

## Similarity results using CrossSim

Number of projects in the database: **580**

The projects belong to the following categories:

- SPARQL, Apache Jena, DBpedia, and RDF Tools
- PDF Reader and Converter
- Selenium Web Browser Test
- Database ORM
- MVC framework
- Web browser
- Compression
- Music Player
- JSON library
- Elastic Search, Web Search
- Others

Number of queries: **50**

Legend:

- **Text in light blue:** Results that are not relevant
- **Text in red:** The features that identify the tasks and functionalities of a project
- **Text in dark blue:** Comments, explanations
- **Text in black:** relevant projects to a query

Remark: In this configuration, the following relationships are considered: **stars, committers and dependencies**, most frequent dependencies are removed.

**1. Query 1: neo4j-contrib/sparql-plugin:** This is a Neo4j Server plugin, providing Sparql Neo4j Server. This effectively turns the Neo4j Server into a Triple Store.

No.	Top-5 retrieved project	Description	Score (Given by human)
1	niclashoyer/neo4j-sparql-extension	Neo4j unmanaged extension for RDF storage and SPARQL 1.1 query features	
2	claudiomartella/dbpedia4neo	A set of hacks to setup a dbpedia endpoint through neo4j	
3	jbarrasa/neosemantics	Experiments bridging the RDF and LPG worlds.	
4	streamreasoning/CSPARQL-engine	Reasoning over RDF stream made easy. The project contains parent pom in the root and a number of module that inherit from parent pom. To install the csparql-core jar, run mvn install on parent pom.	
5	<a href="#">neo4j-contrib/cypher-dsl</a>	<a href="#">A Java DSL for the Cypher Query Language and an optional Query DSL mode</a>	

**2. Query 2: AskNowQA/AutoSPARQL:** AutoSPARQL TBSL is a graphical user interface, which allows to answer natural language queries over RDF knowledge bases. It is based on algorithms implemented in the DL-Learner Semantic Web machine learning framework.

No.	Top-5 retrieved project	Description	Score
<a href="#">1</a>	AKSW/SPARQL2NL	SPARQL2NL - the SPARQL to natural language converter	
<a href="#">2</a>	<a href="#">socialsignin/spring-social-security-demo</a>		
<a href="#">3</a>	<a href="#">yhegde/facebook-page-scraper</a>	<a href="#">Facebook Page Scraper</a> is a suite of tools for collecting data from public Facebook Pages using Facebook's Graph API	
4	pyvandenbussche/sparqls	SPARQL Endpoint Status. SPARQL Endpoint Status project aims at monitoring SPARQL Endpoints	
<a href="#">5</a>	<a href="#">sphas/messic</a>		

**3. Query 3: AKSW/Sparqlify:** Sparqlify is a scalable SPARQL-SQL rewriter whose development began in April 2011 in the course of the LinkedGeoData project.

No.	Top-5 retrieved project	Description	Score
1	nkons/r2rml-parser	R2RML Parser is an award-winning tool that can export relational database contents as RDF graphs, based on an R2RML mapping document.	
2	pyvandenbussche/sparqls	SPARQL Endpoint Status. SPARQL Endpoint Status project aims at monitoring SPARQL Endpoints	
3	rdfhdt/hdt-java	HDT-lib is a Java Library that implements the W3C Submission ( <a href="http://www.w3.org/Submission/2011/03/">http://www.w3.org/Submission/2011/03/</a> ) of the <b>RDF</b> HDT (Header-Dictionary-Triples) binary format for publishing and exchanging RDF data at large scale.	
4	AKSW/SPARQL2NL	SPARQL2NL - the SPARQL to natural language converter	
5	AKSW/RDFUnit	RDFUnit is implemented on top of the Test-Driven Data Validation Ontology and designed to read and produce RDF that complies to that ontology only. The main components that RDFUnit reads are TestCases (manual & automatic), TestSuites, Patterns & TestAutoGenerators. RDFUnit also strictly defines the results of a TestSuite execution along with different levels of result granularity.	

**4. Query 4: AKSW/SPARQL2NL: SPARQL2NL - the SPARQL to natural language converter.**

No.	Top-5 retrieved project	Description	Score
1	AskNowQA/AutoSPARQL	AutoSPARQL TBSL is a graphical user interface, which allows to answer natural language queries over RDF knowledge bases. It is based on algorithms implemented in the DL-Learner Semantic Web machine learning framework	
2	AKSW/jena-sparql-api		
3	<a href="#">otto-de/jsonhome</a>	<a href="#">Libraries to publish and consume json-home documents for RESTful web applications.</a>	
4	<a href="#">omerio/appstart</a>	<a href="#">Appstart is an opinionated Java boilerplate template for rapid development of multi-module Google App Engine applications based on Google Guice, a lightweight dependency injection framework. Appstart uses the App Engine Maven plugin, it has a parent pom and 3 maven projects. 2 of the projects (appstart-front and appstart-backend) are App Engine modules.</a>	
5	AKSW/Sparqlify	Sparqlify is a scalable SPARQL-SQL rewriter whose development began in April 2011 in the course of the LinkedGeoData project.	

5. **Query 5: pyvandenbussche/sparqls:** SPARQL Endpoint Status. SPARQL Endpoint Status project aims at monitoring SPARQL Endpoints based on 4 aspects:

- Discoverability analyses how SPARQL endpoints can be located, what meta-data are available for them, etc.
- Interoperability identifies which features of SPARQL 1.0 and SPARQL 1.1 standards are supported by an endpoint
- Performance measures generic performance aspects such as result-streaming, atomic lookups and simple-joins over a HTTP connection.
- Availability monitors the uptimes of a SPARQL endpoint.

No.	Top-5 retrieved project	Description	Score
1	pyvandenbussche/sparqls	SPARQL Endpoint Status. SPARQL Endpoint Status project aims at monitoring SPARQL Endpoints	
2	streamreasoning/CSPARQL-engine	Reasoning over RDF stream made easy. The project contains parent pom in the root and a number of module that inherit from parent pom. To install the csparql-core jar, run mvn install on parent pom.	
3	rdfhdt/hdt-java	HDT-lib is a Java Library that implements the W3C Submission ( <a href="http://www.w3.org/Submission/2011/03/">http://www.w3.org/Submission/2011/03/</a> ) of the <b>RDF</b> HDT (Header-Dictionary-Triples) binary format for publishing and exchanging RDF data at large scale.	
4	AKSW/RDFUnit	RDFUnit is implemented on top of the Test-Driven Data Validation Ontology and designed to read and produce RDF that complies to that ontology only. The main components that RDFUnit reads are TestCases (manual & automatic), TestSuites, Patterns & TestAutoGenerators. RDFUnit also strictly defines the results of a TestSuite execution along with different levels of result granularity.	
5	<a href="#">thundernet8/AlipayOrdersSupervisor-GUI</a>	<a href="#">GUI of AlipayOrdersSupervisor, implemented in Java and Swing</a>	

