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]	255467	74.136510
1	[Java, ByteCode, Member, Method]	13204	3.831800
2	[Java, ByteCode, Parameter]	13029	3.781015
3	[Git, Commit]	12364	3.588032
4	[File, Git]	10767	3.124583
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.082475
6	[Java, ByteCode, Bound]	7118	2.065643
7	[Java, ByteCode, Member, Field]	3505	1.017151
8	[Java, ByteCode, Bound, WildcardType]	2908	0.843901
9	[Java, Value, ByteCode, Annotation]	2822	0.818944
10	[Xml, Element]	2144	0.622189
11	[Java, ByteCode, Member, Method, Constructor]	2050	0.594910
12	[Xml, Text]	1436	0.416727
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.318059
14	[Java, ByteCode, Member, Method, Lambda]	942	0.273368
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.260890
16	[Type, File, Java, ByteCode, Class]	864	0.250733
17	[Type, File, Java, ByteCode, JavaType]	748	0.217070
18	[Java, Value, ByteCode, Primitive]	652	0.189210
19	[Java, ByteCode, GenericDeclaration, Member, M...	569	0.165124
20	[Json, Key]	560	0.162512
21	[Value, Json, Scalar]	544	0.157869
22	[Type, File, Java, ByteCode, ExternalType]	396	0.114919
23	[Author, Git, Person]	298	0.086480
24	[Value, Array]	267	0.077483
25	[Committer, Git, Person]	252	0.073130
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.068487
27	[Value, Property]	202	0.058620
