

SharePoint 2013 Connector Introduction

The SharePoint 2013 connector will crawl content from any SharePoint 2013 site collection URL.

On this page

- [Features](#)
- [Content Retrieved](#)
 - [Include](#)
- [Limitations](#)
- [Future Development Plan](#)
- [SharePoint Architecture](#)
 - [Summary of SharePoint organization](#)

Features

Some of the features of the SharePoint 2013 connector include:

- Performs incremental crawling (so that only new/updated documents are indexed) using SharePoint's change log timestamp or Aspire Snapshots.
- Fetches access control lists (ACLs) for document level security
- Is search engine independent
- Runs from any machine with access to the given SharePoint URLs
- Supports NTLM and HTTPs
- Support for BCS external lists
- Designed for supporting early binding mechanisms
- Runs without installing anything on SharePoint
- Regular expression patterns for files to include / exclude

Content Retrieved

The SharePoint 2013 connector retrieves several types of documents. Listed below are the inclusions and exclusions of these documents.

Include

- Sites
- Lists
- External Lists (BCS)
- Folders
- Documents or List Items
- Attachments

ListItems can take a number of different formats. For example, documents (pdf, doc, ppt, etc), calendar events or announcements. For more info on how ListItems content types work go to the MSDN [article](#).

Limitations

Due to API limitations, SharePoint 2013 connector has the following limitations:

- The connector uses the REST API to access SharePoint database(s) directly; it doesn't do web crawling.
- Crawling is only supported using a Site or a List as a root url.

Future Development Plan

The following features are not currently implemented, but are on the development plan:

- Index and support people search

Anything we should add? Please [let us know](#).

SharePoint Architecture

Find detailed information on MSDN [article](#).

Summary of SharePoint organization

This is the hierarchy of processes/applications/sites/sub-sites/libraries/folders/documents within SharePoint.

- SharePoint Server
 - SharePoint Web Application Pool
 - SharePoint Web Application (single web application)
 - Main Site Collection (the primary or main site created for the web application, associated with the primary http://xyz.server.com URL)
 - Sub Sites
 - Document Libraries
 - Folders
 - Documents
 - Attachments
 - Other Site Collections
 - Sub Sites
 - Document Libraries
 - Folders
 - Documents
 - Attachments