#### Similarity results using RepoPal

Date: 20/10/2017

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

Remark: Juri performed new RepoPal computation.

**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	claudiomartella/dbpedia4neo	A set of hacks to setup a dbpedia endpoint through neo4j	
2	niclashoyer/neo4j-sparql-extension	Neo4j unmanaged extension for RDF storage and SPARQL 1.1	
		query features	
3	castagna/jena-examples	A collection of ready to use, small and self contained examples on	
		how to use Apache Jena. This is a collection of small and simple	
		examples on how to use Apache Jena	
		to handle RDF data.	
4	eclipse/rdf4j	This is the main code repository for the Eclipse RDF4J project.	
		Please see RDF4J.org for detailed information about RDF4J,	
		including user documentation and downloads of the latest release.	
5	jbarrasa/neosemantics	Experiments bridging the RDF and LPG worlds.	

**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
1	AKSW/SPARQL2NL	SPARQL2NL - the SPARQL to natural language converter	
2	AKSW/Sparqlify	Sparql -> SQL Rewriter enabling virtual RDB -> RDF mappings	
3	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	castagna/jena-examples	This is a collection of small and simple examples on how to use	
		Apache Jena to handle RDF data.	
5	rdfhdt/hdt-java	HDT-lib is a Java Library that implements the W3C Submission	
		(http://www.w3.org/Submission/2011/03/) of the RDF HDT	
		(Header-Dictionary-Triples) binary format for publishing and	
		exchanging RDF data at large scale.	

**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	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/SPARQL2NL	SPARQL2NL - the SPARQL to natural language converter	
3	caskdata/cdap	The Cask <sup>™</sup> Data Application Platform (CDAP)is an integrated,	
		open source application development platform for the Hadoop	
		ecosystem that provides developers with data and application	
		abstractions to simplify and accelerate application development,	
		address a broader range of real-time and batch use cases, and	
		deploy applications into production while satisfying enterprise	
		requirements.	
4	castagna/jena-examples	This is a collection of small and simple examples on how to use	
		Apache Jena to handle RDF data.	
5	nkons/r2rml-parser	R2RML Parser is an award-winning tool that can export relational	
		database contents as RDF graphs.	

# **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/Sparqlify	Sparql -> SQL Rewriter enabling virtual RDB -> RDF mappings	
3	stardog-union/pellet	Pellet is an OWL 2 reasoner in Java; open source (AGPL) and	
		commercially licensed, commercial support available.	
		http://clarkparsia.com/pellet	
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	eclipse/rdf4j	This is the main code repository for the Eclipse RDF4J project.	
		Please see RDF4J.org for detailed information about RDF4J,	
		including user documentation and downloads of the latest release.	

- **5. Query 5: pyvandenbussche/sparqles**: 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	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.	
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. Apache Jena is a Java library, currently in	
		incubation, which can help you parsing, storing and querying RDF	
		data. This is to be considered experimental and work in progress. If	
		you want	
		to share your experience on processing RDF data with any of the	
		Hadoop ecosystem's projects	
3	nkons/r2rml-parser	R2RML Parser is an award-winning tool that can export relational	
		database contents as RDF graphs.	
4	jsonld-java/jsonld-java	JSON-LD implementation for Java	
5	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.	

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

No.	Top-5 retrieved project	Description	Score
1	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
2	arquillian/arquillian-graphene	Robust Functional Tests leveraging WebDriver with flavour of neat	
		AJAX-ready API	
3	conductor-framework/conductor	The Selenium framework that takes you where you want to go	
4	webdriverextensions/	Make your WebDriver based Selenium tests more readable,	
	webdriverextensions	reusability and maintainable by using WebDriver Extensions!	
		http://webdriverextensions.github.io	
5	bonigarcia/webdrivermanager	Automatic Selenium WebDriver binaries management in runtime	
		for Java	

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

No.	Top-5 retrieved project	Description	Score
1	spring-projects/spring-mvc- showcase	Demonstrates the features of the Spring MVC web framework	
2	sikuli/sikuli-slides	Sikuli-Slides is a visual automation tool that enables users to automate and test Graphical User Interfaces (GUIs) using presentation slides.	
3	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
4	astefanutti/metrics-cdi	CDI extension for Dropwizard Metrics. Metrics CDI activates the Metrics AOP Instrumentation for beans annotated with Metrics annotations and automatically registers the corresponding Metric instances in the Metrics registry resolved for the CDI application. The registration of these Metric instances happens each time such a bean gets instantiated. Besides, Metric instances can be retrieved from the Metrics registry by declaring metrics injection points.	
5	fhoeben/hsac-fitnesse-fixtures	An environment to define and run integration tests. It contains Fitnesse fixture (base) classes and a baseline Fitnesse installation.	