28	[Type, File, Java, ByteCode, Interface]	188	0.054558
29	[Java, Value, ByteCode, Class]	172	0.049914

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)

74.14% (255467)

0.59% (2050)
0.62% (2144)
0.82% (2822)
0.84% (2908)
1.02% (3505)
2.07% (7118)
2.08% (7176)
3.12% (10767)
3.49% (12036)
3.59% (12364)
3.78% (13029)
3.83% (13204)

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']
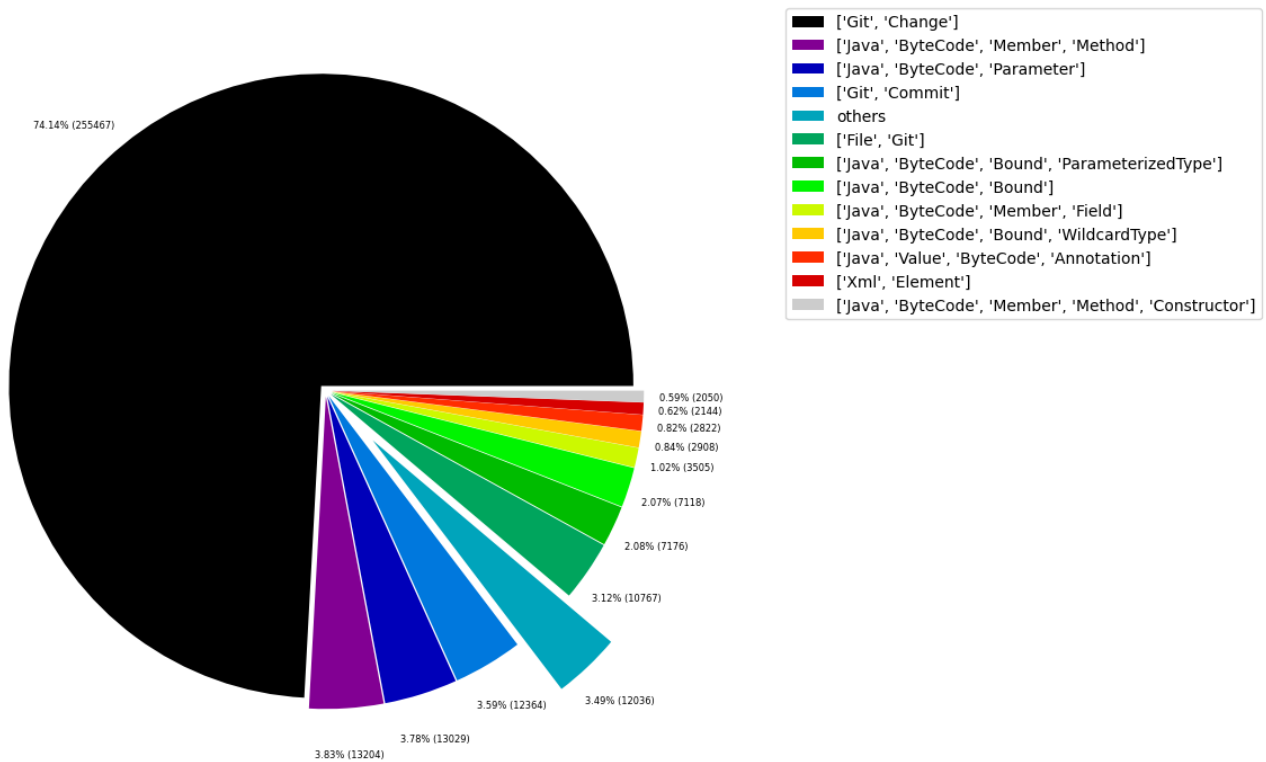- ['Java', 'ByteCode', 'Member', 'Method', 'Constructor']

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	0.000290
1	[Repository, File, Git]	1	0.000290
2	[File, Json]	2	0.000580
3	[File]	3	0.000871
4	[Java, ByteCode, GenericDeclaration, Member, M...	4	0.001161
5	[Maven, Exclusion]	5	0.001451
6	[Value, Array, Json]	6	0.001741
7	[Type, File, Java, ByteCode, Void]	9	0.002612
8	[File, Maven, Xml, Pom, Document]	9	0.002612
9	[Java, ManifestSection]	9	0.002612
10	[File, Java, Manifest]	9	0.002612
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002612
12	[File, Java, ServiceLoader]	10	0.002902
13	[File, Java, Properties]	12	0.003482
14	[Maven, ExecutionGoal]	16	0.004643
15	[Maven, PluginExecution]	16	0.004643
16	[Xml, Attribute]	18	0.005224
17	[jQAssistant, Rule, Concept]	19	0.005514
18	[Maven, Plugin]	21	0.006094
19	[Maven, Configuration]	21	0.006094
20	[Type, File, Java, ByteCode, Enum]	28	0.008126