**6. Query 6: sayems/java.webdriver:** Browser test automation using Selenium WebDriver in Java.

No.	Top-5 retrieved project	Description	Score
1	tarun3kumar/ seleniumtestsframework	STF is a webdriver based test automation library which generates comprehensible tests reports comparable to BDD framework, and yet better since no extra test steps / configuration are required. Not just this, STF has tons of other useful features described at - <a href="http://www.seleniumtests.com/2013/10/...">http://www.seleniumtests.com/2013/10/...</a>	
2	<a href="#">jprante/elasticsearch-plugin-rdf-jena</a>	<a href="#">Using Apache Jena API to store and retrieve RDF triples in Elasticsearch</a>	
3	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
4	seleniumQuery/seleniumQuery	jQuery-like cross-driver interface in Java for Selenium WebDriver	
5	<a href="#">noveogroup-amorgunov/spring-mvc-react</a>	Example of using Spring 4 (rest full api with hibernate) + react.js (client)	

**7. Query 7: xebia/Xebium:** Xebium provides Selenium (webdriver) bindings for FitNesse, with Selenium-IDE support.

No.	Top-5 retrieved project	Description	Score
1	fhoeben/hsac-fitness-fixtures	An environment to define and run integration tests. It contains Fitnessse fixture (base) classes and a baseline Fitnessse installation.	
2	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
3	<a href="#">sikuli/sikuli-slides</a>	<a href="#">Sikuli-Slides is a visual automation tool that enables users to automate and test Graphical User Interfaces (GUIs) using presentation slides.</a>	
4	cukespace/cukespace	Cukes in Space! allows you to deploy and run Cucumber features in the application server of your choice using the Arquillian test framework.	
5	<a href="#">ModelDriven/fUML-Reference-Implementation</a>	<a href="#">Open-source implementation of the Foundational Semantics for Executable UML Models (Foundational UML) specification.</a>	



**8. Query 8: webdriverextensions/webdriverextensions:** Make your WebDriver based Selenium tests more readable, reusability and maintainable by using WebDriver Extensions! <http://webdriverextensions.github.io>.

No.	Top-5 retrieved project	Description	Score
1	seleniumQuery/seleniumQuery	jQuery-like cross-driver interface in Java for Selenium WebDriver	
2	Cognifide/bobcat	Bobcat is an automated testing framework for functional testing of web applications. It wraps Selenium Browser Automation with a handy set of tools (accelerators). Since using Selenium (Webdriver), it aims to mimic the behaviour of a real user, and as such interacts with the HTML of the application.	
3	vmi/selenese-runner-java	This is selenese script interpreter implemented by Java. It supports test-case and test-suite which are Selenium IDE's native format.	
4	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
5	conductor-framework/conductor	The Selenium framework that takes you where you want to go	

**9. Query 9: testIT-WebTester/webtester-core:** testIT WebTester is a web-application UI test automation framework based on **Selenium** (<http://www.seleniumhq.org>). It is the product of years of consulting experience in various projects and aims at providing a very intuitive, declarative and extendable API for writing programmatic UI tests in Java.

No.	Top-5 retrieved project	Description	Score
1	dgageot/simplelelenium	Simple layer on top of Selenium and PhantomJs	
2	<a href="#">bluebreezecf/SparkJobServerClient</a>	<a href="#">Java Client of the Spark Job Server implementing the arranged Rest APIs</a>	
3	brucezee/jspider	A simple, scalable, and highly efficient web crawler framework for Java. Use <b>Selenium</b> to control the browser request data.	
4	Cognifide/bobcat	Bobcat is an automated testing framework for functional testing of web applications. It wraps Selenium Browser Automation with a handy set of tools (accelerators). Since using Selenium (Webdriver), it aims to mimic the behaviour of a real user, and as such interacts with the HTML of the application.	
5	<a href="#">ModelDriven/fUML-Reference-Implementation</a>	<a href="#">Open-source implementation of the Foundational Semantics for Executable UML Models (Foundational UML) specification.</a>	

**10.Query 10: seleniumQuery/seleniumQuery: jQuery-like cross-driver interface in Java for Selenium WebDriver.**

No.	Top-5 retrieved project	Description	Score
1	webdriverextensions/ webdriverextensions	Make your WebDriver based Selenium tests more readable, reusability and maintainable by using WebDriver Extensions! <a href="http://webdriverextensions.github.io">http://webdriverextensions.github.io</a>	
2	conductor-framework/conductor	The Selenium framework that takes you where you want to go	
3	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
4	brucezee/jspider	A simple, scalable, and highly efficient web crawler framework for Java. Use <b>Selenium</b> to control the browser request data.	
5	tarun3kumar/ seleniumtestsframework	STF is a webdriver based test automation library which generates comprehensible tests reports comparable to BDD framework, and yet better since no extra test steps / configuration are required. Not just this, STF has tons of other useful features described at - <a href="http://www.seleniumtests.com/2013/10/...">http://www.seleniumtests.com/2013/10/...</a>	



**11. Query 11: bonigarcia/webdrivermanager:** Automatic Selenium WebDriver binaries management in runtime for Java.

No.	Top-5 retrieved project	Description	Score
1	seleniumkit/selenograph	Powered version of Selenium Grid Router	
2	webdriverextensions/ webdriverextensions	Make your WebDriver based Selenium tests more readable, reusability and maintainable by using WebDriver Extensions! <a href="http://webdriverextensions.github.io">http://webdriverextensions.github.io</a>	
3	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
4	Ardesco/selenium-standalone-server-plugin	A Maven plugin that will download the WebDriver stand alone server executables for use in your mavenised Selenium project	
5	classtag/learn_crawler	Reading notes "write their own web crawler", their own knock code. Mainly recorded the basic realization of the web crawler, the page to the heavy algorithm, the page fingerprint algorithm, text information mining	

**12.Query 12: selenium-cucumber/selenium-cucumber-java:** jQuery-like cross-driver interface in Java for Selenium WebDriver.

No.	Top-5 retrieved project	Description	Score
1	cukespace/cukespace	Cukes in Space! allows you to deploy and run Cucumber features in the application server of your choice using the Arquillian test framework.	
2	jacum/icefaces-archetype	Maven webapp project with JSF/ICEFaces, MyFaces, Jetty, Spring Core, Spring Security, Selenium, JRebel for quick start.	
3	conductor-framework/conductor	The Selenium framework that takes you where you want to go	
4	seleniumQuery/seleniumQuery	jQuery-like cross-driver interface in Java for Selenium WebDriver	
5	Cognifide/bobcat	Bobcat is an automated testing framework for functional testing of web applications. It wraps Selenium Browser Automation with a handy set of tools (accelerators). Since using Selenium (Webdriver), it aims to mimic the behaviour of a real user, and as such interacts with the HTML of the application.	

