Valuating Code at The Apache Software Foundation

Using a weighted application of the Constructive Cost Model (CoCoMo) to put a price tag on software efforts at The Apache Software Foundation in 2022

By Sean Palmer and Daniel Gruno, DinoSource ApS

DINOSOURCE
Open Source Software Consultancy

Table of Contents:

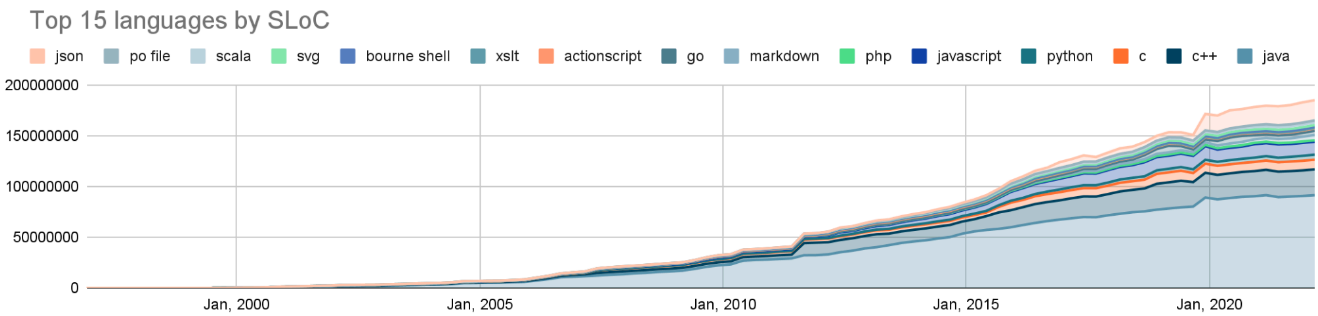
1. Executive Summary

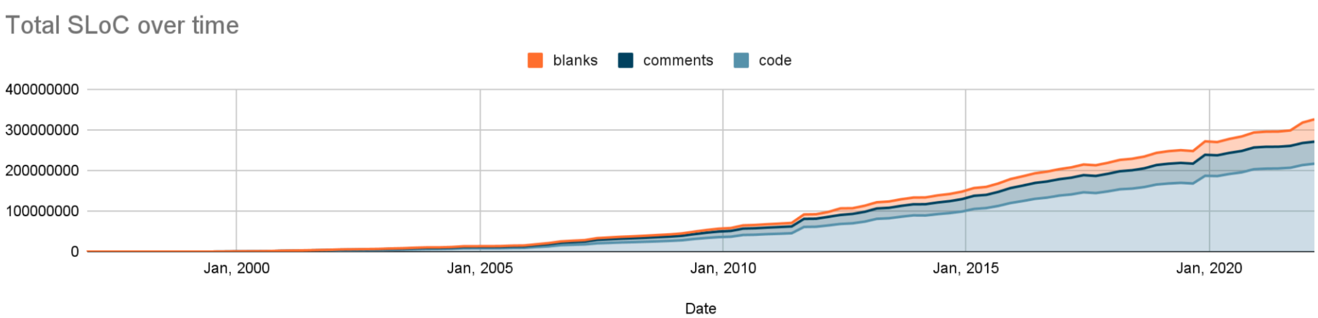
In this report, we describe the steps taken and the considerations behind valuing the combined software catalog stewarded by The Apache Software Foundation (ASF) at approximately $30B. We also explore the limitations and constraints faced when trying to value software at a massive scale.

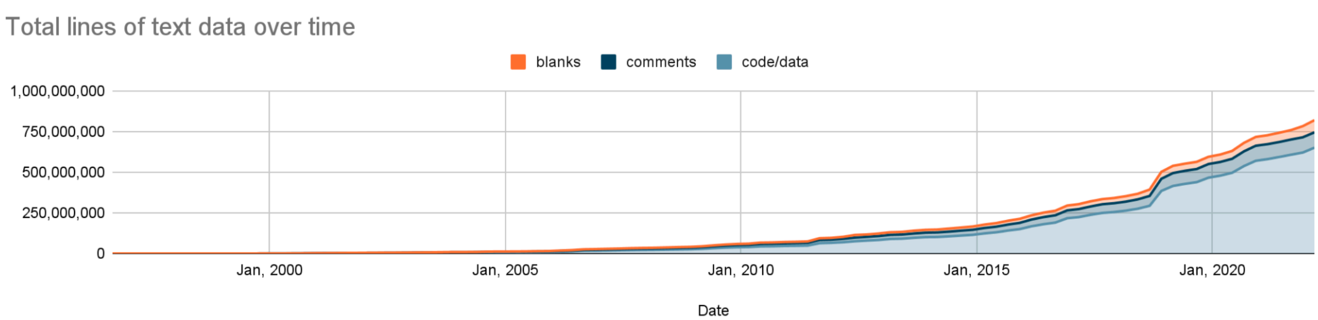
2. History of code at The Apache Software Foundation