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008416
22	[Git, Branch]	33	0.009577
23	[Xml, Namespace]	36	0.010447
24	[Type, File, Java, ByteCode, Annotation]	43	0.012479
25	[File, Directory]	44	0.012769
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024667
27	[Artifact, Maven]	92	0.026698
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.026989
29	[Value, ManifestEntry]	120	0.034824

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']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Java', 'ByteCode', 'GenericDeclaration', 'Member', 'Method']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- others
- ['Author', 'Git', 'Person']
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'GenericDeclaration']
- ['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', 'ExternalAnnotation', 'ExternalType']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
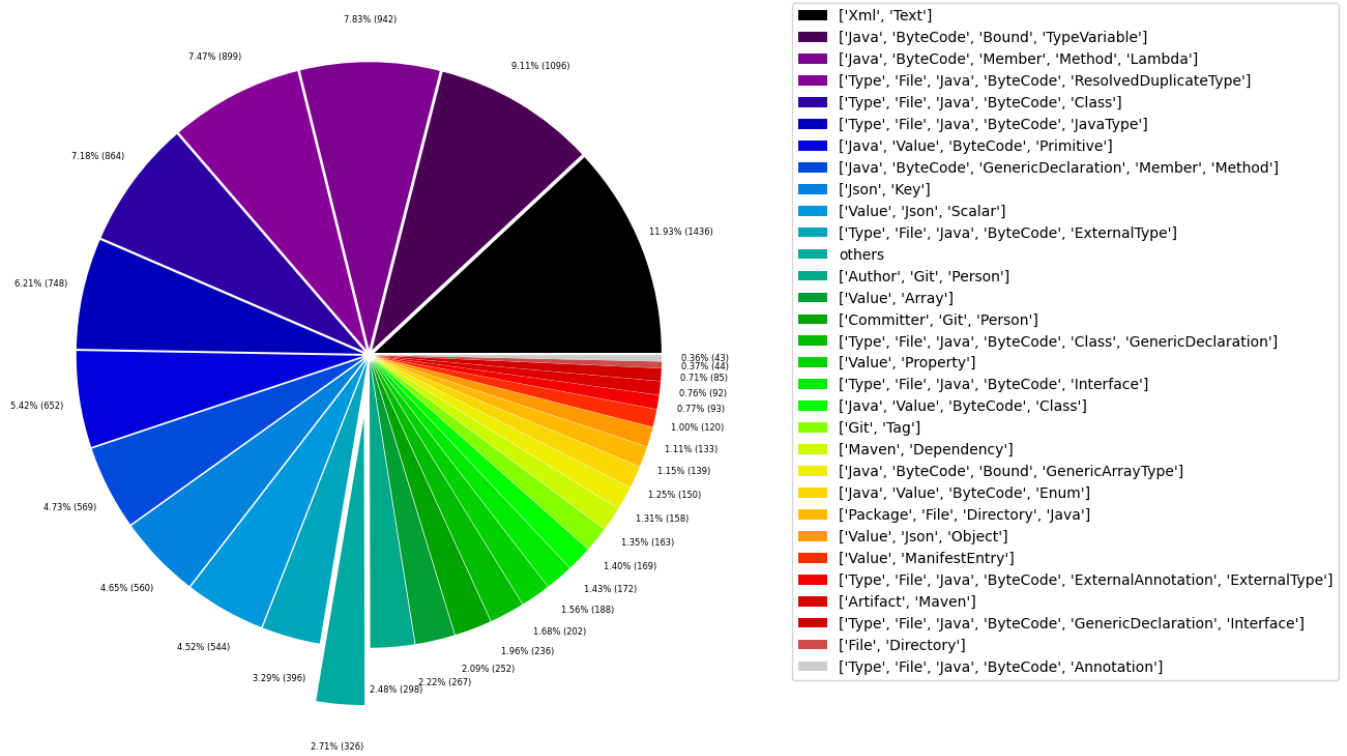- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

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.
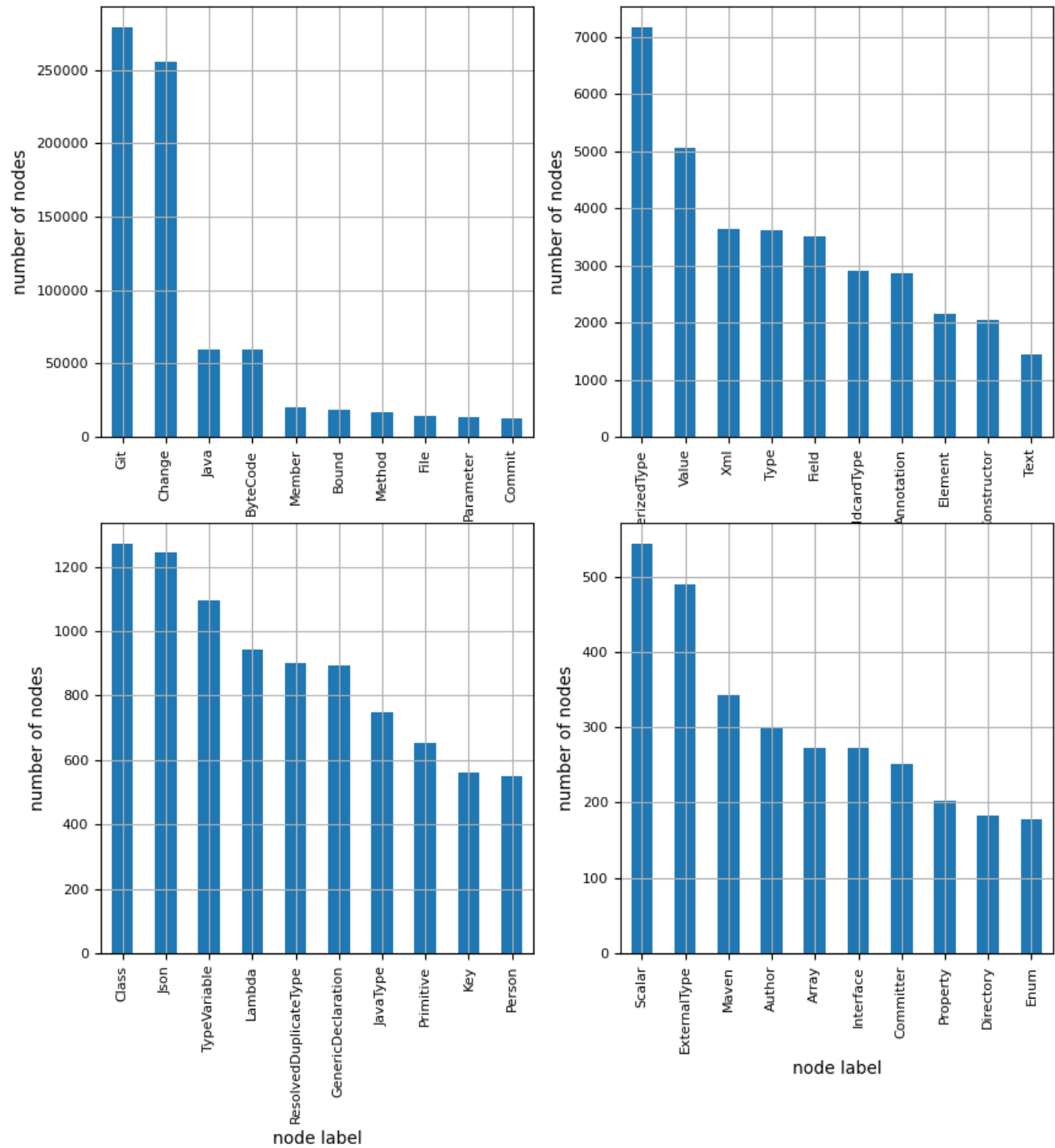
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	279351	81.067646
1	Change	255467	74.136510
2	Java	59361	17.226559
3	ByteCode	59173	17.172002
4	Member	20274	5.883514
5	Bound	18456	5.355930
6	Method	16769	4.866363
7	File	14623	4.243594
8	Parameter	13029	3.781015
