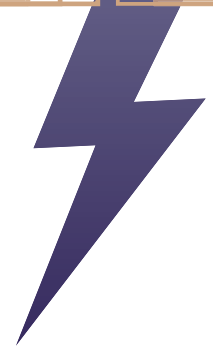


## Webinator as a customizable way to add vertical search engines to multiple industry web portals



*When implementing an optimal solution for the heavy search demands of multiple online properties, a website administrator needs a practical way to easily create and provide a high quality retrieval interface to collections of HTML documents. In this article we review how Trade Press Publishing successfully added powerful and flexible vertical search engines to its popular web portals with the help of Thunderstone's Webinator web index and retrieval system. Webinator serves as an example of the type of applications that can be built around Thunderstone's Taxis RDBMS and Web Script.*

Trade Press Publishing Corporation (<http://www.tradepress.com>) is a privately-held company based in Milwaukee and a leading provider of market intelligence to the facilities management, building service contractor, housekeeping, cleaning supplies distribution and railroad industries. In addition to publishing business-to-business magazines and eNewsletters, it also produces trade shows and conferences, as well as offering variety of related educational and marketing opportunities.

Jesus Carrillo, Director of Information Technology, joined the company's Pre-Press

Division more than 16 years ago. According to him, "I started in the Desktop Publishing Department at an entry-level position that was my first job out of college. And I've been at the same place ever since. The company grew. About ten years ago they dissolved the pre-press part of the business to focus on educational media products and business-to-business publishing. They wanted someone to lead their technology efforts, and they asked me to do that. So, I stayed around and have continued to search out technology applications in the b-to-b publishing space."

### Special Requirements to Index and Search Industry-Focused Web Content

Trade Press Publishing Corporation uses Webinator on four "vertical portal" web sites, including two in the facility management space and two in the sanitary distribution/cleaning space. The main site is at FacilityZone.com (<http://www.facilityzone.com>.) Carrillo said the biggest reason he selected Webinator as the indexing and searching tool for Trade Press had to do primarily with Webinator's open-ended customizability.

According to Carrillo, “Probably the single identifying characteristic of the Webinator software, for us, was the ability to get to the source code. And that allowed us the flexibility to put it to work to do the things that we wanted to accomplish with its back-end. For example, we were indexing over six thousand web sites, which is quite a bit of data. And the first results that came up were kind of cool. We could see how, out of these millions of pages, you do a search, and there’s some logic in there that says ‘these are the ten best ones’ out of the millions of pages I’ve got.

“Taking a closer look at them, we felt that to really bring it to a marketplace and have it be as meaningful as possible to our end users—we needed to go in and play with a lot of the settings to get the search engine to produce the particular kind of results we believed that our users would typically want to find. The straight-out-of-the-box algorithm for searching didn’t have an immediate correlation to precisely what we thought our users would be looking for.

“We spent some pretty significant time working with Thunderstone’s tech support, doing tests and evaluations and changes and modifications, trial and error, to get things to the point where now it seems on a regular basis the terms we’re punching in are getting the types of results that we know will make our users happy,” Carrillo explained.

Webinator’s user features include:

- simple navigation

- intelligent query capabilities with natural language processes, special pattern matchers (regular expressions, numeric quantities, fuzzy patterns,) document similarity searches, in-context result listings, link reference reports, proximity controls and set logic

Carrillo continued,” The setup and deployment of Webinator is extremely easy and straightforward. All the core functionality is there plus the ability to access the source code and be as creative and as customized as your capabilities will let you be. In other words, Thunderstone doesn’t hold you back. Thunderstone lets you take the product to whatever level you’re ready, willing and able to take it. For that reason we’ve stuck with it, we’ve used it, and it’s been great in that regard. That’s not something you’re going to get from the Googles of the world.”

### **‘Locked Box’ Approach of Others Inadequate**

“We took a look at the Google appliance. It was brand new at the time. And the reason we didn’t go with the Google appliance was we had no control over it. No flexibility. No ability to customize. Basically it was a ‘locked box’ sitting in our office, you know? And that’s really not the way we wanted to go about it. We’ve got technical expertise on staff. We can go in. We can study and learn the scripts. We can make our modifications.

“For instance, when you execute a search on [facilityzone.com](http://facilityzone.com)—it executes a search first

off of a SQL Server database that we've got on our end. Then it goes and executes it against the Webinator database and combines the two sets of results. So, we've got results that are built into a page that kind of fall on top of the results that come out of the search engine. There's no way that you're going to be able to do that with the Google Search Appliance. You just won't be able to.

"The access to the source code and the flexibility of Webinator were definitely both something of value to us. Basically, we could not have done what we did without it. Working with the tech support team at Thunderstone, we have access to people who will actually call you back and work with you on some crazy questions."

"We're hoping that Thunderstone will continue to be a leader and help pave the way for how search technology is going to evolve. We'd like to take advantage of the new applications that Thunderstone develops and apply them to our industries and to our users," Carrillo said.

Web portal administrators looking for a web walking and indexing package to help them add vertical search engines to multiple online properties will appreciate the fact that Thunderstone's Webinator:

- indexes multiple sites into one common index
- offers administrators detailed verification and logging of document linkages

- can index/update documents while the database is in use
- permits multiple databases at a site
- features a simple browser interface
- is written in Taxis Web Script for complete flexibility
- provides an SQL query interface to the database for maintenance and reports
- allows remote sites to be copied to the local file system
- lets multiple index engines run concurrently against a common database

### **About Thunderstone**

Thunderstone Software LLC

(<http://www.thunderstone.com>) pioneered simultaneous searching of both structured and unstructured data with the Taxis relational database optimized for full text search. Since 1981 Thunderstone has continued to develop its global reputation as provider of the world's most powerful, scalable and flexible enterprise search solutions.