**13.Query 13: conductor-framework/conductor:** The Selenium framework that takes you where you want to go.

No.	Top-5 retrieved project	Description	Score
1	seleniumQuery/seleniumQuery	jQuery-like cross-driver interface in Java for Selenium WebDriver	
2	selenium-cucumber/selenium-cucumber-java	Automation Testing Using Java. selenium-cucumber is a behavior driven development (BDD) approach to write automation test script to test Web. It enables you to write and execute automated acceptance/unit tests. It is cross-platform, open source and free. Automate your test cases with minimal coding. More Details	
3	testIT-WebTester/webtester-core	testIT WebTester is a web-application UI test automation framework based on Selenium ( <a href="http://www.seleniumhq.org">http://www.seleniumhq.org</a> ). It is the product of years of consulting experience in various projects and aims at providing a very intuitive, declarative and extendable API for writing programmatic UI tests in Java	
4	webdriverextensions/webdriverextensions	Make your WebDriver based Selenium tests more readable, reusability and maintainable by using WebDriver Extensions! <a href="http://webdriverextensions.github.io">http://webdriverextensions.github.io</a>	
5	sayems/java.webdriver	Browser test automation using Selenium WebDriver in Java	

**14.Query 14: caelum/vraptor:** Simple Java Web MVC Framework. A web MVC framework for fast and maintainable development.

No.	Top-5 retrieved project	Description	Score
1	caelum/restfulie-java	restfulie implementation in java, with support to vraptor. CRUD through HTTP is a good step forward to using resources and becoming RESTful, another step further into it is to make use of hypermedia based services and this gem allows you to do it really fast. Inner libraries: In its Java version, Restfulie uses by default: <ul style="list-style-type: none"> <li>- VRaptor:<a href="http://www.vraptor.org">http://www.vraptor.org</a> as the server-side controller</li> <li>- XStream:<a href="http://xstream.codehaus.org">http://xstream.codehaus.org</a> as its serialization library</li> <li>- java.net api for http requests</li> <li>- Spring IoC for dependency injection</li> </ul>	
2	caelum/vraptor4	A web MVC action-based framework, on top of CDI, for fast and maintainable Java development.	
3	skprasadu/spring-security-examples		
4	<a href="#">brucezee/jspider</a>	<a href="#">A simple, scalable, and highly efficient web crawler framework for Java. Use Selenium to control the browser request data.</a>	
5	ExampleDriven/cxf-example	This project demonstrates how to build a simple CXF webservice using SOAP (JAX-WS) or REST (JAX-RS). Technology Used: <ul style="list-style-type: none"> <li>- CXF 2.5.2 SOAP(JAX-WS), REST(JAX-RS) client and endpoint</li> <li>- Spring, Spring MVC 3.1.1</li> <li>- Spring Security 3.10JSR 303 Bean validation</li> </ul>	



**15.Query 15: caelum/vraptor4:** A web MVC action-based framework, on top of CDI, for fast and maintainable Java development.

No.	Top-5 retrieved project	Description	Score
1	caelum/vraptor	Simple Java Web MVC Framework. A web MVC framework for fast and maintainable development	
2	livrospringmvc/lojasadocodigo	home code version of the code using Spring MVC as framework	
3	<a href="#">stephenh/joist</a>	Joist is a ORM based on code generation. The goal is to provide Rails-like "empty domain objects" in an ORM that is simple, pleasant to use, and, if needed, scales nicely to really large schemas.	
4	caelum/restfulie-java	restfulie implementation in java, with support to vraptor. CRUD through HTTP is a good step forward to using resources and becoming RESTful, another step further into it is to make use of hypermedia based services and this gem allows you to do it really fast. Inner libraries: In its Java version, Restfulie uses by default: <ul style="list-style-type: none"> <li>- VRaptor:<a href="http://www.vraptor.org">http://www.vraptor.org</a> as the server-side controller</li> <li>- XStream:<a href="http://xstream.codehaus.org">http://xstream.codehaus.org</a> as its serialization library</li> <li>- java.net api for http requests</li> </ul> Spring IoC for dependency injection	
5	noveogroup-amorgunov/spring-mvc-react	Example of using Spring 4 (rest full api with hibernate) + react.js (client)	

**16.Query 16: KEN-LJQ/WMS:** A warehouse management system implement with Spring MVC, Spring Framework, MyBatis, Shiro and MySQL <http://139.199.179.164:8080/WMS>.

No.	Top-5 retrieved project	Description	Score
1	swapnildipankar/ java_jersey_spring_hibernate_maven	Archetype starter project to create RESTful WS using Java, Jersey WS, Spring, Hibernate, Maven, Liquibase, MySQL and MongoDB. Code for integrating with Hibernate/MySQL and MongoDB/NoSQL.	
2	white-cat/jeeweb-mybatis	JEEWEB is based on SpringMVC + Spring + Hibernate JAVA WEB agile development system; it is a code generation function with intelligent fast development platform; is Spring Framework as the core container, Spring MVC model view controller, Hibernate for data Access layer, Apache Shiro for the authority of the authorized layer, Ehcache commonly used data cache, Disruptor as a concurrent framework, Bootstrap as a front-end framework of the excellent open source system	
3	<a href="#">brucezee/jspider</a>	<a href="#">A simple, scalable, and highly efficient web crawler framework for Java. Use Selenium to control the browser request data.</a>	
4	snakerflow/snaker-web	The snaker-web project is a web application developed based on the snaker process engine, which includes process deployment, design, startup, task execution, dismissal, withdrawal, forwarding, delegate agents, and so on. Based on SpringMVC, Spring3, Hibernate3, Shiro, Snaker to build the basis of demonstration applications	
5	white-cat/jeeweb	JEEWEB is based on SpringMVC + Spring + Hibernate JAVA WEB agile development system; it is a code generation function with intelligent fast development platform	

**17.Query 17: white-cat/jeeweb:** JEEWEB is based on SpringMVC + Spring + Hibernate JAVA WEB agile development system; it is a code generation function with intelligent fast development platform; is Spring Framework as the core container, Spring MVC model view controller, Hibernate for data Access layer, Apache Shiro for the authority of the authorized layer, Ehcache commonly used data cache, Disruptor as a concurrent framework, Bootstrap as a front-end framework of the excellent open source system.

No.	Top-5 retrieved project	Description	Score
1	white-cat/jeeweb-mybatis	JEEWEB is based on SpringMVC + Spring + Hibernate JAVA WEB agile development system; it is a code generation function with intelligent fast development platform; is Spring Framework as the core container, Spring MVC model view controller, Hibernate for data Access layer, Apache Shiro for the authority of the authorized layer, Ehcache commonly used data cache, Disruptor as a concurrent framework, Bootstrap as a front-end framework of the excellent open source system	
2	snakerflow/snaker-web	The snaker-web project is a web application developed based on the snaker process engine, which includes process deployment, design, startup, task execution, dismissal, withdrawal, forwarding, delegate agents, and so on. Based on SpringMVC, Spring3, Hibernate3, Shiro, Snaker to build the basis of demonstration applications	
3	noveogroup-amorgunov/spring-mvc-react	Example of using Spring 4 (rest full api with hibernate) + react.js (client)	
4	<a href="#">LinuxSuRen/phoenix.platform</a>	<a href="#">WebUI automated test platform</a>	
5	KEN-LJQ/WMS	A warehouse management system implement with Spring MVC, Spring Framework, MyBatis, Shiro and MySQL <a href="http://139.199.179.164:8080/WMS">http://139.199.179.164:8080/WMS</a> .	

