Kansas Library Association College and University Libraries Section Proceedings

Volume 6
Number 1 *Open Access, Open Borders:
Networking with Colleagues across State Lines*

Article 6

2016

Workflow Management System: A Customized Solution for Project Automation and Tracking

Amanda Harlan
Kansas State University, aharlan@ksu.edu

Dhanushka Samarakoon
Kansas State University, dhanu@ksu.edu

Follow this and additional works at: <https://newprairiepress.org/culsproceedings>

Part of the [Computer and Systems Architecture Commons](#), [Digital Communications and Networking Commons](#), [Management Information Systems Commons](#), and the [Scholarly Communication Commons](#)

This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](#)

Recommended Citation

Harlan, Amanda and Samarakoon, Dhanushka (2016) "Workflow Management System: A Customized Solution for Project Automation and Tracking," *Kansas Library Association College and University Libraries Section Proceedings*: Vol. 6: No. 1. <https://doi.org/10.4148/2160-942X.1057>

This Article is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Kansas Library Association College and University Libraries Section Proceedings by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Workflow Management System: A Customized Solution for Project Automation and Tracking

Abstract

K-State Libraries have many complex digital workflows that involve multiple people and departments. One, in particular, had the ongoing issues of tracking different stages, individual responsibilities, information gathering, notifications, and uploads done manually. Due to the inefficiency that this caused, it led to the development of an open source workflow management system (WMS)

Keywords

Workflow Management Systems, Open Source, Automation

Workflow Management System: A Customized Solution for Project Automation and Tracking

Amanda Harlan

Assistant Professor, Metadata Librarian

Kansas State University

Dhanushka Samarakoon

Programmer

Kansas State University

Abstract

K-State Libraries have many complex digital workflows that involve multiple people and departments. One, in particular, had the ongoing issues of tracking different stages, individual responsibilities, information gathering, notifications, and uploads done manually. Due to the inefficiency that this caused, it led to the development of an open source workflow management system (WMS).

Literature Review

Workflow management systems have been used in industries for decades to manage various workflows, such as manufacturing and supply chains (Reijers, Vanderfeesten, & van der Aalst, 2016). They are computer systems that manage available resources and their dependencies by processing activities executed by processing entities to reach a targeted objective (Sanghamitra, 2003).

Reijers, Vanderfeesten, & van der Aalst's (2016) case study investigated the efficiency of workflow management systems in corporate organizations. The results showed a reduction in processing times when a workflow management system was introduced. Most workflow management systems are constructed out of existing processes and systems (Shrivastava & Wheeler, 1999).

Mainly focused on workflow management systems related to corporate organizations, few publications focused on workflow management systems and academic libraries. One exception is the University of Utah who developed a workflow management system, University Scholarly Knowledge Inventory System (U-SKIS). U-SKIS assisted staff to keep "track of files, communications, and publishers' archiving policies to determine inclusion in their institutional repository (IR) USPACE" (Morrow & Mower, 2009, p. 286).

Kansas State University (K-State) experienced the same issues as the University of Utah. The workflow that the University of Utah created for ingesting the university's scholarly output into their IR was the impetus to create an in-house system as it was for K-State to manage the various nuances of this type of workflow.

Background

In the early 2000s K-State Libraries decided to expand the scope of K-REx, K-State's IR, to include all K-State scholarship in addition to electronic theses and dissertations (ETDRs). With the scope of K-REx expanded, the Libraries' decided to offer a new service of depositing K-State's scholarly output into K-REx.

With the establishment of this service, a workflow was created that involved multiple systems and library staff from different departments. This workflow was time-intensive a large part due to the manual nature of it. Below are the workflow steps for depositing scholarly output into K-REx.

