

PluginCentral

Plugins provide a large part of the functionality of nutch. This page acts as an up-to-date resource for supported plugins in Nutch. **N.B.** There is a wealth of information regarding pre-Nutch 1.3 plugin development available [here](#)

- [WhyNutchHasAPluginSystem](#)
- [AboutPlugins](#) - General information on what plugins are and how they work.
- [Technical Concepts Behind the Nutch Plugin System](#)
- [Problems with Plugins and Class-Loading](#)
- [WritingPluginExample](#) - A step-by-step example of how to write a plugin using the 1.x API.
- [Writing a plugin to add dates](#) by Ryan Pfister
- [PluginGotchas](#) - Yep there are some Gotchas you need to consider.
- [TikaPlugin](#) - Comments on the Tika integration and differences with existing parse plugins

Plugins You can Download

- [LanguageIdentifierPlugin](#)
- [XMLParser_Plugin](#) (**parse-xml** : parse xml documents using XPath and namespaces)
- [DocumentationTemplate](#)
- [GeoPosition](#)
- [German](#)
- [index-extra](#) - Adds user-configurable fields to the index.
- [protocol-smb](#) - Allows Nutch to crawl MS Windows Shares folder.
- [Index HTML Metatags](#): allows to parse HTML metatags and store them in separate index fields
- [mimetype-filter](#) - Allows Nutch to filter crawled documents before indexing by the extracted MIME type.
- [links-extractor](#) - Allows Nutch to index the inlinks and outlinks of any Web page.