Overview in General

This file contains a general overview of the data in the graph including node labels and relationships types.

References

- jqassistant
- Neo4j Python Driver

Node Labels

Table 1a - Highest node count by label combination

Lists the 30 label combinations with the highest number of nodes. The labels with the lowest node count are listed in table 1b. The total list would sum up to the total number of labels (100%).

The whole table can be found in the CSV report `Node_label_combination_count`.

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	259562	74.361773
1	[Java, ByteCode, Member, Method]	13197	3.780801
2	[Java, ByteCode, Parameter]	13029	3.732671
3	[Git, Commit]	12708	3.640708
4	[File, Git]	10810	3.096951
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.055848
6	[Java, ByteCode, Bound]	7118	2.039232
7	[Java, ByteCode, Member, Field]	3502	1.003286
8	[Java, ByteCode, Bound, WildcardType]	2908	0.833111
9	[Java, Value, ByteCode, Annotation]	2822	0.808473
10	[Xml, Element]	2144	0.614233
11	[Java, ByteCode, Member, Method, Constructor]	2039	0.584152
12	[Xml, Text]	1436	0.411399
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.313992
14	[Java, ByteCode, Member, Method, Lambda]	942	0.269873
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	877	0.251251
16	[Type, File, Java, ByteCode, Class]	810	0.232056
17	[Java, Value, ByteCode, Primitive]	652	0.186791
18	[Type, File, Java, ByteCode, JavaType]	644	0.184499
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.163012
20	[Json, Key]	560	0.160434
21	[Value, Json, Scalar]	544	0.155850
22	[Type, File, Java, ByteCode, ExternalType]	380	0.108866
23	[Author, Git, Person]	298	0.085374
24	[Value, Array]	267	0.076493
25	[Committer, Git, Person]	252	0.072195
26	[Type, File, Java, ByteCode, GenericDeclaratio...	236	0.067612
27	[Value, Property]	202	0.057871
28	[Type, File, Java, ByteCode, Interface]	188	0.053860
29	[Java, Value, ByteCode, Class]	172	0.049276

Chart 1a - Highest node count by label combination

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in Chart 1b.

<Figure size 640x480 with 0 Axes>

Nodes per label combination (more than 0.5% overall)

Legend:
- ['Git', 'Change']
- ['Java', 'ByteCode', 'Member', 'Method']
- ['Java', 'ByteCode', 'Parameter']
- ['Git', 'Commit']
- others
- ['File', 'Git']
- ['Java', 'ByteCode', 'Bound', 'ParameterizedType']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
- ['Xml', 'Element']
- ['Java', 'ByteCode', 'Member', 'Method', 'Constructor']

