



**Microsoft FAST Search Server 2010
for SharePoint
Evaluation Guide**

The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This white paper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in, or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

Unless otherwise noted, the example companies, organizations, products, domain names, e-mail addresses, logos, people, places, and events depicted herein are fictitious, and no association with any real company, organization, product, domain name, e-mail address, logo, person, place, or event is intended or should be inferred.

© 2010 Microsoft Corporation. All rights reserved.

Microsoft, Windows, SharePoint, FAST, Visual Studio, SQL Server, OneNote, PowerPoint, Outlook, and Excel are trademarks of the Microsoft group of companies. All other trademarks are property of their respective owners.

Contents

| | |
|---|----|
| Abstract..... | 1 |
| How to Use This Document..... | 2 |
| Introduction | 3 |
| Technical Problems Solved by FAST Search Server 2010 for SharePoint..... | 4 |
| Microsoft Enterprise Search Products Overview..... | 6 |
| Microsoft Server-Side Search Products | 6 |
| Capabilities Comparison..... | 6 |
| Content Repositories..... | 7 |
| Indexing Scale | 8 |
| Developer Information..... | 9 |
| Enterprise Search Features in FAST Search Server 2010 for SharePoint..... | 10 |
| End-User Perspective..... | 10 |
| IT Professional Perspective | 11 |
| Developer Perspective..... | 11 |
| FAST Search Server 2010 for SharePoint..... | 12 |
| Enterprise Search Enhancements in FAST Search Server 2010 for SharePoint..... | 12 |
| Visual Search Capabilities | 12 |
| Document Thumbnails..... | 12 |
| Scrolling PowerPoint previews..... | 12 |
| Visual Best Bets..... | 12 |
| Conversational Search Capabilities | 13 |
| Sort Results on Managed Properties | 13 |
| Deep Results Refinement..... | 13 |
| Similar Results | 14 |
| Result Collapsing..... | 14 |
| Contextual Search Capabilities..... | 14 |

| | |
|--|----|
| Relevancy Tuning by Document or Site Promotions | 14 |
| Synonyms..... | 14 |
| Managed Properties and Metadata creation..... | 15 |
| Property Extraction..... | 15 |
| Rank Profiles | 16 |
| Linguistics | 17 |
| FAST Search Server 2010 for SharePoint for End Users | 18 |
| FAST Search Server 2010 for SharePoint User Experience..... | 18 |
| Enhanced Search Results..... | 21 |
| Social Search..... | 22 |
| Finding People | 22 |
| Mining and Discovering Expertise | 23 |
| Improving Search based on Social Behavior..... | 23 |
| Custom Search Solutions | 24 |
| FAST Search Server 2010 for SharePoint for IT Professionals..... | 25 |
| FAST Search Server 2010 Architecture Overview | 25 |
| Modular and Scalable Architecture | 26 |
| Integration with Search Center | 27 |
| IT Professional Experience | 27 |
| Windows PowerShell Support..... | 27 |
| Categories of cmdlets..... | 28 |
| Developer Experience..... | 28 |
| FAST Search Server 2010 for SharePoint —Services and Components | 28 |
| FAST Search Server 2010 for SharePoint consists of a number of different services and components. The following sections describe the services and components.... | 28 |
| FAST Search Connector | 28 |
| Web Link Analysis (Web Analyzer)..... | 28 |
| Item Processing | 29 |

| | |
|---|----|
| Indexing..... | 30 |
| Query Matching..... | 30 |
| Query Processing | 30 |
| Connector Framework..... | 31 |
| New Connector Features..... | 31 |
| Creating Connectors | 31 |
| Search Administrator Walkthroughs..... | 32 |
| Administering FAST Search Server 2010 for SharePoint Managed Properties..... | 32 |
| Search Administration at the Site Collection Level..... | 35 |
| FAST Site Collection Settings..... | 37 |
| Appendix A: Search Terminology..... | 42 |
| Appendix B: Feature Comparison between Search Products..... | 46 |
| Appendix C: Resources Available for Evaluating SharePoint Server 2010..... | 48 |
| Appendix D: Feature differences between SharePoint Server 2010 and FAST Search Server 2010 for SharePoint..... | 49 |
| Feature differences from an end-user perspective..... | 49 |
| Search alerts..... | 49 |
| Social tagging and its effect on relevance ranking..... | 49 |
| Definition extraction | 50 |
| Author, date, and title extraction from Word and PowerPoint documents..... | 50 |
| Feature differences from an IT professional perspective | 50 |
| Crawl rules for document formats..... | 50 |
| Mirroring indexes across multiple data centers, backup and recovery functionality | 51 |
| Minimum size footprint | 51 |
| Monitoring capability without using System Center Operations Manager | 52 |
| Feature differences from a developer perspective | 52 |
| Query integration and query-side interfaces..... | 52 |
| Customized ranking | 53 |

Programmatically administering the enterprise search solution.....53
Custom security trimming53

Abstract

This evaluation guide is designed to give you an understanding of the design goals and the details of the enterprise search features provided by Microsoft® FAST™ Search Server 2010 for SharePoint®. This guide is also designed to give you a familiarity with how to implement enterprise search by using FAST Search Server 2010 for SharePoint.

This guide includes:

- Descriptions of the enterprise search features and technologies provided by FAST Search Server 2010 for SharePoint.
- Details of the additional features that FAST Search Server 2010 for SharePoint provides over and above SharePoint Server 2010.
- Details on the indexing and query architecture implemented by FAST Search Server 2010 for SharePoint.
- Tours and walkthroughs of the main search features provided by FAST Search Server 2010 for SharePoint.

This guide is designed for technical decision makers, IT professionals, and developers. The overall goal of this guide is to help you perform a thorough and effective evaluation of the enterprise search features provided by FAST Search Server 2010 for SharePoint.

For the latest information about SharePoint 2010 products, please visit [SharePoint Products and Technologies](#).

How to Use This Document

This document has been designed to enable you to learn about and evaluate the enterprise search features provided by FAST Search Server 2010 for SharePoint.

NOTE: FAST Search Server 2010 for SharePoint shares much in common with the enterprise search features of SharePoint Server 2010. If you are not already familiar with the enterprise search features of SharePoint Server 2010, you should also review the [SharePoint Server 2010 Enterprise Search Evaluation Guide](#).

The document provides comprehensive information that can be used to evaluate all of the enterprise search features and components provided by FAST Search Server 2010 for SharePoint, from the perspective of the following roles:

- **Technical Decision Makers.** You can use this guide to gain an understanding of the business requirements that are met by enterprise search solutions. You can also learn how specific aspects of the search technologies provided by FAST Search Server 2010 for SharePoint work together to fulfill business and technical requirements for a successful enterprise search solution.
- **IT Professionals.** You can use this guide to gain an understanding of how to configure, administer, and manage the enterprise search features of FAST Search Server 2010 for SharePoint. You should pay particular attention to the walkthroughs throughout this guide.
- **Developers.** You can use this guide to gain an understanding of the features of FAST Search Server 2010 for SharePoint. You should read all of the sections in this document so that you gain an insight into the platform on which you will develop solutions. You should also refer to the [SharePoint Server 2010 SDK](#) (which includes the FAST capabilities) for more detailed developer guidance, walkthroughs, and samples.

Introduction

Welcome to this evaluation guide for Microsoft FAST Search Server 2010 for SharePoint. The goal of this guide is to help you gain sufficient knowledge and understanding of FAST Search Server 2010 for SharePoint to evaluate how it can fulfill your organization's enterprise search requirements.

NOTE: This guide describes the enterprise search features FAST Search Server 2010 for SharePoint. If you would like to find out more about the enterprise search features of SharePoint Server 2010, review the [SharePoint Server 2010 Enterprise Search Evaluation Guide](#).

In this section, you will learn about typical business and technical requirements for enterprise search solutions, and how FAST Search Server 2010 for SharePoint fulfills those requirements.

Search connects people to the information they need to get their jobs done. General productivity search solutions increase employee efficiency by connecting a broad set of people to a broad set of information, the most common examples being intranet and people search. In comparison, search-driven applications drive measurable return on investment by helping a well-defined set of people accomplish specific business tasks more efficiently. Search-driven applications, such as research portals and 360° customer insight solutions, aggregate information from a defined set of content repositories, add structure to unstructured information, and provide a contextual, interactive, and actionable experience.

Until now, organizations have had to purchase multiple products to address both the scenarios of general productivity search and search-driven applications—no single search solution was able to deliver the right functionality at the right price point for both scenarios, forcing organizations to incur unnecessary costs, such as:

- Software costs—licenses, support, and maintenance costs were required for multiple search products.
- Headcount costs—dedicated positions were required to design, deploy, and manage different search solutions.

- Opportunity costs—mediocre general productivity search and limited deployment of high-value search applications did not help information worker productivity.

FAST Search Server 2010 for SharePoint offers a new choice in enterprise search. Combining the power of FAST with SharePoint, FAST Search Server 2010 for SharePoint delivers Microsoft's best general productivity search experience *and* a platform for building search-driven applications.

Technical Problems Solved by FAST Search Server 2010 for SharePoint

Every organization has unique search requirements. FAST Search Server 2010 for SharePoint has the capability to customize the search experience so that it fits how your business and your people work. Specifically, with FAST Search Server 2010 for SharePoint, you can:

- Deliver results that are contextually relevant.
- Search in the language of your business.
- Tune relevancy to improve accuracy.
- Customize the search platform to meet your specific indexing and search requirements.
- Configure the user interface to customize the search experience for information workers.

Enterprise search solutions from Microsoft enable you to:

- Provide users with results that are meaningful and dynamically tailored to their jobs, roles, and functions within the organization. This means that your sales teams will be quickly able to find product information, collateral, and answers to RFP questions, while your engineering teams will see specifications and requirements documents at the top of their results sets. Site administrators can tailor search quickly and easily to deliver contextually relevant results the first time.
- Give users the ability to use terms and languages that are unique to your business. Most organizations frequently use a set of internal names, acronyms, or code words. These words can be confusing to different groups, outsiders, or new members of your organization. Users will be able to use their own terminology to

sort, refine, and query your content. Furthermore, advanced language support provides your employees the ability to find content written in its native language.

- Ensure that searches provide accurate ranking for relevant results. The major reason that a user continues to use a search engine is if it returns relevant information near the top of the search results. Microsoft search gets better with social ranking capabilities by promoting popular documents. Site administrators will quickly and easily be able to create and deploy new custom ranking algorithms that are tuned meet multiple business demands simultaneously.
- Provide a great out-of-the-box experience to get search up and running quickly. Additionally, provide a platform that grows with your business needs so that you can:
 - Quickly access and crawl new content repositories.
 - Add your users and business partners to the lists of extracted entities.
 - Perform custom content processing such as sentiment analysis or machine translation.
 - Tailor the user interface with custom SharePoint Web Parts or extend the ones that are available out of the box.

FAST Search Server 2010 for SharePoint provides an enterprise search platform for fulfilling these aims. As a brief overview, FAST Search Server 2010 for SharePoint includes a connector framework that enables the crawler to index files and metadata from various types of content repositories. It also provides an indexing engine that stores the crawled data in an efficient manner in index files, and it provides query servers, query object models, and user interfaces for performing searches on the indexed data.

You will learn more about each of these components later in this guide, but for now be aware that these components all work together to fulfill the aims and meet the requirements of enterprise search solutions.

Microsoft Enterprise Search Products Overview

There are various search products available from Microsoft, so before delving into the details of enterprise search for FAST Search Server 2010 for SharePoint, it will be useful for you to become familiar with all of the products in the enterprise search portfolio.

Microsoft Server-Side Search Products

The following products all provide varying degrees of indexing and search features.

- Microsoft SharePoint Foundation 2010 search
- Microsoft Search Server 2010 Express
- Microsoft Search Server 2010
- Microsoft SharePoint Server 2010
- FAST™ Search Server 2010 for SharePoint*

This guide will delve into the features of these products in later sections.

Capabilities Comparison

The following table compares general enterprise search capabilities that were previously provided in Microsoft Office SharePoint Server 2007 with the new and enhanced features in SharePoint Server 2010 and FAST Search Server 2010 for SharePoint.

| X = available in 2007 O = available in 2010 † = improved in 2010 | | | |
|---|-------------------------------------|---------------------------|--|
| Features and Capabilities | Office SharePoint Server 2007 | SharePoint Server 2010 | FAST Search Server 2010 for SharePoint |
| People and expertise search Capture knowledge not found in documents by searching for people and expertise using SharePoint products. | X | O† | O† |
| SharePoint 2010 connector framework Securely connect to content in SharePoint sites and from sources across your enterprise. Use the Business Data Catalog to easily create your own connectors that work just like those available out of the box. | X | O† | O† |
| 100m content volume with sub-second query response time Meet the scale and performance needs of your entire organization or the specialized needs of individual departments. | X | O† | O† |

| Features and Capabilities | Office SharePoint Server 2007 | SharePoint Server 2010 | FAST Search Server 2010 for SharePoint |
|--|-------------------------------|------------------------|--|
| <p>Search from Windows 7 & Windows Mobile Search beyond the search center. Conduct searches from the Windows 7 desktop and on your Windows mobile device.</p> | X | O† | O† |
| <p>Taxonomy tag integration Bring the power of taxonomy into search. Tag metadata is shown in results, and users can refine by taxonomy-based tags.</p> | | O | O |
| <p>Metadata-driven refinement panel With the new refinement panel in SharePoint Server 2010 and FAST Search Server 2010 for SharePoint, users can narrow the results of their searches and navigate to the right content faster.</p> | | O | O |
| <p>Relevance improves with social behavior The click-through behavior of similar search queries affects the rank that documents receive. The more users click on a certain item, the higher its ranking for related queries.</p> | | O | O |
| <p>Phonetic and nickname search Confidently search for a person's name as it sounds - without worrying about the exact spelling.</p> | | O | O |
| <p>Contextual search Tailor different results and refinement options based on the profile of the user or audience.</p> | | | O |
| <p>Thumbnails and previews Thumbnails and previews make the results of a search query visual, allowing users to recognize the right content quickly.</p> | | | O |
| <p>> 500m content volume with sub-second query response time Scale to extremes with FAST Search Server 2010 for SharePoint while maintaining sub-second query response times.</p> | | | O |
| <p>Advanced content processing with advanced linguistics Extract and create metadata latent in documents to improve search results, sorting capabilities, and the refinement panel.</p> | | | O |
| <p>Search-driven applications Meet all the search application needs you have across your business. Common examples include 360° Customer Insight, Research and Development Innovation Portal, and Product Support.</p> | | | O |
| <p>Business Intelligence Indexing Connector Crawl Microsoft Excel® workbooks and Reporting Services Reports with improved results, descriptions, thumbnails, and refiners. Discover your business intelligence (BI) assets quickly and easily, navigate not only the document, but also the data behind the scenes, and access the information you need quickly and easily.</p> | | | O |

Content Repositories

In addition to the feature comparisons, you should also consider the types of content repositories that can be crawled by each product:

- SharePoint Foundation 2010 can only crawl SharePoint sites in the same farm.
- All the other products in the above table can crawl the following types of content repository:
 - SharePoint sites (in the same farm, or in external farms)
 - Windows file shares
 - Microsoft Exchange public folders
 - Web sites that are not SharePoint sites
 - People Profiles
 - External line-of-business applications
 - Structured content in databases
 - Content returned by Web services
 - Third-party products and solutions including Lotus Notes and Documentum

Indexing Scale

Although there are no hard-coded limits for the number of items that can be indexed by any of the products listed, there are some practical guidelines based on feasibility and performance:

- Search Server 2010 Express and SharePoint Foundation 2010 can index and search up to 300,000 items if they are used with SQL Server® Express; otherwise they can index and search up to 10 million items if they are used with a full edition of SQL Server 2008.
- A scaled-out Search Server 2010 farm can index and search up to 100 million items.
- A scaled-out SharePoint Server 2010 farm can index and search up to 100 million items.
- A FAST Search Server 2010 for SharePoint installation can support extreme scale, and can index and search over a billion items.

