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]	241390	73.194176
1	[Java, ByteCode, Method, Member]	13200	4.002499
2	[Java, ByteCode, Parameter]	13000	3.941855
3	[Git, Commit]	11882	3.602855
4	[File, Git]	10651	3.229592
5	[Java, ByteCode, Bound, ParameterizedType]	7166	2.172872
6	[Java, ByteCode, Bound]	7112	2.156498
7	[Java, ByteCode, Member, Field]	3495	1.059752
8	[Java, ByteCode, Bound, WildcardType]	2906	0.881156
9	[Java, Value, ByteCode, Annotation]	2821	0.855382
10	[Xml, Element]	2144	0.650103
11	[Java, ByteCode, Constructor, Method, Member]	2038	0.617962
12	[Xml, Text]	1436	0.435423
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.332329
14	[Java, ByteCode, Method, Member, Lambda]	936	0.283814
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.272898
16	[Type, File, Java, ByteCode, Class]	861	0.261072
17	[Type, File, Java, ByteCode, JavaType]	746	0.226202
18	[Java, Value, ByteCode, Primitive]	650	0.197093
19	[Java, ByteCode, Method, Member, GenericDeclar...	569	0.172532
20	[Json, Key]	555	0.168287
21	[Value, Json, Scalar]	539	0.163435
22	[Type, File, Java, ByteCode, ExternalType]	396	0.120075
23	[Author, Git, Person]	290	0.087934
24	[Value, Array]	267	0.080960
25	[Committer, Git, Person]	244	0.073986
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.071560
27	[Value, Property]	202	0.061250
28	[Type, File, Java, ByteCode, Interface]	187	0.056702
29	[Java, Value, ByteCode, Class]	172	0.052154

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', 'Method', 'Member']
- ['Java', 'ByteCode', 'Parameter']
- others
- ['Git', 'Commit']
- ['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', 'Constructor', 'Method', 'Member']

73.19% (241390)
0.62% (2038)
0.65% (2144)
0.86% (2821)
0.88% (2906)
1.06% (3495)
2.16% (7112)
2.17% (7166)
3.23% (10651)
3.60% (11882)
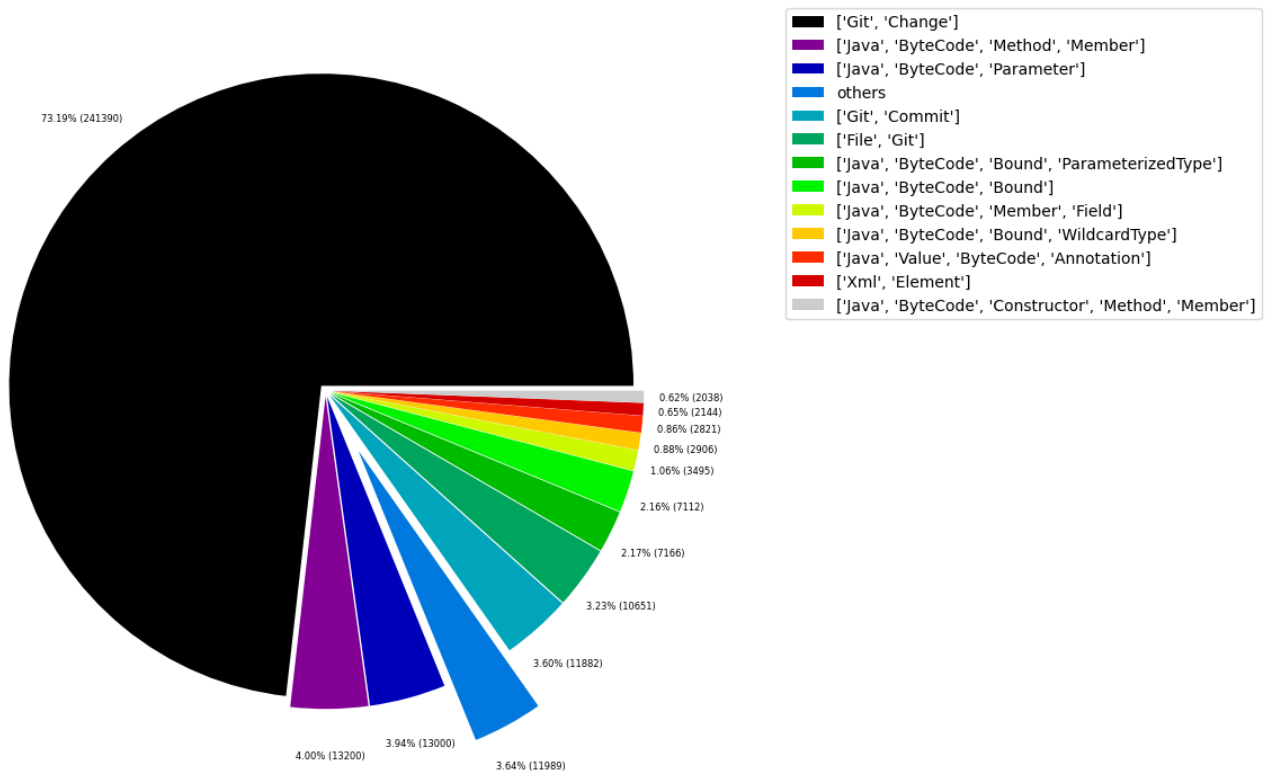3.64% (11989)
3.94% (13000)
4.00% (13200)