**18.Query 18: livrospringmvc/lojacasadocodigo:** home code version of the code using Spring MVC as framework.

No.	Top-5 retrieved project	Description	Score
1	noveogroup-amorgunov/spring-mvc-react	Example of using Spring 4 (rest full api with hibernate) + react.js (client)	
2	swapnildipankar/ java_jersey_spring_hibernate_maven		
3	<a href="#">LinuxSuRen/phoenix.platform</a>	<a href="#">WebUI automated test platform</a>	
4	sroysf/webapp-springmvc-jpa-hibernate	Demonstrates Spring MVC + Spring ORM + Hibernate + JPA + PostgreSQL	
5	socialsignin/spring-social-security-demo		

**19.Query 19: spring-projects/spring-mvc-showcase:** Demonstrates the capabilities of the Spring MVC web framework through small, simple examples. After reviewing this showcase, you should have a good understanding of what Spring MVC can do and get a feel for how easy it is to use. Includes project code along with a supporting slideshow and screen cast.

No.	Top-5 retrieved project	Description	Score
1	sroysf/webapp-springmvc-jpa-hibernate	Demonstrates Spring MVC + Spring ORM + Hibernate + JPA + PostgreSQL	
2	rwinch/spring-modern-templating		
3	ExampleDriven/cxf-example	This project demonstrates how to build a simple CXF webservice using SOAP (JAX-WS) or REST (JAX-RS). Technology Used: <ul style="list-style-type: none"><li>- CXF 2.5.2 SOAP(JAX-WS), REST(JAX-RS) client and endpoint</li><li>- Spring, Spring MVC 3.1.1</li><li>- Spring Security 3.10JSR 303 Bean validation</li></ul>	
4	livrospringmvc/lojacasadocodigo	home code version of the code using Spring MVC as framework	
5	mraible/boot-makeover	A Webapp Makeover with Spring 4 and Spring Boot	

**20.Query 20: sonian/elasticsearch-jetty:** The elasticsearch-jetty plugin brings full power of Jetty and adds several new features to **elasticsearch**. With this plugin elasticsearch can now handle SSL connections, support basic authentication, and log all or some incoming requests in plain text or json formats.

No.	Top-5 retrieved project	Description	Score
1	salyh/elasticsearch-security-plugin	Kerberos, LDAP, Active Directory, PKI/SSL/TLS and host/ip based ACL coarse-grained and document level security for <b>elasticsearch</b> (Authentication, Authorization, Auth, Spnego, ACL, Mutual authentication)	
2	jprante/elasticsearch-oai	The Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) client allows to harvest metadata into <b>Elasticsearch</b> . This application for Elasticsearch can run as a feeder in a standalone JVM, and connects to a remote cluster. It harvests DC / XML / RDF formats from OAI data providers (OAI servers). A list of OAI servers can be found here The metadata is internally represented as resources, using Resource Description Framework (RDF) of the W3C Semantic Web Initiative, before the resources are serialized into JSON for Elasticsearch schema-less indexing.	
3	javanna/elasticsearch-river-solr	The Solr River plugin allows to import data from Apache Solr to elasticsearch.	
4	searchisko/elasticsearch-river-jira	JIRA River Plugin for Elasticsearch	
5	spinscale/elasticsearch-suggest-plugin		

**21.Query 21: dadoonet/spring-elasticsearch:** Spring factories for elasticsearch. Welcome to the **Spring** factories for Elasticsearch project. Actually, since version 1.4.1, this project has been split in two parts:

- Elasticsearch Beyonder which find resources in project classpath to automatically create indices, types and templates.
- This project which is building Client beans using Spring framework.

From 5.0, this project provides 2 implementations of an elasticsearch Client:

- The REST client.

No.	Top-5 retrieved project	Description	Score
1	jprante/elasticsearch-plugin-rdf-jena	Using Apache Jena API to store and retrieve RDF triples in <b>Elasticsearch</b>	
2	SpringDataElasticsearchDevs/spring-data-elasticsearch	Spring Data implementation for ElasticSearch.	
3	hubrick/vertx-rest-client	A fully functional Vert.x REST client with RxJava support and request caching. The design is inspired by Spring's RestTemplate. The handling of MimeTypes and HttpMessageConverters is taken directly from Spring.	
4	ameizi/elasticsearch-jest-example		
5	jprante/elasticsearch-client	Elasticsearch Client is a project for generating a modularized codebase for clients that are accessing <b>Elasticsearch</b> from remote.	

**22.Query 22: elastic/elasticsearch-metrics-reporter-java:** This is a reporter for the excellent Metrics library, similar to the Graphite or Ganglia reporters, except that it reports to an Elasticsearch server. In case, you are worried, that you need to include the 20MB elasticsearch dependency in your project, you do not need to be. As this reporter is using HTTP for putting data into elasticsearch, the only library needed is the awesome Jackson JSON library, more exactly the Jackson Databind library to easily serialize the metrics objects. If you want to see this in action, go to the samples/ directory and read the readme over there, to get up and running with a sample application using the Metrics library as well as a dashboard application to graph..

No.	Top-5 retrieved project	Description	Score
1	jprante/elasticsearch-plugin-rdf-jena	Using Apache Jena API to store and retrieve RDF triples in Elasticsearch	
2	javanna/elasticsearch-river-solr	The Solr River plugin allows to import data from Apache Solr to elasticsearch.	
3	opendatasoft/elasticsearch-plugin-geoshape	This geo shape plugin can be used to index and aggregate geo shapes in elasticsearch.	
4	alexcojocaru/elasticsearch-maven-plugin	Maven plugin to run a single node Elasticsearch cluster during the integration test phase of a build	
5	<a href="#">mraible/boot-makeover</a>	<a href="#">A Webapp Makeover with Spring 4 and Spring Boot</a>	



**23.Query 23: elastic/elasticsearch-support-diagnostics:** Support diagnostics utility for elasticsearch and logstash. The support diagnostic utility is a Java executable that will interrogate the node on the the host it is running on to obtain data and statistics on the running cluster. It will execute a series of REST API calls to the running cluster, run a number of OS related commands(such as top, netstat, etc.), and collect logs and configuration, then bundle them into one or more archives.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">noveogroup-amorgunov/spring-mvc-react</a>	Example of using spring 4 (rest full api with hibernate) + react.js (client)	
2	OpenTSDB/opentsdb-elasticsearch	Search plugin for OpenTSDB	
3	<a href="#">brucezee/jspider</a>	A simple, scalable, and highly efficient web crawler framework for Java. Use Selenium to control the browser request data.	
4	<a href="#">crazyacking/zeekEye</a>	A Fast and Powerful Scraping and Web Crawling Framework.	
5	pitchpoint-solutions/sfs	The distributed object storage server used by PitchPoint Solutions to securely store billions of large and small files using minimal resources. Object data is stored in replicated volumes implemented like Facebooks Haystack Object Store. Object metadata which essentially maps an object name to a volume position is stored in an elasticsearch index. Object metadata is indexed in elasticsearch which allows for quick reads, updates, deletes and listings even when containers contain millions of objects.	