One of the general aims of enterprise search with FAST Search Server 2010 for SharePoint is to implement sub-second query latencies for all searches. To achieve this, you must ensure that no query server deals with more than ten million items. You can achieve this by adding multiple query servers to your farm, and therefore by taking advantage of the new index partitioning features of FAST Search Server 2010 for SharePoint. Index partitioning enables administrators to spread the load for queries across multiple query servers. This is achieved by creating subsets of an index, and propagating individual subsets to different query servers. FAST Search Server 2010 for SharePoint uses a hash of each document's ID to determine in which partition the index entries for a specific document should be stored. At query time, the query object model contacts each query server needed to satisfy the search so that all results to be returned to the user are included.

For more comparison data between the server-side search products from Microsoft, see [Search Technologies for SharePoint 2010 Products](#).

Developer Information

All of the products described above provide a unified [query object model](#). The result is that if you develop a custom solution that uses the query object model for SharePoint Foundation 2010, for example, then it will continue to work if you upgrade to SharePoint Server 2010, or if you migrate your code to FAST Search Server 2010 for SharePoint.

Enterprise Search Features in FAST Search Server 2010 for SharePoint

FAST Search Server 2010 for SharePoint builds on SharePoint Server 2010, and provides significant enhancements to the Enterprise Search capabilities.

Search in SharePoint Server 2010 is targeted at “general productivity search.” General productivity search solutions increase employee efficiency by connecting a broad set of people to a broad set of information. Intranet search is the most common example.

FAST Search Server 2010 for SharePoint can be used on a variety of search problems, providing enhanced “general productivity search” but also uniquely addressing “high-value” search applications. High-value search applications drive measurable return on investment by helping a specific set of people make the most of a specific set of information. Common examples include product support applications, research portals, and customer record locators.

The frameworks and tools used by IT professionals and developers are common across the product line (as much as possible, given the additional capabilities in FAST Search Server 2010 for SharePoint). In understanding the benefits of FAST Search Server 2010 for SharePoint, it is useful to consider:

- Capabilities that are the *same* as those in SharePoint Server 2010 search.
- Capabilities that are *better* than those in SharePoint Server 2010 search.
- Capabilities that are *unique* to FAST Search Server 2010 for SharePoint.

A high level summary of the capabilities is shown in the tables below.

End-User Perspective

| Capability | Difference |
|--|------------|
| One-stop search center to find answers quickly | Better |
| Refinement to explore information quickly | Better |
| Social search to connect with people and expertise | Same |
| Search gets better with use | Same |
| Visual cues for rapid recognition of information | Unique |

| | |
|---|--------|
| Contextual search to meet the needs of diverse groups | Unique |
|---|--------|

IT Professional Perspective

| Capability | Difference |
|---|-------------------|
| Scale and performance | Better |
| Ease of deployment | Same |
| Enterprise-class manageability | Better |
| Secure, broad connectivity | Better |
| Advanced content processing out-of-the-box | Unique |
| Ease of configuring high-end user experiences | Unique |

Developer Perspective

| Capability | Difference |
|--|-------------------|
| Customize the out-of-the-box user experience by using Web Parts | Same |
| Extend connectivity by using the Business Connectivity Services and Federation | Same |
| Combine search with other SharePoint capabilities | Better |
| Use familiar tools built for developer productivity | Same |
| Leverage advanced content processing | Unique |
| Customize relevance | Unique |
| Use advanced query capabilities to create powerful applications | Unique |

FAST Search Server 2010 for SharePoint

FAST Search Server 2010 for SharePoint provides enhancements and new capabilities to the enterprise search features of SharePoint Server 2010. You can use this section to learn more about the enhanced enterprise search features provided by FAST Search Server 2010 for SharePoint. If you are not familiar with the new enterprise search capabilities in SharePoint Server 2010, refer to the [SharePoint Server 2010 Enterprise Search Evaluation Guide](#).

Enterprise Search Enhancements in FAST Search Server 2010 for SharePoint

This section provides a summary of the new and enhanced capabilities that FAST Search Server 2010 for SharePoint provides compared to the other search products from Microsoft. You can use this section to gain an overview of the value of implementing enterprise search solutions based on FAST Search Server 2010 for SharePoint.

Visual Search Capabilities

The visual search capabilities provide an engaging, useful, and efficient way for information workers to interact with search results.

Document Thumbnails

Word documents and PowerPoint presentations can be previewed directly in search results. A thumbnail image is displayed along with the search results to provide rapid recognition of information. This feature is part of the Search Core Results Web Part for FAST Search Server 2010 for SharePoint, and the feature can be configured in that Web Part.

Scrolling PowerPoint previews

The PowerPoint document preview enables an information worker to browse the actual slides in the presentation.

Visual Best Bets

SharePoint Server 2010 Search keywords can have definitions, synonyms, and Best Bets associated with them. FAST Search Server 2010 for SharePoint adds the ability for you to define Visual Best Bets for keywords.

These visual search elements are unique to FAST Search Server 2010 for SharePoint.

Conversational Search Capabilities

The conversational search capabilities provide ways for information workers to interact with and refine their search results, so that they can quickly find the information they require.

Sort Results on Managed Properties

With FAST Search Server 2010 for SharePoint, users can sort results on any managed properties, such as sorting by Author, Document Size, or Title. Relevance ranking profiles can also be surfaced as sorting criteria, allowing end users to pick a different relevance ranking as desired.

This sorting is considerably more powerful than sorting in SharePoint Server 2010 search. By default, SharePoint Server 2010 sorts results on each document's relevance rank. Information workers can re-sort the results by date modified, but these are the only two sort options in SharePoint Server 2010.

Deep Results Refinement

Refinement with FAST Search Server 2010 for SharePoint is considerably more powerful than refinement in SharePoint Server 2010.

SharePoint Server 2010 automatically generates 'shallow' refinement for search results that enable a user to apply additional filters to their search results based on the values returned by the query. 'Shallow' refinement is based on the managed properties returned from the first 50 results by the original query.

FAST Search Server 2010 for SharePoint enables you to specify whether a managed property can be used in a 'shallow' or 'deep' refinement. 'Deep' refinement is based on statistical aggregation of managed property values within the entire result set; 'shallow' refinement is just based on, by default, the first 50 results returned by the query. Using 'deep' refinement, you can find exactly what you are looking for, such as a person who has written a document about a subject, even if this document would otherwise appear down the result list. 'Deep' refinement can also display counts, and lets the user see the number of results in each refinement category.

You can also use the statistical data returned for numeric refinements in other types of analysis.

Similar Results

With FAST Search Server 2010 for SharePoint, results returned by a query include links to 'Similar Results'. When a user clicks on the link, the search is re-defined and re-run to include documents that are similar to the result in question.

Result Collapsing

FAST Search Server 2010 for SharePoint documents that have the same checksum stored in the index will be collapsed as one document in the search result. This means that documents stored in multiple locations in a source system would only be displayed once during search with usage of the collapse search parameter. Collapsed results include links to 'Duplicates'. When a user clicks on the link, the search result displays all versions of this document.

Similar results and result collapsing are unique to FAST Search Server 2010 for SharePoint and are not provided in SharePoint Server 2010 search.

Contextual Search Capabilities

FAST Search Server 2010 for SharePoint allows you to associate Best Bets, Visual Best Bets, document promotions, document demotions, site promotions, and site demotions with defined user contexts in order to personalize the experience for information workers. You can use the FAST Search User Context link in the Site Collection Settings pages to define user contexts for these associations.

Relevancy Tuning by Document or Site Promotions

SharePoint Server 2010 enables you to identify varying levels of authoritative pages that help you tune relevancy ranking by site. FAST Search Server 2010 for SharePoint adds the ability for you to specify individual documents within a site for promotion, and furthermore enables you to associate each promotion with user contexts.

Synonyms

SharePoint Server 2010 keywords can have one-way synonyms associated with them. With one-way synonyms when a query includes a synonymous term for a keyword, items that contain the keyword are returned. However, if a query includes the keyword, then items that contain the synonymous terms are not returned (unless they also contain the keyword).

FAST Search Server 2010 for SharePoint extends synonyms by enabling you to implement both two-way and one-way synonyms. With two-way synonyms, when a query includes a synonymous term for a keyword, items that contain the keyword are returned (just as for one-way synonyms). Furthermore, if a search expression includes the keyword, then items that contain the synonymous terms are returned, regardless of whether they also contain the keyword.

Managed Properties and Metadata creation

SharePoint Server 2010 enables you to create metadata property mappings (known as managed properties). FAST Search Server 2010 for SharePoint adds the ability for you to:

- Enable stemming support and word forms for managed property values, when they are used by an information worker in a query.
- Choose between static and dynamic summaries for display in search results. A dynamic summary will only display a hit-highlighted summary of the specific managed property in the result.
- Define whether information workers can sort results on the managed property in question.
- Define whether information workers can use the managed property in query operators or filters.
- Define whether search results page can use the managed property as a query refiner or deep query refiner in results pages.
- Define the priority associated with the managed property. The priority is one of the inputs into the ranking algorithm, and defines how documents with the search term in this property should be ranked against other documents that may have the search term in other properties.
- Define how managed properties can be grouped into one or more full-text search-enabled indexes.

Property Extraction

Property extraction identifies key information such as people, companies, and locations in documents. The properties can then be used to enhance the search experience, for instance by providing search result refinement based on the properties. You can

improve the precision of the property extraction by editing the include lists and exclude lists for each property extractor. Excluded items are removed immediately, while included items take effect the next time the content is indexed. You can also create custom property extractors based on your organization's specific content using Windows PowerShell and SharePoint administration. Dictionary or Taxonomy based extractors, also called verbatim extractors, will allow you to extract managed properties based on a fixed list of known terms. Developers will also be able to create more dynamic extractors based on the FAST matcher framework and will be able to extend the document pipeline with specialized classifiers, entity extractors, or other processing can be used to support specialized scenarios.

Rank Profiles

The index schema in FAST Search Server 2010 for SharePoint includes rank profiles, which control how relevancy ranking is calculated for each item in search results.

A rank profile defines how relevancy calculations are performed when you search a full-text index. A rank profile consists of several components which are weighted when calculating an item's relevance. You can adjust the weights of a profile's components to improve search result relevance. Rank profile components include the following:

- Freshness. This component manages how the age of an item affects rank.
- Proximity. This component manages how the distance between query terms affects rank.
- Authority. This component manages how links between Web documents affect rank.
- Query authority. This component manages how user selections in previous query results affect rank.
- Context. This component manages how different managed properties within the associated full-text indexes contribute to the rank.
- Managed properties directly impacting the rank. You can specify that the value of a numeric managed property is added to the rank, or you can specify that certain values of a managed property impact the rank. In the latter case you can, for example, define that documents of a given type (as defined by a specific managed property) will get a relevancy boost in the results.

- You can tailor different rank profiles to different use cases, or you can enable advanced information workers to select different rank profiles for different queries.

Custom Rank Profiles are created with Windows PowerShell cmdlets. Refer to the [FAST Search Server 2010 for SharePoint Windows PowerShell Cmdlet Overview](#) and the [FAST Search Server 2010 for SharePoint Windows PowerShell Cmdlet Help](#) guides for more details.

Linguistics

In search, linguistics is defined as the use of information about the structure and variation of languages so that users can more easily find relevant information. The item's relevancy with regard to a query is not necessarily decided based on words common to both query and document, but instead depends on the extent that its content satisfies the user's need for information. Examples of linguistic processing in the item and query processing include character normalization, normalization of stemming variations and suggested spelling corrections. FAST Search Server for SharePoint 2010 performs linguistic processing for items returned by the crawl process before those items are indexed, as well as for the queries before the actual matching occurs.

FAST Search Server 2010 for SharePoint for End Users

This section provides information about how the search experience is blended with the SharePoint experience, and how it has been enhanced for end users when they use FAST Search Center.

FAST Search Server 2010 for SharePoint User Experience

Figure 1 shows the search user interface provided by the FAST Search Center site template. Note that the user interface looks and feels very similar to Search Center sites in SharePoint Server 2010. This provides information workers with a consistent search experience.