2.1 Evolution of the code-base over time

The combined code-base (when excluding generated content such as web site repositories) at the ASF has grown steadily over its 23 year lifespan, from around 100,000 Source Lines of Code (SLoC) at the foundation's inauguration in mid 1999 to around 271,000,000 SLoC at present day, or 768,000,000 lines of text data if you include websites, tests and documentation:







The growth in code is the result of a mix between in-house development and adoption of existing code-bases through either new or existing projects joining the foundation.

3. Development Cost Metrics

Development costs can generally be split into two main determining factors:

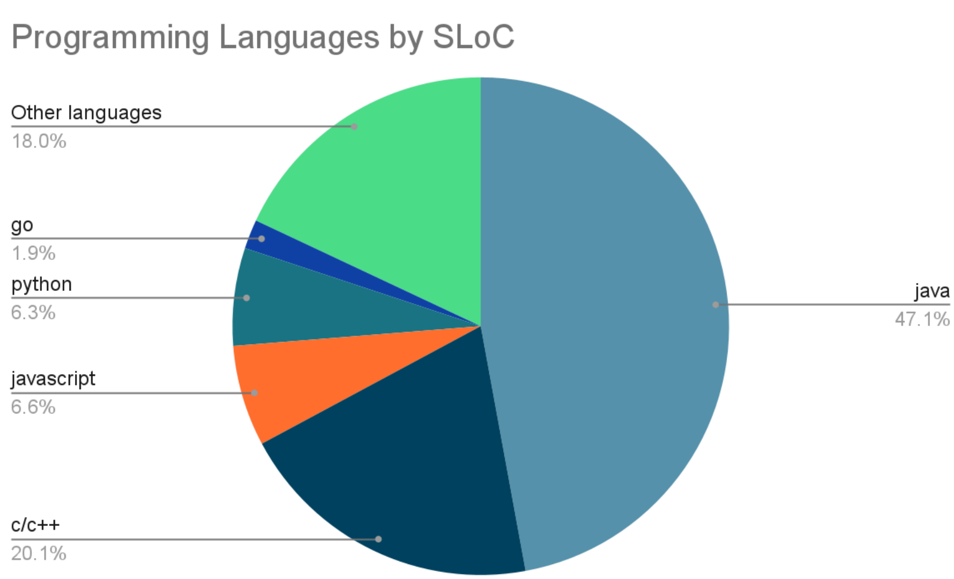
1. The effort (in time) required to develop a product.
2. The cost (in US Dollars per year in this report) of the developer(s) writing the code.

To assess these two factors, we will examine the following in this chapter:

- Language breakdown of the overall codebase, to help determine salary costs per person-year.
- A cost model to help determine the effort, in time, required to develop the code-base.
- Economic adjustment factors (EAF) such as software complexity and external constraints, which aid the cost model in calculating the effort required to produce the software.

3.1 Language distribution at the ASF

The chart below shows the overall distribution of programming languages employed at the ASF as a percentage of the total lines of code. Languages that are assumed to be mostly auto-generated, such as HTML, JSON and XML, have been stripped from this equation.



As evident, the ASF is largely a java oriented community, with C and C++ coming in at a combined second place, and JavaScript and Python almost tying for third place. The heavy reliance on Java sets the ASF apart from most organizations in 2022, as seen in the comparison chart below:

	World ranking	**ASF ranking**
C	1st	4th
Java	2nd	1st
Python	3rd	5th
C++	4th	2nd
JavaScript	5th	3rd

3.2 Calculated weighted average cost of a person-year

Taking into account the aforementioned distribution of languages, the generally perceived average lines of code written per year by an average developer focusing on a specific language, and the generally accepted style and syntax norms of different languages, we have arrived at the following estimated average person-year costs for development at the ASF per:

Language	Yearly salary avg (2022 estimates)	Weight at ASF	Cost part
Java	$72,500	40.00%	$29,000
C/C++	$92,000	25.00%	$21,160
Python	$85,000	11.00%	$9,350
JavaScript	$66,000	8.00%	$5,040
Go	$90,000	6.00%	$5,400
Other Languages (estimated average)	$75,000	10.00%	$9,000
Weighted salary in 2022 US Dollars:			$79,530

The weighted average above is an estimation founded in conservative North American (US/Canada) salary ranges, and could arguably be either higher or lower, depending on the origin and CoLA (Cost of Living Allowance) of the developers behind the software. The aim of this valuation is not to discover the actual expenditure of the software development, as such a task would be near impossible for several reasons[1], but rather to estimate what the expected costs would be for a software company situated in North America at the present time, should there be a need to recreate the software from scratch.

