Purdue University

Purdue e-Pubs

License Management: Making It Fun and Flexible with CORAL

Andrea Langhurst
University of Notre Dame, alanghurst1@gmail.com

Xan Arch
Reed College, archx@reed.edu

LICENSE MANAGEMENT: MAKING IT FUN AND FLEXIBLE WITH CORAL

Andrea Langhurst (langhurst.1@nd.edu) *Licensing/Acquisitions Librarian, University of Notre Dame*

Xan Arch (xanadu@reed.edu) *Collection Development Librarian, Reed College*

ABSTRACT

Do you have license agreements with publishers languishing in your file cabinets? Or have you implemented an ERMS but struggled to manage the overload of data? The University of Notre Dame, Hesburgh Libraries, faced these challenges and decided to build their own solution. Their CORAL product is a suite of interoperable and independent modules designed around the core components of managing electronic resources, and the first module to be rolled out for external use was Licensing. The Licensing module was released to the public under a GNU GPL license in April 2010 and the first institution to implement outside of Notre Dame was Stanford University. This session will compare and contrast two institutions experiences implementing and using CORAL-Licensing. Hear how each institution initiated implementation of CORAL, learn how they determined which licensing terms to track, and find out how they kept the process of electronic resource management flexible, manageable and (almost) fun! We will discuss the adoption process, including internal communication and decision-making as well as hurdles and successes of CORAL-Licensing. For more information on CORAL: http://erm.library.nd.edu/.

INTRODUCTION

With the ever growing number of electronic resources in libraries, the ability of the institution to effectively manage the license agreements associated with the orders becomes more and more vital. The University of Notre Dame, Hesburgh Libraries, made the commitment to build an Electronic Resource Management System (ERMS), called CORAL (Centralized Online Resources Acquisitions and Licensing) in 2009. The CORAL-Licensing module was the first to be released as an open source product, with Stanford University the first institution to install and implement the program in early 2010. While using the same product, both Notre Dame and Stanford made unique decisions based on institutional needs for license management. Making decisions based on what pieces of information each institution wanted to track, and how they wanted to track and use the information, both Notre Dame and Stanford learned valuable lessons about choosing, implementing and using an ERMS.

NOTRE DAME

The University of Notre Dame, Hesburgh Libraries, decided to undertake a project to build a system to assist in the management of electronic resources after extensive review of available ERMS solutions as well as a review of industry standards for managing electronic resources. Initially we had looked at vended, third-party solutions available, but determined that by building our own ERM product, CORAL, we could build in functionality that could make the management of electronic resources more robust to meet the needs of Notre Dame as an individual institution

and at the same flexible enough to be changed to incorporate the needs of different institutions as well. The Licensing Module was the first piece of CORAL to be completed and released for public implementation under a GNU GPL license, in April 2010.

There are many parts of the CORAL-Licensing module that I appreciate as Licensing/Acquisitions Librarian at Notre Dame. Previously our license management workflow involved keeping scanned agreements in a departmental share drive, with a very small number of older paper agreements stored in file cabinets; we had spreadsheets to assist in document management as well. Building CORAL-Licensing gave us the opportunity to stop and think about how we license library resources. We considered things like what kinds of license documents were we signing and how the various documents related to each other. Asking ourselves, how we would want to use a search interface to find such a variety of documents and how we wanted to be able to use the pieces of information recorded in license agreements. Our goal was to build a tool that wouldn't require us to add more information than we wanted to, but instead easily allowed us to select what information we wanted to manage. We wanted to do this not just because we were able to make these decisions while developing our own ERMS, but also in order to produce a system that could be implemented and used by non-developers; to share the flexibility of CORAL with others.