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	0.000303
1	[File, TS, Scan]	1	0.000303
2	[File, Json]	2	0.000606
3	[File]	3	0.000910
4	[Java, ByteCode, Constructor, Method, Member, ...	4	0.001213
5	[Maven, Exclusion]	5	0.001516
6	[Value, Array, Json]	6	0.001819
7	[File, Maven, Xml, Pom, Document]	9	0.002729
8	[Type, File, Java, ByteCode, Void]	9	0.002729
9	[Java, ManifestSection]	9	0.002729
10	[File, Java, Manifest]	9	0.002729
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002729
12	[File, Java, ServiceLoader]	10	0.003032
13	[File, Java, Properties]	12	0.003639
14	[Maven, ExecutionGoal]	16	0.004852
15	[Maven, PluginExecution]	16	0.004852
16	[Xml, Attribute]	18	0.005458
17	[jQAssistant, Rule, Concept]	19	0.005761
18	[Maven, Plugin]	21	0.006368
19	[Maven, Configuration]	21	0.006368
20	[Git, Branch]	28	0.008490
21	[Type, File, Java, ByteCode, Enum]	28	0.008490
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008793
23	[Xml, Namespace]	36	0.010916
24	[Type, File, Java, ByteCode, Annotation]	43	0.013038
25	[File, Directory]	44	0.013342
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025774
27	[Artifact, Maven]	92	0.027896
28	[Type, File, Java, ByteCode, ExternalType, Ext...	93	0.028199
29	[Value, ManifestEntry]	120	0.036386

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)