3.3 Estimating total effort

3.3.1 Choosing a cost model

In this report we will take a look at the two most often used cost models for software development, the Constructive Cost Models, or CoCoMo for short.

The first CoCoMo model, CoCoMo 1 (also known as CoCoMo 81), was developed by Barry W. Boehm in 1981 and was based on qualitative studies of 63 software projects, making it one of the most reliable software cost models we have.

CoCoMo 2 is an extension of version 1 that was developed at the University of Southern California in 1995 and finalized in 2000. It is backed by studies of 163 software projects and extends the logic to include circular production methods with more scaling factors and reliance on recycling old codebases. It was furthermore extended to attempt estimating cost from a function point perspective instead of only focusing on lines of code.

In short, we can describe and distinguish the two models as follows:

CoCoMo 1:

- Focused on linear development processes
- Size of software stated in terms of Lines of code
- Effort is defined using three standard complexity levels, each with four pre-set constants
- Skews towards higher per-line cost for larger projects

CoCoMo 2:

- Focused on circular development processes
- Size of software stated in terms of Object points, function points and/or lines of code
- Effort is defined using five distinct scales, with up to 22 sub-scales
- Skews towards higher per-line cost for smaller projects

[1] Software at the ASF is developed through a mix of paid and volunteer efforts by more than 50,000 contributors around the world. Putting a price on an unknown distribution of free/paid work and estimating salaries of people whose CoLA values are hidden from us would not be an easy feat.

Having considered the adjusting factors within each model, we decided to use CoCoMo 1 for this report, as it provides a more solid result given the constraints in this report and with software at such a massive scale. For more information about the constraints we have considered, please see section 3.5 in this report.

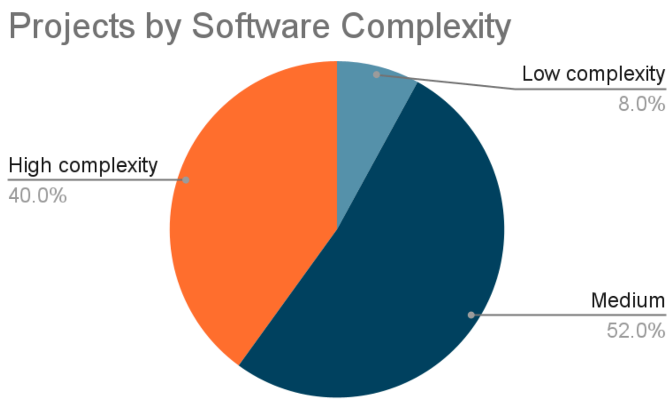
For the sake of completeness, we have also analyzed each project individually using CoCoMo 2 (see data table 1 in the appendix for both CoCoMo 1 and 2 estimates for each project), and arrived at roughly the same estimate as with CoCoMo 1 within a, considered by us, acceptable 12% margin of error, meaning the results in section 3.4 could skew by up to $3.5B depending on the model used.

3.3.2 Code complexity factors and defined weights

Using the CoCoMo 1 model, we can estimate the overall time required to replicate a collection of projects the size of what the ASF stewards. As complexity and criticality increases for certain software projects, the amount of time (in person months) required to develop, test and adjust the software rises exponentially, meaning we must use a weighted estimation, taking into account the complexity of project at the ASF and the basis for the underlying code-base, both in terms of programming languages and expertise required. It is important to emphasize that, as with most software valuations, the below metrics and weighting are based on an estimate of the complexity of products at the ASF. In this valuation, we have split the software into three main categories:

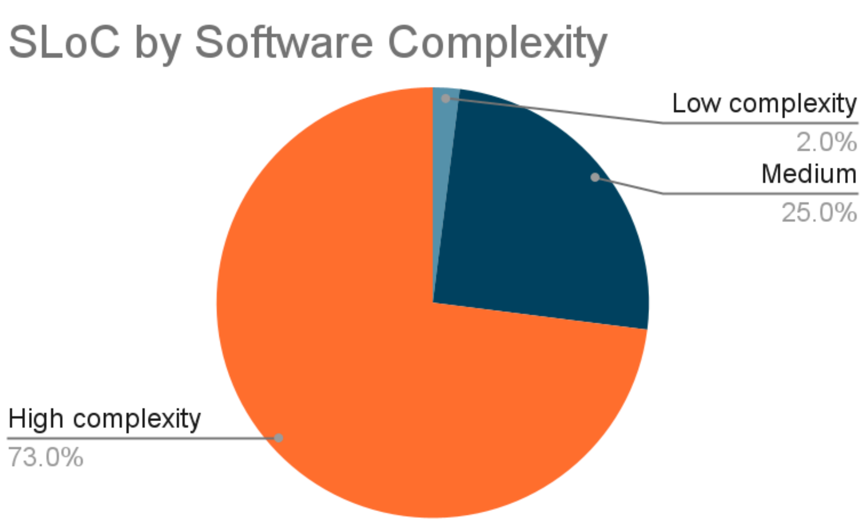
Complexity level:	**Description:**	**CoCoMo constants (a/b/c/d):**
Low complexity	Small, organic projects with less emphasis on mission criticality and testing/QA	2.4 / 1.05 / 2.5 / 0.38
Medium complexity	A intermediately complex project with a mix of rigid and less rigid constraints and tests	3.0 / 1.20 / 2.5 / 0.35
High complexity	Highly complex projects with very stringent hardware, software and test constraints	3.6 / 1.20 / 2.5 / 0.32

With these three base constant groups defined, we have assessed the complexity of projects at the ASF by looking at the *size* (in SLoC and number of developers attached), *age, purpose and total changes* *relative to overall size of the codebase* of each project[2] and estimated the levels to be in the order of 8/52/40 respectively, meaning 8% can be considered low complexity projects, 52% medium complexity and 40% high complexity. It is worth noting at this stage that these assessments should generally be considered *rough estimates*, as they are based on quantitative and not qualitative studies.



3.3.3 Code complexity as weighted per SLoC

When estimating the overall value of the code-base at the ASF, we must be careful to treat each project as an individual project. It is easy to jump to conclusions and enter all 271 million lines of code into a single equation, but that would skew the results greatly, as that would be interpreted as a single, very complex project with high maintenance costs involved. Furthermore, each level of complexity has its own code-base size, which must first be accounted for. We can easily do this by counting the lines of code that belong to a project of each of the three respective complexity categories. Doing that, we get the following distribution:



[2] See data table 1+2 on pages 11 through 19 for background data. Note that data for some software projects could not be calculated due to missing/invalid SLoC data from our code indexing provider.

This result clearly shows a major discrepancy between complexity by project volume and complexity by SLoC volume. Thus, to get a more exact result, each project must be individually calculated, as done in table 1 in the appendix.

3.3.4 Weighted effort results

Applying the three levels of complexity and their associated CoCoMo constants to the aforementioned weighted salary and size of the codebase, we are able to arrive at an estimated time to complete, and ultimately (in the next section) an estimated valuation in US Dollars.

Using the CoCoMo 1 model, $E(k) = a(\frac{k}{1000})^{b}$ where E is the effort (time required in person-months), and k is the SLoC for a single project, we can estimate the total effort required to develop the code-base, by applying the equation to each project individually (with a, b, c and d set to their respective values according to the complexity of the project, as described in section 3.3.2) and adding the results together.

Applying this approach to our data, as found in appendix 5.1, we get the following results:

	Person-months	Weight by SLoC	Weight by Projects
Low complexity portion	16,710	2%	8%
Medium complexity portion	490,436	25%	52%
High complexity portion	3,965,530	73%	40%
Combined effort:	4,472,675	100%	100%

3.4 Estimated combined costs

Once effort and average salary data has been compiled, we can combine it to get the final estimation result in US Dollars:

	Person-months	Approximate USD cost	Weight by SLoC	Weight by Projects
Low complexity portion	16,710	$111M	2%	8%
Medium complexity portion	490,436	$3,250M	25%	52%
High complexity portion	3,965,530	$26,282M	73%	40%
Combined results:	4,472,675	$29,643M	100%	100%

3.5 Known constraints in cost estimation

When valuating the code-base, we have considered the following constraints, which may affect the overall real value:

- The code-base measurements have been measured against the default branch of projects' repositories. While this is in part mitigated by incorporating all commits to all branches into the complexity levels, it cannot be ruled out that there is hidden or unpublished work in specific branches that could increase the value of a given project.

- Assumed automatically generated code such as HTML, JSON, XML has been devalued by a factor of ten so as to not skew the results, but we cannot truly know the extent of manually crafted code within these languages, and this may affect the end result.