A screenshot of how CORAL helps organize our Elsevier Journal agreements. The main or master agreement is at the top, with amendments or child documents easily viewable. Archived documents are also accessible, anticipating that while new agreements can be signed, the older agreements might need to be referenced in the future.

agreement records as attachments (instead of keeping related documents separate, often in a single person's email). In terms of agreement types, we started with the more common document types of Agreement (requiring just a Notre Dame signature) or Countersigned Agreement (requiring signatures from multiple parties). Additionally CORAL's Licensing Module allows the institution the ability to identify its own list of agreement types and statuses, like many other of the information fields tracked in CORAL. In addition to types like Agreement and Countersigned Agreement, we created a type, SERU, to begin to track how often we were able to reference SERU terms when licensing resources, without looking back through emails to find previous reference of the information. Similarly, we decided to also have a type called Order Form, often a one or two page document list basic transaction details, requiring a signature but invoking earlier negotiated terms – especially useful for tracking items like the many EBSCO and ProQuest Order Forms we process.
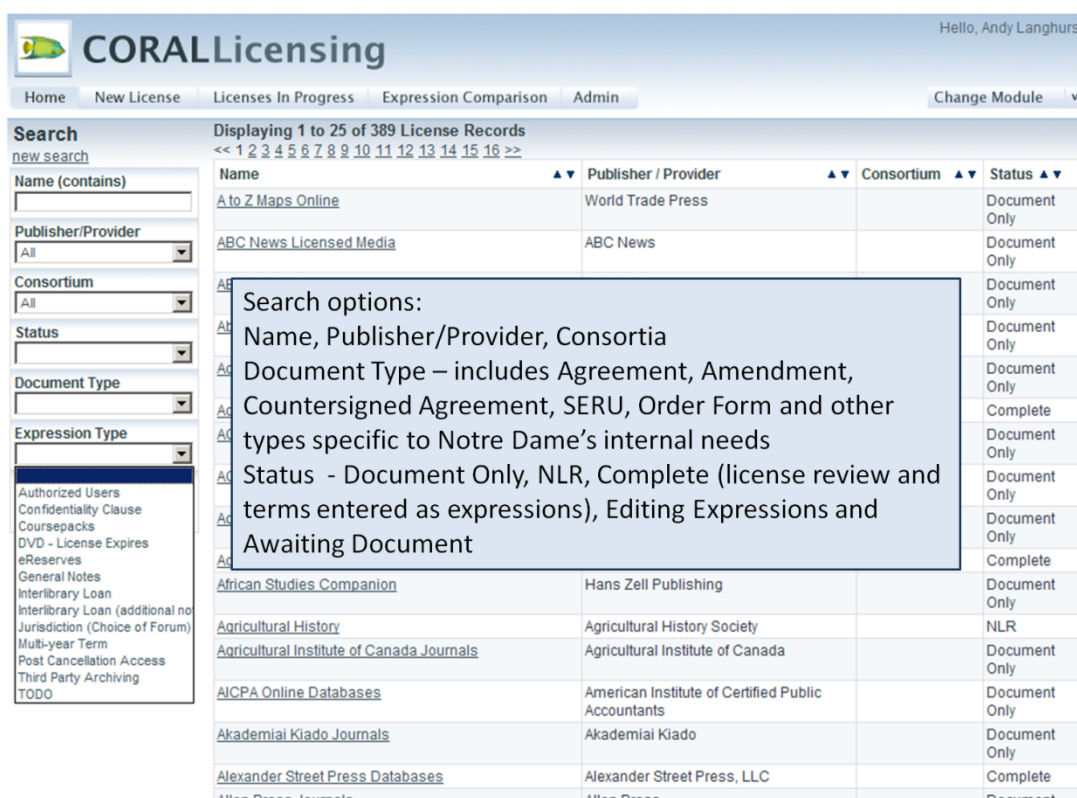
A screenshot of the initial starting screen for CORAL-Licensing. Users will see agreements listed alphabetically and have the option to search by a variety of fields based on the institution's own configuration of the system.