Slice labels:
74.36% (259562)
0.58% (2039)
0.61% (2144)
0.81% (2822)
0.83% (2908)
1.00% (3502)
2.04% (7118)
2.06% (7176)
3.10% (10810)
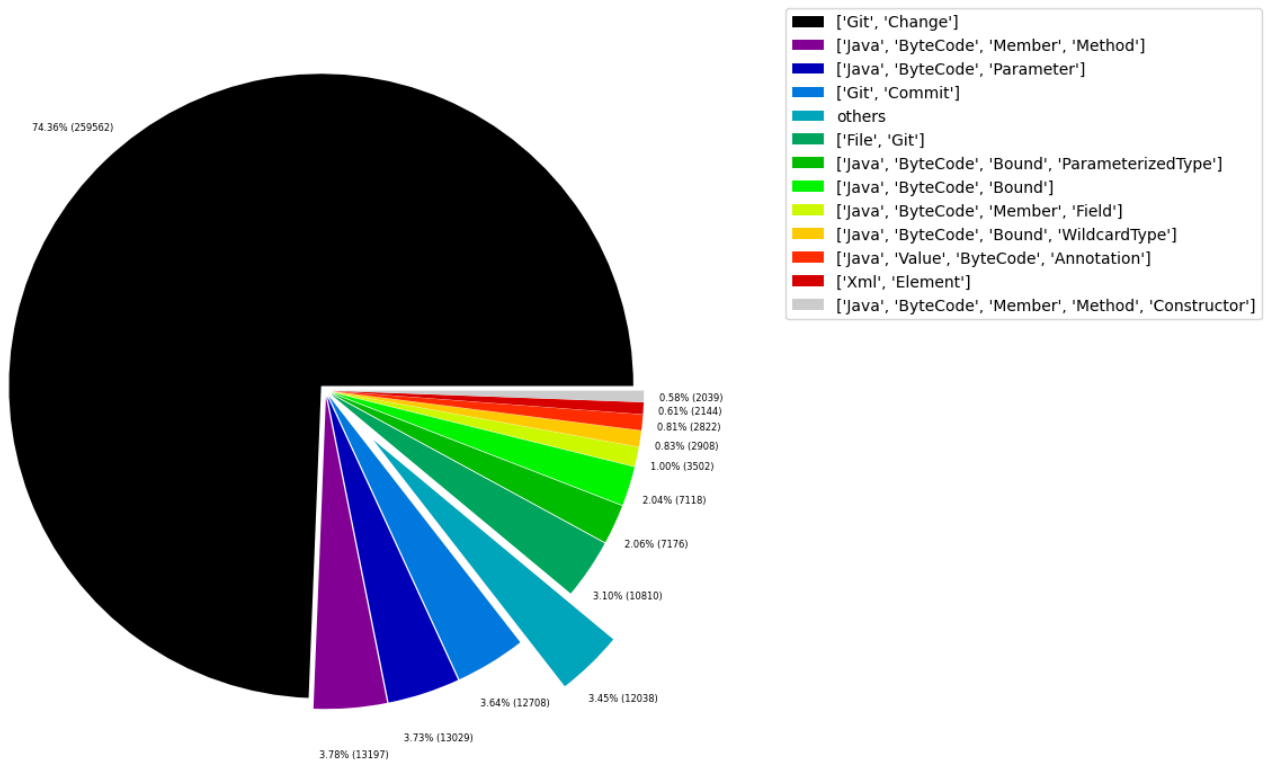3.45% (12038)
3.64% (12708)
3.73% (13029)
3.78% (13197)

Table 1b - Lowest node count by label combination

Lists the 30 label combinations with the lowest number of nodes until they reach 0.5% of the total node count, which are shown above.

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Analyze, Task, jQAssistant]	1	0.000286
1	[Repository, File, Git]	1	0.000286
2	[File, Json]	2	0.000573
3	[File]	3	0.000859
4	[Java, ByteCode, Member, Method, GenericDeclar...	4	0.001146
5	[Maven, Exclusion]	5	0.001432
6	[Value, Array, Json]	6	0.001719
7	[Type, File, Java, ByteCode, Void]	9	0.002578
8	[File, Maven, Xml, Pom, Document]	9	0.002578
9	[Java, ManifestSection]	9	0.002578
10	[File, Java, Manifest]	9	0.002578
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002578
12	[File, Java, ServiceLoader]	10	0.002865
13	[File, Java, Properties]	12	0.003438
14	[Type, File, Java, ByteCode, Throwable, Extern...	16	0.004584
15	[Maven, ExecutionGoal]	16	0.004584
16	[Maven, PluginExecution]	16	0.004584
17	[Xml, Attribute]	18	0.005157
18	[jQAssistant, Rule, Concept]	19	0.005443
19	[Maven, Plugin]	21	0.006016
20	[Maven, Configuration]	21	0.006016
21	[Type, File, Java, ByteCode, Throwable, Resolv...	22	0.006303
22	[Type, File, Java, ByteCode, Enum]	28	0.008022
23	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008308
24	[Git, Branch]	35	0.010027
25	[Xml, Namespace]	36	0.010314
26	[Type, File, Java, ByteCode, Annotation]	43	0.012319
27	[File, Directory]	44	0.012606
28	[Type, File, Java, ByteCode, Class, Throwable]	54	0.015470
29	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024352

Chart 1b - Lowest node count by label combination

Shows the lowest (less than 0.5% overall) node count label combinations. Therefore, this plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

