



IT06 **Search in SharePoint Server 2007: A Technical Drilldown**

European Microsoft SharePoint Conference
2007

February 12th to 14th, 2007
Convention Center
Hotel Estrel, Berlin, Germany

 Microsoft®
**Office SharePoint®
Server 2007**

IT06

Search in SharePoint 2007: A Technical Drilldown

Lars Fastrup
Chief Software Engineer
Mondosoft, Denmark
<http://blog.mondosoft.com/ontolica>



Agenda

- SharePoint Version Overview
- Improvements to Search in MOSS 2007
- Finding People and Expertise
- Indexing Unstructured Data
- Indexing Structured Data
- Customizing the Search UI
- Search Usage Reporting

SharePoint Version Overview

SharePoint Products and Technologies

- Search features differ with specific version
 - SharePoint Server 2007
 - Enterprise CAL
 - Standard CAL
 - SharePoint Server 2007 for Internet Sites
 - SharePoint Server for Search Standard Edition
 - SharePoint Server for Search Enterprise Edition
 - Windows SharePoint Services 3.0
- A details of the features found in each version can be found on www.microsoft.com/office
- This session is based on functionality in SharePoint Server 2007 (Enterprise CAL)

MOSS 2007 vs. WSS V3

- Windows SharePoint Services 3.0
 - Subset of SharePoint Server search features
 - No index rebuild during upgrade to SharePoint Server
 - Indexing and query always on the same machine
 - Search scope limited to site and sub site content only
 - Only content stored in WSS 3.0 can be indexed
- Microsoft Office SharePoint Server 2007
 - Substantial additional features and functionality (e.g. Indexing BDC)
 - Indexer and query server roles can be distributed
 - Custom Search scopes across multiple content sources
 - Managed via a Shared Service Provider
 - Local and remote content sources can be indexed



Relevancy Improvements

- Substantially improved relevancy was the number one goal for this release
- New algorithms for Web and Enterprise Content
- New relevancy elements including
 - Click Distance
 - Anchor Text
 - URL Depth
 - Metadata Extraction from documents
- Testing has shown significantly better relevancy
 - 500% improvement on common queries
 - 100% improvement on all queries

Security Improvements

- Query-time security trimming (Same as SPS 2003)
 - Custom security trimming API.
- Minimum crawler permission is "Full Read"
 - Still provides the same security trimming functionality
 - Automatically configured for new sites
- Search visibility options
 - Prevent sites or specific lists appearing in search results
- "Security only" crawl and single item removal

Scalability and Performance Improvements

- Tested to about 50 million documents
- Shared Services Provider Model
 - SharePoint Sites within the farm can consume centrally managed Search Services.
 - Configurable at the site collection level to add scopes, keywords and best bets.
 - One index per SSP.
- Fast incremental crawls using the new change log
- Continuous Index Propagation
 - Items appear in results minutes after being indexed
- Crawler Impact Rules
 - Control the impact of crawling on remote sources

Improved Management

- Simplified and streamlined Admin UI
- Improved management and mapping of crawled properties
- Multiple crawler start addresses for each content source
- Improved management of search scopes
 - Scopes decoupled from Content Sources
 - Scopes can span multiple content sources
 - Scope by Property, Site, Content Source, URL
 - Granular rules for scope creation
- Web Based Index Log
 - Browse and filter crawler log
- Complete admin API
- Microsoft Operations Manager Management pack

Demo

Managing Search

 Microsoft®
Office SharePoint®
Server 2007

Improved End User Experience

- Search Center
 - Clean UI, Search Site Template, Web Parts
- New Search Syntax
 - Implicit AND, Property Search, +/-, Advanced Search
- New Results Features
 - Context Summary, "Did you mean", Duplicate Collapsing, Hit Highlighting, Best Bets, Auto Definitions
- Improved Alerting and Search RSS Feeds
- Customization
 - Web Parts, ASP.NET 2.0, Results with XSLT

Demo

Improved End User Experience

 Microsoft®
Office SharePoint®
Server 2007

Finding People and Expertise

- SharePoint Profile data
 - Built from AD, BDC, LDAP v3 and manually entered
 - Privacy controls
- My Colleagues
 - Stored in a list in My Site
 - Outlook 2007 suggested or manually entered
 - Privacy controls
- Results grouped by social relevancy
 - My colleagues, my colleagues colleagues, everyone else

Demo

People and Expertise Search

 Microsoft®
Office SharePoint®
Server 2007

Indexing Unstructured Data

Extensible Connectivity

- Protocol Handler
 - Connects to a content source and enumerates the documents
 - Ships with support for
 - NTFS File System, Exchange Public Folders, SharePoint, HTTP, Lotus Notes
 - Partners providing support for
 - Documentum, Hummingbird, FileNet, Interwoven and others
- iFilter
 - Reads a document and extracts the plain text content and properties for the indexer to index
 - Ships with support for over 200 file types
 - Partners/Vendors providing support for many more e.g. PDF

Demo

Indexing Unstructured Data

 Microsoft®
Office SharePoint®
Server 2007

Indexing Structured Data

Business Data Catalog

- Index and Search Structured Data
- NO CODE! XML Application Definition
 - Sample SAP and Siebel definitions included OOB
- Access to data through ADO.NET or Web Service
- Render BDC Data in SharePoint Server
 - Business DataView Web Part
 - SharePoint Designer
- Surface Data in SharePoint Lists
 - Business Data Column
- Create Custom Actions
- Import data into User Profiles

Demo

Indexing Structured Data

 Microsoft®
Office SharePoint®
Server 2007

Customizing the Search UI

- Customize the Query
 - Query Web Part Properties
 - Custom Search Scopes
 - Custom Search Center Tabs
 - Customize the Advanced Search Page
 - Object Model and Search Web Service
- Customize the Results
 - Core Results Web Part Properties
 - ASP.NET 2.0 Master Pages and Page Layouts
 - Apply custom XSL to results set
 - Custom page for BDC Search results
 - Object Model and Search Web Service

Demo

Customizing the Search UI

 Microsoft®
Office SharePoint®
Server 2007

Search Usage Reporting

- MOSS 2007 now includes usage reporting out of the box.
- Best way to improve search is to understand current usage.
- Several built-in search activity reports
 - Search query reports: Number of queries and top queries.
 - Search result reports: Top Results, Queries With Zero Results, Results with low click-through and a few more.
- Available at both site collection level and SSP level.
- Export data for extended reporting in Excel.
- Respond to feedback with configuration changes or editorial results (Best Bets).

Demo

Usage Reporting



Microsoft®
Office SharePoint®
Server 2007

Q&A