**24.Query 24: SpringDataElasticsearchDevs/spring-data-elasticsearch:** Spring Data implementation for ElasticSearch. Spring Data makes it easier to build Spring-powered applications that use new data access technologies such as non-relational databases, map-reduce frameworks, and cloud based data services as well as provide improved support for relational database technologies. The Spring Data Elasticsearch project provides integration with the elasticsearch search engine.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">bluebreezecf/SparkJobServerClient</a>	Java Client of the Spark Job Server implementing the arranged Rest APIs	
2	<a href="#">JMaNGOS/JMaNGOS</a>	JMaNGOS is a full featured World of Warcraft server, including authentication, client updates, and world content serving	
3	<a href="#">searchisko/elasticsearch-river-jira</a>	The JIRA River Plugin allows index Atlassian JIRA issues and issue comments into Elasticsearch. It's implemented as Elasticsearch river plugin and uses JIRA REST API to obtain issues from JIRA instance.	
4	<a href="#">javanna/elasticsearch-river-solr</a>	The Solr River plugin allows to import data from Apache Solr to elasticsearch.	
5	<a href="#">elaatifi/orika</a>	<a href="#">simpler, better and faster Java bean mapping framework</a>	

**25.Query 25: javanna/elasticshell:** The elasticshell is a javascript shell written in Java. It allows to interact with a running elasticsearch cluster using the Java API.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">javanna/elasticsearch-river-solr</a>	The Solr River plugin allows to import data from Apache Solr to elasticsearch.	
2	<a href="#">searchisko/elasticsearch-river-jira</a>	JIRA River Plugin for Elasticsearch	
3	<a href="#">pitchpoint-solutions/sfs</a>	The distributed object storage server used by PitchPoint Solutions to securely store billions of large and small files using minimal resources. Object data is stored in replicated volumes implemented like Facebooks Haystack Object Store. Object metadata which essentially maps an object name to a volume position is stored in an elasticsearch index. Object metadata is indexed in elasticsearch which allows for quick reads, updates, deletes and listings even when containers contain millions of objects.	
4	<a href="#">jprante/elasticsearch-oai</a>	The Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) client allows to harvest metadata into <b>Elasticsearch</b> . This application for Elasticsearch can run as a feeder in a standalone JVM, and connects to a remote cluster. It harvests DC / XML / RDF formats from OAI data providers (OAI servers). A list of OAI servers can be found here The metadata is internally represented as resources, using Resource Description Framework (RDF) of the W3C Semantic Web Initiative, before the resources are serialized into JSON for Elasticsearch schema-less indexing.	
5	<a href="#">msgpack/msgpack-hadoop</a>	<a href="#">MessagePack-Hadoop integration provides an efficient schema-free data representation for Hadoop and Hive.</a>	

**26.Query 26: mariamhakobyan/elasticsearch-river-kafka:** The Kafka River plugin allows you to read messages from Kafka and index bulked messages into elasticsearch. The bulk size (the number of messages to be indexed in one request) and concurrent request number is configurable. The Kafka River also supports consuming messages from multiple Kafka brokers and multiple partitions.

No.	Top-5 retrieved project	Description	Score
1	endgameinc/elasticsearch-river-kafka	Kafka River Plugin for ElasticSearch	
2	reachkrishnaraj/kafka-elasticsearch-standalone-consumer	Kafka Standalone Consumer [Indexer] will read messages from Kafka, in batches, process and bulk-index them into ElasticSearch.	
3	srecon/elasticsearch-cassandra-river		
4	<a href="#">flipkart-incubator/hbase-object-mapper</a>		
5	searchisko/elasticsearch-river-jira	The JIRA River Plugin allows index Atlassian JIRA issues and issue comments into Elasticsearch. It's implemented as Elasticsearch river plugin and uses JIRA REST API to obtain issues from JIRA instance.	

**27.Query 27: OpenTSDB/opentsdb-elasticsearch:** Search plugin for OpenTSDB.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">jprante/elasticsearch-plugin-rdf-jena</a>	Using Apache Jena API to store and retrieve RDF triples in Elasticsearch	
2	<a href="#">CUTR-at-USF/SiriRestClient</a>	An Android library for interacting with a RESTful SIRI interface for real-time transit data, such as that currently deployed by MTA BusTime API	
3	<a href="#">noveogroup-amorgunov/spring-mvc-react</a>	Example of using Spring 4 (rest full api with hibernate) + react.js (client)	
4	<a href="#">hubrick/vertx-rest-client</a>	A fully functional Vert.x REST client with RxJava support and request caching. The design is inspired by Spring's RestTemplate. The handling of MimeTypes and HttpMessageConverters is taken directly from Spring.	
5	<a href="#">martypitt/JsonViewExample</a>	Support for custom Jackson @JsonView within Spring MVC	

**28.Query 28: codelibs/elasticsearch-cluster-runner:** This project runs Elasticsearch cluster on one JVM instance for your development/testing easily. You can use Elasticsearch Cluster Runner as Embedded Elasticsearch in your application.

No.	Top-5 retrieved project	Description	Score
1	jprante/elasticsearch-plugin-rdf-jena	Using Apache Jena API to store and retrieve RDF triples in Elasticsearch	
2	visallo/vertexium	Vertexium is an API to manipulate graphs, similar to blueprints. Unlike blueprints, every Vertexium method requires authorizations and visibilities. Vertexium also supports multivalued properties as well as property metadata. You'll need a running Accumulo and Elastic Search cluster to try out the Accumulo implementation of Vertexium. Please see the Accumulo installation docs and Elastic Search setup guide for setting up the respective clusters.	
3	jloisel/elastic-crud	Simple yet elegant Elasticsearch Crud Repository.	
4	codelibs/fess	Fess is very powerful and easily deployable Enterprise Search Server. <a href="http://fess.codelibs.org/">http://fess.codelibs.org/</a>	
5	vhyza/elasticsearch-analysis-lemmagen	Elasticsearch lemmatizer for 15 languages	

**29.Query 29: opendatasoft/elasticsearch-plugin-geoshape:** This geo shape plugin can be used to index and aggregate geo shapes in elasticsearch. This plugin adds a geo mapping type, a geoshape aggregation, a geohash\_clustering aggregation and a geo REST entry point. geo mapping can replace geo\_shape type. It parses GeoJSON shapes, and indexes them in elasticsearch.

No.	Top-5 retrieved project	Description	Score
1	javanna/elasticsearch-river-solr	The Solr River plugin allows to import data from Apache Solr to elasticsearch.	
2	jprante/elasticsearch-plugin-rdf-jena	Using Apache Jena API to store and retrieve RDF triples in Elasticsearch	
3	rmagen/elastic-gremlin		
4	codelibs/elasticsearch-cluster-runner	This project runs Elasticsearch cluster on one JVM instance for your development/testing easily.	
5	eBay/cassandra-river	Cassandra river for Elastic search.	