<Figure size 640x480 with 0 Axes>

Nodes per label combination (less than 0.5% overall)

Legend:
- ['Xml', 'Text']
- ['Java', 'ByteCode', 'Bound', 'TypeVariable']
- ['Java', 'ByteCode', 'Member', 'Method', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Type', 'File', 'Java', 'ByteCode', 'Class']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'ByteCode', 'Member', 'Method', 'GenericDeclaration']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- others
- ['Author', 'Git', 'Person']
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Class']
- ['Value', 'Property']
- ['Type', 'File', 'Java', 'ByteCode', 'Interface']
- ['Java', 'Value', 'ByteCode', 'Class']
- ['Git', 'Tag']
- ['Maven', 'Dependency']
- ['Java', 'ByteCode', 'Bound', 'GenericArrayType']
- ['Java', 'Value', 'ByteCode', 'Enum']
- ['Package', 'File', 'Directory', 'Java']
- ['Value', 'Json', 'Object']
- ['Value', 'ManifestEntry']
- ['Type', 'File', 'Java', 'ByteCode', 'Throwable', 'JavaType']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType', 'ExternalAnnotation']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'Throwable']
- ['File', 'Directory']
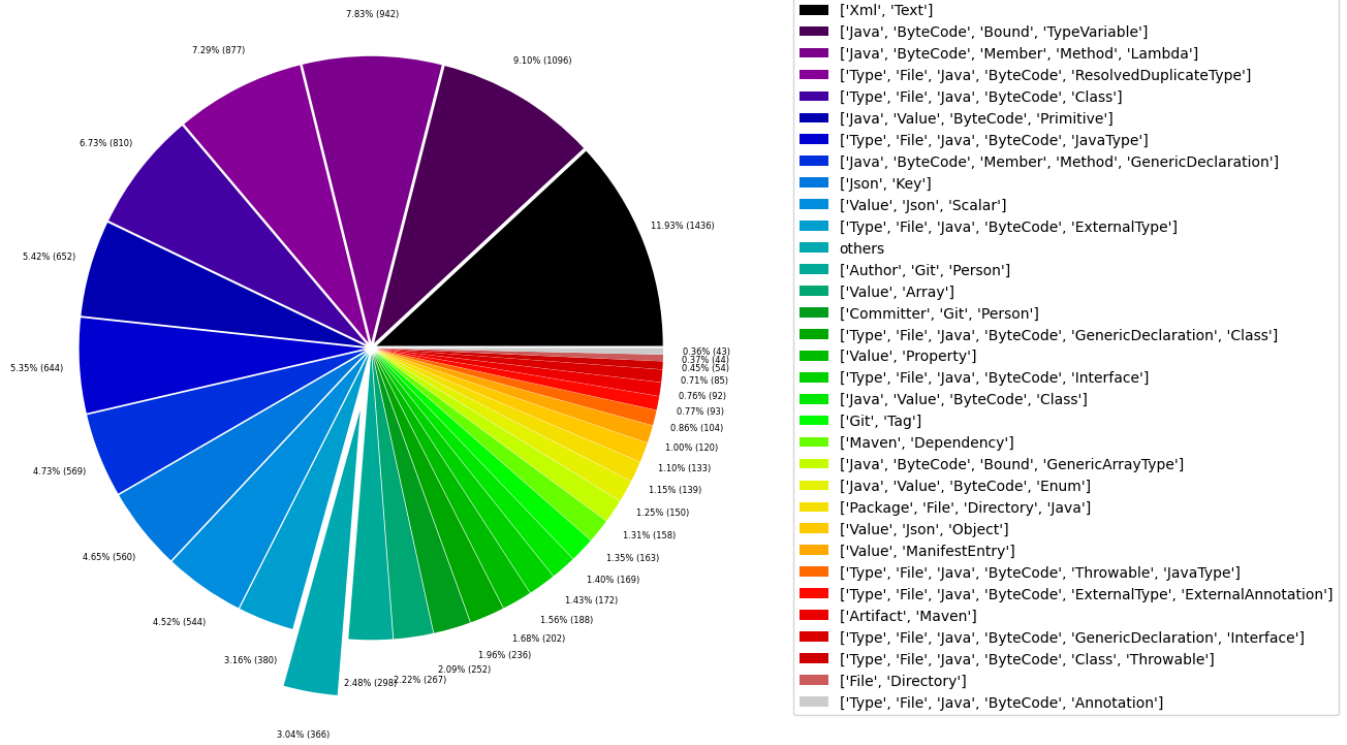- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