- Requests to authors for permission to deposit their works in K-REx;
- Citation information pulled from RefWorks;
- Additional metadata added to the citation information;
- Publishers identified, and their policies concerning self-archiving checked;
- Requests to publishers asking their permission to deposit works done by K-State authors into K-REx when applicable;
- Requests to authors asking for their manuscripts to be deposited into K-REx when applicable;
- Embargos added when applicable;
- Quality assurance of metadata and files checked; and
- Files and metadata were submitted to the IR; and
- K-State authors notified that their publication(s) were available in K-REx.

As this service grew and the complexity of the workflow above increased, it became evident that the workflow was not sustainable. It consumed too much staff time and the volume of works published by K-State authors every month was increasing. The concept of a home-grown workflow management system was born.

Workflow Management System

In February 2011, the Library IT team met with the K-REx manager and other involved library staff to discuss the feasibility of automating the workflow above as much as possible. Below were some of the main issues that the existing workflow created.

- The collecting of citations needs to be expedited and made scalable for mass harvesting;
- The tracking of tasks and who they are assigned to needs to be centralized in one system;
- The accessibility of publisher policies needs to improve so that they can be cross-referenced from tasks easily;
- The communication between library staff and K-State authors needs to be centralized for better management and tracking purposes;
- The creation of a comprehensive reporting system is needed to assess and measure the efficiency of the workflow.

A Web application called Workflow Management System (WMS) was developed to manage from start to end the workflow of depositing K-State scholarly output into K-REx.

System Infrastructure

K-State Libraries server¹ infrastructure consists fully of an array of Sun servers using a Solaris operating system. The WMS needed to be compatible with this hardware and operating system. An Apache web server² was selected, because it ran on a variety of operating systems, Solaris one of them. In November 2015 it was estimated that 50% of all websites were being served by an Apache web server (Netcraft, 2015). The scripting language³ PHP would be utilized since it was widely used at K-State Libraries for some projects and well supported on an Apache web server. For the backend database management system, MySQL was selected. After careful consideration of multiple PHP frameworks⁴, it was decided that CodeIgniter was going to be the framework for the WMS for the following reasons:

- Its lean approach and performance; and
- Ability to isolate each part of a system and make modifications with minimal effects to system

CodeIgniter includes only the bare minimum set of libraries⁵ that is needed to get a system up and running, and also provides code level application security through its security features. Given that this system was only intended for internal use only, it was deployed on a server that was exclusively used by staff.

Development

The development of the WMS followed an iterative and incremental⁶ software development model. Small sections of the system were implemented in each cycle and presented to the stakeholders for feedback. After feedback, the system was improved, and more requirements implemented with the final product going live in April 2011. With the system in production, stakeholders requested additional features created through multiple development iterations.

System Overview

Task creation. Tasks can be created in two ways: manually through the system (Figure 1) or automatically through bulk import.

The screenshot shows the 'Workflow Management System' interface. At the top, there is a navigation bar with links: 'Tasks', 'Publishers', 'Reports', 'RefWorks', and 'Log Out'. The main content area is divided into three panels. The left panel, titled 'Bibliographic Data', contains various input fields for publication information. The middle panel, titled 'Processing Notes', contains fields for user selection, status, and other task-related details. The right panel, titled 'Abstract/Keywords', contains fields for the abstract and keywords. A green 'Save Task' button is positioned at the bottom of the middle panel.

Figure 1. Form that can be filled out with publication information to create a task.

The bulk import method imports citations from RefWorks creating tasks in the WMS. Library staff searched and collected relevant citations and added additional metadata in RefWorks then downloaded the citation information in an XML file (Figure 2).

Figure 2. RefWorks has an option to export citation information into an XML file.

The XML file then was uploaded to the WMS ready for processing. One or more files could be selected, and clicking on the “Create Tasks” button triggered the bulk creation of tasks (Figure 3).

Figure 3. List of XML files that can be selected to create tasks in bulk.

Task processing. The newly created tasks located on the WMS homepage (Figure 4) contained the following information: task ID, author name, publication title, task status, system user, date initiated, and date last updated. By default, new tasks always appear first, but a system user can sort by status, user, date initiated, and date last updated. A user can also search tasks by task ID, author name, and title.