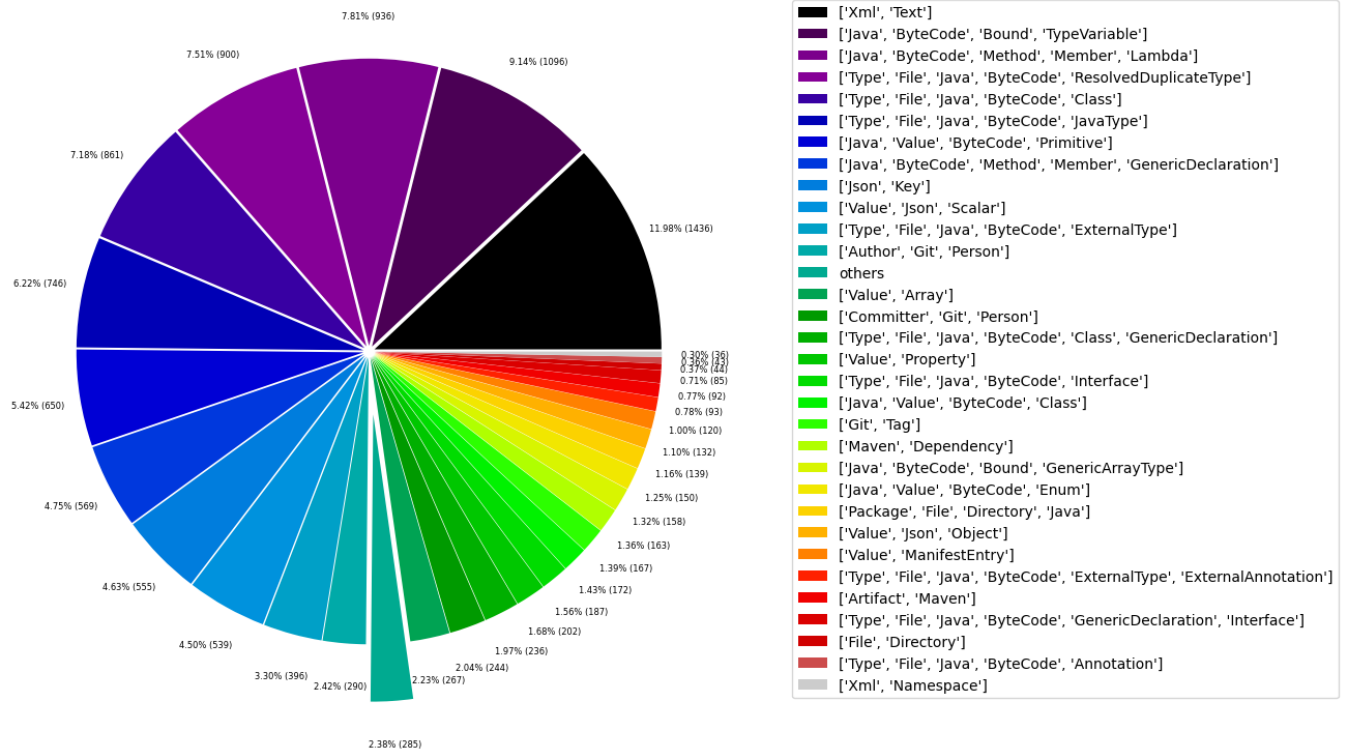| Legend	
■	['Xml', 'Text']
■	['Java', 'ByteCode', 'Bound', 'TypeVariable']
■	['Java', 'ByteCode', 'Method', 'Member', 'Lambda']
■	['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
■	['Type', 'File', 'Java', 'ByteCode', 'Class']
■	['Type', 'File', 'Java', 'ByteCode', 'JavaType']
■	['Java', 'Value', 'ByteCode', 'Primitive']
■	['Java', 'ByteCode', 'Method', 'Member', 'GenericDeclaration']
■	['Json', 'Key']
■	['Value', 'Json', 'Scalar']
■	['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
■	['Author', 'Git', 'Person']
■	others
■	['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', 'ExternalType', 'ExternalAnnotation']
■	['Artifact', 'Maven']
■	['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
■	['File', 'Directory']
■	['Type', 'File', 'Java', 'ByteCode', 'Annotation']
■	['Xml', 'Namespace']