**30.Query 30: huangchen007/elasticsearch-rest-command:** elasticsearch-rest-command Restful pipeline command support plugin for Elasticsearch <http://www.elasticsearch.org>

No.	Top-5 retrieved project	Description	Score
1	karussell/ElasticSearchExample	Example using ElasticSearch with a bit Wicket	
2	jprante/elasticsearch-oai	The Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) client allows to harvest metadata into <b>Elasticsearch</b> . This application for Elasticsearch can run as a feeder in a standalone JVM, and connects to a remote cluster. It harvests DC / XML / RDF formats from OAI data providers (OAI servers). A list of OAI servers can be found here The metadata is internally represented as resources, using Resource Description Framework (RDF) of the W3C Semantic Web Initiative, before the resources are serialized into JSON for Elasticsearch schema-less indexing.	
3	searchisko/elasticsearch-river-jira	The JIRA River Plugin allows index Atlassian JIRA issues and issue comments into Elasticsearch. It's implemented as <b>Elasticsearch</b> river plugin and uses JIRA REST API to obtain issues from JIRA instance.	
4	msathis/SQLToNoSQLImporter	<a href="#">Solr like data import handler to migrate data from sql systems to nosql</a>	
5	lintool/clueweb	<a href="#">Hadoop tools for manipulating ClueWeb collections</a>	



**31.Query 31: pitchpoint-solutions/sfs:** The distributed object storage server used by PitchPoint Solutions to securely store billions of large and small files using minimal resources. Object data is stored in replicated volumes implemented like Facebooks Haystack Object Store. Object metadata which essentially maps an object name to a volume position is stored in an **elasticsearch** index. Object metadata is indexed in **elasticsearch** which allows for quick reads, updates, deletes and listings even when containers contain millions of objects.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">hubrick/vertx-rest-client</a>	<a href="#">A fully functional Vert.x REST client with RxJava support and request caching. The design is inspired by Spring's RestTemplate. The handling of MimeTypes and HttpMessageConverters is taken directly from Spring.</a>	
2	<a href="#">searchisko/elasticsearch-river-jira</a>	The JIRA River Plugin allows index Atlassian JIRA issues and issue comments into Elasticsearch. It's implemented as Elasticsearch river plugin and uses JIRA REST API to obtain issues from JIRA instance.	
3	<a href="#">jprante/elasticsearch-plugin-rdf-jena</a>	Using Apache Jena API to store and retrieve RDF triples in <b>Elasticsearch</b>	
4	<a href="#">msathis/SQLToNoSQLImporter</a>	<a href="#">Solr like data import handler to migrate data from sql systems to nosql</a>	
5	<a href="#">brucezee/jspider</a>	<a href="#">A simple, scalable, and highly efficient web crawler framework for Java. Use Selenium to control the browser request data.</a>	

**32.Query 32: javanna/elasticsearch-river-solr:** The Solr River allows to query a running **Solr** instance and index the returned documents in elasticsearch. It retrieves documents via json response writer, through http get requests (solrj is not used anymore to communicate with **Solr**). All the common query parameters are supported. The solr river is not meant to keep solr and elasticsearch in sync, that's why it automatically deletes itself on completion, so that the river doesn't start up again at every node restart. This is the default behaviour, which can be disabled through the close\_on\_completion parameter.

No.	Top-5 retrieved project	Description	Score
1	jprante/elasticsearch-plugin-rdf-jena	Using Apache Jena API to store and retrieve RDF triples in Elasticsearch	
2	<a href="#">msokolov/lux</a>	<a href="#">Lux is an open source XML search engine formed by fusing two excellent technologies: the Apache Lucene/Solr search index and the Saxon XQuery/XSLT processor. At its core, Lux provides XML-aware indexing, an XQuery 1.0 optimizer that rewrites queries to use the indexes, and a function library for interacting with Lucene via XQuery. These capabilities are tightly integrated with Solr, and leverage its application framework in order to deliver a REST service, application server, and supporting tools.</a>	
3	jprante/elasticsearch-client	Elasticsearch Client is a project for generating a modularized codebase for clients that are accessing <b>Elasticsearch</b> from remote.	
4	javanna/elasticshell	The elasticshell is a javascript shell written in Java. It allows to interact with a running elasticsearch cluster using the Java API	
5	eBay/cassandra-river	Cassandra river for Elastic search.	

**33.Query 33: mesos/hadoop: Hadoop** on Mesos. To run Hadoop on Mesos you need to add the hadoop-mesos-0.1.0.jar library to your **Hadoop** distribution (any distribution that uses protobuf > 2.5.0) and set some new configuration properties. Read on for details.

No.	Top-5 retrieved project	Description	Score
1	miniway/kafka-hadoop-consumer	Another kafka-hadoop-consumer	
2	castagna/jena-grande	Jena Grande is a collection of utilities, experiments and examples on how to use MapReduce, Pig, HBase or Giraph to process data in RDF format.	
3	marginweb/Wikipedia-noSQL-Benchmark	A simple benchmark of noSQL databases for both read/update and MapReduce performances <a href="http://www.nosqlbenchmarking.com">http://www.nosqlbenchmarking.com</a>	
4	<a href="#">Netflix/ocelli</a>	<a href="#">In biology Ocelli is a simple eye as opposed to a complex eye. Ocelli the library is a simple load balancer implementation as opposed to very complex load balancers that do too many via magic or are too tightly coupled with the client transport.</a>	
5	<a href="#">huangchen007/elasticsearch-rest-command</a>	<a href="#">Restful pipeline command support plugin for Elasticsearch</a>	

**34.Query 34: pentaho/big-data-plugin:** Kettle plugin that provides support for interacting within many "big data" projects including Hadoop, Hive, HBase, Cassandra, MongoDB, and others. The Pentaho Big Data Plugin Project provides support for an ever-expanding Big Data community within the Pentaho ecosystem. It is a plugin for the Pentaho Kettle engine which can be used within Pentaho Data Integration (Kettle), Pentaho Reporting, and the Pentaho BI Platform.

No.	Top-5 retrieved project	Description	Score
1	pentaho/pentaho-hadoop-shims	Hadoop Configurations, also known and shims and the Pentaho Big Data Adaptive layer, are collections of Hadoop libraries required to communicate with a specific version of Hadoop (and related tools: Hive, HBase, Sqoop, Pig, etc.). They are designed to be easily configured.	
2	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
3	msgpack/msgpack-hadoop	MessagePack-Hadoop integration provides an efficient schema-free data representation for Hadoop and Hive.	
4	Knewton/KassandraMRHelper	Library for processing Cassandra SSTables with Hadoop MapReduce.	
5	<a href="#">qubole/quark</a>	<a href="#">Quark is a data virtualization engine over analytic databases.</a>	

**35.Query 35: asakusafw/asakusafw:** Asakusa is a full stack framework for distributed/parallel computing, which provides with a development platform and runtime libraries supporting various distributed/parallel computing environments such as Hadoop, Spark, M3 for Batch Processing, and so on. Users can enjoy the best performance on distributed/parallel computing transparently changing execution engines among MapReduce, SparkRDD, and C++ native based on their data size.