Table 1c - Highest node count by single label

Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.
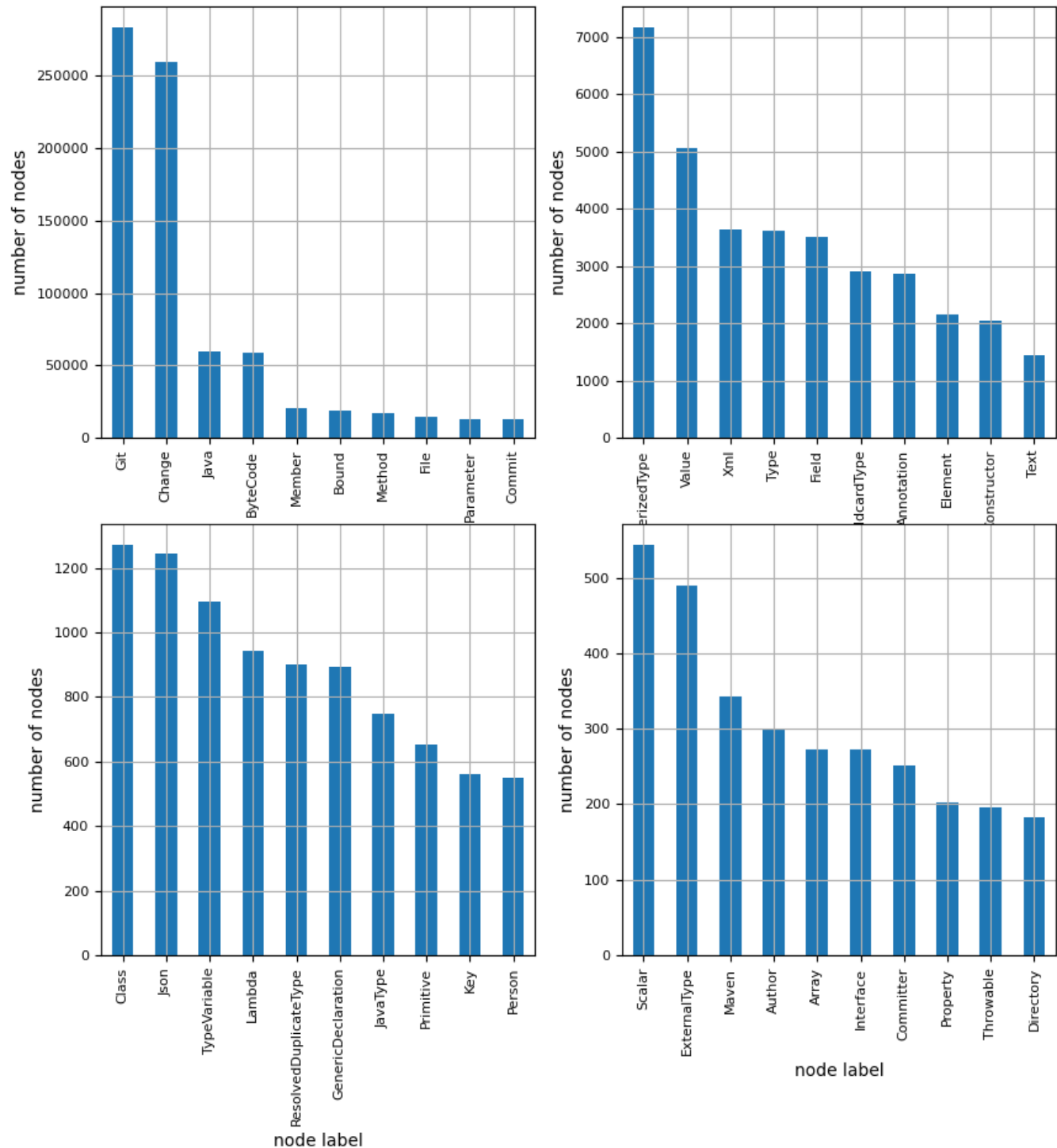
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	283835	81.315731
1	Change	259562	74.361773
2	Java	59340	17.000284
3	ByteCode	59152	16.946424
4	Member	20253	5.802271
5	Bound	18456	5.287449
6	Method	16751	4.798985
7	File	14666	4.201654
8	Parameter	13029	3.732671
9	Commit	12708	3.640708
10	ParameterizedType	7176	2.055848
11	Value	5068	1.451929
12	Xml	3643	1.043681
13	Type	3618	1.036519
14	Field	3502	1.003286
15	WildcardType	2908	0.833111
16	Annotation	2865	0.820792
17	Element	2144	0.614233
18	Constructor	2043	0.585298
19	Text	1436	0.411399
20	Class	1272	0.364415
21	Json	1245	0.356679
22	TypeVariable	1096	0.313992
23	Lambda	942	0.269873
24	ResolvedDuplicateType	899	0.257554
25	GenericDeclaration	894	0.256122
26	JavaType	748	0.214294
27	Primitive	652	0.186791
28	Key	560	0.160434
29	Person	550	0.157569
30	Scalar	544	0.155850
31	ExternalType	489	0.140093
32	Maven	343	0.098266
33	Author	298	0.085374
34	Array	273	0.078212
35	Interface	273	0.078212
36	Committer	252	0.072195
37	Property	202	0.057871
38	Throwable	196	0.056152
39	Directory	183	0.052428

Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

```
<Figure size 640x480 with 0 Axes>
```

Node count by label

Relationship Types

Table 2a - Highest relationship count by type

Lists the 30 relationship types with the highest number of occurrences. The whole table can be found in the CSV report `Relationship_type_count`.

```
Total number of relationships: 1084008
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	259562	23.944657
1	MODIFIES	259562	23.944657
2	UPDATES	170922	15.767596
3	CREATES	61205	5.646176
4	DELETES	37987	3.504310
5	INVOKES	35947	3.316119
6	COMMITTED	25416	2.344632
7	DEPENDS_ON	21931	2.023140
8	OF_TYPE	21394	1.973602
9	DECLARES	20698	1.909396
10	OF_RAW_TYPE	17093	1.576833
11	HAS_PARENT	15391	1.419823
12	HAS	14125	1.303035
13	HAS_COMMIT	12708	1.172316
14	RETURNS	12578	1.160324
15	HAS_FILE	10810	0.997225
16	RENAMES	10552	0.973425
17	READS	9161	0.845104
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.764570
19	HAS_NEW_NAME	6237	0.575365
20	OF_GENERIC_TYPE	5906	0.544830
21	RESOLVES_TO	5215	0.481085
22	SIMILAR	3977	0.366879
23	WRITES	3818	0.352211
24	CONTAINS	3814	0.351842
25	RETURNS_GENERIC	3545	0.327027
26	ANNOTATED_BY	2810	0.259223
27	REQUIRES	2174	0.200552
28	HAS_FIRST_CHILD	2144	0.197785
29	HAS_LAST_CHILD	2144	0.197785

Chart 2a - Highest relationship count by type

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in the second chart.

