About Nutch

Table of contents

1. Overview

Nutch is open source web-search software. It builds on Lucene (http://jakarta.apache.org/lucene/) , adding web-specifics, such as a crawler, a link-graph database, parsers for HTML and other document formats, etc.

For more information about Nutch, please see the Nutch wiki. (http://www.nutch.org/cgi-bin/twiki/view/Main/Nutch)

2. Incubation Disclaimer

The Nutch project is an effort undergoing incubation at the Apache Software Foundation (ASF). Incubation is required of all newly accepted projects until a further review indicates that the infrastructure, communications, and decision making process have stabilized in a manner consistent with other successful ASF projects. While incubation status is not necessarily a reflection of the completeness or stability of the code, it does indicate that the project has yet to be fully endorsed by the ASF.

The incubation proposal for Nutch is here (http://wiki.apache.org/incubator/NutchProposal) .

The status of the Nutch incubation process is here (http://incubator.apache.org/projects/nutch.html) .