No.	Top-5 retrieved project	Description	Score
1	msgpack/msgpack-hadoop	MessagePack-Hadoop integration provides an efficient schema-free data representation for Hadoop and Hive.	
2	yc-huang/Hive-mongo	hive storage handler for connecting with MongoDB (Hadoop)	
3	lintool/clueweb	Hadoop tools for manipulating ClueWeb collections	
4	mozilla-metrics/akela	A bunch of utility classes for Java, Hadoop, HBase, Pig, etc.	
5	emsixteen/IterativeReduce	IterativeReduce - parallel iterative YARN framework for Hadoop	

**36.Query 36: klarna/HiveRunner:** An Open Source unit test framework for Hadoop hive queries based on JUnit4.

No.	Top-5 retrieved project	Description	Score
1	yc-huang/Hive-mongo	hive storage handler for connecting with MongoDB (Hadoop)	
2	sonalgoyal/hiho	Hadoop Data Integration with various databases, ftp servers, salesforce. Incremental update, dedup, append, merge your data on Hadoop.	
3	<a href="#">qubole/quark</a>	<a href="#">Quark is a data virtualization engine over analytic databases.</a>	
4	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
5	alexholmes/hadoop-utils	A set of Hadoop utilities to make working with Hadoop a little easier.	

**37.Query 37: sonalgoyal/hiho:** Hadoop Data Integration with various databases, ftp servers, salesforce. Incremental update, dedup, append, merge your data on Hadoop.

No.	Top-5 retrieved project	Description	Score
1	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
2	emsixteen/IterativeReduce	IterativeReduce - parallel iterative YARN framework for Hadoop	
3	<a href="#">qubole/quark</a>	<a href="#">Quark is a data virtualization engine over analytic databases.</a>	
4	pranab/visitante	Set of Hadoop and Storm based tools for web analytic <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
5	castagna/jena-grande	Jena Grande is a collection of utilities, experiments and examples on how to use MapReduce, Pig, HBase or Giraph to process data in RDF format.	

**38.Query 38: pranab/beymani:** Hadoop and Storm based outlier analysis implementations for cyber security and fraud detection <http://pkghosh.wordpress.com/>.

No.	Top-5 retrieved project	Description	Score
1	pranab/visitante	Set of Hadoop and Storm based tools for web analytic <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
2	pranab/chombo	Big Data ETL and Utilities for Hadoop Map Reduce, Spark and Storm <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
3	pranab/avenir	Set of Machine Learning and Stochastic Optimazion tools based on Hadoop, Spark and Storm <a href="https://pkghosh.wordpress.com/">https://pkghosh.wordpress.com/</a>	
4	jt6211/hadoop-dns-mining	This is a small framework for performing large amounts of DNS lookups using Hadoop.	
5	pranab/sifarish	Content based and collaborative filtering based recommendation and personalization engine implementation on Hadoop and Storm <a href="http://pkghosh.wordpress.com">http://pkghosh.wordpress.com</a>	



**39.Query 39: lintool/Ivory:** A Hadoop toolkit for web-scale information retrieval research. Ivory includes features presented in many academic research papers:

- A scalable inverted indexing algorithm implemented in Hadoop (TREC 2009), designed to work "out of the box" with the ClueWeb09 collection web crawl.
- A retrieval engine that implements the Markov Random Field retrieval model (SIGIR 2005), with additional extensions such as the Weighted Sequential Dependence model (WSDM 2010) and Latent Concept Expansion (SIGIR 2007).

No.	Top-5 retrieved project	Description	Score
1	lintool/clueweb	Hadoop tools for manipulating ClueWeb collections	
2	lintool/Cloud9	Cloud9 is a Hadoop toolkit for working with big data <a href="http://cloud9lib.org/">http://cloud9lib.org/</a>	
3	castagna/jena-grande	Jena Grande is a collection of utilities, experiments and examples on how to use MapReduce, Pig, HBase or Giraph to process data in RDF format.	
4	jpatanooga/KnittingBoar	Parallel Iterative Algorithm (SGD) on Hadoop's YARN framework	
5	uma-pi1/mgfsn	Sequential pattern mining aims to discover hidden patterns and relationships in collections of sequential data; it has been successfully applied in a number of data mining tasks including text mining (e.g., finding frequent phrases), market basket analysis (e.g., finding frequent sequences of product sales), and web usage mining (e.g., finding frequent sequences of page visits). In particular, in the context of text mining, frequent phrases (commonly referred to as n-grams) are widely used in applications such as machine translation, speech recognition and information extraction. The current version of MG-FSM is tested with Hadoop	

		2.5.0 and Hadoop 2.6.0.	
--	--	-------------------------	--

**40.Query 40: GoogleCloudPlatform/bigdata-interop:** Libraries and tools for interoperability between Hadoop-related open-source software and Google Cloud Platform. The Google Cloud Storage connector for Hadoop enables running MapReduce jobs directly on data in Google Cloud Storage by implementing the Hadoop FileSystem interface.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">emsixteen/IterativeReduce</a>	IterativeReduce - parallel iterative YARN framework for Hadoop	
2	<a href="#">alexholmes/hadoop-utils</a>	A set of Hadoop utilities to make working with Hadoop a little easier.	
3	<a href="#">classtag/learn_crawler</a>	Reading notes "write their own web crawler", their own knock code. Mainly recorded the basic realization of the web crawler, the page to the heavy algorithm, the page fingerprint algorithm, text information mining	
4	<a href="#">Azure/azure-documentdb-hadoop</a>	Microsoft Azure DocumentDB Hadoop Connector: This project provides a client library in Java that allows Microsoft Azure DocumentDB to act as an input source or output sink for MapReduce, Hive and Pig jobs.	
5	<a href="#">pyvandenbussche/sparqls</a>	<p>SPARQL Endpoint Status. SPARQL Endpoint Status project aims at monitoring SPARQL Endpoints based on 4 aspects:</p> <ul style="list-style-type: none"> <li>- Discoverability analyses how SPARQL endpoints can be located, what meta-data are available for them, etc.</li> <li>- Interoperability identifies which features of SPARQL 1.0 and SPARQL 1.1 standards are supported by an endpoint</li> <li>- Performance measures generic performance aspects such as result-streaming, atomic lookups and simple-joins over a HTTP connection.</li> </ul> <p>Availability monitors the uptimes of a SPARQL endpoint</p>	

**41.Query 41: Conductor/kangaroo:** Kangaroo is Conductor's collection of open source Hadoop Map/Reduce utilities. Currently, Kangaroo includes:

- A scalable Kafka input format
- Several FileInputFormats optimized for S3 input data.
- A input format for distributed task execution
- A compression codec for the framing format of Snappy.

