATIVE DOCUMENTS

### OSS Contributor Agreements

University of California	
Guidelines for Contributing to Open/Community Source Software .....	70
DuraSpace	
Contributor License Agreements .....	74
Hydra Project Intellectual Property Licensing and Ownership .....	76
The Apache Software Foundation	
Individual Contributor License Agreement (“Agreement”) V2.0 .....	78

### OSS Licenses

University of Colorado	
Technology Transfer Office. The Software Process .....	82
Indiana University	
Research & Technology Corporation. Process. Software Licensing .....	83
Open Source Initiative	
Open Source Licenses .....	84
The Apache Software Foundation	
Apache License Version 2 .....	86
Free Software Foundation	
GNU General Public License Version 3 .....	90
Open Source Initiative	
The MIT License .....	102

## **OSS Copyright Notices**

DuraSpace	
The Hydra Project Code Copyright Statement.....	104
George Washington University	
Software Copyright Notice .....	106
Georgetown University Libraries	
Software Copyright Notice .....	107
North Carolina State University	
Software Copyright Notice .....	108

## **OSS Adoption Decision**

Auburn University	
Customizing VuFind .....	110

## **Job Descriptions of OSS Contributors**

Auburn University	
Information Technology Specialist IV/V .....	122
North Carolina State University Libraries	
Digital Technologies Development Librarian .....	123
University of Notre Dame	
Sr. Software Engineer .....	124
UI Software Engineer.....	125
Ohio State University	
Systems Developer/Engineer .....	127
University of Rochester	
Java/Web Application Developer .....	130
University of Saskatchewan	
Programmer Analyst .....	132
University of Tennessee	
Digital Initiatives Programmer 1 .....	134
Digital Initiatives Programmer 2 .....	141
Systems Development Librarian .....	148
University of Virginia	
Software Engineer IV.....	150
Systems Engineer V .....	158

## Organization Charts

Brown University Library	
Digital Technologies .....	168
University of California, Irvine Libraries	
Libraries Information Technology .....	169
Duke University Libraries	
Org Chart.....	170
Indiana University Libraries Bloomington	
Library Technologies .....	171
University of New Mexico Libraries	
Overview-UL&LS Organizational Structure .....	172
North Carolina State University Libraries	
Digital Library Initiatives .....	173
Information Technology .....	174
User Experience .....	175
University of Pennsylvania Library	
Penn Libraries Reporting Organization .....	176
University of Rochester Libraries	
River Campus Libraries Information Technology Services .....	177
York University Libraries	
Library Computing Services .....	178

## SELECTED RESOURCES

<b>Books and Journal Articles.....</b>	<b>181</b>
<b>Technology Transfer and Intellectual Property Policies .....</b>	<b>182</b>
<b>Organization Forge URIs.....</b>	<b>183</b>
<b>Project Forge URIs .....</b>	<b>183</b>





**SURVEY RESULTS**





---

## EXECUTIVE SUMMARY

---

### Introduction

Open source software (OSS) “licenses must permit non-exclusive commercial exploitation of the licensed work, must make available the work’s source code, and must permit the creation of derivative works from the work itself.” [St. Laurent, Andrew M. (2008). *Understanding Open Source and Free Software Licensing*. O’Reilly Media, p 8. ISBN 9780596553951].

The emergence of OSS increases collaboration among research libraries, providing greater control of library tools, as well as improving usability and quality of library resources. This collaborative approach fits neatly with the knowledge and resource sharing ideology of libraries. While OSS is ostensibly “free,” adoption of OSS within an organization is not without significant support, integration, and development costs.

The purpose of this survey was to study ARL member libraries’ adoption and/or development of OSS for functions such as an integrated library system (ILS), discovery layer, electronic resource management, inter-library loan, digital asset management, institutional repository, course reserve, streaming media, study room scheduler, digital preservation, publishing, floor maps, data warehouse, and other library-related purposes. We wanted to understand organizational factors that affect decisions to adopt OSS, the cost of OSS, and the awareness of OSS systems already in use. With regard to development of OSS, we wanted to understand: 1) research libraries’ policies and practices on open sourcing their code; 2) the frequency of research library contributions to open source projects; 3) the reluctance of research libraries to make their code openly available; and 4) the

most common benefits and challenges encountered when research libraries open source their code.

