VeriCite: an Affordable, Integrated Solution for Content Matching/Anti-Plagiarism in Canvas

Kevin Nolan, Melody Brake and Pat Burns
Colorado State University
Unizin Innovation Summit
April 19, 2017
2:15-3:00 PM
Outline

• Background and history
• “The $64,000 Question” – access to licensed content
• Progression/quality of VeriCite finding plagiarism
• Relationship with LongSight
• Demonstration Q&A
Background

• CSU started on WebCT in 1998, and was on Blackboard Learn when CSU joined Unizin early, June 2014
• Established a two-year replacement cycle for Blackboard, starting July 1, 2014
• Needed an Anti-Plagiarism tool to replace SafeAssign, bundled with Blackboard, critical in Composition, Lab classes
  • Educational – cite properly
  • Deterrent
Quest for an Anti-Plagiarism (AP) Tool

• SafeAssign
  • Could not unbundle from BB (can you imagine Canvas LTI for a BB product?)
  • Also, BB would not release “our data”
    • Need to start over with an AP database

• Most Unizin sites were using Turnitin (TII)
  • Was relatively expensive at the time
  • Most sites not all that enamored with the Company
  • Policy that TII owns “all of the data” – FERPA violation, in addition to a very poor business practice
  • Decided to look elsewhere
Suggestion from Lois Brooks at OSU

• LongSight ([http://www.longsight.com/](http://www.longsight.com/))
  • Support for Open Source software (Sakai, Drupal, DSpace,...)
  • Scott Siddall was the CEO (Lois knows him)
    • CIO at Kenyan College
    • Vice Provost at Denison University
    • EDUCAUSE Board of Directors

• New AP tool
  • VeriCite ([http://vericite.com/](http://vericite.com/))
  • Very “fresh” at the time, but technology showed great promise
  • Scott willing to work with us to improve the tool
  • Unizin explored a Master Services Agreement, but Scott would not escrow the code as open source
Elements of the Exploration

• Effectiveness in finding plagiarism
  • Side-by-side testing of VeriCite, SafeAssign, and Turnitin
  • Detailed exploration of “the $64,000 question” – next slides
• Quality of the user interface, and faculty feedback
• Integration with Canvas
• Development trends
• Relationship with the company
Faculty Questions

• Faculty – “Why does the AP tool not detect plagiarism inherent in this article that I just found online in journal xx? It’s online!”

• Answer - AP tools don’t search our particular library-licensed content
  • We make it very easy for faculty to search licensed content, usually by IP address range
  • They expect all tools to have the same visibility they do!
“The $64,000 Question”

• “Can’t we just get an AP tool that searches our licensed content?”

(Most dreaded question for IT always starts with “Can’t you just?”)

• Scope for CSU: $7.8M annual costs
  • ~50,000 journal titles (out of ~100,000 total)
  • ~ 1,000 online databases
  • ~500,000 eBooks

• Ask a Librarian!!!
Exploration Team: March-Sep. 2015

• CSU
  • Pat Burns, VP for IT
  • Kevin Nolan and Melody Brake, LMS support team
  • Elaine Green, Academic Integrity
  • Meg Brown-Sica, Asst. Dean, CSU Libraries
  • Rachel Erb, CSU Libraries contracts

• Internet2 (later on)
  • Shel Waggoner

• OSU
  • Shannon Riggs, Director of Course Development
  • Lynn Greenough, Instructional Technology Manager

• Unizin
  • Amin Qazi
  • Robin Littleworth

• LongSight
  • Scott Siddall
  • Brian Holladay
Licensed Content: Massive Complexity

- Hundreds of contracts
  - Metadata for all contracts
- Database of databases
- Link resolver
  - Multiple ways of getting to the same licensed content
- Manually constructed
- Varying access control methodologies
Selection of Representative Journals

Does the contract permit “mining” content?

- American Chemical Society: No
- Elsevier license: Elsevier will review requests
- JSTOR (a non-profit aggregator) “encourages text mining”
- Nature: Yes, for the purposes of academic research
- Taylor and Francis: silent in agreement
- Wiley: silent in agreement
Our Proposal to Publishers

• Allow Longsight to crawl our content, provided Longsight would:
  • Be bound contractually to our terms and conditions
  • Crawl and index content during off hours
  • Store our content in a private area
  • Encrypt our content
  • Purge all content after 30 days
The Answers

• Except for JSTOR, no Publisher would agree to our proposal

• JSTOR was amenable, but indicated that, as an aggregator, they would have to get permission from all of their publishers!
CrossRef

• http://www.crossref.org/
• “We are a not-for-profit membership organization for scholarly publishing working to make content easy to find, cite, link, and assess.”
  • Collaboration – better together
  • Global community of members who publish

• Anyone can subscribe
• Special site of licensed materials
• Can be crawled by agreement
• Nothing can be downloaded
Finding Effectiveness

• CSU piloted Canvas and VeriCite fall 2015/spring 2016.
• LTI external tool – functional but not intuitive.
  • February 20, 2015: OSU CIO got a commitment from Instructure for native integration of VeriCite into Canvas
• Monthly meetings between CSU, OSU, and VeriCite began in March 2015
  • Unizin (Amin Qazi and Robin Littleworth) joined the activity in April 2015
• Matching algorithm needed work.
• An 8-month exploration
  • OSU and CSU: test cases, side-by-side testing using TII, SA and VC
  • VeriCite room for improvement, but they were the only company willing to work with us
Working out the Kinks
But It’s Not as good as SafeAssign!
• Part true, part “halo effect”.

• Matching algorithm needed work. Instructors reported instances where VeriCite did not catch items easily found on Internet.

• LTI interface had its challenges.

• VeriCite LTI strips out formatting. Some students thought something went wrong and would resubmit.

• Initial and final similarity score could be confusing. Why did it change?
Relationship with Longsight

• Passed along examples of gaps in VeriCite to their support.
  • Very responsive – adjustments made quickly to algorithm.

• Educate instructors and students on interface.

• Business and Natural Sciences expressed specific concerns which led to...
Road Trip! VeriCite Comes to CSU

Onsite visit June 2016
• Focus Groups with faculty, college coordinators, LMS admins.

Results:
• Our users felt they had been heard.
• More open access journal sites indexed.
• Bibliography excluded from content matching.
• “Draft” version – not placed in CSU VeriCite repository.
• Can exclude/include self plagiarism.
The Grail! Full integration with Canvas

• Became available for testing in November 2016.
• Tested – worked well.
• Turned on in Dec., after finals. Ready for spring semester.
• Publicized, offered documentation and training.

Instructors pleased. This was one of the main features they had waited for.
2017 What’s New:  
Instructor Ability to View “Other Student” Info

• Similarity Report showed paper matched another CSU student’s paper but instructor could not see other student’s information.  
  • Instructors could see this info in SafeAssign

• Discussed internally – within FERPA to enable instructors to see other student’s information and be able to contact other instructor.

• Raised question with VeriCite.

• VeriCite added functionality. Institution accepts responsibility for FERPA.

• Can now see other student’s name, assignment, course and instructor of other course.
Testing New User Interface

• Piloting view paper and source document side by side instead of top and bottom.

In 40 fields of study, in addition to a professional degree in veterinary medicine, in fiscal year 2012, CSU spent $375.9 million on research and development, ranking 60th in the nation overall and 34th when excluding medical school spending.

• Feb 18 Canvas update impacted VeriCite advance settings.
  • VeriCite implemented short term fix while we await Canvas resolution.
Demo
Any Questions?