Workflow Management System							
Tasks - Publishers - Reports - RefWorks - Log Out							
10	records per page			Search:			
ID	KSU Contact	Title	Status	User	Initiated On	Last Updated	
4013	Xia, Tian tianxia	U.S. implications of the Smithfield acquisition by Shuanghui.	New Upload	None	2015-05-11	2015-05-11	
4012	Miller, Chad T. ctmiller	Using google maps web-application to create virtual plant maps for use as an online study tool in plant identification courses	New Upload	None	2015-05-11	2015-05-11	
4011	Wietzel, David L. dwietzel	Forensic Spectroscopic Chemical Fingerprinting of Fingerprints	New Upload	None	2015-05-11	2015-05-11	
4010	Rhodes, Davina H. davinahrhodes	Persistent organic pollutants and biomarkers of diabetes risk in a cohort of Great Lakes sport caught fish consumers	New Upload	None	2015-05-11	2015-05-11	
4009	Riding, Kyle A. kriding	Calcined kaolinite-bentonite clay blends as supplementary cementitious materials	New Upload	None	2015-05-11	2015-05-11	
4008		Prevalence, diversity and characterization of enterococci from three coraciform birds	New Upload	None	2015-05-11	2015-05-11	
4007	Poland, Jesse A. jpoland	Efficient Use of Historical Data for Genomic Selection: A Case Study of Stem Rust Resistance in Wheat	New Upload	None	2015-05-11	2015-05-11	
4006	Bouchet, Sophie sbouchet	Dent and Flint maize diversity panels reveal important genetic potential for increasing biomass production	New Upload	None	2015-05-11	2015-05-11	
4005		Spiked GBS: A unified, open platform for single marker genotyping and whole-genome profiling	New Upload	None	2015-05-11	2015-05-11	
4004		Domestic Pigs Are Susceptible to Infection with Influenza B Viruses	New Upload	None	2015-05-11	2015-05-11	
4003	Shi, Yong Cheng ycshi	Formation of vitamin E emulsion stabilized by octenylsuccinic starch: factors affecting particle size and oil load.	New Upload	None	2015-05-11	2015-05-11	
4002	Paul, Gregory D. gregpaul	Predicting participation in a victim-offender conference	New Upload	None	2015-05-11	2015-05-11	
4001	Kukanich, Butch P. kukanich	Pharmacokinetics of meloxicam in mature swine after intravenous and oral administration	New Upload	None	2015-05-11	2015-05-11	
4000		High prevalence of "Candidatus Rickettsia andeanae" and apparent exclusion of rickettsia parkeri in adult Amblyomma maculatum (Acar: Ixodidae) from Kansas and Oklahoma	New Upload	None	2015-05-11	2015-05-11	
3999		Pharmacokinetics of oral amantadine in greyhound dogs	New Upload	None	2015-05-11	2015-05-11	

Showing 1 to 15 of 592 entries

Previous 1 2 3 4 5 Next

Figure 4. List of tasks on the WMS' homepage that is ready to be processed.

To edit a task, a system user needs to click on the left blue button (Figure 4). On the task edit page, several workflow actions can be accomplished or tracked (Figure 5).

Bibliographic Data Title: U.S. implications of the Smithfield acquisition by Shuanghui Authors: Xia, T. Journal Title: Choices: The Magazine of Food, Farm, and Resources Book Title: Conference Title: Publisher: American Agricultural Economics Association Volume: 30 Issue: 1 Start Page: unassigned End Page: Article ID: Publication Year: 2015 DOI: ISSN / ISBN: 0896-5558 Type: Article (author version) Language:	Processing Notes KSU Contact: Xia, Tian tianxia KSU Faculty/Staff Authors and eIDs: Xia, Tian tianxia KSU Student Author eIDs: Status: New Upload Assigned User: None Notes: Collection: None Published Citation: Manuscript Citation: Copyright Statement: Publishers Link: Embargo Lift Date: FileNames: <input type="button" value="Save Task"/> Inserted Via: RefWorks Task Created On: 2015-05-11 Last Updated On: 2015-05-11	Abstract/Keywords Abstract: Keywords: Communication <input type="button" value="Compose Email"/>
--	--	--

Figure 5. Task form where information is added or edited.

