

# SharePoint 2010 Introduction

The SharePoint 2010 connector will crawl content from SharePoint site collection URL that you specify.

## On this page

- [Features](#)
- [Content Retrieved](#)
  - [Include](#)
- [Limitations](#)
- [Future Development Plan](#)

## Features

---

Some of the features of the SharePoint 2010 connector include:

- Performs incremental crawling (so that only new/updated documents are indexed)
- 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
- Supports site discovery
- Supports Claims Users and Groups
- Designed for supporting early binding mechanisms
- Optionally, can run without installing anything on SharePoint (with important limitations)
- Regular expression patterns for files to include / exclude

## Content Retrieved

---

The SharePoint 2010 connector retrieves several types of documents. Listed below are the inclusions of these documents.

### Include

- Sites
- Lists
- 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 2010 connector has the following limitations:

- The connector uses web services to access SharePoint database(s) directly; it doesn't do web crawling.

## Future Development Plan

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

- Automatic metadata propagation  
From site "about" pages to all of the files within the site
- Index and support people search

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