

SolrProposal

Proposal for new project Solr

Yonik Seeley – yonik at apache dot org

(0) rationale

Solr is a search server based on Apache Lucene and focused on full-text search, relevancy, and performance.

Some of the server features include:

- Query and update interfaces over HTTP and XML.
- Replication based on index snapshots and rsync.
- A schema allowing declarative specification of fields and their types, including specification of text analysis filter chains.
- External configuration for Lucene parameters, stopword lists, and synonym lists.
- A web based admin interface with statistics and debugging.
- An extensive caching framework for filters, queries, documents, and user caches.
- Plugins for customizing query handling, caching and many other server aspects.

Solr is currently an internal CNET project, and is used to power search and faceted browsing in a number of CNET sites.

(0.1) criteria

Meritocracy:

Solr's developers are comfortable operating as a meritocracy.

Community:

Establishing an active open source community will be the top priority.

Core Developers:

Solr starts with the committers listed below, several of whom have worked on the project in the past at CNET.

Alignment:

Solr currently uses the following Apache projects: Ant, Lucene.

(0.2) warning signs

Orphaned products:

Solr is not an orphan, use within CNET is growing rapidly.

Inexperience with open source:

Solr's committers are experienced with open source and have made contributions to other open source projects.

Homogenous developers:

Solr's initial committers include several CNET employees since it's currently an internal project, as well as other Apache Software Foundation committers with interest in the project who do not have any CNET affiliation.

Reliance on salaried developers:

Some committers are currently paid to make contributions to Solr.

No ties to other Apache products:

Solr has strong ties to Lucene.

A fascination with the Apache brand:

The committers respect the Apache brand and feel that Apache is the right place to generate a strong open source community.

(1) scope of the subprojects

The scope of the project is to provide a stand-alone full-text search server. The proposal is that this project become a sub-project of Apache Lucene.

(2) identify the initial source from which the subproject is to be populated

CNET will donate the initial source code for this project.

(3) identify the ASF resources to be created

(3.1) mailing list(s)

- solr-dev
- solr-user

(3.2) Subversion or CVS repositories

- <https://svn.apache.org/repos/asf/incubator/solr>

(3.3) Jira

- Solr (SOLR)

(4) identify the initial set of committers

- Yonik Seeley, CNET (Lucene committer)
- Bill Au, CNET
- Chris Hostetter, CNET
- Yoav Shapira, Apache Software Foundation (no CNET affiliation, ASF committer for Tomcat, log4j, Jakarta...)
- Doug Cutting, ASF
- Otis Gospodneti, ASF

(5) identify apache sponsoring individual

- Doug Cutting, Champion and Mentor
- Erik Hatcher, Mentor
- Ian Holsman, Mentor (CNET Employee, but not directly involved in that department)
- Lucene PMC, Sponsor
(as defined in http://incubator.apache.org/incubation/Roles_and_Responsibilities.html)