Tracking workflow actions is done by selecting from a drop-down menu a task status (Figure 6). Each status represents a specific part of the workflow. A task will stay open until all workflow actions are completed and the task is closed. The WMS retains all closed tasks, which system users can sort or search mainly for troubleshooting reasons.

Field	Value
Status	New Upload
Assigned User	Assigned
Notes	Assigned to Metadata Team
Collection	Standing Permission
Published Citation	Closed - deposited in K-REx

Figure 6. List of statuses a task is assigned during its life cycle.

Publisher information. To expedite the task processing activity, the WMS includes publisher information for ease of policy checking. Publisher information can enter the system via two methods: manually using a form (Figure 7) or automatically during the RefWorks bulk import process.

Field	Value
Publisher Name	
Policy Link	
Policy Text	
What we can put up	
What we need to add	
Embargo	
Notes	

Figure 7. Publisher form that is filled out with information about the publisher's self-archiving policy.

During the bulk import process, a script looks for publisher names and cross-references those names with a master publisher table (Figure 8) in the WMS. Existing publisher records are linked to a task, and new records are created for publishers not found in the master table. The creation of new publisher records are blank, so information about the publisher's self-archiving policy will need to be added to make the record complete.

ID	Name	View	Edit	Remove
0	Select a Publisher			
637	A V S AMER INST PHYSICS			
511	AACC International			
35	Academia Press			
632	Academic Press			
619	Academic Press Inc.			
537	AChe/FoTNS			

Mark all selected publisher names as alternate names for the following publisher:
 Academia Press
 Merge

1. Tasks under alternate names would be automatically merged.
2. Alternate names cannot contain any meta data.
3. Alternate name should not have been marked as a master name previously.

Figure 8. Master table (list) of publishers and their self-archiving policies.

To avoid the creation of duplicate records for the same publisher during the bulk import process, a system user selects the preferred publisher name from a drop-down list and marks its name variations, which triggers a merge sequence in the WMS. The system moves the name variations to the master publisher record and assigns the tasks associated with the name variations to the master record. The next time the RefWorks bulk import script encounters existing name variations, they will be replaced with their preferred publisher names that are accessible from a drop-down list on the task edit page.

Statistics and reports. To track and assess tasks, two reports are available in the WMS: completion tasks and filter by status. Completion tasks report produces success/fail numbers of manually and bulk import created tasks within a given date range (Figure 9). Included in the report is the number of items deposited, denied permission, with no manuscript, or not deposited for other reasons. This report allows stakeholders to assess patterns, and make strategic decisions concerning the workflow.

Start Date:		End Date:		Generate Report
-------------	--	-----------	--	-----------------

Report from 2013-01-01 to 2013-05-20					
--------------------------------------	--	--	--	--	--

Tasks created by Form		
Month	Year	Tasks
May	2013	41
April	2013	37
March	2013	30
February	2013	31
January	2013	17

Closed Tasks created by Form					
Year-Month	Deposited	Denied	No Manuscript	Other	Total
2013 - May	17	2	6	0	25
2013 - April	21	1	1	0	23
2013 - March	33	4	0	4	41
2013 - February	7	1	1	2	11
2013 - January	12	0	15	3	30

Tasks created by Refworks		
Month	Year	Tasks
May	2013	25
April	2013	79
March	2013	79
February	2013	20
January	2013	52

Closed Tasks created by Refworks					
Year-Month	Deposited	Denied	No Manuscript	Other	Total
2013 - May	16	3	22	1	42
2013 - April	53	2	7	16	78
2013 - March	27	3	1	8	39
2013 - February	24	4	2	1	31
2013 - January	17	1	20	5	43

Figure 9. Report that lists the number of created and closed tasks in a given month along with the number of tasks deposited or not deposited.