```
<Figure size 640x480 with 0 Axes>
```

Relationship types (more than 0.5% overall)

Legend:
- CONTAINS_CHANGE
- MODIFIES
- UPDATES
- CREATES
- others
- DELETES
- INVOKES
- COMMITTED
- DEPENDS_ON
- OF_TYPE
- DECLARES
- OF_RAW_TYPE
- HAS_PARENT
- HAS
- HAS_COMMIT
- RETURNS
- HAS_FILE
- RENAMES
- READS
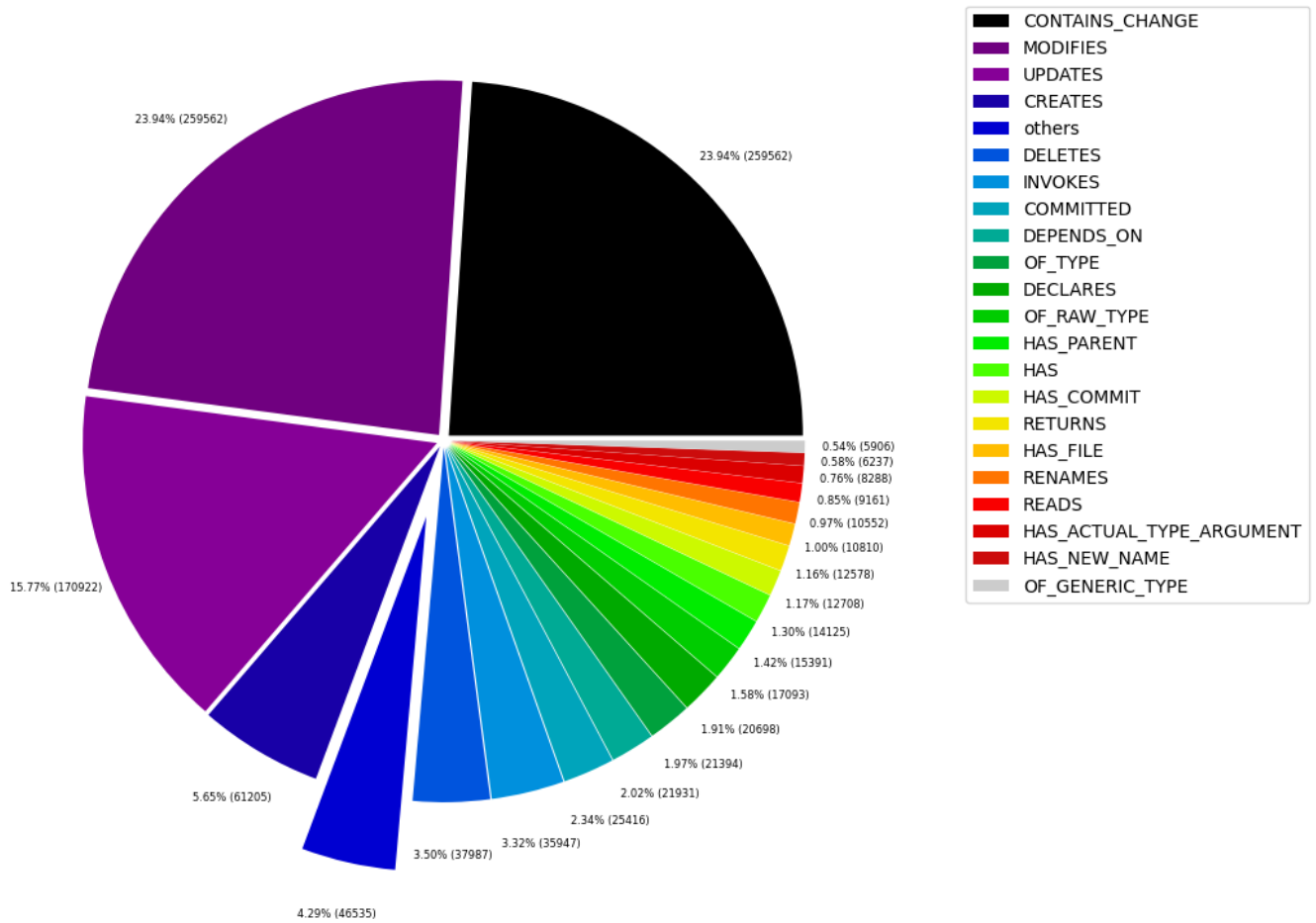- HAS_ACTUAL_TYPE_ARGUMENT
- HAS_NEW_NAME
- OF_GENERIC_TYPE

Table 2b - Lowest relationship count by type

Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.
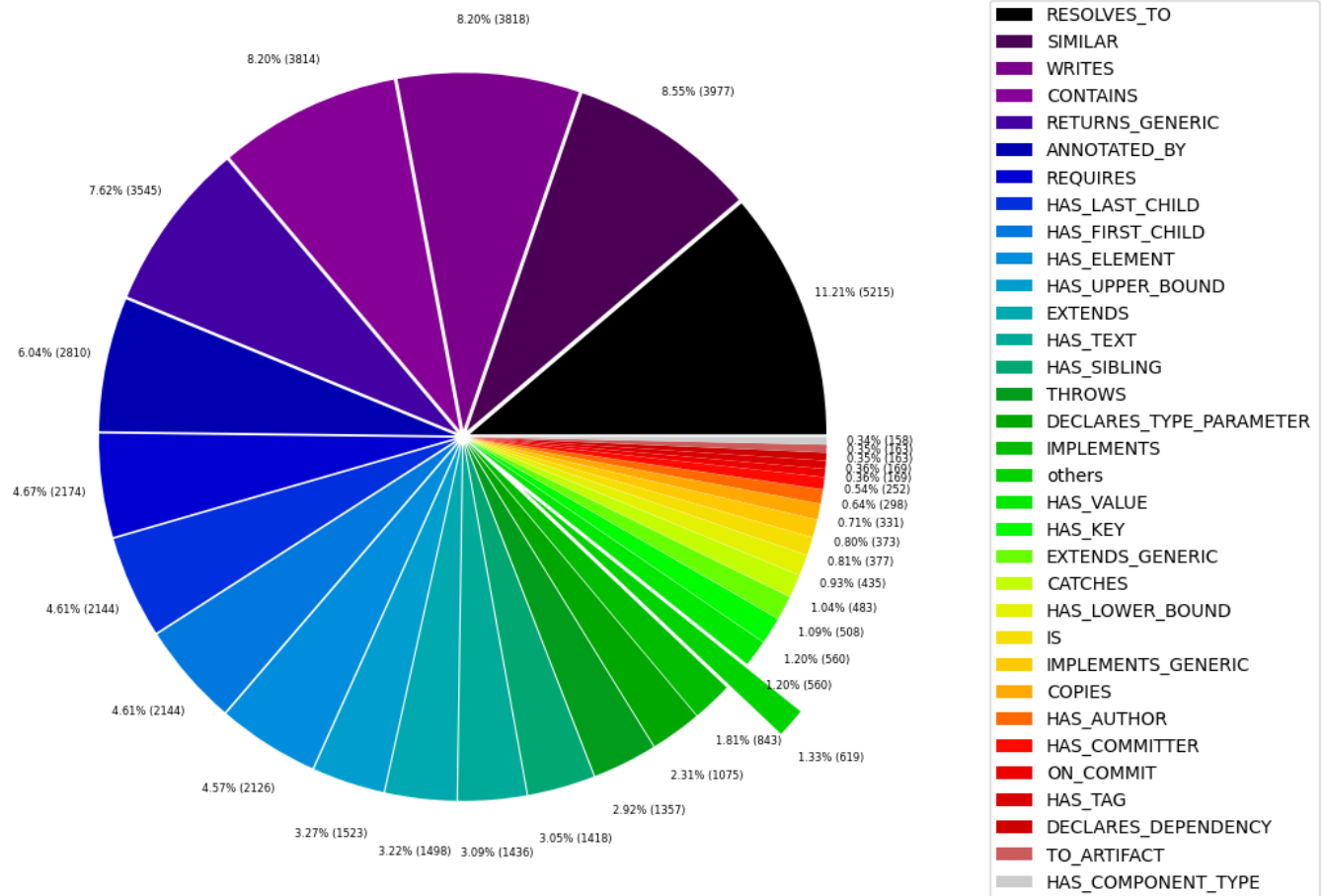
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000092
1	THROWS_GENERIC	5	0.000461
2	EXCLUDES	5	0.000461
3	DESCRIBES	9	0.000830
4	HAS_ROOT_ELEMENT	10	0.000923
5	HAS_GOAL	16	0.001476
6	HAS_EXECUTION	16	0.001476
7	OF_NAMESPACE	18	0.001661
8	HAS_ATTRIBUTE	18	0.001661
9	INCLUDES_CONCEPT	19	0.001753
10	USES_PLUGIN	21	0.001937
11	IS_ARTIFACT	21	0.001937
12	HAS_CONFIGURATION	21	0.001937
13	REQUIRES_TYPE_PARAMETER	21	0.001937
14	REQUIRES_CONCEPT	28	0.002583
15	HAS_BRANCH	35	0.003229
16	DECLARES_NAMESPACE	36	0.003321
17	HAS_HEAD	36	0.003321
18	HAS_DEFAULT	36	0.003321
19	CONTAINS_VALUE	121	0.011162
20	COPY_OF	126	0.011624
21	HAS_COMPONENT_TYPE	158	0.014576
22	DECLARES_DEPENDENCY	163	0.015037
23	TO_ARTIFACT	163	0.015037
24	ON_COMMIT	169	0.015590
25	HAS_TAG	169	0.015590
26	HAS_COMMITTER	252	0.023247
27	HAS_AUTHOR	298	0.027491
28	COPIES	331	0.030535
29	IMPLEMENTS_GENERIC	373	0.034409