- The analysis and complexity levels are primarily done using quantitative measurements. As such, there will surely be outliers among projects where the complexity level is off by one or more levels which may be discovered using a qualitative approach to each project. As such an endeavor comes at a high financial cost, we have chosen not to undertake it for this report.

- While we are using average salaries from 2022 data to estimate the development cost, it does not take into account the real salaries of the developers working on projects at the ASF, nor does it take into account the complexity requirements for the individual projects. We consider our estimations to be fair and balanced, but acknowledge that accounting for average salaries on a per-project and per-complexity level would yield a more precise and fine-grained result.

- As with most Open Source projects, there will be code within project repositories that have been modified or copied verbatim from other projects, both within and outside of the ASF. Such additions will invariably increase the valuation without having an actual value-add.

- Specific hardware-driven economic factors are not known. Projects that rely heavily on specific hardware and their specifications plus availability, such as Apache PLC4X or Apache Hadoop, may have a considerably higher associated cost than is given in this report.

- The Constructive Cost Models (CoCoMo 1 + 2) were not designed with software systems of this size in mind (CoCoMo for instance is based on software projects in the range 2,000 to 100,00 SLoC), and as such may either skew towards too small or large an effort required as the code-base increases by several orders of magnitude in these calculations. COSYSMO was also considered but was not found to be a viable alternative for this task.

4. Conclusion

We have reviewed the code-base under stewardship at the ASF and examined how the general principles of cost modeling under the Constructive Cost Model could best be applied, given our previously listed constraints. We have examined and assessed the individual cost aspects and levels of complexity for both each project and as an overall figure, and reached an estimated value of the code-base of $30 billion US Dollars. This covers more than 372,000 person-years spent developing the code.

5. Appendix

5.1 Data table 1: Software complexity across project at the ASF

The following data table was provided using aggregate data from Snoot (www.snoot.io)
This data has been cross-checked with the Apache Kibble (kibble.apache.org) instance running at The Apache Software Foundation. While this list covers the bulk of the project at the ASF, there are projects not covered by this analysis yet. At present, we have detailed data on 224 out of 237 individual projects at the ASF. As such, the totals in these rows do not equal the total of software at the ASF.

Project	Primary language	Age (in years)	No. of Authors	Commits	SloC	Line Changes	Complexity (estimated)	CoCoMo 1 estimate	CoCoMo 2 estimate
accumulo	java	10	212	9138	796034	5687102	2	5323	5702
activemq	java	16	411	29814	2080546	22015160	3	34524	30298
age	c	2	32	367	58004	294106	1	171	197
airavata	java	10	155	19482	2570641	25394220	2	19786	20698
airflow	python	8	2531	25370	2471219	218914556	3	42442	36610
allura	python	12	118	13746	146189	1415568	2	798	884
annotator	typescript	4	13	855	5680	127260	2	21	24
ant	java	22	129	22443	796560	8917923	3	10908	10541
any23	java	13	40	2031	75595	669954	2	381	428
apisix	go	2	382	4978	163524	866803	2	904	1000
apr	c	22	74	15280	284601	2133827	2	1682	1840
archiva	java	16	59	13721	1672400	52630088	2	12225	12901
aries	java	16	112	22868	477432	5906350	2	3003	3250
arrow	c++	6	878	12602	1182155	30536949	3	17519	16271
asterixdb	java	11	122	10161	997792	23009719	3	14294	13503
atlas	java	7	191	7748	491081	3445759	2	3099	3352
attic	java	15	634	49679	2581090	23215697	3	44716	38404
avro	java	12	289	3952	238770	2269414	3	2570	2802
axis	java	21	161	64516	2011445	113537620	3	33152	29193
bahir	java	9	134	542	43924	129035	2	207	235