The filter by status report (Figure 10) produces a list of dormant tasks associated with specific statuses and not updated since selected date. This information allows stakeholders to decide whether to revive these tasks or close them. Additionally, an internal reminder system sends periodic emails containing tasks in a given status not updated since a predefined number of days to an administrator. A cron job⁷ can be configured to trigger this functionality.

ID	KSU Contact	Title	Status	User	Initiated On	Last Updated	
3496	White, Bradley J. whiteb	Effect of ambient temperature on viral replication and serum antibody titers following administration of a commercial intranasal modified-live infectious bovine rhinotracheitis/parainfluenza-3 virus vaccine to beef cattle housed in high- and moderate-ambient temperature environments	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-12-11	2015-04-13	✓
3484	Baker, Luan M. lbaker	The new agent: A qualitative study to strategically adapt new agent professional development	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-12-11	2015-04-08	✓
3405	Phromm, Peter H. phromm	Analysis of Atomic Force Microscopy Phase Data To Dynamically Detect Adsorbed Hydrogen under Ambient Conditions.	Contacted Publisher, Awaiting feedback	Amanda Harlan	2016-10-31	2015-05-08	✓
3333	Reinhardt, Christopher D. cd3	Feedlot nutritionist Boot Camp: An intensive short-course for commercial agriculture graduate students	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-10-13	2014-12-03	✓
3328	Prasad, Vani vprasad	Characterization of a spring wheat association mapping panel for root traits.	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-10-13	2014-11-11	✓
3327	Procter, Sandra B. sprocter	Positive quality of life factors identified from EFNEP participant stories	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-10-13	2014-11-11	✓
3312		Influence of nitrogen fertilizer on growth and yield of grain sorghum hybrids and inbred lines	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-10-01	2014-11-03	✓
3302	Thomson, Daniel U. dthomson	Perception of lameness management, education, and effects on animal welfare of feedlot cattle by consulting nutritionists, veterinarians, and feedlot managers.	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-09-29	2014-10-10	✓
3290	Thomson, Daniel U. dthomson	Implementation of industry-oriented animal welfare and quality assurance assessment in Kansas cattle feeding operations.	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-09-29	2014-10-10	✓
3285		Staphylococcus aureus secretes a unique class of neutrophil serine protease inhibitors	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-09-24	2014-09-30	✓
3244	Assafa Mulisa, Yared yareda	Winter canola yield and survival as a function of environment, genetics, and management	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-09-12	2014-10-15	✓
3232	Riding, Kyle A. kriding	Diffusivity of Binary and Ternary Concrete Mixture Blends	Contacted Publisher, Awaiting feedback	Marty Courtois	2014-09-05	2014-09-29	✓

Figure 10. A report that lists tasks not updated by a given date.

Communication management. The WMS allows system users to initiate communication through email with either authors or publishers within the WMS through the task edit page by selecting the “compose email” button located at the bottom of the page (Figure 11).

Figure 11. The “compose email” allows a system user to select a predefined email template.

The WMS opens a blank email pre-populated with author email addresses, and the system user can select from a drop-down list an email subject and predefined email template triggering the body of the email to be pre-populated with coined text associated with the selected email template that can be modified (Figure 12) before sending the email. The system maintains a log of all sent emails for troubleshooting and management reasons.

U.S. implications of the Smithfield acquisition by Shuanghui.