**8. Query 8: 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>.

No.	Top-5 retrieved project	Description	Score
1	bonigarcia/webdrivermanager	Automatic Selenium WebDriver binaries management in runtime	
		for Java	
2	sayems/java.webdriver	Browser test automation using Selenium WebDriver in Java	
3	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
4	dsyer/spring-boot-ratpack	This project provudes autoconfiguration and a CLI extension for	
		Spring Boot with Ratpack. Most of the features enabled here are	
		actually already in Ratpack, e.g. see @EnableRatpack. Using this	
		project as well as ratpack-spring-boot means lets you write an	
		application using Ratpack and Spring with slightly less overhead,	
		just by adding dependencies.	
5	zserge/anvil-examples	Short and descriptive examples for the Anvil MVC framework	
		http://anvil.site	

**9. Query 9: 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.

No.	Top-5 retrieved project	Description	Score
1	hitherejoe/AndroidTvBoilerplate	Android TV Boilerplate Project	
2	AndroidBootstrap/android-	A template/bootstrap/boilerplate application that includes tons of	
	bootstrap	great open source tools and frameworks.	
3	florent37/Freezer	A simple & fluent Android ORM, how can it be easier? RxJava2	
		compatible	
4	crazyhitty/Munch	A minimalistic Rss Reader for Android devices	
5	jloisel/elastic-crud	Simple yet elegant ElasticSearch Crud Repository.	

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

No.	Top-5 retrieved project	Description	Score
1	MarkusBernhardt/robotframework-	Java port of the Selenium 2 (WebDriver) Python library for Robot	
	selenium2library-java	Framework	
2	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
3	tarun3kumar/	STF is a webdriver based test automation library which generates	
	seleniumtestsframework	comprehensible tests reports comparable to BDD framework, and	
		yet better since no extra test steps / configuration are required.	
4	MachinePublishers/jBrowserDriver	A programmable, embeddable web browser driver compatible with	
		the Selenium WebDriver spec headless, WebKit-based, pure Java	
5	sayems/java.webdriver	Browser test automation using Selenium WebDriver in Java	

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

No.	Top-5 retrieved project	Description	Score
1	webdriverextensions/webdriverextension	Make your WebDriver based Selenium tests more readable,	
	S	reusability and maintainable by using WebDriver Extensions! <a href="http://webdriverextensions.github.io">http://webdriverextensions.github.io</a>	
2	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
3	sayems/java.webdriver	Browser test automation using Selenium WebDriver in Java	
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	FINRAOS/JTAF-ExtWebDriver	Extensions for WebDriver is an enhancement to the powerful WebDriver API, with robust features that keep your browser automation running smoothly. It includes a widget library, improved session management and extended functions over the existing WebDriver API.	

# **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	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
2	bonigarcia/webdrivermanager	Automatic Selenium WebDriver binaries management in runtime	
		for Java	
3	zanata/zanata-platform	Zanata is a web-based system for translators to translate	
		documentation and software online using a web browser.	
		http://zanata.org/	
4	jenkinsci/acceptance-test-harness	Acceptance tests cases for Jenkins and its plug-ins based on	
		selenium and docker.	
5	MachinePublishers/jBrowserDriver	A programmable, embeddable web browser driver compatible with	
		the Selenium WebDriver spec headless, WebKit-based, pure Java	

### 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	tarun3kumar/	STF is a webdriver based test automation library which generates			
	seleniumtestsframework	comprehensible tests reports comparable to BDD framework, and			
		yet better since no extra test steps / configuration are required.			
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	sayems/java.webdriver	Browser test automation using Selenium WebDriver in Java			
4	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and			
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>			
5	LinuxSuRen/	WebDriver WebUI automated testing framework			
	phoenix.webui.framework	http://surenpi.com			

