

FAST Search for the Enterprise Speed your way to key data

©2011 Ramona Maxwell – all rights reserved

The July 14th, 2011 presentation to the [San Francisco SharePoint User Group](#) by Carl Grimm of Avenade provided a great overview of the challenges large organizations face implementing a search solution that can search vast document stores and return accurate results at lightening speed. Aside from user feedback he discussed methods for evaluating search (such as the 'churn rate' or number of times the user inputs an adjusted search term before they hit on a satisfactory result) before moving into the key question: *Why is it so difficult to implement an effective search solution?* One of the causes mentioned was 'contributory negligence' where documents being placed in the search store are not properly tagged, for instance when someone reuses a favorite document with 'save as' but doesn't update the metadata to reflect its new content. Problems also result from regional language differences as even within the same country common items can have localized names. The speaker provided a charming example of folks from Pittsburgh referring to rubber bands as gum bands. User education is key, as one quote estimated that in 25% of SharePoint search installs considered a failure the causes were traced to frustration and a lack of training among users. Additionally, some business users aren't clear on the distinction between an intranet and the internet, searching their company site for services like Facebook and Google. The speaker noted that a #1 search in some companies is 'what's for lunch?' Perhaps some serious bandwidth could be saved by making that high-priority piece of information obvious!

As the speaker began to delve into the specific feature-set of [FAST Search Server 2010 for SharePoint](#) (available in Enterprise Edition only and with an additional license for the search component) the topic quickly became engaging. FAST uses entity extraction, in other words it looks through every document for relevant chunks of information like dates, company names, addresses – things that you could identify with your own red pen if someone handed you a document and said 'find me everything that looks like a phone number' and isolates them as attributes of the document. It then creates a virtual reference to the document based on its characteristics. You've seen this in action if you use Gmail and you open a message containing event information. Often the right panel will pop up a suggested calendar entry with the details pre-filled from your email. The entity extractor in FAST Search Server 2010 is reputed to be very accurate and able to produce good quality results.

Mr. Grimm offered a peek at what the man behind the curtain is really doing when you type in a request for 'courage' or 'brain.' He gave an example of someone performing medical research searching for 'protein affecting lupus.' In a properly designed search system the term 'protein' would actually lead to a table listing specific protein names, thereby greatly increasing the likelihood of finding pertinent information. FAST may substitute other search terms, for instance replacing 'affecting' with the more direct 'cause.' But what is really helpful in returning results having to do with metabolic proteins as opposed to those in tuna fish is the fact that FAST knows how to promote terms based on the context of the requester. Continuing with a medical facility analogy, it can promote results for insurance codes to a user searching from the billing department and promote scientific terms for a user in a laboratory. The user context can flow from Active Directory or from the user's audience profile that FAST can help you extract from existing corporate resources such as org charts and corporate databases living outside of SharePoint.

FAST also has methods of weighting terms to avoid typical search problems like finding a hundred documents that are *not* the one you want, simply because they cite the author of the one you're looking for in the bibliography. Terms that appear in near proximity to each other in a document will increase relevance, as will terms found in a title. FAST also has support for a couple hundred document types out of the box, its own web crawler, unlimited metadata types and it can ingest custom taxonomies for indexing. FAST also has a more flexible security access model than standard SharePoint search and can handle single sign-on systems easily. FAST supports multiple levels of tuning based on whether speed or cardinality should take precedence in a particular application.

What really seemed significant regarding FAST to me is its potential to provide for general business documents what PowerPivot does for analysis in Excel: that is to centralize and control the document store while still providing access. Users are able to share and manipulate documents within the FAST interface without the need to download them to personal media. This seemed very similar to PowerPivot document hosting that allows permission based sharing, version control and even usage metrics, all of which are a long overdue improvement to the practice of emailing business documents.

FAST may not suit every business, it's expensive and can really benefit from a dedicated search server. However for an organization whose success depends on search technology that is swift and accurate it is a worthwhile investment. To learn more about the benefits of FAST please visit [Microsoft's case study portal](#).

Please visit [MaxwellArts.com](#), [SharePointSolver.com](#) and [SQLSolver.com](#)