beam	java	7	1312	29516	1576191	80199758	3	24742	22326
bigtop	groovy	10	216	3687	181945	5017648	3	1855	2078
bloodhound	python	10	20	1501	404055	1632346	1	1309	1665
bookkeeper	java	10	204	3905	440948	17615553	3	5365	5501
brooklyn	java	10	140	29516	1165389	7011308	2	8158	8671
brpc	c++	4	142	1996	251461	799065	2	1464	1605
buildr	ruby	14	26	4323	42982	2125630	2	202	230
bval	java	12	23	1352	35753	449648	2	165	188
calcite	java	9	449	8204	893652	32154529	3	12523	11962
camel	java	15	1317	133036	3736167	52615732	3	69697	57678
carbondata	scala	6	249	5199	471665	5172986	2	2962	3207
cassandra	java	13	597	21668	1064414	7990926	3	15447	14498
cayenne	java	15	66	7628	459824	4386293	2	2879	3118
celix	c	11	42	2705	211207	1177465	2	1204	1325
clerezza	java	12	22	4013	36391	4243990	3	269	354
cloudstack	java	11	795	51084	2097867	54139514	3	34869	30575
cocoon	java	22	76	22408	442978	5758714	2	2761	2993
comdev	javascript	11	69	6071	69355	865136	2	346	389
commons	java	20	783	88947	2341711	28161926	3	39787	34505
cordova	javascript	13	1306	38057	1523682	100155262	3	23756	21510
couchdb	javascript	16	1203	40622	493514	10370102	3	6141	6226
crail	java	5	16	420	23796	79089	1	67	74
creadur	java	14	23	2507	46451	2048864	2	221	250
ctakes	java	13	49	4400	555905	152879544	3	7084	7097
curator	java	10	147	2708	86703	1192273	2	444	497
cxf	java	13	271	37773	1161064	5348744	2	8124	8636
daffodil	scala	10	74	3766	318639	23629913	3	3633	3848
datafu	java	10	40	516	53979	792040	2	261	295
datalab	java	5	77	10194	210444	1663242	2	1200	1320
datasketches	java	6	58	5483	290781	1551357	2	1723	1884
db	java	17	18	2133	253640	6588854	2	1479	1621
deltaspike	java	10	84	2439	153024	380296	2	840	930
directory	java	18	124	36874	1432033	28172427	3	22052	20091
diversity	python	2	4	13	177	58611	2	0	0
dolphinschedule	java	3	386	5762	154737	3506721	2	850	941

r

doris	c++	4	313	5166	851273	3173368	2	5738	6139
drill	java	9	273	6391	1607594	53228391	3	25335	22816
druid	java	9	620	22539	2247231	13200572	2	17020	17853
dubbo	java	10	920	17686	648197	19636450	3	8518	8403
echarts	javascript	8	277	10310	2054332	31124839	3	34002	29878
empire-db	java	13	15	1556	76135	839305	2	384	431
felix	java	16	106	87680	795603	59007763	3	10893	10527
fineract	java	9	270	7176	365886	4642913	2	2229	2425
flagon	javascript	7	46	1354	305734	2451973	2	1823	1990
flex	mxml	12	150	38002	10640127	72962820	3	244703	182326
flink	java	11	1601	43264	2751988	37377020	3	48292	41209
flume	java	10	88	2629	608425	1002537	1	2012	2612
fluo	java	8	69	1657	74360	318912	2	374	420
freemarker	java	12	49	4380	63746	11418515	3	527	655
geode	java	7	297	16445	2236662	103286734	3	37655	32807
geronimo	java	18	101	29879	1575375	13638579	2	11434	12080
giraph	java	11	55	1153	179522	580558	2	1004	1108
gobblin	java	8	192	4610	398444	7628312	2	2452	2664
gora	java	11	74	1444	89741	478298	2	462	517
griffin	scala	5	84	1236	63517	7967340	3	525	653
groovy	java	18	445	34479	623927	7564885	3	8137	8058
guacamole	java	11	141	8615	244726	1436649	2	1421	1558
hadoop	java	16	758	108552	19026986	264369006	3	415952	379743
hawq	c	9	197	3715	2439425	15171450	2	18659	19539
hbase	java	14	737	58460	4914475	487821804	3	96845	77971
helix	java	10	109	5290	262993	3515410	2	1540	1687
heron	java	6	169	6145	310897	28955476	3	3527	3745
hive	java	13	522	20922	2610691	34274618	3	45332	38888
hivemall	java	8	36	1935	121112	1163503	2	646	719
hop	java	2	50	2863	897787	10791328	2	6091	6509
httpcomponents	java	17	157	9045	223623	2048721	2	1284	1411
httpd	c	26	169	124414	518951	75165519	3	6523	6580
hudi	java	5	327	7332	604506	6446912	2	3911	4213
iceberg	java	4	278	5352	665463	20909490	3	8791	8649