# **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	Custom Twitter Search via ElasticSearch&Wicket	
2	hibernate/hibernate-orm	Hibernate's core Object/Relational Mapping functionality	
		http://hibernate.org	
3	heroku/heroku.jar	Java Wrapper for Heroku API	
4	Salaboy/emergency-service-drools-	This application was built to show how to architect applications	
	app	using BPM, a Rule Engine and Complex Event Processing. The	
		aim of the application is to provide a generic architecture to	
		implement flexible solution t hat uses the previous mentioned	
		topics as basic principles.	
5	karussell/Jetwick	Custom Twitter Search via ElasticSearch&Wicket	

# **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	spring-projects/spring-mvc-	Demonstrates the features of the Spring MVC web framework	
	showcase		
3	Dreampie/Resty	The minimalist framework of RESTful(server and client) - Resty	
4	hibernate/hibernate-orm	Hibernate's core Object/Relational Mapping functionality	
		http://hibernate.org	
5	blad/solid-android	Android TheNounProject Thumbnail Browser.	

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

No.	Top-5 retrieved project	Description	Score
1	white-cat/jeeweb-mybatis	SpringMVC + Spring + Mybatis + Mybatis Plus JAVA WEB agile	
		development system; it is a code generation function with	
		intelligent fast development platform; is the Spring Framework as	
		the core container, Spring MVC	
2	LinuxSuRen/	WebDriver WebUI automated testing framework	
	phoenix.webui.framework	http://surenpi.com	
3	h819/spring-boot	A Spring Boot Framework based JavaEE application, a Java	
		Commons Utils, and more java practice.	
4	shuaiweili/Dubbo-Zookeeper-Netty-	Use dubbo registration service, netty do server, springmvc	
	SpringMVC	provide restful interface	
5	bupt1987/JgFramework	Netty-based powerful game server framework	

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, Ehcahe 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, Ehcahe commonly used data cache, Disruptor as	
		a concurrent framework, Bootstrap as a front-end framework of the	
		excellent open source system	
2	biezhi/blade	Lightning Fast and Elegant Mvc Framework for Java8	
3	b3log/latke	A Spring-like, JSON based Java Web Framework	
4	allwefantasy/ServiceFramework	Java MVC framework, agile, fast, rich domain model, made	
		especially for server side of mobile application	
5	hzxie/voj	Cross-platform online judge system based on Spring MVC	
	_	framework and ActiveMQ.	

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

No.	Top-5 retrieved project	Description	Score
1	ebean-orm/ebean	Ebean ORM (JPA lite)	
2	thymeleaf/thymeleafexamples-	This is an example application showing how to compose and send	
	springmail	dynamic e-mails using Spring and Thymeleaf.	
3	alexnederlof/Jasper-report-maven-	A fast jasper report maven plugin	
	plugin		
4	thymeleaf/thymeleaf-spring	Thymeleaf integration module for Spring	
5	hantsy/angularjs-springmvc-	A RESTful sample using Spring Boot, Spring MVC, Spring Data	
	sample-boot	and Angular/Bootstrap.	

**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-	Demonstrates Spring MVC + Spring ORM + Hibernate + JPA +	
	hibernate	PostgreSQL	
2	sps/mustache-spring-view	java spring framework mvc view for Mustache.js templates	
3	xebia/Xebium	: Xebium provides Selenium (webdriver) bindings for FitNesse,	
		with Selenium-IDE support	
4	Salaboy/emergency-service-drools-	This application was built to show how to architect applications	
	app	using BPM, a Rule Engine and Complex Event Processing. The	
		aim of the application is to provide a generic architecture to	
		implement flexible solution t hat uses the previous mentioned	
		topics as basic principles.	
5	heroku/heroku.jar	Java Wrapper for Heroku API	

**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	Aconex/scrutineer	Compares a source of truth sorted stream with another to find	
		mismatches. Designed for verifying indexes such as ElasticSearch	
		& Solr are synchronized with their source of data (usually a DB)	
2	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser. <a href="http://fluentlenium.org">http://fluentlenium.org</a>	
3	salyh/elasticsearch-security-plugin	Kerberos, LDAP, Active Directory, PKI/SSL/TLS and host/ip	
		based ACL coarse-grained and document level security for	
		elasticsearch (Authentication, Authorization, Auth, Spnego, ACL,	
		Mutual authentication)	
4	searchbox-io/Jest	Elasticsearch Java Rest Client.	
5	spinscale/elasticsearch-suggest-	Plugin for elasticsearch which uses the lucene FSTSuggester	
	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 elastic search Client:

- The REST client.

