

bin/nutch plugin

Plugin is an alias for org.apache.nutch.plugin.[PluginRepository](#)

This command can be used to load a plugin from the repository and execute its class main(). The plugin repository is a registry of all plugins. At system boot up the plugin repository is built by parsing the manifest files of all plugins. Plugins that are required which do not exist under other plugins are not registered. For each plugin a plugin descriptor instance will be created. The descriptor represents all meta information about a plugin so a plugin instance will be created later when it is required, this allows so called *lazy* plugin loading.

When loading plugins and building them into our working Nutch distribution we need to be aware of various configuration files, just as we would be if we were crawling with Nutch. The main files to be aware of are:

- nutch-default.xml (from which to copy properties)
- nutch-site.xml (from which to copy properties to)

Usage:

```
bin/nutch plugin <pluginId> <className> [args ...]
```

<pluginId>: The id of the plugin you wish to execute. e.g. the COMMAND

<className>: The class with the main() function you wish to execute.

[args ...]: 0 ...n arguments to pass to the plugin. This is sparsely documented as arguments are plugin specific as well as dependent.

[CommandLineOptions](#)