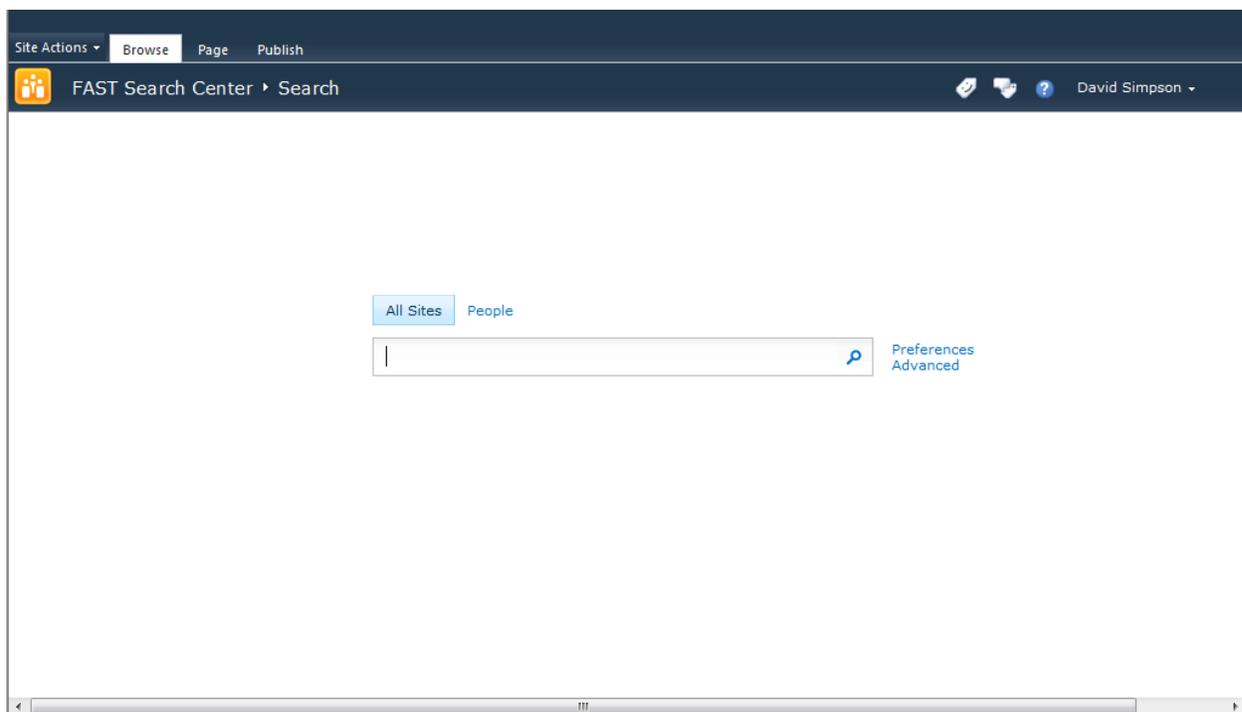


Figure 1. Consistent Search User Interface

Figure 2 shows the advanced search user interface provided by the FAST Search Center site template. Note that the advanced user interface also looks and feels similar to that provided by Search Center sites in SharePoint Server 2010.

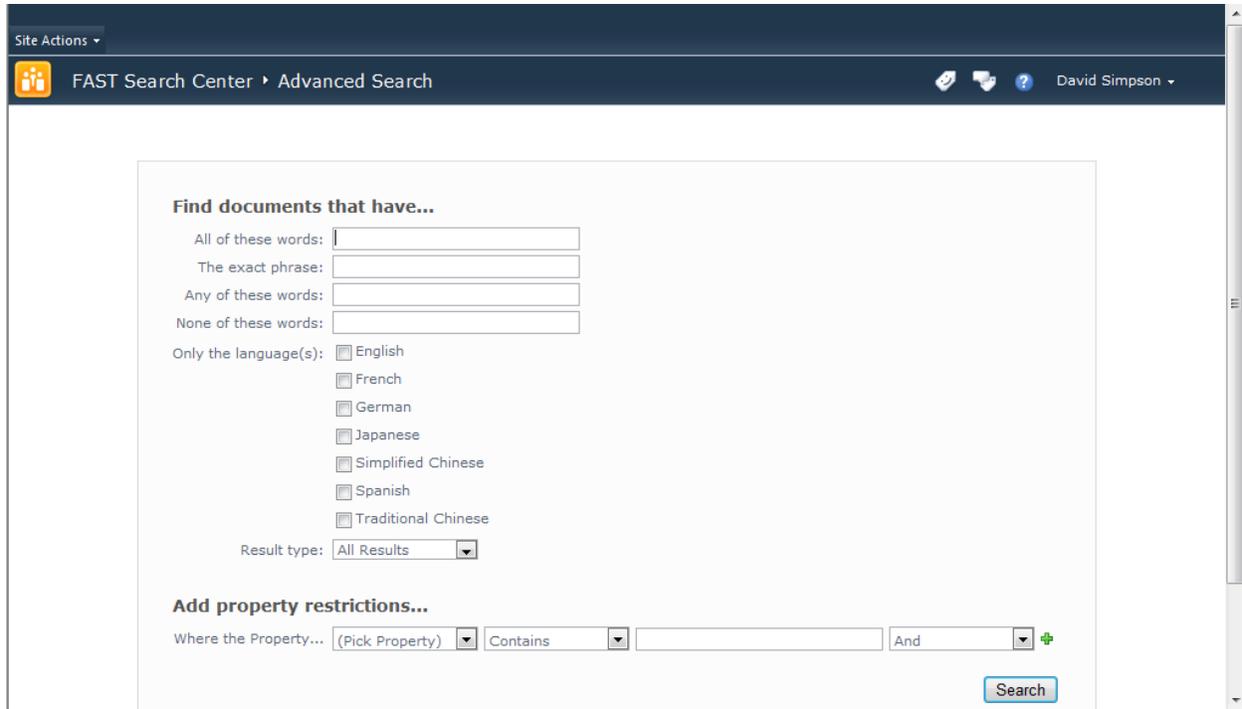


Figure 2. Consistent Advanced Search

Figure 3 shows the user preferences provided by the FAST Search Center site template. Note that the user preferences interface also looks and feels similar to that provided by Search Center sites in SharePoint Server 2010. This provides information workers with a consistent search experience.

Search Suggestions

Show search suggestions as you type in the search box (if enabled by administrator).

Show search suggestions in the search box

Search Language

Choose whether search should prefer certain languages. The languages you select will appear in a dropdown on the results page. Search will use this language to understand your keywords and give preference to results of the language.

Search using my browser's language

Search using the following languages (maximum of five):

| | | |
|--|--|---|
| <input type="checkbox"/> Arabic | <input type="checkbox"/> Hindi | <input type="checkbox"/> Punjabi |
| <input type="checkbox"/> Bengali | <input type="checkbox"/> Hungarian | <input type="checkbox"/> Romanian |
| <input type="checkbox"/> Bulgarian | <input type="checkbox"/> Icelandic | <input checked="" type="checkbox"/> Russian |
| <input type="checkbox"/> Catalan | <input type="checkbox"/> Indonesian | <input type="checkbox"/> Serbian (Cyrillic) |
| <input type="checkbox"/> Chinese (Simplified) | <input type="checkbox"/> Italian | <input type="checkbox"/> Serbian (Latin) |
| <input type="checkbox"/> Chinese (Traditional) | <input type="checkbox"/> Japanese | <input type="checkbox"/> Slovak |
| <input type="checkbox"/> Croatian | <input type="checkbox"/> Kannada | <input type="checkbox"/> Slovenian |
| <input type="checkbox"/> Czech | <input type="checkbox"/> Korean | <input type="checkbox"/> Spanish (Spain) |
| <input type="checkbox"/> Danish | <input type="checkbox"/> Latvian | <input type="checkbox"/> Spanish (Mexico) |
| <input type="checkbox"/> Dutch | <input type="checkbox"/> Lithuanian | <input type="checkbox"/> Swedish |
| <input checked="" type="checkbox"/> English | <input type="checkbox"/> Malay | <input type="checkbox"/> Tamil |
| <input type="checkbox"/> Finnish | <input type="checkbox"/> Malayalam | <input type="checkbox"/> Telugu |
| <input type="checkbox"/> French | <input type="checkbox"/> Marathi | <input type="checkbox"/> Thai |
| <input checked="" type="checkbox"/> German | <input type="checkbox"/> Norwegian | <input type="checkbox"/> Turkish |
| <input type="checkbox"/> Greek | <input type="checkbox"/> Polish | <input type="checkbox"/> Ukrainian |
| <input checked="" type="checkbox"/> Gujarati | <input type="checkbox"/> Portuguese (Brazil) | <input type="checkbox"/> Urdu |
| <input type="checkbox"/> Hebrew | <input type="checkbox"/> Portuguese (Portugal) | <input type="checkbox"/> Vietnamese |

Default language:

Figure 3. Consistent User Preferences

Enhanced Search Results

Figure 4 shows a typical results page in a FAST Search Center.

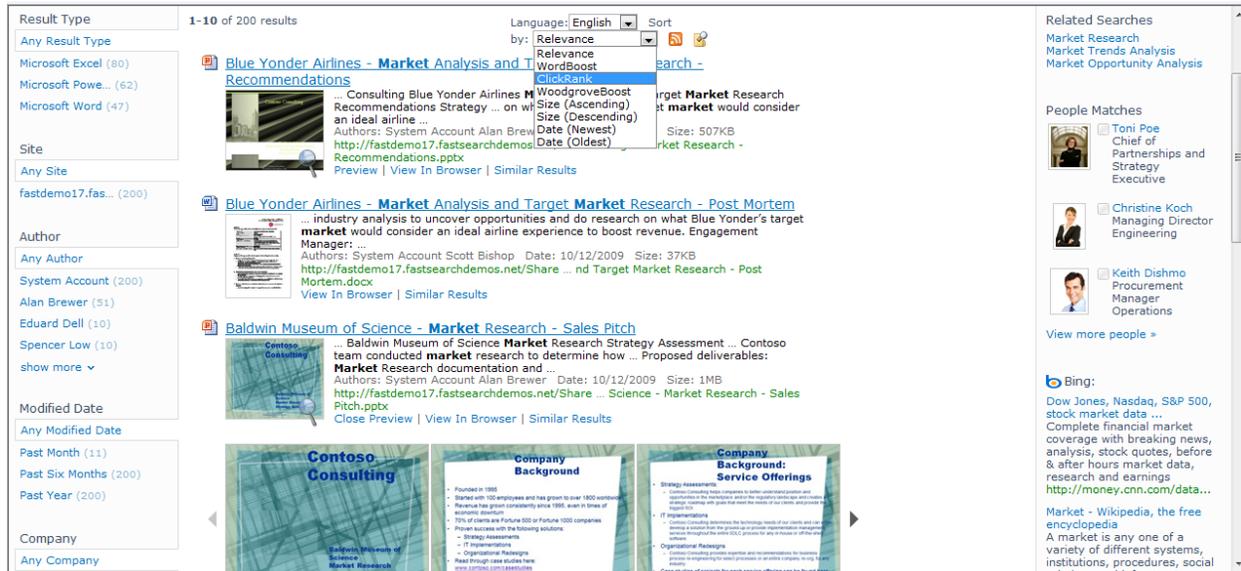


Figure 4. Enhanced Search Results Page

Notice how the overall results experience is very similar to, but more powerful than the results provided by SharePoint Server 2010. The refiners are to the left, the search results are down the middle, and the federated results (from People Profiles and other repositories) are on the right, along with related searches. However, there are some key capabilities that are only available with FAST Search Server 2010 for SharePoint. The most noticeable difference is the document previews and thumbnails for the search results, and the Visual Best Bet at the top of the results. Microsoft Word and Microsoft PowerPoint® documents have thumbnails, and PowerPoint documents can be previewed in the results. Along the left hand side, all of the refiners have exact counts. This reflects the total number of documents in the entire result set that contain this value. Additionally, you can see how results can be sorted by managed properties and ranking profiles.

Figure 5 shows another results page that includes a Visual Best Bet.

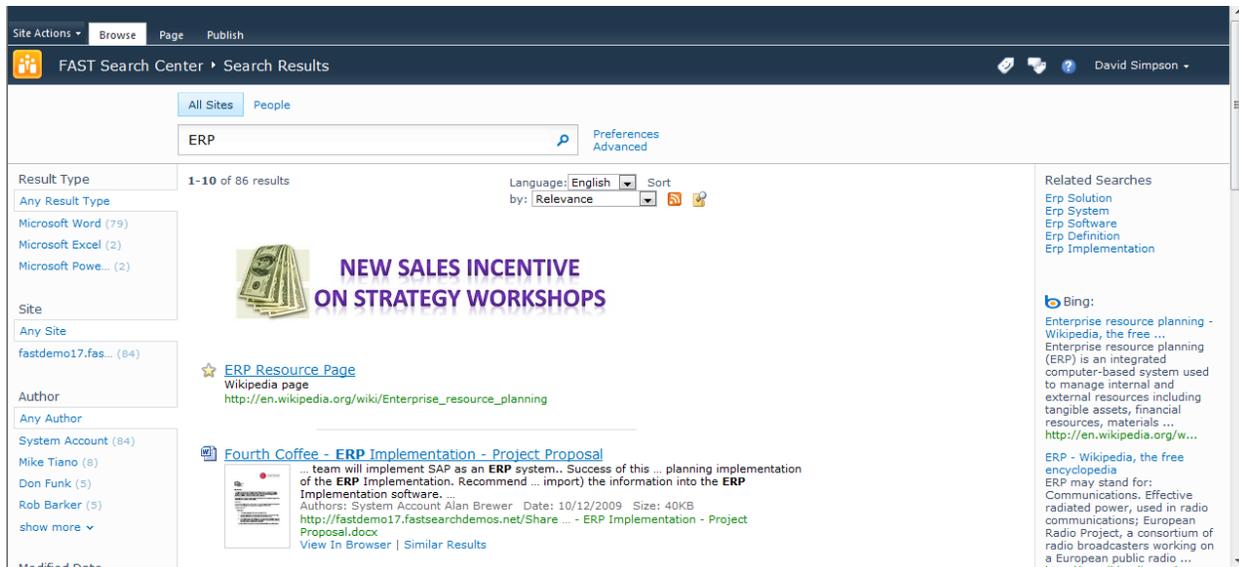


Figure 5. Search Results with a Visual Best Bet

Social Search

A significant aspect to people's work in an organization is interacting with other people and finding the right people to connect with who have specific skills and talents. This can be a daunting challenge in a large organization. FAST Search Server 2010 for SharePoint addresses this challenge through search, and connects this search to the social capabilities in SharePoint Server 2010. A people search center provides specific capabilities for connecting with people. FAST Search Server 2010 for SharePoint provides a user experience for social search that is consistent with SharePoint Server 2010.