No.	Top-5 retrieved project	Description	Score
1	karussell/Jetwick	Custom Twitter Search via ElasticSearch&Wicket	
2	FluentLenium/FluentLenium	FluentLenium helps you writing readable, reusable, reliable and	
		resilient UI functional tests for the browser.	
		http://fluentlenium.org	
3	hibernate/hibernate-orm	Hibernate's core Object/Relational Mapping functionality	
		http://hibernate.org	
4	FDVSolutions/DynamicJasper	Dynamic Reports using Jasper Reports <a href="http://dynamicjasper.com">http://dynamicjasper.com</a>	
5	hibernate/hibernate-search	This project provides synchronization between entities managed	
		by Hibernate ORM and full-text indexing services like Apache	
		Lucene and Elasticsearch.	

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	davidB/metrics-influxdb	A reporter for metrics which announces measurements to an	
		InfluxDB server.	
2	palantir/stash-codesearch-plugin	Provides global repository, commit, and file content search for	
		Atlassian Stash instances. Stash Codesearch is a service for	
		searching and analyzing files and commits in Atlassian Stash Git	
		repositories. It is backed by ElasticSearch (v1.3.3).	
3	vhyza/elasticsearch-analysis-	Elasticsearch lemmatizer for 15 languages	
	lemmagen		
4	searchbox-io/Jest	Elasticsearch Java Rest Client.	
5	lintool/Cloud9	Cloud9 is a Hadoop toolkit for working with big data	
		http://cloud9lib.org/	

**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	codelibs/elasticsearch-cluster-	This project runs Elasticsearch cluster on one JVM instance for	
	runner	your development/testing easily. You can use Elasticsearch Cluster	
		Runner as Embedded Elasticsearch in your application	
2	searchbox-io/Jest	Elasticsearch Java Rest Client.	
3	hibernate/hibernate-search	This project provides synchronization between entities managed by	
		Hibernate ORM and full-text indexing services like Apache	
		Lucene and Elasticsearch.	
4	naman14/Timber	Material Design Music Player	
5	klarna/HiveRunner	An Open Source unit test framework for Hadoop hive queries	
		based on JUnit4	

**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	dadoonet/spring-elasticsearch	Spring factories for elasticsearch. Welcome to the Spring factories	
		for Elasticsearch project.	
2	elastic/elasticsearch-metrics-	This is a reporter for the excellent Metrics library, similar to the	
	reporter-java	Graphite or Ganglia reporters, except that it reports to an	
		Elasticsearch server	
3	Aconex/scrutineer	Compares a source of truth sorted stream with another to find	
		mismatches. Designed for verifying indexes such as ElasticSearch	
		& Solr are synchronized with their source of data (usually a DB)	
4	jaibeermalik/searchanalytics-	Customer Product search clicks analytics using big data Hadoop,	
	bigdata	Hive, Oozie, ElasticSearch, Akka, Spring Data	
5	raphw/byte-buddy	Runtime code generation for the Java virtual machine.	
		http://bytebuddy.net	

**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	searchbox-io/Jest	Elasticsearch Java Rest Client.	
2	ncolomer/elasticsearch-osmosis-	An Osmosis plugin that index OpenStreetMap data into	
	plugin	elasticsearch	
3	jprante/elasticsearch-plugin-rdf-	Using Apache Jena API to store and retrieve RDF triples in	
	jena	Elasticsearch	
4	jprante/elasticsearch-oai	The Open Archives Initiative - Protocol for Metadata Harvesting	
		(OAI-PMH) client allows to harvest metadata into Elasticsearch.	
		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	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.	

