

# About Nutch

## Table of contents

1 Overview.....	2
-----------------	---

## 1. Overview

Nutch is open source web-search software. It builds on [Lucene and Solr](#), adding web-specifics, such as a crawler, a link-graph database, parsers for HTML and other document formats, etc.

Nutch can run on a single machine, but gains a lot of its strength from running in a [Hadoop cluster](#)

The system can be enhanced (eg other document formats can be parsed) using a plugin mechanism.

For more information about Nutch, please see the [Nutch wiki](#).