Finding People

FAST Search Server 2010 for SharePoint provides an address book-style name lookup experience with better name matching, making it easier to find people by name, title and organizational structure. This includes phonetic name matching that will return names that sound similar to what the user has typed in a query. It will also return all variations of common names, including nicknames (for supported languages).

The refiners provided on the core search results are also provided with people search results—exploring results via name, title, and various fields in a user's profile enable quick browsing and selection of people. People search results also include real-time presence through Microsoft Office Communications Server, making it easy to

immediately connect with people once they are found through search. Figure 6 shows a People Search results page.

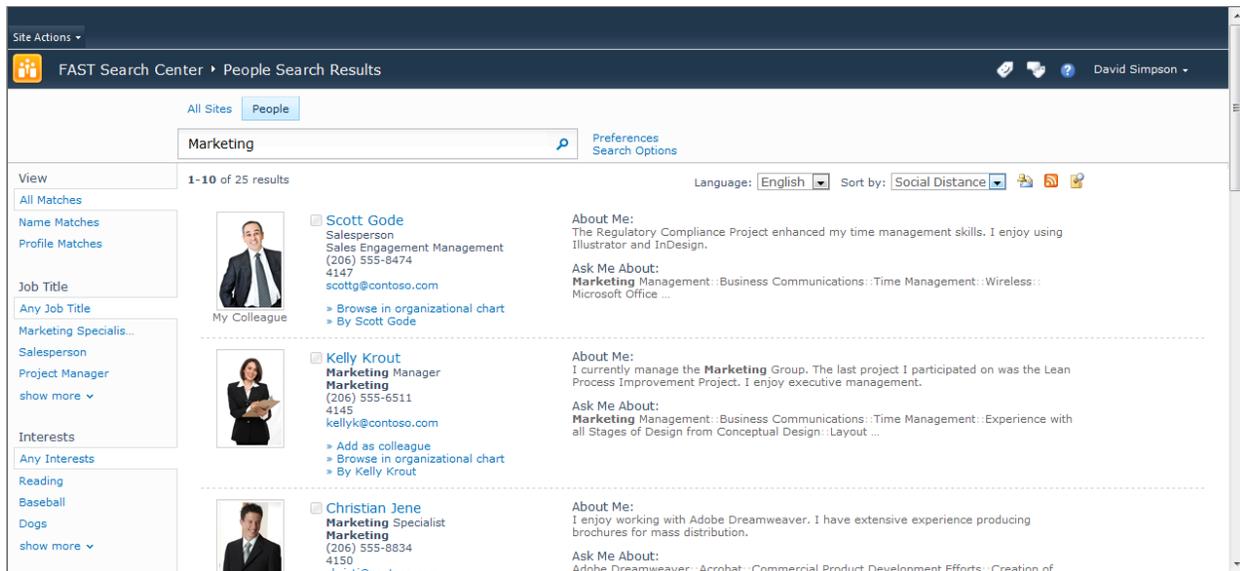


Figure 6. People Search Result Page

Mining and Discovering Expertise

Users can manually submit or automatically generate a list of colleagues mined from Outlook®. Automatically generated lists of colleagues are a way of rapidly inferring social relationships throughout the organization, which speeds the adoption and usefulness of people search results. FAST Search Server 2010 for SharePoint also infers expertise by automatically suggesting topics mined from the user’s Outlook inbox and suggesting additions to their expertise profile in their My Site. This makes it easy to populate My Site profiles and means that more people have well-populated profiles and get the benefits of this in both search and communities.

Improving Search based on Social Behavior

For many organizations, SharePoint sites have become gathering places where people create, share and interact with information. Social behavior is taken into account in order to provide high quality search results in several ways. The relevance ranking for people search takes social distance into account. A direct colleague will appear before someone three degrees removed. Second, FAST Search Server 2010 for SharePoint supports social tagging of content, and this feedback can influence the relevance of content in search results. People’s day-to-day usage of information in FAST Search Server 2010 for SharePoint and Microsoft Office can have a measurable impact on

search relevance, thereby helping the organization harness the collective wisdom of its people.

Custom Search Solutions

FAST Search Server 2010 for SharePoint is also the ideal platform for developing search-based applications and custom solutions. Figure 7 illustrates an example of a Customer Relationship Management application based on search. In this scenario, a sales executive can find related content from previous engagements with a customer or with a particular service offering that is very similar to something in her pipeline. This helps the organization secure new business by leveraging the organization's past successes.

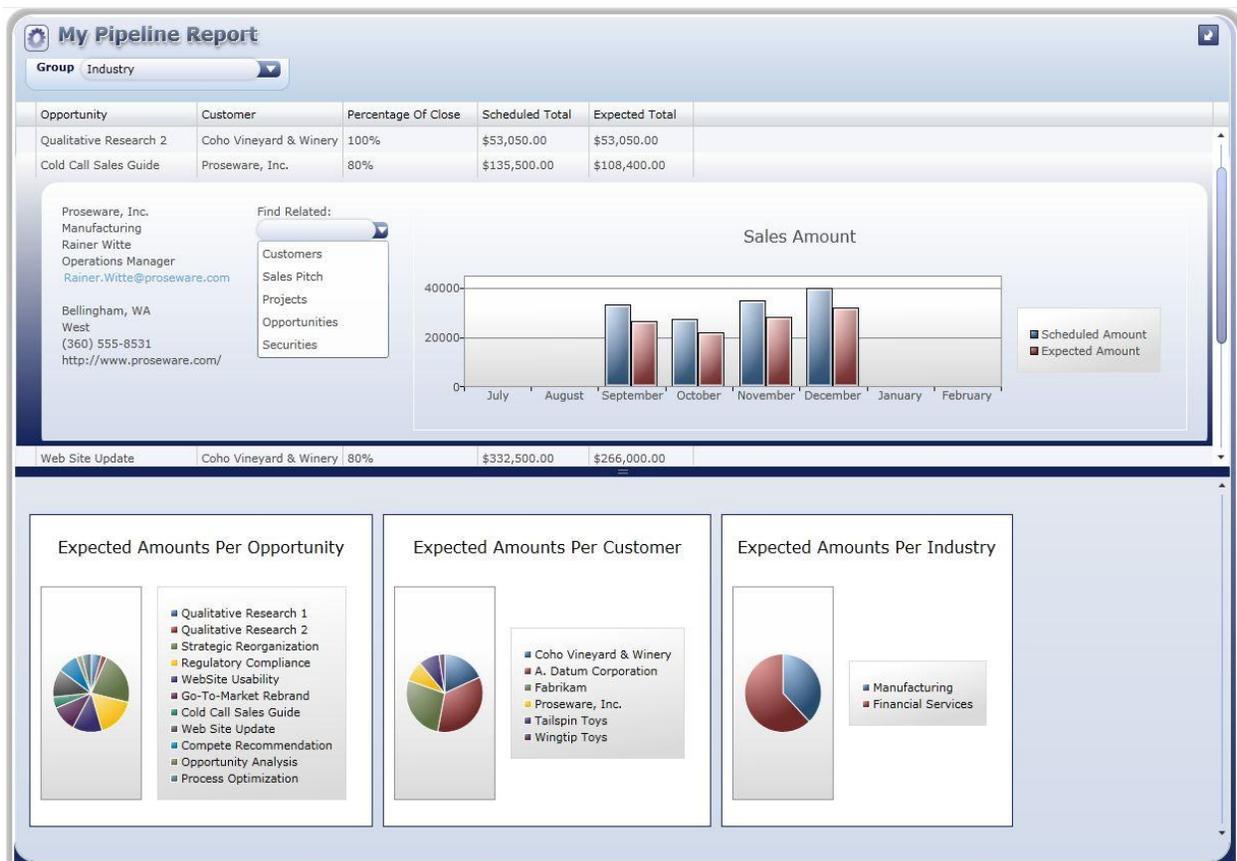


Figure 7. Search-based Applications

FAST Search Server 2010 for SharePoint for IT Professionals

FAST Search Server 2010 Architecture Overview

Figure 8 shows the FAST Search Server 2010 for SharePoint system as part of an overall SharePoint deployment.

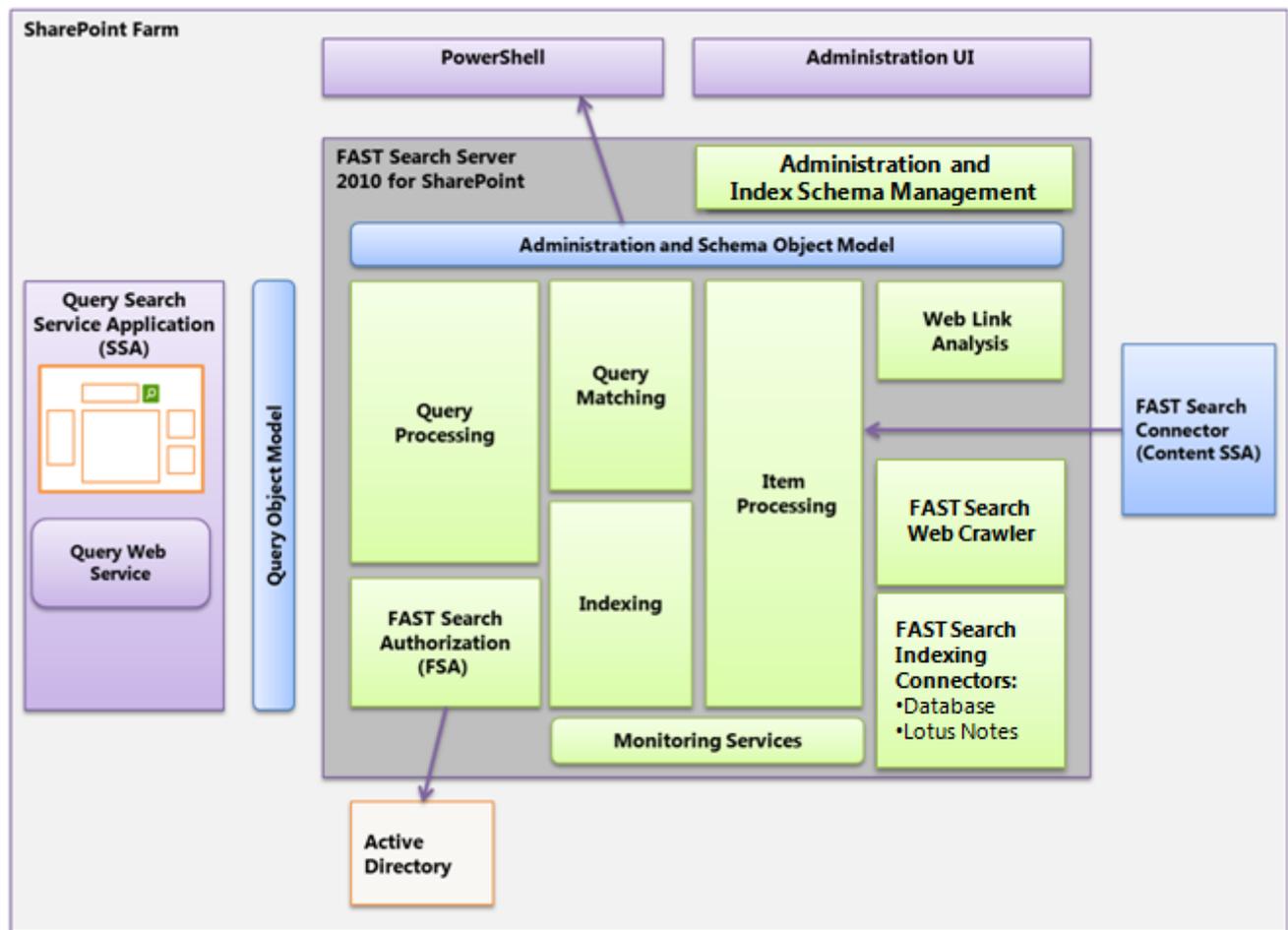


Figure 8. FAST Search Server 2010 for SharePoint in an Overall SharePoint Farm Architecture

- **FAST Search Server 2010 for SharePoint.** The server infrastructure that provides processing, indexing, and query capabilities, and the administration of these features.
- **FAST Query Search Service Application.** The Query Search Service Application provides the query Web front-end capabilities. You install and deploy this Search

service application on SharePoint Web servers, and it hosts Search Web Parts and customized query integration capabilities for your search solution.

- **FAST Search Connector Content Search Service Application.** The Content Search Service Application retrieves content for indexing from SharePoint farms and other content repositories. You install and deploy the Content Search Service Application on a SharePoint application server. The Content Search Service Application includes an indexing connector that can retrieve content from a variety of external repositories, including SharePoint farms, internal and external Web servers, Exchange public folders, line-of-business data and file shares. FAST Search Server 2010 for SharePoint also provides a set of additional indexing connectors for advanced content retrieval use cases. Configuration of the additional indexing connectors is performed via XML files and through Windows PowerShell cmdlets and then operated by command-line operations, as opposed to the other supported indexing connectors integrated in SharePoint Server, which are administered directly from Central Administration.

Modular and Scalable Architecture

FAST Search Server 2010 for SharePoint is built on a highly modular architecture where the services can be scaled individually to achieve the desired performance with respect to:

- Amount of indexed content. By partitioning into multiple indexes, you can index over a billion documents within a single farm.
- Query load. You can scale the query matching components in a row/column matrix where the columns reflect the index partitioning, and the rows add query performance and fault-tolerance for query evaluation.
- Freshness (indexing latency). FAST Search Server 2010 for SharePoint enables you to optimize for low latency from the moment a document is changed in the source repository to the moment it is searchable. This can be done by proper dimensioning of the crawling, item processing, and indexing to fulfill your requirements. These three parts of the system can be scaled independently through the modular architecture.

Integration with Search Center

When you create a search site based on the FAST Search Center template, the resulting site will look and feel very similar to a SharePoint Server 2010 Search Center site.

However, the Web Parts that are included are extended versions of the SharePoint Server 2010 Search Web Parts. They include the additional capabilities provided by FAST Search Server 2010 for SharePoint, such as sorting by managed properties, previewers, deep refiners, similar results links, and so on.