**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	srecon/elasticsearch-cassandra-river	Elasticsearch river for Cassandra 2.* version with CQL support,	
		based on data pull method from Cassandra cluster. Project Cloned	
		and modified from the origin https://github.com/eBay/cassandra-	
		river.	
2	reachkrishnaraj/kafka-elasticsearch-	Kafka Standalone Consumer [Indexer] will read messages from	
	standalone-consumer	Kafka, in batches, process and bulk-index them into ElasticSearch.	
3	javanna/elasticsearch-river-solr	The Solr River allows to query a running Solr instance and index	
		the returned documents in elasticsearch.	
4	endgameinc/elasticsearch-river-	Kafka River Plugin for ElasticSearch	
	kafka		
5	confluentinc/kafka-connect-	Kafka Connect Elasticsearch connector	
	elasticsearch		

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

No.	Top-5 retrieved project	Description	Score
1	confluentinc/kafka-connect-hdfs	Kafka Connect HDFS connector	
2	codelibs/elasticsearch-cluster-	This project runs Elasticsearch cluster on one JVM instance for	
	runner	your development/testing easily. You can use Elasticsearch Cluster	
		Runner as Embedded Elasticsearch in your application	
3	searchbox-io/Jest	Elasticsearch Java Rest Client.	
4	Netflix/spectator	Client library for collecting metrics.	
5	codelibs/fess	Fess is very powerful and easily deployable Enterprise Search	
		Server. <a href="http://fess.codelibs.org/">http://fess.codelibs.org/</a> . Fess is Elasticsearch based search	
		server, but knowledge/experience about Elasticsearch is NOT	
		needed because of All-in-One Enterprise Search Server.	

**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	spinscale/elasticsearch-suggest-	Plugin for elasticsearch which uses the lucene FSTSuggester	
	plugin		
2	codelibs/fess	Fess is very powerful and easily deployable Enterprise Search	
		Server. <a href="http://fess.codelibs.org/">http://fess.codelibs.org/</a>	
3	hibernate/hibernate-search	This project provides synchronization between entities managed by	
		Hibernate ORM and full-text indexing services like Apache	
		Lucene and Elasticsearch.	
4	elastic/elasticsearch-support-	Support diagnostics utility for elasticsearch and logstash.	
	diagnostics		
5	alexcojocaru/elasticsearch-maven-	A Maven plugin to run a single node Elasticsearch cluster during	
	plugin	the integration test phase of a build	

**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	sirensolutions/siren-join	SIREn Plugin to add relational join capabilities to Elasticsearch	
		http://siren.solutions/searchplugins/	
2	pranab/sifarish	Content based and collaborative filtering based recommendation	
		and personalization engine implementation on Hadoop and Storm	
		http://pkghosh.wordpress.com	
3	rmagen/elastic-gremlin	TinkerPop 3 implementation on Elasticsearch backend	
4	owlike/genson	Genson a fast & modular Java <> Json library	
		http://owlike.github.io/genson/	
5	amplab/succinct	Succinct is a data store that enables queries directly on a	
		compressed representation of data. This repository maintains the	
		Java implementations of Succinct's core algorithms, and	
		applications that exploit them, such as a Apache Spark binding for	
		Succinct.	

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

No.	Top-5 retrieved project	Description	Score
1	karussell/ElasticSearchExample	Example using ElasticSearch with a bit Wicket	
2	ameizi/elasticsearch-jest-example	ElasticSearch Java Rest Client Examples	
3	jaibeermalik/searchanalytics-	Customer Product search clicks analytics using big data Hadoop,	
	bigdata	Hive, Oozie, ElasticSearch, Akka, Spring Data	
4	salyh/elasticsearch-imap	IMAP and POP3 email importer for Elasticsearch (no river	
		anymore)	
5	allwefantasy/ServiceFramework	Java MVC framework, agile, fast, rich domain model, made	
		especially for server side of mobile application	