While I knew I wanted to track the types of licensing arrangements we were entering into, we also began to realize that there were additional pieces of information we wanted to leverage. Another benefit to building our own system was the ability to decide that we would associate a status with each particular license record. A status of Document Only, Complete, Awaiting Document, or Editing, would help us to continue to manage the license management workflow going forward. Building CORAL-Licensing also presented an opportunity to better track

situations where we were informed that a formal license was not required - we created a status of NLR, or No License Required. I knew we were often receiving emails from providers that stated no license terms existed or were required, but we had no way to record how often this was happening, outside a spreadsheet, and no way to associate the email in a place many people would be able to easily find it later on. The example of our NLR status demonstrates the flexibility of a creating a status to meet an unmet need for our particular institution with the functionality to choose to store additional information with a license record. CORAL-Licensing has given us the ability to associate a product with an email attachment and to record it in a way that many people can see both pieces of information.   The attachment feature can be used to store items like copies of email negotiations, invoices specifically referenced in agreements, PDFs of relevant webpages, or internal communications specific to the individual license record as well.

One area for improvement we saw from the systems we had looked at previously was the inflexibility in working with license clauses or what we called Expressions. We built the licensing module to be flexible in how many clauses we wanted to track (for Notre Dame, not many) and in how much information could be included in each text field (unlimited). In going through the steps of evaluating existing agreement to identify goals and requirements for our ERMS, we began to see (like many have realized) that using an ERMS for license management could actually add more work. It was important to me to try and prevent the license management process and system from become overwhelming and unsustainable – so we began to ask questions about what we needed to track and store, to try to keep our plan for managing the documents from getting out of control. We decided to only track a short list of expressions, including definitions and information regarding Authorized Users, Coursepack, eReserve and ILL use, Post Cancellation Access and Third-Party Archiving, among others. CORAL-Licensing provided a variety of ways of using the Expression information, for example through a Terms Tool to push ILL and coursepack information outside CORAL and an Expression Comparison feature, to look for trends in particular expressions/clauses. The goal of our Terms Tool is to assist in providing ILL and coursepacks specifics to authorized Notre Dame users for ILL & coursepacks decision-making. The ability to "push" reviewed/approved license terms to external users through our terms tool with our ILL terms tool for example, helps our Hesburgh Libraries ILL team in decision-making by displaying ILL language and allows the institution to take advantage of the permissions granted to us in existing license agreements. The Expression Comparison feature allows CORAL users to look at language by expression type – giving me a better idea of what sort of language we have been agreeing to across many agreements. When negotiating new license agreements for example, the Expression Comparison feature allows me to quickly identify ideal language from past agreements, to be able to request that language when needed in new negotiations. CORAL is helping us to take better advantage of the permissions granted to us in license agreements – allowing us to get a better idea of what we've agreed to in the past, and making that information more widely available to interested internal users.

A screenshot of the initial expressions view for an individual license agreement.

Notre Dame's experience building CORAL, with additional modules of Organizations, Cancellations, Usage Statistics and Resources (an acquisitions module currently in development), has been an extraordinary lesson in the importance of understanding the data when trying to manage it. CORAL-Licensing has given me a way to internally communicate information about our existing license agreements and a place for people to go to see old and new agreements, once they are loaded. Beyond being a better way to manage and then later find agreements, the Licensing module is giving Notre Dame the ability to better appreciate the similarities and differences among our various license agreements and to search and manage the over 600 PDF agreements now loaded in Notre Dame's instance of the CORAL-Licensing module. Attempting to build a licensing module with enough flexibility to change as needed is a great lesson in prioritizing needs and balancing an idea of "oh, wouldn't that be nice" with a question of "is it a requirement and are we committed to tracking it in that way". CORAL-Licensing has been a rewarding experience in taking an active role in deciding how we want to manage electronic resources in the future, in a way that makes managing the license agreements (almost) fun!