FAST Search Centers, like those of SharePoint Server 2010, also include people search capabilities. In fact, people search is provided natively by SharePoint Server 2010, and FAST Search Server 2010 for SharePoint simply federates people search into the Search Center user interface.

IT Professional Experience

IT professionals work with FAST Search Server 2010 for SharePoint in a similar way to the way that they administer the enterprise search features of SharePoint Server 2010. For example, you can use the Site Settings page to administer FAST Search keywords, FAST Search site promotion and demotion, and FAST Search user contexts, much like you would administer SharePoint Server 2010 keywords at the site collection level. Furthermore, you can use Central Administration to work with FAST managed properties much like you would administer SharePoint Server 2010 at the Search Service Application level.

Windows PowerShell Support

Windows PowerShell is a powerful command-line shell and scripting language that helps you perform administrative tasks. FAST Search Server 2010 for SharePoint includes over 80 Windows PowerShell cmdlets.

Windows PowerShell cmdlets for FAST Search Server 2010 for SharePoint are specialized .NET classes that implement specific system administration actions in the areas of search management, index schema management, security, and administration. Typical tasks include uploading a custom dictionary, mapping crawled properties to managed properties, and customizing the index schema.

FAST Search Server 2010 for SharePoint uses a master Windows PowerShell snap-in called Microsoft.FASTSearch.PowerShell.dll. This DLL includes all the code for over 80 cmdlets.

Categories of cmdlets

FAST Search Server 2010 for SharePoint cmdlets cover four basic areas:

- Administration
- Security
- Index schema
- Spell tuning

For more information, refer to the overview page for [FAST Search Server 2010 for SharePoint](#).

Developer Experience

FAST Search Server 2010 for SharePoint Web Parts use the same unified object model as SharePoint Server 2010 and the other search platforms from Microsoft. The result is that if you develop a custom solution that uses the query object model for SharePoint Server 2010, for example, then it will continue to work if you migrate your code to FAST Search Server 2010 for SharePoint.

FAST Search Server 2010 for SharePoint —Services and Components

FAST Search Server 2010 for SharePoint consists of a number of different services and components. The following sections describe the services and components.

FAST Search Connector

The FAST Search Connector enables you to retrieve content for indexing from SharePoint farms, public Exchange folders, Web sites, databases and file shares.

You install and deploy the FAST Search Connector as a Search Service Application (SSA) on a SharePoint application server, and it shares the same crawler/connector framework as SharePoint Server 2010. This includes the Business Data Catalog functionality, which is substantially improved from Office SharePoint Server 2007.

FAST Search Server 2010 for SharePoint also includes additional indexing connectors for more targeted content retrieval needs. This includes connectors for database access (JDBC), Lotus Notes, and a dedicated Web crawler.

Web Link Analysis (Web Analyzer)

The Web Analyzer has two main functions: It analyzes search click-through logs and hyperlink structures. Both contribute to better-ranked search results. Items that show

many clicks in the search click-through log are popular and therefore receive better rank scores than less-viewed items. Items that are linked to from many other items are also perceived to be more relevant for the user and therefore receive better rank scores.

The Web Analyzer scales up to many nodes to reduce the total time that is needed for the analysis.

Item Processing

The item processing service receives items to be indexed from indexing connectors. The item processing service extracts content from source documents in various formats, discovers and sets managed properties, and performs linguistic processing on the content. The item processing service then sends the processed items to the indexing service.

Key features of the item processing service are as follows:

- Mapping from crawled properties to managed properties. Managed properties contain the content that will be indexed including metadata associated with the items. You will first perform a crawled property discovery based on an initial set of crawled items. Based on this discovery you can change the mapping to managed properties.
- Parsing of document formats such as Office and PDF. This includes extracting searchable text and metadata from these formats.
- Extracting properties from the retrieved content. The property extraction can detect various properties such as names and dates from the documents, and maps them into managed properties. In this manner, you can query these properties, and also change query refinement based on these properties. It is also possible to create custom property extractors using, for example, a dictionary of product names relevant to your organization.
- Linguistic processing of items before indexing. In search, linguistics is defined as the use of information about the structure and variation of languages so that users can more easily find relevant information. The item's relevancy with regard to a query is not necessarily decided based on words common to both query and document, but instead the extent that its content satisfies the user's need for

information. Examples of linguistic processing in the item processing include character normalization and normalization of stemming variations.

Indexing

The indexing service creates searchable indexes based on the processed items. No external database is needed for the indexing or the generated indexes.

The indexing service supports a seamless content partitioning into index columns in order to handle large content volumes.

The indexing is controlled by an index schema object model, which defines the mapping of managed properties into the searchable index structures, and how relevancy ranking is to be performed.

Query Matching

The query matching service uses the indexes created by the indexing service to retrieve the items that match a query and then return these items as a query result set.

A query usually contains several terms combined with query operators, such as AND and OR. The query matching service looks up each term in the index and retrieves a list of items in which that term appears. The order of the returned items is based on the requested sorting mechanism, which is usually the relevance ranking that is calculated from various item properties, or a sort based on one or more of the item properties.

The query matching service can also return a hit-highlighted summary for each item in the query hit list. A hit-highlighted summary consists of a fragment of the original item in which the matching query terms are highlighted.

The query matching service is responsible for the deep refinement that is associated with query results. Query refinement enables drilling down into a query result by using aggregated statistical data that was computed for the query result. The query matching service maintains aggregation data structures to enable deep refinement across large result sets.

Query Processing

The query processing service performs the processing of queries and results that can be performed without access to the index. Query processing includes query-language parsing, linguistic processing, and item-level security processing. Result processing

includes merging the results from multiple index columns, formatting the query hit list, formatting the query refinement data, and removing duplicates.

It is the responsibility of the query processing service to ensure that the user performing a query sees only the results that he or she is authorized to see. The query processing service therefore validates the user's permissions and rewrites the incoming query with an access filter that corresponds to the current user and group membership.

Connector Framework

SharePoint Server 2010 provides a new framework for connecting to and crawling content repositories. Connectors for SharePoint sites, Web sites, file shares, custom databases and Web services (via Business Connectivity Services), Exchange public folders, and Lotus Notes databases are provided with the product.

New Connector Features

The connector framework provides improvements over the protocol handlers in previous versions of SharePoint Server. For example, attachments as well as the content in e-mail messages can now be crawled. Also, item-level security descriptors can now be retrieved for external data exposed by Business Connectivity Services. Furthermore, when crawling a Business Connectivity Services entity, additional entities can be crawled via its entity relationships. Connectors also perform better than previous versions of protocol handlers, by implementing concepts such as inline caching and batching.

Connectors support richer crawl options than the protocol handlers in previous versions of SharePoint Server. For example, they support the full crawl mode that was implemented in previous versions, and they support timestamp-based incremental crawls. However, they also support change log crawls that can remove items that have been deleted since the last crawl.

Creating Connectors

In previous versions of SharePoint Server, it was very difficult to create protocol handlers for new types of external systems. Protocol handlers were required to be coded in unmanaged C++ code, and typically took a long time to test and stabilize.

With FAST Search Server 2010 for SharePoint, you have many more options for crawling external systems. You can choose to:

- Use SharePoint Designer 2010 to create external content types and entities for databases or Web Services, and then simply crawl those entities
- Use Visual Studio® 2010 to create external content types and entities for databases or Web Services, and then simply crawl those entities
- Use Visual Studio 2010 to create .NET types for Business Connectivity Services (typically for back-end systems that implement dynamic data models, such as document management systems). Then use either SharePoint Designer 2010 or Visual Studio 2010 to create external content types and entities for the .NET type.

NOTE: You can still create protocol handlers (as in previous versions of SharePoint Server) if you need to.

Search Administrator Walkthroughs

The enterprise search features provided by FAST Search Server 2010 for SharePoint can be administered at the site collection level and at the Search Service Application level.

Many of the operations for administering FAST Search Server 2010 for SharePoint are identical to those same operations in SharePoint Server 2010. You should refer to the [SharePoint Server 2010 Enterprise Search Evaluation Guide](#) for walkthroughs that are common to both SharePoint Server 2010 and FAST Search Server 2010 for SharePoint. The following walkthroughs are unique to FAST Search Server 2010 for SharePoint.

Administering FAST Search Server 2010 for SharePoint Managed Properties

This section provides information for how administrators work with FAST Search Server 2010 for SharePoint settings at the Central Administration level. Use the following procedure to start administering managed properties for FAST Search Server 2010 for SharePoint.

Creating Managed Properties with FAST Search Server 2010 for SharePoint

1. Click **Start>All Programs>Microsoft SharePoint 2010 Products>SharePoint 2010 Central Administration**.
2. On the Quick Launch, click **General Application Settings**.
3. In the **Search** section, click **Farm Search Administration**.
4. In the **Search Service Applications** section, click **FAST Search Query**.

5. On the Quick Launch, in the **Administration** section, click **FAST Search Administration**.

The FAST Search Administration page appears.

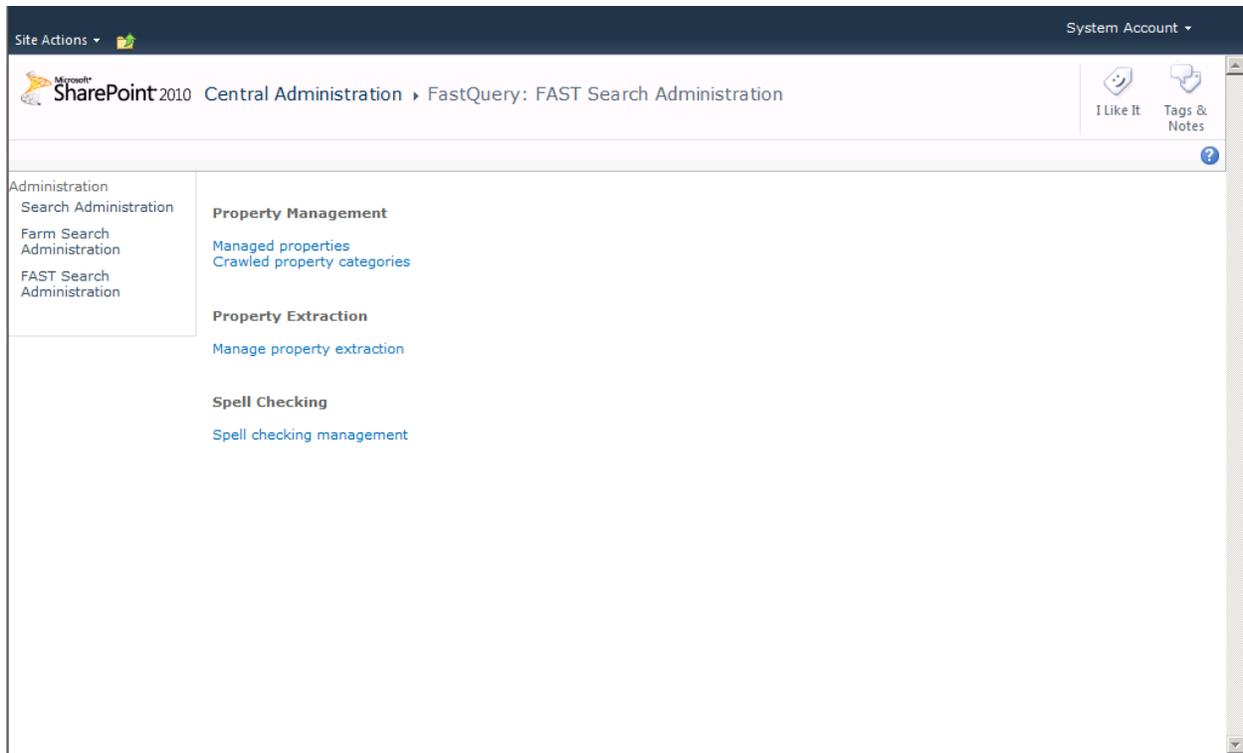


Figure 9. FAST Search Administration

6. On the **Fast Search Administration** page, click **Managed properties**.
7. Click **Add Managed Property**.
The **New Managed Property** page appears.

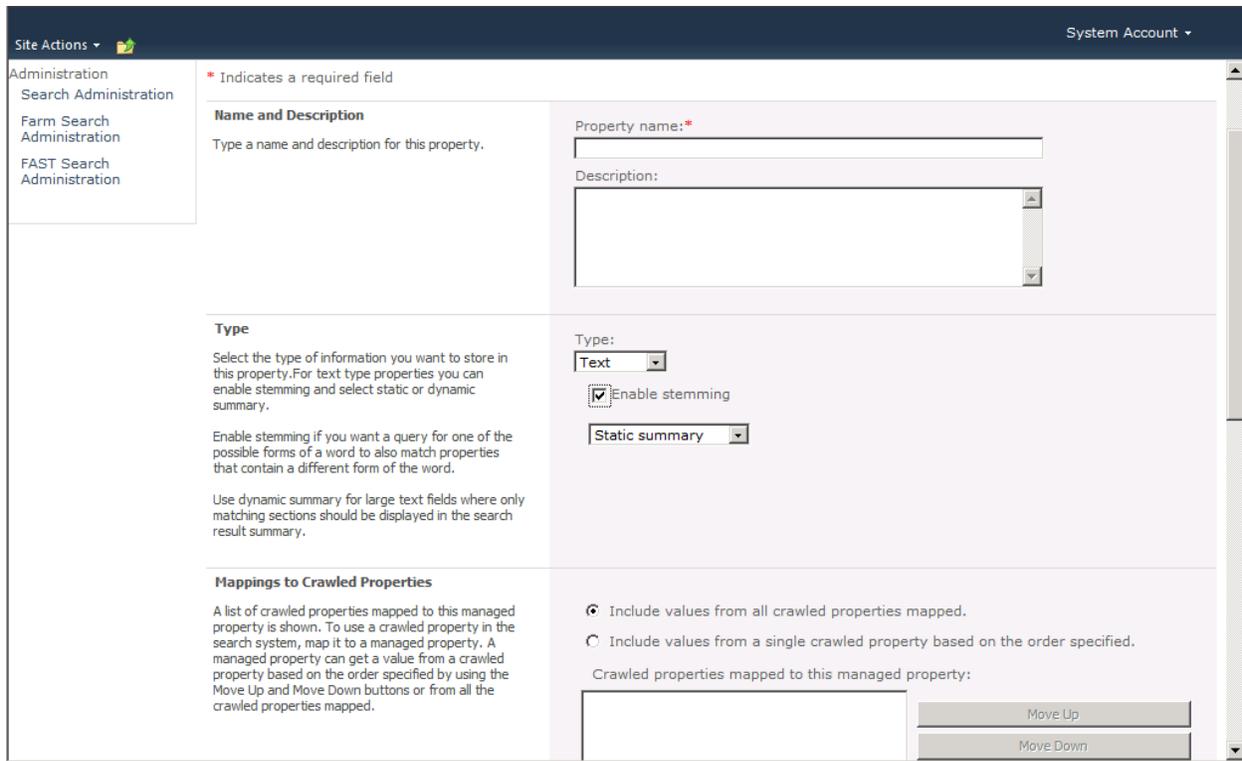


Figure 10. New Managed Property Page

FAST managed properties include the stemming and summary behaviors that are not provided by SharePoint Server 2010 managed properties.