**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	sonian/elasticsearch-jetty	The elasticsearch-jetty plugin brings full power of Jetty and adds	
		several new features to elasticsearch.	
2	sirensolutions/siren-join	SIREn Plugin to add relational join capabilities to Elasticsearch	
		http://siren.solutions/searchplugins/	
3	jprante/elasticsearch-plugin-rdf-	Using Apache Jena API to store and retrieve RDF triples in	
	jena	Elasticsearch	
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	codelibs/elasticsearch-cluster-	This project runs Elasticsearch cluster on one JVM instance for	
	runner	your development/testing easily. You can use Elasticsearch Cluster	
		Runner as Embedded Elasticsearch in your application	

**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	srecon/elasticsearch-cassandra-river	Elasticsearch river for Cassandra 2.* version with CQL support,	
		based on data pull method from Cassandra cluster. Project Cloned	
		and modified from the origin <a href="https://github.com/eBay/cassandra-">https://github.com/eBay/cassandra-</a>	
		river	
2	mariamhakobyan/elasticsearch-	The Kafka River plugin allows you to read messages from Kafka	
	river-kafka	and index bulked messages into elasticsearch.	
3	ncolomer/elasticsearch-osmosis-	An Osmosis plugin that index OpenStreetMap data into	
	plugin	elasticsearch	
4	karussell/ElasticSearchExample	Example using ElasticSearch with a bit Wicket	
5	datasalt/pangool	Pangool is a framework on top of Hadoop that implements Tuple	
		MapReduce	

**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	mayconbordin/streaminer	A collection of algorithms for mining data streams	
2	mozilla-metrics/akela	A bunch of utility classes for Java, Hadoop, HBase, Pig, etc.	
3	blad/solid-android	Android TheNounProject Thumbnail Browser.	
4	spring-projects/spring-mvc-showcase	Demonstrates the features of the Spring MVC web framework	
5	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.	

**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	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. Apache Jena is a Java library, currently in	
		incubation, which can help you parsing, storing and querying RDF	
		data. This is to be considered experimental and work in progress. If	
		you want	
		to share your experience on processing RDF data with any of the	
		Hadoop ecosystem's projects, please, do fork this.	
2	sonalgoyal/hiho	Hadoop Data Integration with various databases, ftp servers,	
		salesforce. Incremental update, dedup, append, merge your data on	
		Hadoop.	
3	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.).	
4	lintool/Cloud9	Cloud9 is a Hadoop toolkit for working with big data	
		http://cloud9lib.org/	
5	caskdata/cdap	The Cask <sup>™</sup> Data Application Platform (CDAP)is an integrated,	
		open source application development platform for the Hadoop	
		ecosystem that provides developers with data and application	
		abstractions to simplify and accelerate application development,	
		address a broader range of real-time and batch use cases, and	

deploy applications into production while satisfying enterprise	
requirements.	

**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.	<b>Top-5 retrieved project</b>	Description	Score
1	datasalt/pangool	Pangool is a framework on top of Hadoop that implements Tuple	
		MapReduce	
2	mozilla-metrics/akela	A bunch of utility classes for Java, Hadoop, HBase, Pig, etc.	
3	marginweb/Wikipedia-noSQL-	A simple benchmark of noSQL databases for both read/update and	
	Benchmark	MapReduce performances <a href="http://www.nosqlbenchmarking.com">http://www.nosqlbenchmarking.com</a>	
4	sonalgoyal/hiho	Hadoop Data Integration with various databases, ftp servers,	
		salesforce. Incremental update, dedup, append, merge your data on	
		Hadoop.	
5	ogrisel/pignlproc	Apache Pig utilities to build training corpora for machine	
		learning / NLP out of public Wikipedia and DBpedia dumps	
		(Inside: 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	dianping/wormhole	Wormhole is dianping massive data transfer tool, it currently support data source and destination like hdfs, hive, hbase, mysql, greenplum, sqlserver, mongodb, sftp, salesforce	
2	elastic/elasticsearch-support-diagnostics	Support diagnostics utility for elasticsearch and logstash.	
3	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	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	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.	