No.	Top-5 retrieved project	Description	Score
1	miniway/kafka-hadoop-consumer	Another kafka-hadoop-consumer	
2	Knewton/KassandraMRHelper	Library for processing Cassandra SSTables with Hadoop MapReduce.	
3	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
4	amient/kafka-hadoop-loader	Hadoop Job for schemaless incremental loading of messages from Kafka topics onto hdfs with configurable output partitioning.	
5	pranab/visitante	Set of Hadoop and Storm based tools for web analytic <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	

**42.Query 42: datasalt/pangool:** Pangool is a framework on top of Hadoop that implements Tuple MapReduce.

No.	Top-5 retrieved project	Description	Score
1	pranab/visitante	Set of Hadoop and Storm based tools for web analytic <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
2	Knewton/KassandraMRHelper	Library for processing Cassandra SSTables with Hadoop MapReduce.	
3	<a href="#">msokolov/lux</a>	Lux is an open source XML search engine formed by fusing two excellent technologies: the Apache Lucene/Solr search index and the Saxon XQuery/XSLT processor. At its core, Lux provides XML-aware indexing, an XQuery 1.0 optimizer that rewrites queries to use the indexes, and a function library for interacting with Lucene via XQuery. These capabilities are tightly integrated with Solr, and leverage its application framework in order to deliver a REST service, application server, and supporting tools.	
4	castagna/jena-grande	Jena Grande is a collection of utilities, experiments and examples on how to use MapReduce, Pig, HBase or Giraph to process data in RDF format.	
5	alexholmes/hadoop-utils	A set of Hadoop utilities to make working with Hadoop a little easier.	

**43.Query 43: laserson/avro2parquet:** Hadoop MapReduce tool to convert Avro data files to Parquet format.

No.	Top-5 retrieved project	Description	Score
1	castagna/jena-grande	Jena Grande is a collection of utilities, experiments and examples on how to use MapReduce, Pig, HBase or Giraph to process data in RDF format.	
2	alexholmes/hadoop-utils	A set of Hadoop utilities to make working with Hadoop a little easier.	
3	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
4	alexholmes/hiped2	Source code that accompanies the book "Hadoop in Practice, Second Edition".	
5	Knewton/KassandraMRHelper	Library for processing Cassandra SSTables with Hadoop MapReduce.	

**44.Query 44: Knewton/KassandraMRHelper:** Library for processing Cassandra SSTables with Hadoop MapReduce.

No.	Top-5 retrieved project	Description	Score
1	eBay/cassandra-river	Cassandra river for Elastic search.	
2	<a href="#">P7h/StormTweetsSentimentD3Viz</a>		
3	pranab/visitante	Set of Hadoop and Storm based tools for web analytic <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
4	blackberry/KaBoom	A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths	
5	spotify/hdfs2cass	Hadoop mapreduce job to bulk load data into Cassandra	

**45.Query 45: blackberry/KaBoom:** A High Performance Cluster Consumer for Kafka that creates Avro (boom) files in Hadoop in time based directory paths.

No.	Top-5 retrieved project	Description	Score
1	amient/kafka-hadoop-loader	Hadoop Job for schemaless incremental loading of messages from Kafka topics onto hdfs with configurable output partitioning.	
2	Knewton/KassandraMRHelper	Library for processing Cassandra SSTables with Hadoop MapReduce.	
3	<a href="#">qubole/quark</a>	<a href="#">Quark is a data virtualization engine over analytic databases.</a>	
4	Conductor/kangaroo	Kangaroo is Conductor's collection of open source Hadoop Map/Reduce utilities.	
5	alexholmes/hadoop-utils	A set of Hadoop utilities to make working with Hadoop a little easier.	



**46.Query 46: jt6211/hadoop-dns-mining:** This is a small framework for performing large amounts of DNS lookups using Hadoop.

No.	Top-5 retrieved project	Description	Score
1	alexholmes/hadoop-utils	A set of Hadoop utilities to make working with Hadoop a little easier.	
2	pranab/beymani	Hadoop and Storm based outlier analysis implementations for cyber security and fraud detection <a href="http://pkghosh.wordpress.com/">http://pkghosh.wordpress.com/</a>	
3	<a href="#">searchisko/elasticsearch-river-jira</a>		
4	miniway/kafka-hadoop-consumer	Another kafka-hadoop-consumer	
5	<a href="#">endgameinc/elasticsearch-river-kafka</a>	<a href="#">Kafka River Plugin for ElasticSearch</a>	

**47.Query 47: psaravan/JamsMusicPlayer:** Jams is a free, powerful and elegant music player for Android. Jams used to be a trial/paid app on the Play Store. Due to my lack of adequate free time and Google's new requirement for paid developers to publicly display their home address, I've decided that I can no longer provide the level of support and regular updates that paid users expect from a developer.

No.	Top-5 retrieved project	Description	Score
1	C-Aniruddh/ACEMusicPlayer	A material music player for the android platform.	
2	SubstanceMobile/GEM	It is a local music player, show you how to use MediaPlayer to make a simple MediaPlayer.	
3	markzhai/LyricHere	Local lyric-play support music player, includes a powerful LyricView.	
4	<a href="#">pkliang/gankmaku</a>	Project created for learning how to apply the Android-CleanArchitecture with Reactive Programming and Android good practices to develop a MVP (Model View Presenter) application using <a href="#">gank.io/api</a> as the domain of the application.	
5	architjn/Auro	1st Most Fastest, Latest Designed and open source Music player	

**48.Query 48: TheAndroidMaster/Pasta-Music:** A material design music player for Android, based off Pasta for Spotify.

Pasta Music is a material design music player for android that attempts to create a better user experience than the standard music players which can have too many features or be generally confusing to users. It was created to show an improvement in design and to allow older and slower devices to have quicker access to local music files.

No.	Top-5 retrieved project	Description	Score
1	rohanoid5/Muzesto	Open source android music player with heavy implementations of visualization.	
2	<a href="#">futuresimple/android-autoprovider</a>		
3	<a href="#">Prokky/ClearRecyclerAdapter</a>		
4	PangHaHa12138/MusicPlayerdemo		
5	Freelander/ting	Ting is material design music social media player app	

**49.Query 49: SubstanceMobile/GEM:** It is a local music player, show you how to use MediaPlayer to make a simple MediaPlayer.

No.	Top-5 retrieved project	Description	Score
1	TheAndroidMaster/Pasta-Music	A material design music player for Android, based off Pasta for Spotify	
2	<a href="#">Prokky/ClearRecyclerAdapter</a>		
3	rohanoid5/Muzesto	Open source android music player with heavy implementations of visualization.	
4	<a href="#">crazyhitty/minx</a>	<a href="#">A minimalistic, text based web browser.</a>	
5	Freelander/ting	Ting is material design music social media player app	

**50.Query 50: markzhai/LyricHere:** Local lyric-play support music player, includes a powerful LyricView.

No.	Top-5 retrieved project	Description	Score
1	<a href="#">pkliang/gankmaku</a>	Project created for learning how to apply the Android-CleanArchitecture with Reactive Programming and Android good practices to develop a MVP (Model View Presenter) application using <a href="#">gank.io/api</a> as the domain of the application.	
2	<a href="#">rohanoid5/Muzesto</a>	Open source android music player with heavy implementations of visualization.	
3	<a href="#">Freelander/ting</a>		
4	<a href="#">maxiee/HeartBeat</a>		
5	<a href="#">Prokky/ClearRecyclerAdapter</a>		