FAST managed properties include the ability for you to specify whether the managed property can be used as a sort by field and whether it can be used in as a filter term in search queries. You can also specify whether the managed property should be represented as refiner field. If so, you can also specify whether it is a deep refiner or just a shallow refiner. Furthermore, you can specify the priority level for the managed property. The priority is one of the inputs into the ranking algorithm, and defines how documents with the search term in this property should be ranked against other documents that may have the search term in other properties.

Figure 11 shows the user interface for working with all of these attributes

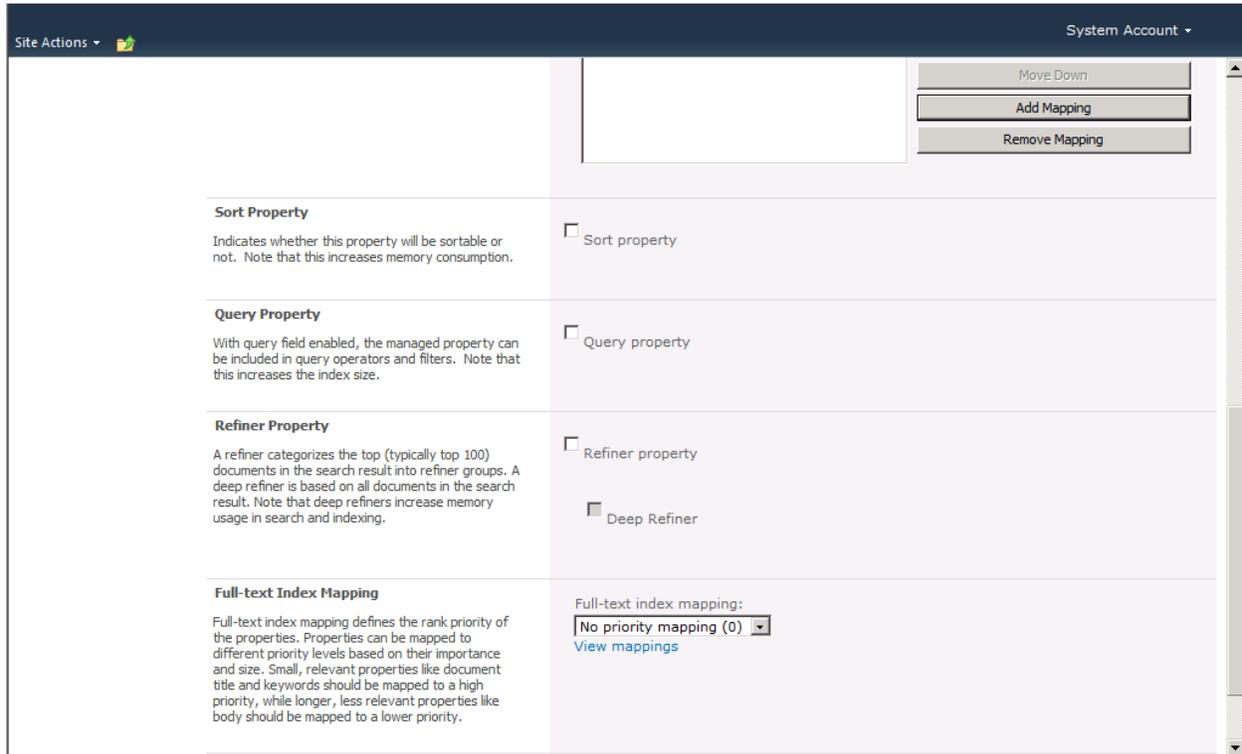


Figure 11. Sorting, Query Filters, and Refinement Attributes

Search Administration at the Site Collection Level

Administrators can use the site collection administration pages to define keywords, Best Bets, synonyms, and definitions. Administrators can also use the site collection administration pages to define search scopes.

NOTE: Any settings created or modified at the site collection level affect only that site collection.

Creating FAST Search Centers

NOTE: The following procedure creates a FAST Search Center at the root Web for a site collection. This is the generally recommended approach and architecture for creating Search Center sites with FAST Search Server 2010 for SharePoint.

1. Click **Start>All Programs>Microsoft SharePoint 2010 Products>SharePoint 2010 Central Administration**.
2. In the **Application Management** section, click **Create site collections**.

3. In the **Web Application** section, use the Web Application changer to select the Web application where you want to create the Search Center.
4. In the **Title** text box, type **FAST Search Center**.
5. In the **Description** text box, type **FAST Search Center for SharePoint 2010**.
6. In the **Web Site Address** section, select **/sites/** in the drop-down list, and then type **fastsearch** in the text box.
7. In the **Template Selection** section, click the **Enterprise** tab.
8. Click **FAST Search Center**.
Note: Do not select *Basic Search Center*, because this template does not include tabs and people search features. Also, do not select **Enterprise Search Center**, because this is the SharePoint Server 2010 enterprise search center, rather than that of FAST Search Server 2010 for SharePoint.
9. In the **Primary Site Collection Administrator** section, type your name in the text box, and then click **Check Names**.
10. Click **OK**.
After a short period of time, the site collection is created and the **Top-Level Site Successfully Created** page appears.
11. Click the hyperlink to the new site collection to start exploring FAST Search Center.

FAST Site Collection Settings

This section provides information for how site collection administrators work with FAST Search Server 2010 for SharePoint settings.

NOTE: Any settings created or modified at the site collection level affect only that site collection.

Exploring FAST Search Settings at the Site Collection Level

As a site collection administrator, you can use the site settings page to start creating FAST Search keywords, site promotions and demotions, and user contexts. Figure 12 shows links to the FAST Search settings in a site collection settings page.

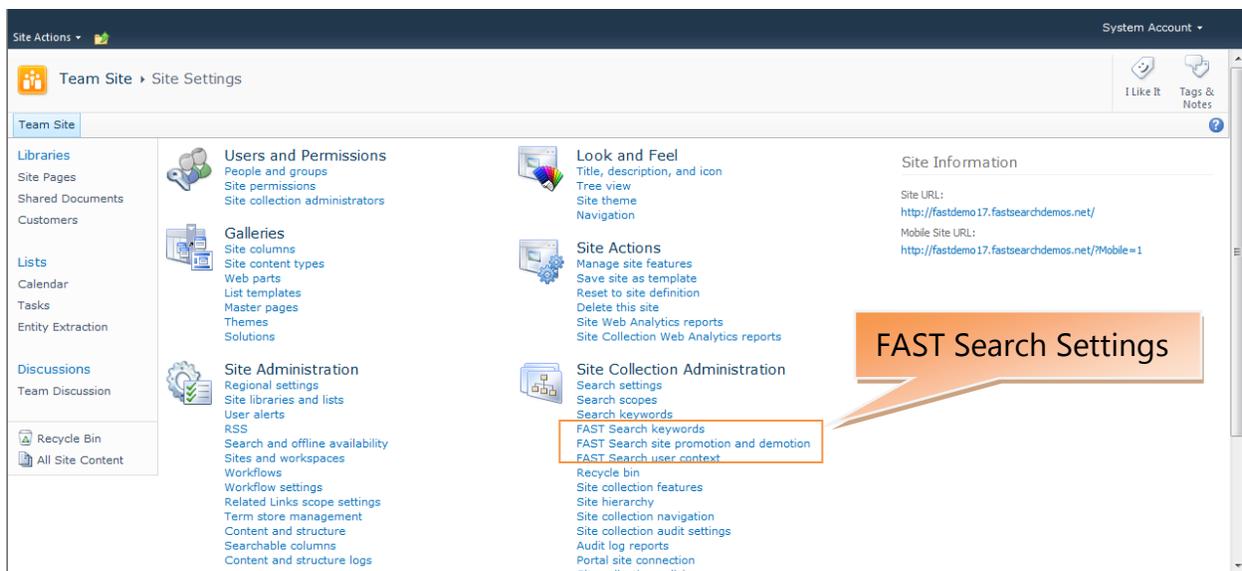


Figure 12. FAST Site Collection Settings

FAST Search Keywords

Figure 13 shows the main page for administering FAST Search keywords.

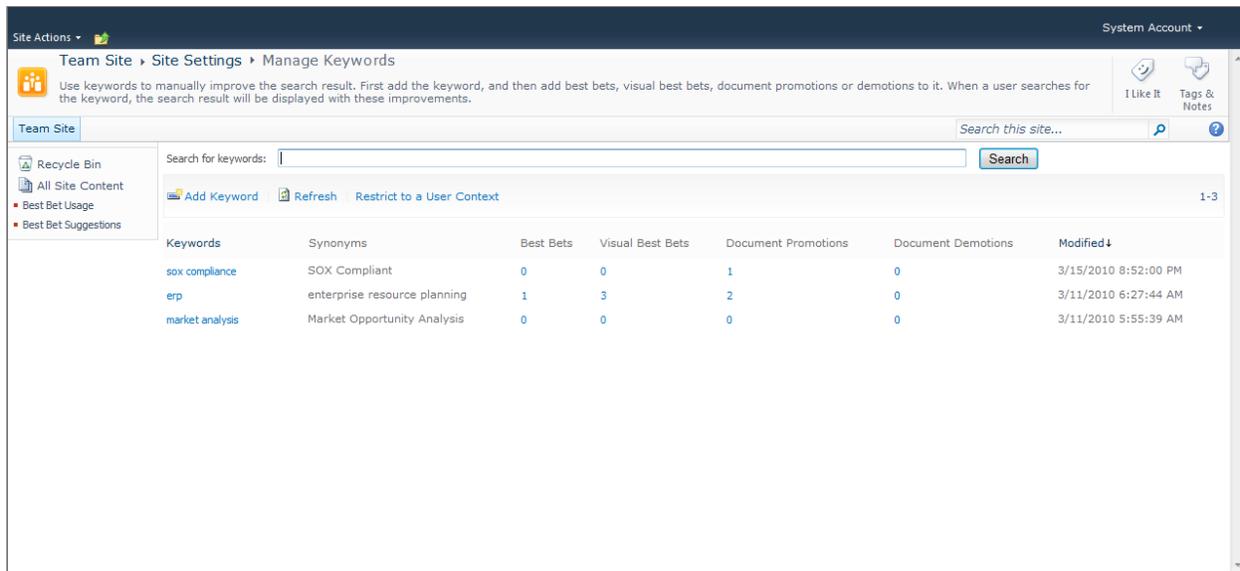


Figure 13. FAST Search Keywords

FAST Search keywords are more powerful than their SharePoint Server 2010 counterparts.

You can create Best Bets, Visual Best Bets, Document Promotions, and Document Demotions for a keyword, and you can also restrict a keyword to a specific user context.

Figure 14 shows the FAST Search keyword creation process, including one-way and two-way synonyms.

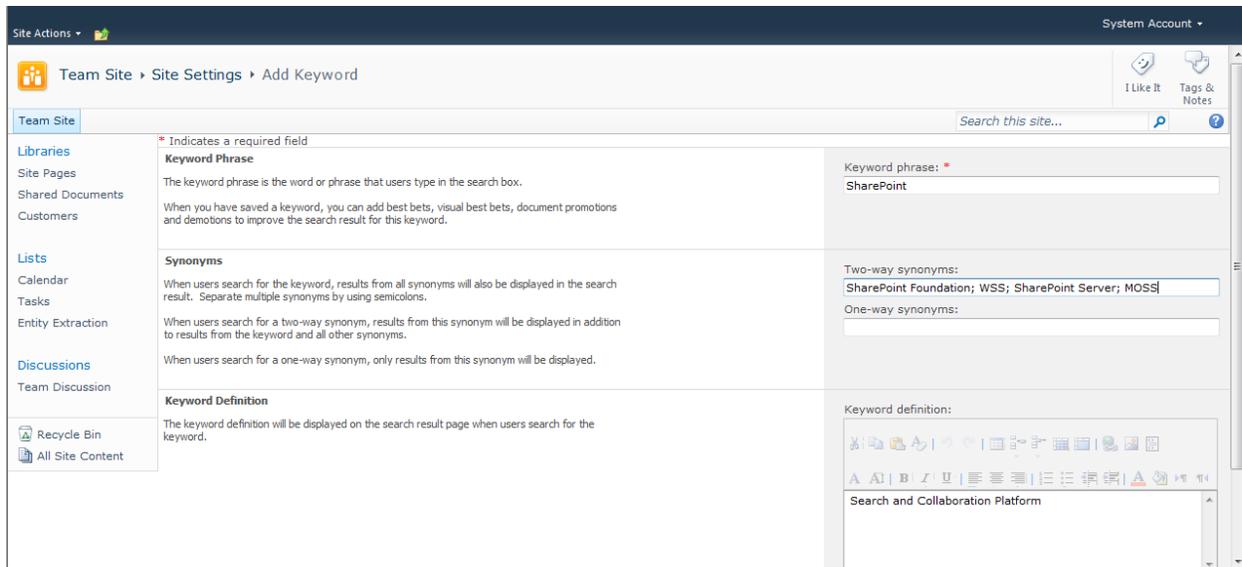


Figure 14. FAST Keyword Creation

FAST Search Site Promotions and Demotions

You can promote or demote sites, which affect the relevance ranking for documents in those locations. Figure 15 shows the promotion page.