9	Commit	12364	3.588032
10	ParameterizedType	7176	2.082475
11	Value	5068	1.470733
12	Xml	3643	1.057198
13	Type	3618	1.049943
14	Field	3505	1.017151
15	WildcardType	2908	0.843901
16	Annotation	2865	0.831423
17	Element	2144	0.622189
18	Constructor	2054	0.596071
19	Text	1436	0.416727
20	Class	1272	0.369134
21	Json	1245	0.361299
22	TypeVariable	1096	0.318059
23	Lambda	942	0.273368
24	ResolvedDuplicateType	899	0.260890
25	GenericDeclaration	894	0.259439
26	JavaType	748	0.217070
27	Primitive	652	0.189210
28	Key	560	0.162512
29	Person	550	0.159610
30	Scalar	544	0.157869
31	ExternalType	489	0.141908
32	Maven	343	0.099539
33	Author	298	0.086480
34	Array	273	0.079225
35	Interface	273	0.079225
36	Committer	252	0.073130
37	Property	202	0.058620
38	Directory	183	0.053107
39	Enum	178	0.051656

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: 1069399
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	255467	23.888838
1	MODIFIES	255467	23.888838
2	UPDATES	167307	15.644956
3	CREATES	60811	5.686465
4	DELETES	37841	3.538530
5	INVOKES	35689	3.337295
6	COMMITTED	24728	2.312327
7	DEPENDS_ON	21931	2.050778
8	OF_TYPE	21394	2.000563
9	DECLARES	20719	1.937443
10	OF_RAW_TYPE	17093	1.598374
11	HAS_PARENT	15025	1.404995
12	HAS	14125	1.320835
13	RETURNS	12578	1.176175
14	HAS_COMMIT	12364	1.156163
15	HAS_FILE	10767	1.006827
16	RENAMES	10492	0.981112