**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	ogrisel/pignlproc	Apache Pig utilities to build training corpora for machine	
		learning / NLP out of public Wikipedia and DBpedia dumps	
		(Inside: Hadoop).	
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. Apache Jena is a Java library, currently in	
		incubation, which can help you parsing, storing and querying RDF	
		data. This is to be considered experimental and work in progress. If	
		you want	
		to share your experience on processing RDF data with any of the	
		Hadoop ecosystem's projects, please, do fork this and send your	
		pull requests.	
3	marginweb/Wikipedia-noSQL-	A simple benchmark of noSQL databases for both read/update and	
	Benchmark	MapReduce performances <a href="http://www.nosqlbenchmarking.com">http://www.nosqlbenchmarking.com</a>	
4	alexholmes/hadoop-book	This repo contains the code, scripts and data files that are	
	_	referenced from the book Hadoop in Practice, published by	
		Manning.	
5	hibernate/hibernate-orm	Hibernate's core Object/Relational Mapping functionality	
		http://hibernate.org	

**38.Query 38: 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>.

No.	Top-5 retrieved project	Description	Score
1	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>	
2	pranab/visitante	Set of Hadoop and Storm based tools for web analytic	
		http://pkghosh.wordpress.com/	
3	pranab/sifarish	Content based and collaborative filtering based recommendation	
		and personalization engine implementation on Hadoop and Storm	
		http://pkghosh.wordpress.com	
4	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>	
5	jt6211/hadoop-dns-mining	This is a small framework for performing large amounts of DNS	
		lookups using Hadoop. This is a work in progress, pull requests are	
		welcome.	

- **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).
  - Implementations of "learning to efficient rank" models: the efficient sequential dependence model (SIGIR 2010), temporally-constrained ranking functions (CIKM 2010), and ranking cascades (SIGIR 2011).
  - The Postings Cartesian Product algorithm for pairwise similarity computations (SIGIR 2009).
  - A cross-language pairwise similarity computation engine based on locality-sensitive hashing (SIGIR 2011).

No.	Top-5 retrieved project	Description	Score
1	lintool/Cloud9	Cloud9 is a Hadoop toolkit for working with big data	
		http://cloud9lib.org/	
2	elasticinbox/elasticinbox	ElasticInbox - scalable, distributed email store	
		http://www.elasticinbox.com	
3	hibernate/hibernate-search	This project provides synchronization between entities managed by	
		Hibernate ORM and full-text indexing services like Apache	
		Lucene and Elasticsearch.	
4	hibernate/hibernate-orm	Hibernate's core Object/Relational Mapping functionality	
		http://hibernate.org	
5	lintool/clueweb	Hadoop tools for manipulating ClueWeb collections	

**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	dsyer/spring-boot-ratpack	This project provudes autoconfiguration and a CLI extension for	
		Spring Boot with Ratpack. Most of the features enabled here are	
		actually already in Ratpack, e.g. see @EnableRatpack. Using this	
		project as well as ratpack-spring-boot means lets you write an	
		application using Ratpack and Spring with slightly less overhead,	
		just by adding dependencies.	
2	avthart/spring-boot-axon-sample	Sample application using Spring Boot, Axon, ElasticSearch,	
		AngularJS and Websockets	
3	awslabs/emr-dynamodb-connector	Implementations of open source Apache Hadoop/Hive interfaces	
		which allow for ingesting data from Amazon DynamoDB	
4	msathis/SQLToNoSQLImporter	Solr like data import handler to migrate data from sql systems to	
		nosql	
5	qubole/quark	Quark is a data virtualization engine over analytic databases.	

**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	amient/kafka-hadoop-loader	Hadoop Job for schemaless incremental loading of messages from	
		Kafka topics onto hdfs with configurable output partitioning.	
2	miniway/kafka-hadoop-consumer	Another kafka-hadoop-consumer	
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	confluentinc/kafka-connect-	Kafka Connect Elasticsearch connector	
	elasticsearch		

### **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	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.	
2	ncolomer/elasticsearch-osmosis-	An Osmosis plugin that index OpenStreetMap data into	
	plugin	elasticsearch	
3	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>	
4	javanna/elasticsearch-river-solr	The Solr River allows to query a running Solr instance and index	
		the returned documents in elasticsearch.	
5	CeON/CoAnSys	Content Analysis System is a framework for mining scientific	
		publications using Apache Hadoop.	