This survey was distributed to the 125 ARL member libraries in February 2014. Seventy-seven libraries (62%) responded to the survey by the March 17, 2014 deadline.

### Library IT Staff

The 66 responding academic and public libraries reported between two and 50 staff with IT responsibilities as all or part of their duties, with an average of 16 and a median of 14. Three national libraries reported between 130 and 350 IT staff. This bimodal distribution is stark, with the national libraries an order of magnitude larger than their university counterparts. Despite this difference in staff size, we find no statistically significant differences in the relative participation in OSS projects.

Seventy respondents (91%) develop software in-house. Of those, the most common software development practices include using version control (86%) and performing usability tests (86%). The least common practices include the use of independent quality assurance (24%), adherence to a formal, written code reuse policy (10%), and the presence of a committee or working group to encourage code reuse (7%). The most common other software practices mentioned by respondents were agile/scrum development methodologies (5 of 15 respondents) and pair programming (2 respondents). Most respondents reported that their library IT staff are encouraged to experiment with new technologies (75 or 99%), and prototype potential projects (62 or 82%).

As expected, we found a strong positive correlation between staff size and support for software development best practices (particularly creation of software documentation and specifications, creation of user documentation, performing code reviews, using version control, practicing casual code reuse, and standardizing development by utilizing a common framework).

When asked how users give feedback to IT staff, several findings emerged:

- Library employees most commonly give feedback through a helpdesk or bug tracking system (69 respondents, or 91%) and by emailing or calling the system manager/developer directly (67 or 88%).
- Employees of the parent institution give feedback through a form on the library website (54 or 71%), through subject librarians (44 or 59%), by emailing or calling the system manager/developer directly (39 or 51%), and through a helpdesk or bug tracking system (35 or 46%).
- In-library patrons most commonly give feedback through a form on the library website (59 or 78%) and through subject librarians (58 or 76%).
- Remote users most commonly give feedback through a form on the library website (60 or 79%), and through subject librarian (49 or 64%)

In-library users and remote users most commonly use the same feedback methods, suggesting that proximity to the physical library may not significantly impact feedback channels.

In our review of organizations that contribute to open source projects, software development staff ranged from one or two to as many as 14. While organizations that contribute to large scale, formal open source projects were clearly investing heavily in programming staff, it was also clear that a few organizations who didn't have resources for large technology staffs could still contribute to projects with as few as one programmer. The median number of staff reported as working on OSS projects was two, with an average of nearly four.

Organizational structures varied considerably. Within smaller organizations, single programmers are often located in library systems or web units. Within larger organizations, software development staff are often clustered together in application development units located in digital library, digital projects, or library technology branches of the organization.

## Library Software

The survey asked respondents to provide information about the type of software used for various library purposes. All 76 respondents use one or more vended products, 72 identified types of open source software used by the library, and 50 identified software that was built in-house. Below are some of the highlights of the range of applications being used.

- Fifty-eight respondents (76%) use a vended, locally hosted integrated library system (ILS). No respondents use an ILS built in house, but five use an open source ILS.
- Forty-five respondents (59%) use a vended, locally hosted interlibrary loan (ILL) system and 29 (38%) license a software as a service (SaaS) ILL system.
- Forty-nine respondents (64%) use a SaaS discovery layer, 17 (22%) use a vended, locally hosted discovery layer, and 10 (13%) use a discovery layer that is built in house. Several respondents indicated that their discovery layer was both a vended, locally hosted system and also built in house, suggesting significant customizations to a vended product.
- Forty-seven respondents (62%) use a locally hosted and supported OSS institutional repository.
- Forty respondents (53%) use a locally hosted and supported OSS digital preservation system.
- Fifty-one institutions (67%) have adopted a system that is open source and supported by a third party.
- The most commonly built in-house systems were floor maps (28 respondents) and digital asset management systems (19 respondents).