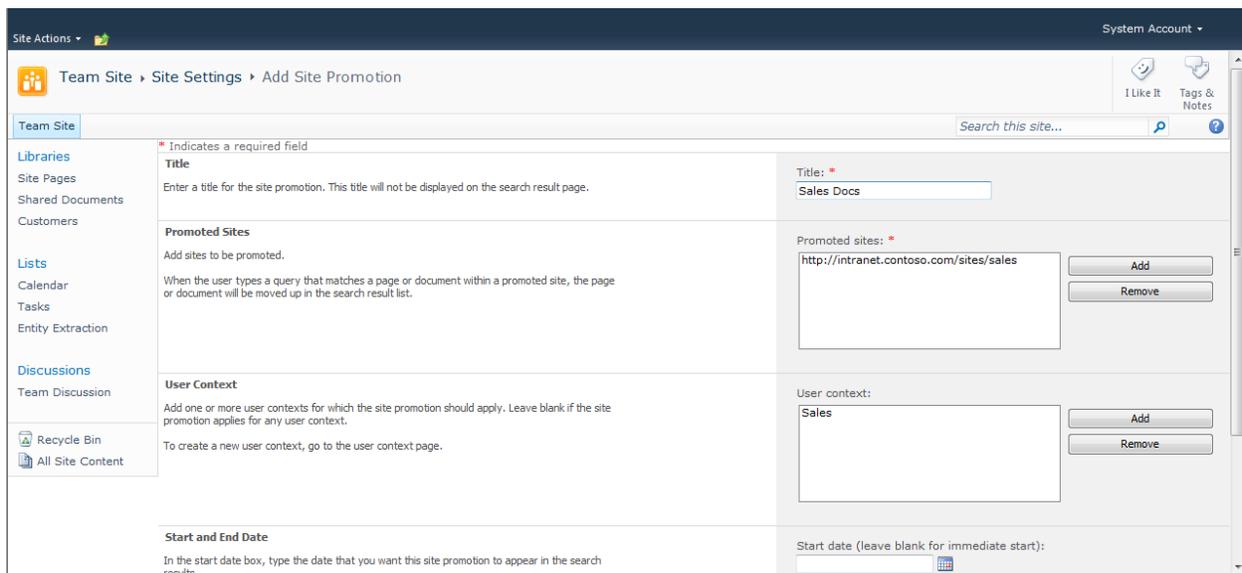


Figure 15. Site Promotion

FAST Search User Contexts

You can create FAST Search user contexts by using the site settings page in a site collection. When you have created a user context, you can associate it with Best Bets, Visual Best Bets, document promotions, document demotions, site promotions, and site

demotions. Figure 16 shows how to start creating a FAST Search user context called *Execs*. The new user context can apply to users with specific office locations or with specific knowledge

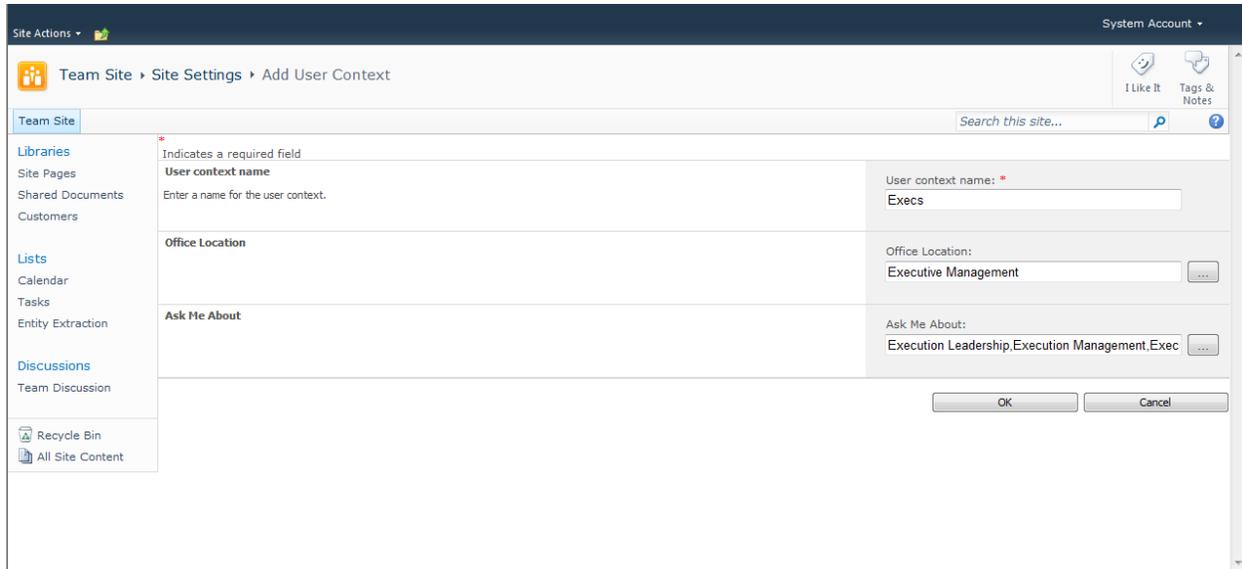


Figure 16. Adding a User Context

The table of keywords can be filtered to display only Best Bets, Visual Best Bets, Document Promotions and Document Demotions that apply to a specific user context. This is useful for a content manager who is responsible for a specific user context, such as employees based in Munich, for example. Use the *Restrict to a User Context* link, as illustrated in Figure 17, to see only Best Bets, Visual Best Bets, Document Promotions and Document Demotions that apply for this user context.

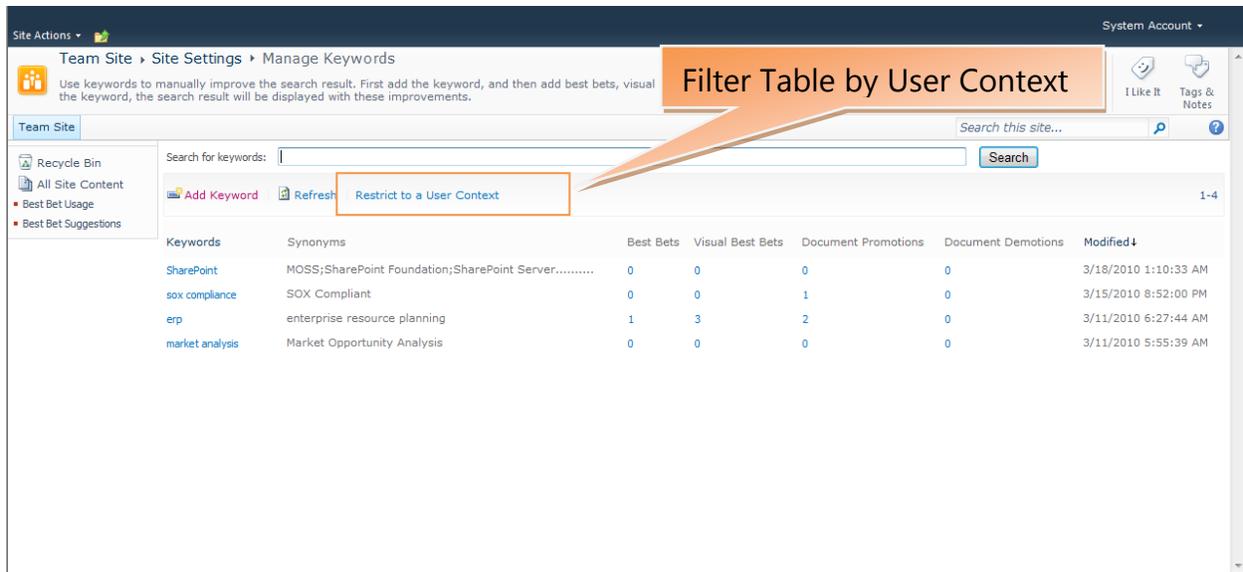


Figure 17. Filtering Best Bets, Visual Best Bets, Document Promotions and Document Demotions for a User Context

You can also associate promotions with a user context as shown in Figure 18.

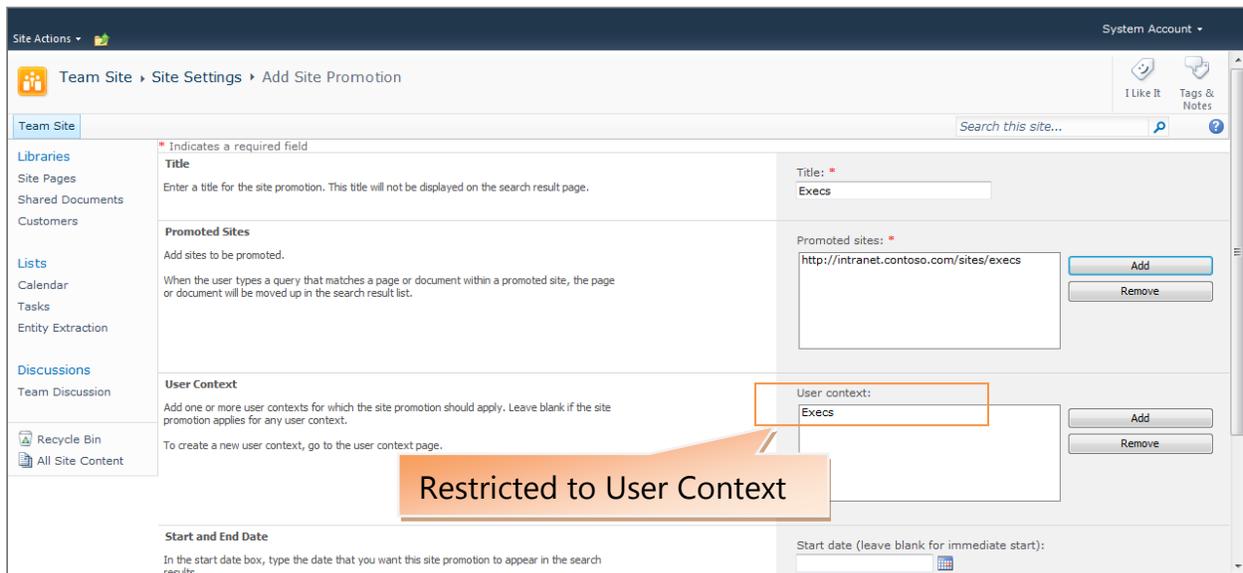


Figure 18. Promotions and User Contexts

Appendix A: Search Terminology

Before delving into the details of enterprise search features provided by FAST Search Server 2010 for SharePoint, it will be useful for you to ensure that you are familiar with search terms and definitions. You can use the following table to review brief descriptions of the terms used later in this guide.

| Term | Definition |
|-------------------------|---|
| Best Bet | Best Bets are URLs to documents that are associated with one or more keywords. Typically these documents or sites are ones that you expect users will want to see at the top of the search results list. Best Bets are returned by queries that include the associated keywords, regardless of whether the URL has been indexed. Site collection administrators can create keywords and associate Best Bets with them. |
| Connector | Connectors are components that communicate with specific types of system, and are used by the crawler to connect to and retrieve content to be indexed. Connectors communicate with the systems being indexed by using appropriate protocols. For example, the connector used to index shared folders communicates by using the FILE:// protocol, whereas connectors used to index Web sites use the HTTP:// or HTTPS:// protocols. |
| Content Source | Content sources are definitions of systems that will be crawled and indexed. For example, administrators can create content sources to represent shared network folders, SharePoint sites, other Web sites, Exchange public folders, third-party applications, databases, and so on. |
| Crawl Rule | Crawl rules specify how crawlers retrieve content to be indexed from content repositories. For example, a crawl rule might specify that specific file types are to be excluded from a crawl, or might specify that a specific user account is to be used to crawl a given range of URLs. |
| Crawl Schedule | Crawl schedules specify the frequency and dates/times for crawling content repositories. Administrators create crawl schedules so that they do not have to start all crawl processes manually. |
| Crawled Property | Crawled properties represent the metadata for content that is indexed. Typically, crawled properties include column data for SharePoint list items, document properties for Microsoft Office or other binary file types, and HTML metadata in Web pages. Administrators map crawled properties to managed properties in order to provide useful search experiences. See <i>Managed Property</i> later in this table for more details. |

| Term | Definition |
|----------------------------|--|
| Crawler | The crawler is the component that uses connectors to retrieve content from content repositories. |
| Crawler Impact Rule | A crawler impact rule governs the load that the crawler places on source systems when it crawls the content in those source systems. For example, one crawler impact rule might specify that a specific content repositories that is not used heavily by information workers should be crawled by requesting 64 documents simultaneously, whereas another crawler impact rule might specify less aggressive crawl characteristics for systems that are constantly in use by information workers. |
| Federation | Federation is the concept of retrieving search results from multiple search providers, based on a single query performed by an information worker. For example, your organization might include federation with Bing.com so that results are returned by SharePoint Server and Bing.com for a given query. |
| IFilter | IFilters are used by connectors to read the content in specific file types. For example, the Word IFilter is used to read Word documents, whereas a PDF IFilter is used to read PDF files. |
| Index | An index is a physical file that contains indexed content, and which is used by query servers to satisfy a query. |
| Indexer | Indexers manage the content to be included in an index, and propagate that content to query servers where they are stored in index files. |
| Indexing Engine | <i>See Indexer</i> |
| Index Partition | <i>See Partitioned Indexes</i> |
| Managed Property | Administrators create managed properties by mapping them to one or more crawled property. For example, an administrator might create a managed property named <i>Client</i> that maps to various crawled properties called <i>Customer</i> , <i>Client</i> , and <i>Cust</i> from different content repositories. Managed properties can then be used across enterprise search solutions, such as in defining search scopes and in applying query filters. |
| OpenSearch | OpenSearch is an industry standard that enables compliant search engines to be used in federated scenarios. <i>See Federation</i> for more details. |