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

No.	Top-5 retrieved project	Description	Score
1	HubSpot/Rosetta	Java library that leverages Jackson to take the pain out of mapping	
		objects to/from the DB, designed to integrate seamlessly with jDBI	
2	spotify/hdfs2cass	Hadoop mapreduce job to bulk load data into Cassandra	
3	confluentinc/kafka-connect-hdfs	kafka-connect-hdfs is a Kafka Connector for copying data between	
		Kafka and Hadoop HDFS.	
4	emsixteeen/IterativeReduce	IterativeReduce - parallel iterative YARN framework for Hadoop	
5	reachkrishnaraj/kafka-elasticsearch-	Kafka Standalone Consumer [Indexer] will read messages from	
	standalone-consumer	Kafka, in batches, process and bulk-index them into ElasticSearch.	

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

No.	Top-5 retrieved project	Description	Score
1	spotify/hdfs2cass	Hadoop mapreduce job to bulk load data into Cassandra	
2	confluentinc/kafka-connect- elasticsearch	Kafka Connect Elasticsearch connector	
3	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	
4	zscott/MultiBitExchange	MultiBit Exchange is an open source platform for building exchanges.	
5	eBay/cassandra-river	Cassandra river for Elastic search	

**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	Conductor/kangaroo	Kangaroo is Conductor's collection of open source Hadoop	
		Map/Reduce utilities	
3	reachkrishnaraj/kafka-elasticsearch-	Kafka Standalone Consumer [Indexer] will read messages from	
	standalone-consumer	Kafka, in batches, process and bulk-index them into ElasticSearch.	
4	confluentinc/camus	Camus is LinkedIn's Kafka->HDFS pipeline. It is a mapreduce job	
		that does distributed data loads out of Kafka.	
5	YahooArchive/samoa	SAMOA (Scalable Advanced Massive Online Analysis) is an	
		open-source platform for mining big data streams.	
		http://samoa.incubator.apache.org	

# **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	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>	
2	endgameinc/elasticsearch-river-	Kafka River Plugin for ElasticSearch	
	kafka		
3	zserge/anvil-examples	Short and descriptive examples for the Anvil MVC framework	
		http://anvil.site	
4	webdriverextensions/	Make your WebDriver based Selenium tests more readable,	
	webdriverextensions	reusability and maintainable by using WebDriver Extensions!	
		http://webdriverextensions.github.io	
5	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.	

**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	tomahawk-player/tomahawk-	Tomahawk's Android Music Player	
	android		
3	naman14/Timber	Material Design Music Player	
4	ryanhoo/StylishMusicPlayer	A stylish music player for android device 16+	
5	SubstanceMobile/GEM	A music player for Android, in stunning Material Design.	

**48.Query 48:** The Android Master/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	marverenic/Jockey	A music player for Android focused on simplicity, performance,	
		and design	
2	OpenSilk/Orpheus	Orpheus Music Player	
3	fastbootmobile/encore	An open source, extendable music player for Android.	
4	SubstanceMobile/GEM	A music player for Android, in stunning Material Design.	
5	SpongeBobSun/Prodigal	Music Player for Android which looks & feels like classic player	

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

No.	Top-5 retrieved project	Description	Score
1	harjot-oberai/MusicDNA	A Music Player for android that renders beautiful	
		DNA(Visualization) of the currently playing music.	
2	C-Aniruddh/ACEMusicPlayer	material music player for the android platform.	
3	marverenic/Jockey	A music player for Android focused on simplicity, performance,	
	-	and design	
4	psaravan/JamsMusicPlayer	Jams is a free, powerful and elegant music player for Android.	
5	architjn/Auro	1st Most Fastest, Latest Designed and open source Music player	

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

No.	Top-5 retrieved project	Description	Score
1	Mavamaarten/vk_music_android	Music An android music player for vk.com, inspired by the Google	
		Play Music app and Material Design.	
2	fastbootmobile/encore	An open source, extendable music player for Android.	
3	architjn/Auro	1st Most Fastest, Latest Designed and open source Music player	
4	psaravan/JamsMusicPlayer	Jams is a free, powerful and elegant music player for Android.	
5	TheAndroidMaster/Pasta-Music	A material design music player for Android, based off Pasta for	
		Spotify	