Geophysical Research Abstracts
Vol. 20, EGU2018-14610, 2018
EGU General Assembly 2018
© Author(s) 2018. CC Attribution 4.0 license.

Exposing relational databases through an emulated SOLR search API

Damian Ulbricht, Kirsten Elger, and Boris Radosavljevic
GFZ German Research Centre for Geosciences, Potsdam, Germany (ulbricht@gfz-potsdam.de)

Programming web-based search engines, populating the search index and showing an asset of information in an intuitive way is challenging. To users, the „big players“ in the search engine market act as reference and people expect similar discovery functionality in scientific search portals. Querying the asset is expected to trigger a full-text search and autocompletion. In addition, search facets are expected to assist the search.
Facetted search and full-text search are often associated with document-based search engines that come as integrated search platform for larger assets and act as database with an own search index. Today, most relational databases provide a toolset to support advanced search. SQL databases can be indexed for full-text search and facets can be created by assembling the required information in a database view that is available for query.
For our applications at the GFZ German Research Centre for Geosciences in Potsdam, specifically the new data discovery portal RI@GFZ [1], developed a search component on top of a mySQL database that exposes a subset of the SOLR search API. We make use of a fully normalised database structure and combine the tables in a label-value style database view. The search component is written in PHP and implements facetting of results, full-text search, counting of facet content and counting of result documents. The browser-based search frontend uses AJAX-SOLR and connects the search component via the SOLR API to provide a search field with auto completion, to offer clickable facets and to render the search results. Frontend and the PHP search component are complemented with an upload script that populates the database via a CSV file.
Our solution is especially useful during prototyping or for smaller assets of information when setting up a SOLR instance and its handlers for data import does not seem to be appropriate but could be an option in the future. In addition, existing relational databases could easily provide a subset of the SOLR search API by creating the label-value style database view. The software is available on GitHub [2].

[1] http://dataservices.gfz-potsdam.de/mesi
[2] http://github.com/ulbricht/risearch