| Term | Definition |
|---------------------------|---|
| Partitioned Index | SharePoint Server 2010 includes a new concept that enables administrators to spread the load for queries across multiple query servers. This is achieved by creating subsets of an index, and propagating individual subsets to different query servers. The subsets are known as <i>partitions</i> . At query time, the query object model contacts each query server that can satisfy the search so that all results to be returned to the user are included. |
| Property Database | Managed properties and security descriptors for search results are not stored in the physical index files. Instead, they are efficiently stored in a database. Query servers typically satisfy a query by retrieving information from both the index file and the property database. |
| Query Object Model | The query object model is responsible for accepting inputs from search user interfaces, and for issuing appropriate queries to query servers. The search Web Parts provided by SharePoint Server 2010 use the query object model to run queries. Developers can also create custom user interfaces and solutions that run queries by using the query object model. |
| Query Server | Query servers query retrieve data from index files and the property databases to satisfy queries. |
| Ranking | Ranking defines the sort order in which results are returned from queries. Typically, results are sorted in order of descending relevance, so that the most relevant documents are presented near the top of the results page. However, information workers might choose to apply a different sort order, such as by date modified. |
| Relevance | Relevance describes how well a given search satisfies a user's information needs. Relevance includes which documents are returned in the results (document recall) and the order of those documents in the results (ranking). |
| Search Center | Search Center is a site based on the Search Center site template. It provides a focused user interface that enables information workers to run queries and work with search results. |
| Search Document | See <i>Search Item</i> |
| Search Item | A search item represents a document, list item, file, Web page, Exchange public folder post, or database row that has been indexed. Search items are sometimes referred to as <i>search documents</i> , but the key point is that these items are returned by search queries. |

| Term | Definition |
|---------------------|---|
| Stemming | Words in each language can have multiple forms, but essentially mean the same thing. For example, the verb <i>To Write</i> includes forms such as <i>writing</i> , <i>wrote</i> , <i>write</i> , and <i>writes</i> . Similarly, nouns normally include singular and plural versions, such as <i>book</i> and <i>books</i> . The stemming feature in enterprise search can increase recall of relevant documents by mapping one form of a word to its variants. |
| Stop Word | Stop words (sometimes known as noise words) are those words for which there is no value in indexing them. For example, "a" , "and", and "the" are listed in the stop word file by default. There is no value in indexing these words as they are likely to be contained in a high percentage of indexed items. Furthermore, information workers rarely search for just these types of terms. |
| Synonym | Synonyms are words that mean the same thing as other words. For example, you might consider <i>laptop</i> and <i>notebook</i> to mean the same thing. Administrators can create synonyms for keywords that information workers are likely to search for in their organization. Additionally, synonyms that can be used to improve recall of relevant documents are stored in thesaurus files. |
| Word Breaker | Streams of data are retrieved from content repositories, and those streams are broken down into discrete words for indexing. Word breakers are the components that break down streams into individual words. Streams to be indexed are normally broken down by identifying spaces, punctuation marks, and the particular rules of each language. Also, when a user enters multiple words into a search box, that query is broken into discrete terms by a word breaker. |

Appendix B: Feature Comparison between Search Products

You can use the following table to make quick comparisons of the search features provided by each product:

| Feature | SharePoint Foundation 2010 | Search Server 2010 Express | Search Server 2010 | SharePoint Server 2010 | FAST Search Server 2010 for SharePoint |
|---|----------------------------|----------------------------|--------------------|------------------------|--|
| Basic site search | Y | Y | Y | Y | Y |
| Best Bets | | Y | Y | Y | Y |
| Visual Best Bets | | | | | Y |
| Similar Results | | | | | Y |
| Duplicate Results | Y | Y | Y | Y | Y |
| Search Scopes | | Y | Y | Y | Y |
| RSS Feeds for Search Results | Y | Y | Y | Y | Y |
| Alerts for Search Results | Y* | Y* | Y* | Y* | |
| Advanced Search Page | | Y | Y | Y | Y |
| Search Enhancement based on user context | | | | | Y |
| Crawled and Managed Properties | | Y | Y | Y | Y** |
| Entity Extraction | | | | | Y |
| Query Federation | | Y | Y | Y | Y |
| Query Suggestions | | Y | Y | Y | Y |
| Sort Results on Managed Properties or Rank Profiles | | | | | Y |
| Relevancy Tuning by Document or Site Promotions | | Y | Y | Y | Y** |

| Feature | SharePoint Foundation 2010 | Search Server 2010 Express | Search Server 2010 | SharePoint Server 2010 | FAST Search Server 2010 for SharePoint |
|----------------------------|-----------------------------------|-----------------------------------|---------------------------|-------------------------------|---|
| Shallow Results Refinement | | Y | Y | Y | Y |
| Deep Results Refinement | | | | | Y |
| Document Preview | | | | | Y |
| Windows 7 Federation | | Y | Y | Y | Y |
| People Search | | | | Y | Y |
| Phonetic Name Search*** | | | | Y | Y |
| Nickname Search*** | | | | Y | Y |
| Self Search | | | | Y | Y |
| Social Search | | | | Y | Y |
| Taxonomy Integration | | | | Y | Y |
| Multi-Tenant Hosting | | | | Y | |
| Rich Web Indexing Support | | | | | Y |

* For NTLM environments

**FAST Search Server 2010 for SharePoint provides enhanced capabilities in these areas. Please refer to the evaluation guide for *FAST Search Server 2010 for SharePoint* for more details.

*** For a subset of the supported languages

Appendix C: Resources Available for Evaluating SharePoint Server 2010

Microsoft encourages you to use the following resources as aids in installing and evaluating SharePoint Server 2010 in addition to this evaluation guide:

- The [product documentation](#) on TechNet will help you install SharePoint Server 2010.
- [SharePoint.microsoft.com](#) offers a variety of white papers and other resources.
- [MSDN SharePoint Server Developer Center](#) contains numerous technical resources from a developer's perspective about Microsoft SharePoint 2010 Products.
- [TechNet](#) contains numerous resources on how to deploy, manage, maintain and support SharePoint Server 2010.
- The [Enterprise Search TechCenter](#) on TechNet has tabs with information about each of the enterprise search products.
- The [Microsoft SharePoint Team Blog](#) is the official blog of the SharePoint Product Group.
- The [Microsoft SharePoint Server 2010 Evaluation Guide](#) gives IT professionals an introduction and overview of the SharePoint Server 2010 features that are most pertinent to installing, managing, and configuring the SharePoint farm.
- The [SharePoint 2010 Developer Reviewers Guide](#) contains an overview of the extensibility and customization points available for developers.

Appendix D: Feature differences between SharePoint Server 2010 and FAST Search Server 2010 for SharePoint

When you upgrade from SharePoint Server 2010 to FAST Search Server 2010 for SharePoint, some features are different for end-users, IT professionals, and developers.

Feature differences from an end-user perspective

Search alerts

By using search alerts in SharePoint Server 2010, users can receive daily or weekly updates on all the new or changed documents for a specific query.

FAST Search Server 2010 for SharePoint does not keep information about when a document was first indexed or when it was modified, and therefore does not provide this feature.

However, similar alerting functionality can be achieved via partners or Microsoft Services. A possible approach would be to save the queries, run them as batches, and then use an external mechanism to track results that have or have not been presented before.

Social tagging and its effect on relevance ranking

Social tagging is provided in SharePoint Server 2010 enterprise search. The feature is available through information that is stored in and extracted from a separate user information database.

FAST Search Server 2010 for SharePoint expects to find all relevant information from the crawled document metadata. Because social tagging is not stored in the metadata, FAST Search Server 2010 for SharePoint cannot return this information. This does not mean that social tagging is removed completely, only its effect on search results and the possibility to present or refine on social tags in search results.

If the information still exists in the user database, it is possible to tag content and have other users see the tags. Social tagging and relevance ranking can then be implemented through custom code created by a partner, Microsoft Services, or on-site developers.

Definition extraction

SharePoint Server 2010 extracts meanings of definitions from indexed text. Definition extraction occurs during a crawl and is a process of associating terms with one or more descriptions through the recognition of cue phrases.

FAST Search Server 2010 for SharePoint does not have a similar feature.

Author, date, and title extraction from Word and PowerPoint documents

Author, date and title extraction in SharePoint Server 2010 is enabled through a plugin that is designed to handle Word and PowerPoint documents (both legacy 97-2003 and the 2007 format).

FAST Search Server 2010 for SharePoint supports author, date, and title extraction from Office documents, but the search experience and relevancy can be affected depending on the metadata of the documents.

FAST Search Server 2010 for SharePoint uses the metadata of Office documents to extract author, date, and title of the indexed document and depends on the metadata being absolutely correct.

The extracted author, date, and title metadata title will be wrong if the following occurs:

- The author did not specify the metadata when creating the document.
- The author changed the author, date, or title in the content, but not in the metadata.
- The author used a template that had a metadata author, date, or title, but did not update the metadata to reflect the data in the new document.

Feature differences from an IT professional perspective**Crawl rules for document formats**

SharePoint Server 2010 specifies the document formats to be crawled by defining a set of File Types to *include* in the content index. For more information, see [Operations – Add a file type to the content index](#).

FAST Search Server 2010 for SharePoint handles this in the opposite way. Instead of defining what to include, you specify the file types to be *excluded* from a crawl. For more information, see [Operations – Include a file type in the content index](#).

Mirroring indexes across multiple data centers, backup and recovery functionality

FAST Search Server 2010 for SharePoint does not offer mirroring and synchronization of indexes across multiple data centers.

The FAST Search Server 2010 for SharePoint property index is stored on a file system and not in SQL Server as for SharePoint Server 2010. Index redundancy or index fault tolerance can be achieved by various ways of indexer backup deployments. In most large deployments, the index will be spread over several columns, and it is usual to have several rows with copies of the index for redundancy.

You can run a full data backup for a FAST Search Server 2010 for SharePoint farm, but this involves a long period with crawling suspended. The recommended approach for index high-availability in FAST Search Server 2010 for SharePoint is the fault-tolerant indexer setup.

In a fault-tolerant indexer setup of FAST Search Server 2010 for SharePoint, the data index is automatically copied to the backup indexers, which reside on the backup indexer row. If one indexer server fails with unrecoverable disk errors, you can either recover the server farm from the latest backup, or manually enable a backup indexer to act as the new primary indexer.

For more information about redundancy and availability, see [Planning and architecture - FAST Search Server farm redundancy and availability](#). For more information about backup and recovery strategy, see [Operations – Plan the system backup and recovery strategy](#).

Minimum size footprint

FAST Search Server 2010 for SharePoint and SharePoint Server 2010 have different architectures. This difference has the following effects on size footprint:

- Running FAST Search Server 2010 for SharePoint on the same servers as SharePoint Server 2010 is not supported. You can only migrate to separate FAST Search Server 2010 for SharePoint servers, and not within the existing SharePoint Server 2010 servers.
- The FAST Search Server 2010 for SharePoint resource usage pattern differs from SharePoint Server, and therefore requires different kinds of server resources (more local disk, less SQL dependencies).

- The FAST Search Server 2010 for SharePoint topology differs from SharePoint Server 2010. Therefore, the number of servers and the configuration of the servers will be different.

FAST Search Server 2010 for SharePoint has a different and larger disk footprint than SharePoint Server 2010 enterprise search. FAST Search Server 2010 for SharePoint uses less disk space in SQL Server, but more disk space locally. The total disk footprint is typically 2.5 times larger with FAST Search Server 2010 for SharePoint. FAST Search Server 2010 for SharePoint can support a larger document capacity per server than SharePoint Server 2010 given sufficient storage space.

For information about performance and capacity effects, see [FAST Search Server 2010 for SharePoint Capacity Planning](#).

Monitoring capability without using System Center Operations Manager

SharePoint Server 2010 offers search health monitoring through a search administration dashboard in Central Administration.

For monitoring FAST Search Server 2010 for SharePoint, we recommend using System Center Operations Manager 2007 R2.

If a System Center Operations Manager solution is not used, the administrator can monitor the system by collecting monitoring data from monitoring interfaces such as the Windows event log, performance counters, Windows Management Instrumentation (WMI), and SharePoint Unified Logging System (ULS) to verify system performance and health, either manually or by using custom automation/scripting or alternative management suites.

For more information about monitoring and configurations with and without System Center Operations Manager, see [Operations - Monitor FAST Search Server 2010 for SharePoint](#).

Feature differences from a developer perspective

Query integration and query-side interfaces

SharePoint Server 2010 and FAST Search Server 2010 for SharePoint support different search syntaxes for building search queries. For example, queries in SharePoint Server 2010 that use SQL syntax must be revisited and changed to use keyword syntax in FAST

Search Server 2010 for SharePoint. For more information about the different query syntaxes, see [Building Search Queries](#).

The same query-side interfaces can be used both with FAST Search Server 2010 for SharePoint and SharePoint Server 2010 enterprise search. If you have a custom search application that is running on SharePoint Server 2010, you should be aware that some advanced query features that are available through the Federation Object Model or the Query Web service will behave differently after you upgrade to FAST Search Server 2010 for SharePoint.

[FAST Search Query Integration Overview](#) describes the feature differences and what you must consider. In particular, note how to handle people search queries in FAST Search Server 2010 for SharePoint.

As an example, the following Query Web Service schema elements are not supported (that is, they are ignored) by FAST Search Server 2010 for SharePoint:

IgnoreAllNoiseQuery, IncludeHighConfidenceResults, HighlightQuerySuggestions and CapitalizeFirstLetters.

Customized ranking

SharePoint Server 2010 and FAST Search Server 2010 for SharePoint have different integration interfaces for handling custom relevance tuning. There are more options available in FAST Search Server 2010 for SharePoint, but the RelevanceModel element is not supported. For more information about how to customize search relevancy, see [Improving Relevance for FAST Search Server 2010 for SharePoint](#).

Programmatically administering the enterprise search solution

The SDK interfaces and administration features are different for SharePoint Server 2010 and FAST Search Server 2010 for SharePoint. For more information about how to administer FAST Search Server 2010 for SharePoint by using the Administration Object Model, see [Programmatically Administering FAST Search Server 2010 for SharePoint](#).

Custom security trimming

SharePoint Server 2010 provides support for custom security trimming of search results through a SecurityTrimmer interface (ISecurityTrimmer2). FAST Search Server 2010 for SharePoint does not support this interface for result-side custom security trimming. All security trimming is performed as part of the query matching, based on ACL information that is stored in the index. Because FAST Search Server 2010 for SharePoint provides query refinement based on all items that match a query, this ensures that refinement counts only reflect the items that the user is entitled to see.

One way to customize security trimming is to write custom claims providers to add principals (groups/users) from other domains into the query rewrite. Another option is to write custom crawlers that provide custom ACL information to the index.