ignite	java	8	469	32131	2365598	16237150	2	18027	18890
impala	c++	10	250	11000	1001062	23411414	3	14350	13552
iotdb	java	4	255	26491	1014990	27433926	2	6988	7449
isis	java	17	85	22893	1729084	104536856	3	27649	24719
jackrabbit	java	17	131	76189	1511218	19289493	2	10914	11540
james	java	18	149	17652	1582840	37959492	3	24867	22430
jclouds	java	12	407	15784	756025	10531353	3	10246	9952
jena	java	19	143	23515	883592	32483759	3	12354	11814
jmeter	java	23	83	19842	304626	3003040	2	1815	1983
johnzon	java	7	34	853	63612	186523	1	188	218
jspwiki	java	20	28	9538	165858	38836821	3	1660	1877
juddi	java	18	22	3561	324102	7066797	2	1946	2122
juneau	java	5	24	2762	475880	5390763	2	2992	3238
kafka	java	10	1161	15645	1483718	8614347	3	23010	20890
karaf	java	16	319	25249	424995	5641800	3	5133	5282
kibble	python	4	18	873	47493	90015	1	138	158
knox	java	9	94	3214	244215	1097045	2	1417	1555
kudu	c++	9	199	10851	1133859	12099849	2	7911	8414
kylin	java	7	337	11237	370659	11682846	3	4356	4544
libcloud	python	12	442	8516	310995	2328449	3	3529	3747
liminal	javascript	2	19	163	19217	49549	1	53	58
livy	scala	6	82	547	50997	132714	2	245	277
logging	java	21	327	21066	1897352	43040292	3	30909	27377
lucene	javascript	20	418	67593	12176	105577002	3	72	106
lucenenet	javascript	16	87	7951	3689158	42006558	2	29654	30794
madlib	c++	11	115	2918	98683	6502572	3	890	1060
mahout	java	14	69	4656	374023	10131065	2	2284	2485
manifoldcf	java	12	51	17091	434137	14636621	2	2699	2927
marvin	python	4	32	463	33101	234966	1	95	106
maven	java	18	856	171620	2357598	28718651	3	40111	34763
mesos	c++	10	455	24239	1008483	33575818	3	14478	13663
milagro	c++	6	49	847	1055364	8283326	2	7300	7775
mina	java	18	124	8965	378014	25328216	3	4460	4644
mnemonic	java	6	36	762	47779	315545	2	228	258
mxnet	c++	6	1025	26991	3008661	189310110	3	53747	45455

myfaces	java	17	111	35399	2167558	41368604	3	36263	31694
mynewt	c++	6	303	19575	4898229	96423956	3	96460	77687
nemo	java	5	58	643	75709	316325	2	382	429
netbeans	java	15	246	16441	8707783	22856819	3	192394	146262
nifi	java	7	629	10150	1753437	32472545	3	28117	25102
nlpcraft	java	2	13	2852	48589	8656557	2	232	263
nutch	java	17	98	3590	98946	1495926	2	515	575
nuttx	c	15	601	49218	3795905	19591378	3	71036	58693
ofbiz	java	15	107	128713	2511660	87758532	3	43277	37269
olingo	java	8	77	2824	523490	2725950	2	3329	3596
odt	java	11	60	2651	245617	12814066	3	2658	2890
oozie	java	10	42	3070	304735	1311141	2	1816	1983
openjpa	java	15	50	7317	721859	5562908	2	4771	5121
openmeetings	java	6	22	5725	242895	13835968	2	1409	1545
opennlp	java	13	78	3416	613344	10570822	2	3975	4281
openoffice	c++	10	118	17515	9287254	926200148	3	207858	157000
openwebbeans	java	13	36	4880	165376	997881	2	916	1012
openwhisk	scala	6	438	26448	709988	4428291	2	4683	5028
orc	java	6	147	2082	176145	4303808	2	983	1085
ozone	java	3	196	5362	479840	2398091	2	3020	3268
pagespeed	c++	11	113	14628	2443195	22192486	2	18691	19572
parquet	java	9	303	3227	191821	5294162	3	1976	2202
pdfbox	java	14	39	30269	286216	20865116	3	3194	3420
pegasus	c++	4	36	1027	65961	529014	2	327	368
petri	svg	2	7	582	1162	20496	1	3	2
phoenix	java	10	272	15818	652631	6688832	2	4261	4583
pig	java	13	38	4950	463883	2319415	2	2907	3149
pinot	java	7	317	11416	509935	7854022	3	6387	6454
pivot	java	13	12	5065	168984	1437591	2	938	1037
plc4x	java	4	83	6443	144677	11279791	2	788	874
poi	java	20	44	10050	485682	9227958	2	3061	3312
ponymail	javascript	6	23	4588	87487	1177101	2	449	502
portals	java	18	37	3161	379740	4738941	2	2324	2526
pulsar	javascript	5	721	16663	2801973	60432974	3	49347	42033
qpid	java	15	185	78576	1503149	125685771	3	23372	21191