Chart 2b - Lowest relationship count by type

Shows the lowest (less than 0.5% overall) relationship types. This plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

<Figure size 640x480 with 0 Axes>

Relationship types (less than 0.5% overall)

■	RESOLVES_TO
■	SIMILAR
■	WRITES
■	CONTAINS
■	RETURNS_GENERIC
■	ANNOTATED_BY
■	REQUIRES
■	HAS_LAST_CHILD
■	HAS_FIRST_CHILD
■	HAS_ELEMENT
■	HAS_UPPER_BOUND
■	EXTENDS
■	HAS_TEXT
■	HAS_SIBLING
■	THROWS
■	DECLARES_TYPE_PARAMETER
■	IMPLEMENTS
■	others
■	HAS_VALUE
■	HAS_KEY
■	EXTENDS_GENERIC
■	CATCHES
■	HAS_LOWER_BOUND
■	IS
■	IMPLEMENTS_GENERIC
■	COPIES
■	HAS_AUTHOR
■	HAS_COMMITTER
■	ON_COMMIT
■	HAS_TAG
■	DECLARES_DEPENDENCY
■	TO_ARTIFACT
■	HAS_COMPONENT_TYPE

Node labels with their relationships

Table 3a - Highest relationship count by node labels and relationship type

Lists the 30 node labels and their relationship types with the highest number of occurrences.

	sourceLabels	relationType	targetLabels	numberOfRelationships	numberOfNodesWithSameLabelsAsSource	numberOfNodesWithS
0	[Git, Change]	MODIFIES	[File, Git]	259562	259562	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	259562	12708	
2	[Git, Change]	UPDATES	[File, Git]	170922	259562	
3	[Git, Change]	CREATES	[File, Git]	61205	259562	
4	[Git, Change]	DELETES	[File, Git]	37987	259562	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21999	13102	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15382	12708	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12708	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12708	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12708	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10810	1	
11	[Git, Change]	RENAMES	[File, Git]	10552	259562	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8363	13102	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8221	13102	
14	[File, Git]	HAS_NEW_NAME	[File, Git]	6237	10810	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6079	13029	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5090	770	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4017	770	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3427	7118	
19	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2657	13029	
22	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Method, Constructor]	WRITES	[Java, ByteCode, Member, Field]	2438	2039	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...]	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2142	770	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2119	2039	
28	[Type, File, Java, ByteCode, GenericDeclaratio...]	DECLARES	[Java, ByteCode, Member, Method]	2106	230	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2080	13102	

Graph Density

```
total_number_of_nodes (vertices): 349053
total_number_of_relationships (edges): 1084008
-> total directed graph density: 8.897151413911938e-06
-> total directed graph density in percent: 0.0008897151413911938
```