- The most frequently adopted OSS systems include institutional repositories (52 respondents), blogging (50 respondents), digital preservation (47 respondents), and publishing (40 respondents).

### OSS Adoption

Seventy-four of the 76 responding libraries (97%) report having adopted open source software. Of these, only five (plus one parent institution) have a formal written policy related to adoption of OSS. Twenty-five libraries (34%) have an informal policy, but the other 43 (59%) have no OSS adoption policy. Several respondents reported that policies were currently being created, but could not be shared at the time of their response.

Most respondents indicated their institution had neither a sustainability strategy (50 of 71 respondents, or 70%) nor an exit strategy (53, or 75%). Reported strategies include minimizing customizations, providing sufficient staffing with needed expertise, and only adopting systems with good documentation and an active community. More than half of the respondents who commented on their exit strategy emphasized the criticality of data migration (8 of 15 relevant comments).

Survey respondents were then asked to identify the system they had most recently adopted and to provide the number of staff and hours required to implement that system. A wide variety of projects were identified, the most common being Drupal, Blacklight, Omeka, and DSpace. Respondents reported from one to eight staff members dedicated to implementation, with a mean and median of three staff. The number of hours required for initial implementation varied dramatically, ranging from 0.75 hours to 9,000 hours with a mean of 573 hours and a median of 160 hours.

Respondents were asked to identify the open source system they most recently adopted that is still in production and to describe the resources needed to support that system. For most respondents, the system referred to in this question was the same system described in the implementation question above. The number of staff required to maintain this system ranges from 0 (for a digital exhibit) to 10 (for a CMS)

with a mean of 2.1 and a median of 2. The number of hours required to support the same system ranged from 0 (for the exhibit) to 512 (for a digital repository) per month, with a mean of 68 hours and a median of 20 hours.

Only 16 of 72 respondents (22%) were able to track the costs of either adopting or contributing to an OSS system. Ten respondents who could track the cost of their most recently adopted OSS system reported that expenses ranged from \$400 to over \$600,000 and, in some cases, represented a multi-year investment. These funds covered a variety of expenses including staff time, hosting, travel, and consulting. The nearly universal source of funding for adopting or contributing to an OSS system was the library's operating budget (69 of 70 respondents, or 99%). A few had additional funding from grants, their university, or a consortium. One ArchivesSpace project received only consortium and grant funding.

The survey asked respondents to describe three benefits and three challenges associated with adopting OSS. The most common benefit is the ability to customize the software (50 responses). Other common themes include low cost or time to implement (27 responses) and the association with an active community (27 responses). The most common challenge was the need for highly skilled staff who could provide support for the OSS system (40 responses). Other commonly cited challenges include poor documentation (19 respondents), a need for additional training or expertise (16 respondents), and substandard development practices (12 respondents).

### OSS Contribution

Fifty-six of the responding libraries (78%) have contributed resources to an open source project. The number of projects contributed to by each library ranges from 1 to 20, with an average of 4.6 and a median of 3. Thirty-two libraries report being the primary code contributor for at least one project; a different set of 32 libraries (with significant overlap) identified themselves as the original developer of an open source project.

Commonly reported examples of projects include DSpace (12 respondents), Fedora (11 respondents), Hydra (9 respondents), Kuali (6 respondents),

Blacklight (5 respondents), and ArchivesSpace (4 respondents). Below are some of the highlights of library contributions to the projects.

- The most common contributions involved code or developer time (47 respondents), funding (36 respondents), hosting (36 respondents), and testing (8 respondents).
- Across all types of contributions, the most common types of projects included institutional repositories (38 respondents), digital preservation (30 respondents), digital asset management (22 respondents), discovery layer (15 respondents), publishing (13 respondents), authentication/identity management (10 respondents), and electronic resource management (10 respondents).
- Code was most commonly contributed to projects on institutional repositories (32 respondents), digital preservation (22 respondents), digital asset management (20 respondents), and discovery layers (11 respondents).
- Digital preservation and institutional repository projects most often received funding via monetary contributions (19 and 18 respondents, respectively), followed by digital asset management projects (8 respondents).
- Hosting was contributed most often to digital preservation projects (9 respondents), followed by repository and publishing projects (5 respondents each).