STANFORD

The Stanford University Libraries purchased an ERMS when this type of tool first came into the market, but we found it difficult to use and incompatible with our ILS. As many other libraries have experienced, the time and staffing needed for data entry was not worth the end result so we discontinued using the product. Since then, we have not had an ERMS but have arranged our workflow to compensate. License management was not an immediate need at first, so it was not addressed. Over time, however, it became obvious that something would need to be done with the piles of license documents in drawers and in shared folders online. When we wondered which electronic resources restricted walk-in users or which ones would allow us to download

an archival copy of the material, we relied on memory, email records, and sometimes reading through the licenses one by one.

As Stanford's Electronic Resources Librarian, I started to investigate solutions that would not involve us purchasing an ERMS. The commercial ERMS on the market would require the same time and staff resources as the product we purchased several years ago and would have the same challenges in integrating with our existing systems. But did we need to integrate? If we only needed a license management system, rather than an entire ERMS, maybe we could look at a different type of product. I considered database systems like Filemaker and open source products like Drupal, but in each case, I would need additional design work to make a usable system and I did not have the staff for this project. After seeing Notre Dame's presentation of CORAL at a conference, I contacted them and asked for a demo version.

In the meantime, I had been working with an intern, Michael Nack, to determine what was important for us in a license tracking system. What would our librarians and staff need for their work with licenses? We thought about two ways of interacting with the system. First, a librarian or staff user would want to know the terms of a particular license. Maybe they were considering purchasing a new product from a vendor and needed to know the vendor's typical restrictions on content. Or they planned to include some content in a coursepack and wanted to know if this was possible for a particular electronic product already in Stanford's holdings. For this interaction, they would need the ability to navigate quickly through the system to a particular license, searching by year or title. The second way of interacting with a license display system would be finding every license that met a certain set of criteria. For example, if our technical team was considering how to prevent walk-in users from accessing licensed electronic material, how could we find all licenses that required this restriction? Or as we start to consider archiving purchased content, how could we find which licenses included our preferred perpetual access language? For this work, we would need to search on a license term or restriction and find all matching documents. As we began to look for license management solutions, I kept these two types of interaction in mind.

What license terms would we need to track internally? I had been focusing on ONIX-PL as a standard that would allow us to mark up our licenses now with the eventual goal of ingesting them into a purchased ERMS. However, the more I looked into ONIX-PL, the more it felt like too much work. It requires a large dictionary of possible license terms and values for completely encoding a license document. While there was a lot of excitement around the standard, it didn't seem close to universal adoption and therefore I was wary of putting in the time if the standard did not live up to its potential. After seeing CORAL presented, I realized that a license management system could be pared down to only the terms that were most important to the library. The advantage of using a lightweight system like CORAL is that we could pick and choose the terms we wanted to track, rather than parsing an entire license. This would make the process of license ingestion simpler and quicker, and put the focus on the most important aspects of the document.

However, choosing these few key terms (or "expressions" in CORAL) was not easy and required input from several parties in the library. One of the most important was archival access terms, since we were starting to think about how we could request and store purchased electronic content. We ultimately coded this issue in four different terms. The perpetual access term captures if the resource is perpetual or subscription access. The archival type term determines if we can request an archival file immediately or only after we stop paying access fees. The preferred perpetual access term indicated if the license agreement includes our preferred terms, allowing for additional rights like text mining. Finally, third-party archiving captures if the content is available in LOCKSS, Portico, or another archiving platform. To

capture what we could do with the electronic content, we added terms for course packs/e-reserves and interlibrary loan (ILL). Finally, we added two terms to capture security questions. The first was the rights of walk-in users and the second was the requirement for immediate notification to the vendor upon any breach of access by an outside party.