17	READS	9164	0.856930
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.775015
19	HAS_NEW_NAME	6219	0.581542
20	OF_GENERIC_TYPE	5906	0.552273
21	RESOLVES_TO	5190	0.485319
22	SIMILAR	3977	0.371891
23	WRITES	3818	0.357023
24	CONTAINS	3814	0.356649
25	RETURNS_GENERIC	3545	0.331495
26	ANNOTATED_BY	2810	0.262764
27	REQUIRES	2174	0.203292
28	HAS_FIRST_CHILD	2144	0.200486
29	HAS_LAST_CHILD	2144	0.200486

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>
```
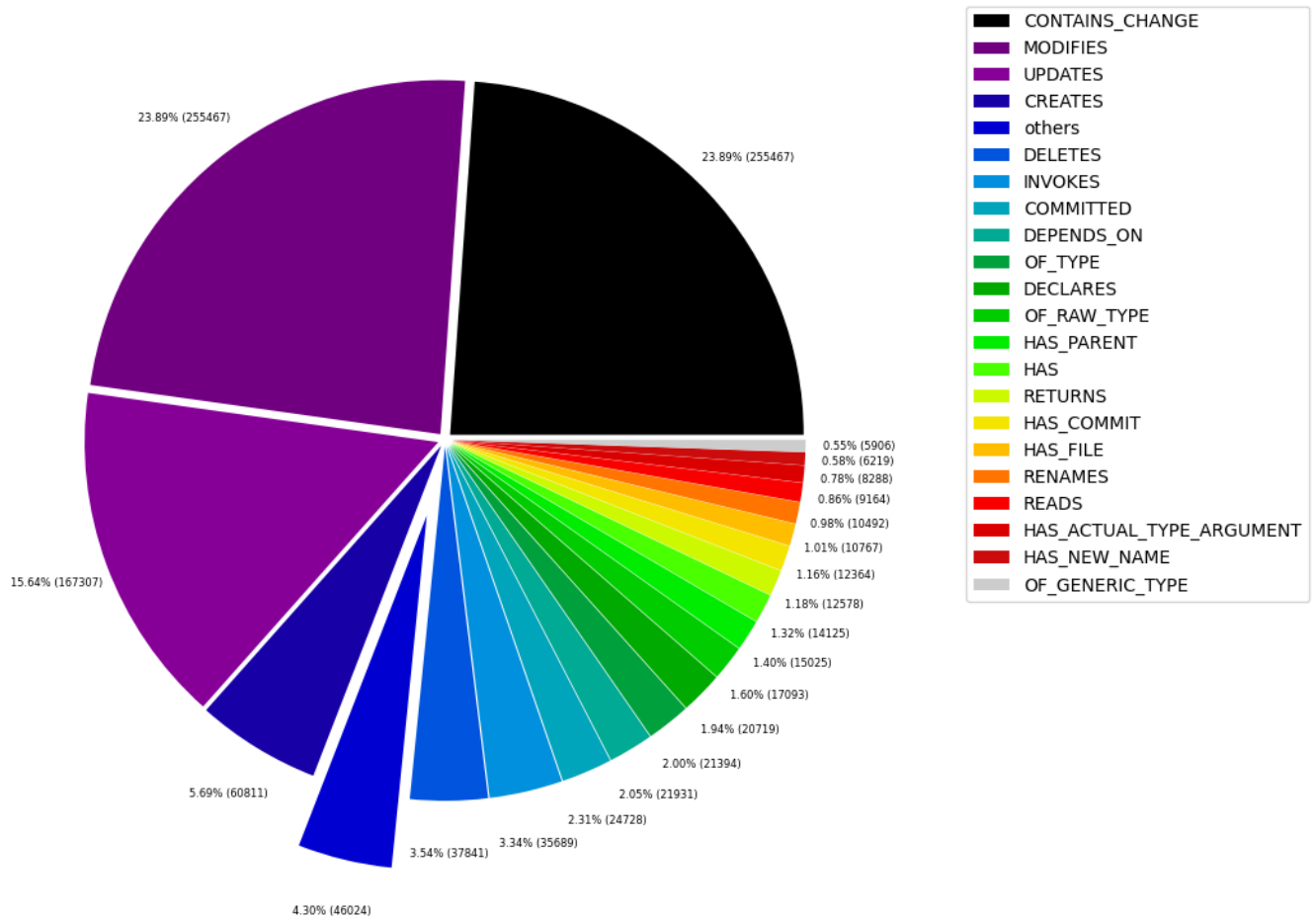
Relationship types (more than 0.5% overall)

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

Pie chart values:
- 23.89% (255467)
- 23.89% (255467)
- 15.64% (167307)
- 5.69% (60811)
- 4.30% (46024)
- 3.54% (37841)
- 3.34% (35689)
- 2.31% (24728)
- 2.05% (21931)
- 2.00% (21394)
- 1.94% (20719)
- 1.60% (17093)
- 1.40% (15025)
- 1.32% (14125)
- 1.18% (12578)
- 1.16% (12364)
- 1.01% (10767)
- 0.98% (10492)
- 0.86% (9164)
- 0.78% (8288)
- 0.58% (6219)
- 0.55% (5906)

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.
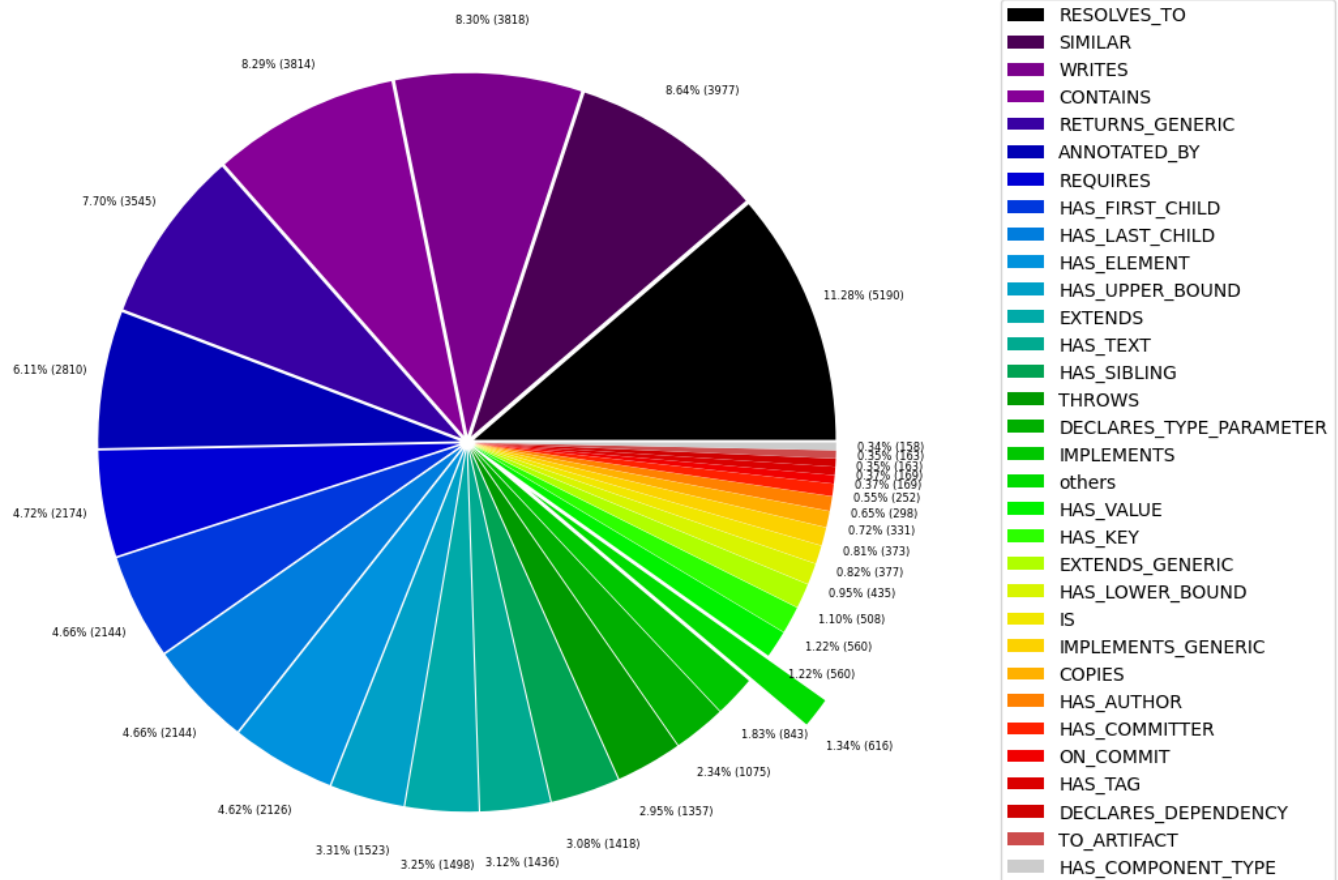
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000094
1	THROWS_GENERIC	5	0.000468
2	EXCLUDES	5	0.000468
3	DESCRIBES	9	0.000842
4	HAS_ROOT_ELEMENT	11	0.001029
5	HAS_GOAL	16	0.001496
6	HAS_EXECUTION	16	0.001496
7	OF_NAMESPACE	18	0.001683
8	HAS_ATTRIBUTE	18	0.001683