When asked about reasons for open sourcing their project, respondents listed the following as being “important” or “very important”: a belief that open sourcing would lead to better software (30 respondents), a desire to contribute to an open source community (29 respondents), and shared effort in development and quality assurance of the project (27 respondents).

Sixty respondents (78%) develop plugins, extensions, or customizations for a library-related proprietary or vended system. Of these, 31 (54%) indicated vendors allowed them to distribute the code under an open source license.

As was the case with OSS adoption policies, 44 respondents indicated their library has no policy in

place for contribution to open source projects, while 20 respondents have an informal policy. Thirty-four respondents stated that they have no tech transfer policy, while 23 respondents indicated that their parent institution has a formal, written tech transfer policy.

Respondents were asked to describe three benefits and three challenges associated with **contributing** to OSS. The benefit most commonly cited was engagement in the open source community (38 respondents). Other common themes included control of product features and direction (25 respondents), and recognition/reputation (14 respondents). The most common challenge was allocating sufficient staff time to make meaningful contributions (24 respondents). Other commonly cited challenges included writing generalized software for use by a larger community, and securing the financial resources needed to support the open source project and community (7 respondents each).

Since open source project members are rarely collocated, a variety of tools have been employed to help coordinate development efforts. Common tools used include shared version control (37 of 45 respondents, or 82%), an issue tracker (36 or 80%), a mailing list, (32 or 71%), and a wiki (25 or 56%). Forty-one respondents (79%) use a public repository or forge to share their open source code; Github was by far the most common (38 of 41 respondents, or 93%).

The most common licenses used by respondents were GNU Public License v3 (16 respondents), Apache, and Creative Commons (15 responses each).

Respondents were asked to rank a set of success indicators in terms of their importance for the respondent’s institution. A significant number (41 or 80%) identified as most important that the functionality better suits their institution’s needs.

Respondents were asked if any of their in-house software could have been, but has not yet been, released under an open source license. The 53 respondents (69%) who answered in the affirmative expressed concerns about the staff time commitment required to support the community (41 or 77%), the readiness of code quality for public adoption (39 or 74%), and dependence on other internal systems (30 or 57%).

## Conclusion

This survey reveals that nearly all of the responding ARL member libraries are developing custom software and/or adopting one or more open source systems. Contribution to OSS projects is also common, with more than three quarters of respondents actively contributing to OSS projects.

Many respondents expressed a desire on the part of their developers to share with and participate in one or more OSS communities. Larger organizations committed more resources to OSS projects than smaller organizations, but we found no significant correlations suggesting a disproportionate level of commitment to OSS projects as a function of IT staff size. The nearly universal adoption of OSS systems and the high level of contribution to OSS projects may suggest that adoption of and contribution to OSS projects has entered the mainstream for libraries. Simply stated, libraries that develop software also predominantly contribute to OSS projects.

The results of this survey suggest that libraries view organizational behaviors surrounding the adoption of open source software separate from

contribution to OSS projects. For example, while respondents view OSS adoption as a means of saving time and resources, contributing to OSS projects is viewed as being advantageous for different reasons, primarily engagement in an OSS community. For developers, the sense of social involvement in a community represented by an OSS project can be a positive source of professional satisfaction, ultimately leading to greater productivity and a return on investment for the library.

Control of software emerged as a theme common to both adopting and contributing to OSS projects. Those adopting OSS systems felt that access to source code gave them greater control, allowing them to change the software as needed, rather than being subject to the whims of a proprietary solution. Those that contributed to OSS projects felt that they gained greater opportunity to influence product direction, especially with respect to product features. In both cases, they perceived a sufficient benefit to their overall productivity to justify the expense of their involvement (as adopters, contributors, or both) in OSS systems.