To

Cc

BCc

Subject

Body

Dear [Publisher]:

On behalf of the authors, I would like to request permission to archive in the K-State Research Exchange the [final published version] [author's accepted manuscript] of this article:

The K-State Research Exchange (K-Rex) is the institutional repository of Kansas State University and contains scholarly work of our faculty and students, including faculty publications and student theses. Items in K-REx are freely available on the web at <http://krex.ksu.edu>.

Figure 12. Predefined email template with coined text for requesting permission from a publisher to deposit a publication Into K-REx.

Conclusion

At the time when the WMS went into production in 2011, it alleviated most of the manual processing especially

- processing new publications;
- checking publisher policies;
- tracking system users' actions;
- reporting of completed and dormant tasks
- communicating with authors and publishers; and
- adding additional information in RefWorks

which allowed for more time spent on outreach publicizing the service to the K-State community.

Even though some parts streamlined the workflow in the WMS, it did not decrease the time spent

- waiting for author and publisher responses
- managing files (i.e., manuscripts, published versions, etc.)
- uploading files and metadata to K-REx;
- adding new publisher policies and updating existing policies;
- creating cover pages for manuscripts;
- addressing table and figure issues in manuscripts; and
- checking name authorities of authors.

Due to the nature of some issues, the WMS could not address them. Others, though, could be addressed and were in the queue for future development.

The WMS is not being utilized to manage the workflow of depositing K-State scholarly output into K-REx, because a more efficient process involving scripts and batch loading automates parts of the workflow. Though the WMS is not active anymore, the flexibility of the system allows for further development without significant impact to the existing code base. Two examples of future developments are the automation of attaching files to tasks through data harvesting, and automating the upload of files and metadata from the WMS to a designated platform with the SWORD protocol⁸.

Since the WMS' inception in 2011, there have been several development cycles that have added more features, demonstrating the versatility of the system. More importantly, the adaptability of the WMS provides opportunities to repurpose it for other uses (i.e., a workflow for local name authorities), making it more appealing to a broader audience. If there is a need to manage a complex workflow with automation options, this system is a viable solution.

References

- Morrow, A., & Mower, A. (2009). University Scholarly Knowledge Inventory System: A workflow system for institutional repositories. *Cataloging & Classification Quarterly*, 47, 286-296. <http://dx.doi.org/10.1080/01639370902737307>
- Netcraft. (2015, November 16). *November 2015 Web Server Survey*. Retrieved from <http://news.netcraft.com/archives/2015/11/16/november-2015-web-server-survey.html>
- Reijers, H. A., Vanderfeesten, I., & van der Aalst, W.M.P. (2016). The effectiveness of workflow management systems: A longitudinal study. *International Journal of Information Management*, 36, 126–141. <http://dx.doi.org/10.1016/j.ijinfomgt.2015.08.003>
- Sanghamitra, D. (2003). *Workflow management systems* (Unpublished doctoral dissertation). Kansas State University, Manhattan, KS.
- Shrivastava, S. K., & Wheeler, S. M. (1999). Guest editor's introduction - Workflow management systems. *IEEE Concurrency*, 7(3), 16-17. <http://dx.doi.org/10.1109/MCC.1999.788775>

Footnotes

¹ Server refers to the physical hardware that an operating system is installed on.

² Web server operates on an operating system and contains the website itself. It is responsible for accepting requests from users and serving the correct web pages.

³ Scripting language is a server-side programming language used to code the website and the database management system (DBMS) that stores user data.

⁴ Framework is a collection of programs that one can use to develop applications.

⁵ A library is a collection of commonly used functions, compiled and stored together, so that other programs can access and use them without having to recreate the logic.

⁶ Iterative and incremental software development model includes the creation of a working model based on the initial user requirements. This initial model goes through multiple iterations of improvements based on continuous user feedback.

⁷ Cron jobs are a type of scheduled routine tasks that one can ask the server to perform at predetermined intervals.

⁸ SWORD (Simple Web-service Offering Repository Deposit) is a protocol that allows the deposit of content into digital repositories, from multiple sources through an application programming interface.