ranger	java	7	152	5107	532051	2411167	2	3390	3661
ratis	java	5	90	1444	78078	302576	2	395	443
reef	java	9	105	3199	179261	1155532	2	1002	1106
rocketmq	go	5	654	5321	1613709	8462333	2	11746	12404
roller	java	16	38	4832	118956	5149322	2	633	705
royale	actionscript	9	104	18759	1757912	8667137	2	12928	13628
rya	java	6	48	1012	533338	1967776	1	1752	2260
samza	java	8	249	2743	240848	756355	2	1395	1531
santuario	java	20	25	7239	216720	12511397	2	1240	1363
sdap	python	4	34	741	84680	210572	1	254	298
sedona	java	6	66	1060	71483	15164175	3	604	743
serf	c	20	18	2990	46010	262933	2	219	248
servicecomb	java	4	261	7861	557554	7557434	2	3572	3854
servicemix	java	16	107	38363	330277	2904938	2	1987	2167
shardingsphere	java	7	564	72935	2341838	162379619	3	39790	34507
shiro	java	16	124	2767	102228	94164945	3	929	1102
singa	c++	6	86	3958	263674	2494041	2	1544	1691
sis	java	11	25	6617	681172	1871881	1	2265	2958
skywalking	java	6	531	9793	328174	5361428	3	3764	3975
sling	java	14	292	122365	2552717	29372298	3	44127	37940
solr	java	20	401	35457	1032263	18282532	3	14888	14017
spamassassin	perl	20	58	67144	148903	16880155	3	1458	1667
spark	scala	11	2647	45159	1429879	23146869	3	22012	20058
spot	javascript	6	60	777	36370	2181315	2	168	191
steve	javascript	7	5	384	29578	76511	1	84	94
storm	java	10	540	10650	2701610	100328300	3	47233	40380
streampipes	java	7	60	8057	210207	4125525	2	1198	1318
streams	java	9	37	1769	107116	4469059	2	563	628
struts	javascript	20	181	12800	1387963	29724665	3	21240	19412
submarine	java	2	82	1196	192501	1007098	1	601	737
subversion	c	22	205	155837	1125754	58702601	3	16521	15420
superset	typescript	6	825	19806	686904	36492924	3	9132	8956
synapse	java	19	77	13209	1093904	10219103	2	7599	8088
syncope	java	11	64	11138	335081	3450652	2	2020	2202
systemds	java	10	188	7671	688920	31115214	3	9164	8985

tapestry	java	14	49	6275	548417	16749984	2	3507	3785
tcl	c	21	20	2678	115519	1015728	2	613	682
teaclave	rust	4	78	1291	280149	7390514	2	1653	1808
tez	java	8	78	4927	394805	1206899	2	2427	2637
thrift	c++	15	532	6381	371841	1571660	3	4373	4560
tika	java	14	158	7006	211784	38113212	3	2225	2455
tinkerpop	java	8	226	15230	339106	3586361	2	2047	2231
tomcat	java	21	167	105534	969925	18091626	3	13816	13089
tomee	java	20	187	30835	3467882	82999078	3	63735	53140
toree	scala	7	109	1095	36944	318287	2	171	194
trafficcontrol	go	6	209	10985	761934	9058547	2	5068	5434
trafficserver	c++	12	492	20230	702071	5439452	3	9375	9174
training	css	3	19	256	26839	250615	1	76	84
turbine	java	20	45	11188	258160	11957189	3	2822	3053
tuweni	java	2	27	1027	1147724	15668928	1	3918	5250
tvm	python	5	803	8829	863853	19507729	3	12023	11524
uima	java	15	55	29012	1626354	97451589	3	25690	23109
unomi	java	7	69	2835	123695	627882	2	662	736
vcl	php	13	17	3390	248268	1032651	1	785	974
velocity	css	21	23	6487	374007	13174834	2	2284	2484
whimsy	ruby	9	49	7218	71545	398337	2	358	403
wicket	java	17	157	33694	492312	25639612	3	6123	6209
ws	java	18	63	9635	408465	5843870	2	2521	2737
xalan	java	22	44	18523	1430170	38154295	2	10260	10861
xerces	c++	22	28	13972	298741	12683167	2	1776	1940
xmlgraphics	java	22	58	14837	964424	15217910	2	6599	7042
yetus	css	12	95	1199	41269	930202	2	193	220
yunikorn	go	3	84	1728	105465	8583803	2	553	617
zeppelin	java	8	512	6439	329678	2498899	2	1983	2163
zookeeper	java	14	269	4339	230523	20238749	3	2464	2695