

Apache Solr - Magnolia integration



✖

Your Rating: Results: 87 rates



Implemented in 1.0

[magnolia-solr-module](#) on Forge. See [documentation](#)

What is Solr?

Solr is a layer of code on top of Lucene that transforms Lucene into an enterprise search platform. [...] Solr provides the following capabilities:

- *Web service: Solr places Lucene over HTTP, allowing programs written in any language to invoke Lucene.*
- *XML based schema for managing indexed fields and their characteristics.*
- *Admin tools for configuration, data loading, index replication, statistics, logging and cache management*
- *Large scale distributed search*
- *Fixed/paid result list placement*
- *Faceting — the dynamic clustering of items or search results into categories that let users drill into search results (or even skip searching entirely) by any value in any field, as seen in popular e-commerce sites such as Amazon or Zappos.*

The idea of integrating Solr and Magnolia came to me a few months ago and already did some work on my own. At first, I just wanted to synchronize the DMS docs with Solr and have some of the goodies the latter provides (faceting search for one) available out-of-the-box, e.g. a stk template subcategory to easily configure a feature-rich search page (something like this for instance <http://www.lucidimagination.com/search/?q=grilli>). However, this integration could actually apply to **all** textual contents which is eventually served by Magnolia, without limitations. One of the most interesting things IMO, besides having a dedicated and feature-rich search server, is the ability to expose Magnolia managed data both to Java and non Java clients (via both http and through specific API, e.g. solr-ruby or solPython).