9	INCLUDES_CONCEPT	19	0.001777
10	USES_PLUGIN	21	0.001964
11	REQUIRES_TYPE_PARAMETER	21	0.001964
12	IS_ARTIFACT	21	0.001964
13	HAS_CONFIGURATION	21	0.001964
14	REQUIRES_CONCEPT	28	0.002618
15	HAS_BRANCH	33	0.003086
16	HAS_HEAD	34	0.003179
17	DECLARES_NAMESPACE	36	0.003366
18	HAS_DEFAULT	36	0.003366
19	CONTAINS_VALUE	121	0.011315
20	COPY_OF	126	0.011782
21	HAS_COMPONENT_TYPE	158	0.014775
22	TO_ARTIFACT	163	0.015242
23	DECLARES_DEPENDENCY	163	0.015242
24	ON_COMMIT	169	0.015803
25	HAS_TAG	169	0.015803
26	HAS_COMMITTER	252	0.023565
27	HAS_AUTHOR	298	0.027866
28	COPIES	331	0.030952
29	IMPLEMENTS_GENERIC	373	0.034879

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)

Legend:
- RESOLVES_TO
- SIMILAR
- WRITES
- CONTAINS
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_FIRST_CHILD
- HAS_LAST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- THROWS
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- others
- HAS_VALUE
- HAS_KEY
- EXTENDS_GENERIC
- 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, Commit]	CONTAINS_CHANGE	[Git, Change]	255467	12364	
1	[Git, Change]	MODIFIES	[File, Git]	255467	255467	
2	[Git, Change]	UPDATES	[File, Git]	167307	255467	
3	[Git, Change]	CREATES	[File, Git]	60811	255467	
4	[Git, Change]	DELETES	[File, Git]	37841	255467	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21799	13110	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15016	12364	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12364	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12364	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12364	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10767	1	
11	[Git, Change]	RENAMES	[File, Git]	10492	255467	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8362	13110	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8218	13110	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6371	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6219	10767	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5105	814	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4480	814	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	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	2050	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2203	814	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2114	2050	
28	[Type, File, Java, ByteCode, Class, GenericDec...]	DECLARES	[Java, ByteCode, Member, Method]	2112	231	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2091	13110	

Graph Density

```
total_number_of_nodes (vertices): 344590
total_number_of_relationships (edges): 1069399
-> total directed graph density: 9.006077789450154e-06
-> total directed graph density in percent: 0.0009006077789450154
```