Expressions	Qualifiers
Perpetual Access	Access Only
	Archival Copy Available
	Silent
Archival Copy Type	At Request
	At Request with Time Delay
	Only if Publisher Fails
	After Subscription Termination
	Silent
Preferred Perpetual Access Terms?	Yes
	No
Third-Party Archiving	LOCKSS
	Portico
	Other
Course Packs and E-reserves	Allowed
	Prohibited
	Silent
Interlibrary Loan	Allowed
	Prohibited
	Silent
Notification on Breach	Take Immediate Action
	Take Reasonable Action
	Silent
Walk-In Patrons	Allowed
	Prohibited
	Silent

With a list of terms established and a basic idea of how we wanted to use the system, we decided to install CORAL on a test server. The first advantage of CORAL for us was its modularity. It was built to handle licenses, unlike Filemaker or Drupal, but did not come with a full suite of other functions that we didn't need. Set-up and administration of CORAL were easy,

and if we did need further customization, the underlying framework was simple for our technical staff to manipulate. The other advantage we saw quickly was the concept that every aspect of a license term did not need to be coded. For example, we did not need to set up the system for print ILL allowed, electronic ILL allowed, print and electronic ILL allowed etc. CORAL is based around entering a snippet of the license text in a free text field, creating a quick view of the important terms of the document. Rather than requiring an exact interpretation of the text, the program allows a user to rely on the language of the document itself to decide the limits of the license.

However, there were a few hurdles. While a program like Filemaker could be set up to retrieve a single record or every record meeting a given set of criteria, CORAL was only designed for the retrieval of an individual license. The search capabilities helped a user narrow down a list of licenses by publisher or year, but did not allow retrieval of all licenses where, for example, ILL was prohibited. The browse capabilities could help somewhat by allowing a user to browse all snippets in the licenses, but since Stanford had roughly 600 licenses, the browse function would quickly become unwieldy. Another issue for us was the ability to make controlled lists of options for any single term. CORAL is split into "expressions" (or license terms) and "qualifiers" such as "allowed" or "prohibited." In some cases, we wanted specific qualifiers to capture nuances like which third-party archiving platforms were permitted, so the qualifiers might be "LOCKSS," "Portico," and "other," for example. This quickly led to a long list of qualifiers and growing challenges in entering a new license accurately. When a librarian input a new license and chose a qualifier, how could we make sure they were choosing from only the available options for that expression?

With these concerns in mind, we started talking to the Notre Dame team and to our internal technical staff about how best to make CORAL fit our needs. Notre Dame was willing for us to make changes if we sent the new code back to them for integration; they liked our ideas and thought they would be good additions to CORAL. Our technical staff agreed to do the work when it would fit their development schedule, but when other CORAL testers made similar suggestions, the Notre Dame team decided to make the changes themselves right away.

With new search capabilities in place, as well as the ability to make a limited list of qualifiers for every expression, Stanford was ready to start entering licenses into a production version of CORAL. We had also purchased Optical Character Recognition (OCR) software so we could scan our licenses into machine-readable PDFs. This allowed us to cut and paste chunks of text into CORAL from the PDF license as well as quickly search the document for the expressions needed.

The process of implementing CORAL was significant in several ways for me. First, the concept behind CORAL was eye-opening. It is built as modules that can handle pieces of the ERMS workflow, as needed by a library. The license management piece can be detailed or minimal, depending on the needs of the users. Instead of relying on an absolute statement of whether a particular use of information is allowed or prohibited, CORAL allows you to rely on the text of the license itself to make a determination. Also, the experience of working with the designers of a new open source system to discover how their system could be extended and modified to meet

our needs was enjoyable. The design of CORAL made it flexible to fit Stanford's needs and the great development team at Notre Dame meant that the process of implementation was truly fun.

CONCLUSION

For electronic resource management, what really matters to your library's workflow? Are all the features of a commercial ERMS needed for your work or could you use a more lightweight flexible product? Notre Dame and Stanford both went through a process of evaluating what their library needed for license management and how these needs could or could not be met by the commercial products on the market. Notre Dame had the resources to build their own product while Stanford was happy to find another library thinking along the same lines. The processes of deciding which license terms to track and how to implement the product internally were also insights into each library's priorities. The end result for both Notre Dame and Stanford was a flexible, and yes, fun product to manage license documents.