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.
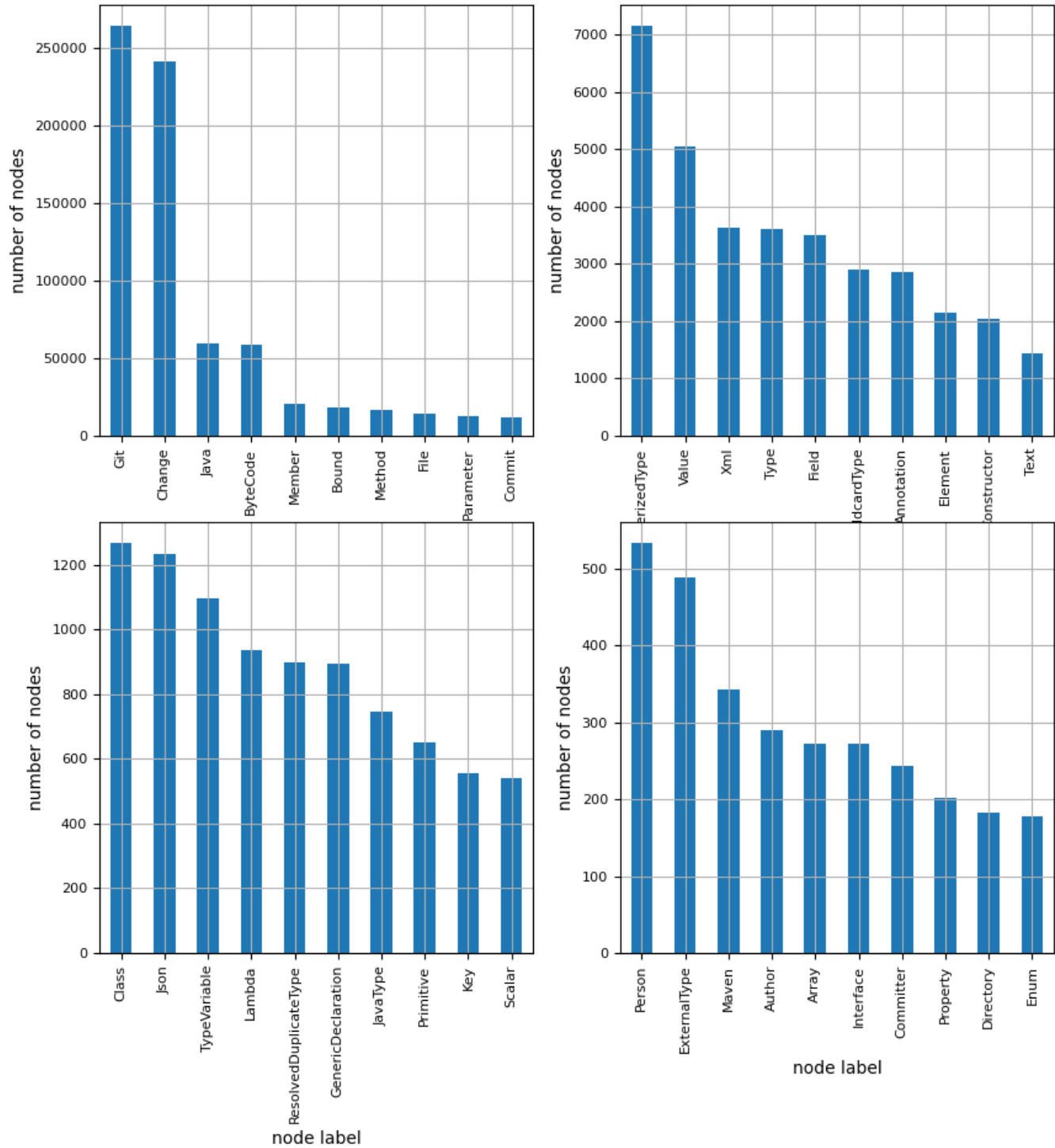
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	264652	80.247670
1	Change	241390	73.194176
2	Java	59274	17.973038
3	ByteCode	59086	17.916032
4	Member	20242	6.137771
5	Bound	18438	5.590763
6	Method	16747	5.078018
7	File	14502	4.397290
8	Parameter	13000	3.941855
9	Commit	11882	3.602855
10	ParameterizedType	7166	2.172872
11	Value	5059	1.533988
12	Xml	3643	1.104629
13	Type	3613	1.095532
14	Field	3495	1.059752
15	WildcardType	2906	0.881156
16	Annotation	2864	0.868421
17	Element	2144	0.650103
18	Constructor	2042	0.619174
19	Text	1436	0.435423
20	Class	1269	0.384786
21	Json	1234	0.374173
22	TypeVariable	1096	0.332329
23	Lambda	936	0.283814
24	ResolvedDuplicateType	900	0.272898
25	GenericDeclaration	894	0.271078
26	JavaType	746	0.226202
27	Primitive	650	0.197093
28	Key	555	0.168287
29	Scalar	539	0.163435
30	Person	534	0.161919
31	ExternalType	489	0.148274
32	Maven	343	0.104004
33	Author	290	0.087934
34	Array	273	0.082779
35	Interface	272	0.082476
36	Committer	244	0.073986
37	Property	202	0.061250
38	Directory	183	0.055489
39	Enum	178	0.053973

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: 999894
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	241390	24.141559
1	MODIFIES	241390	24.141559
2	UPDATES	155832	15.584852
3	CREATES	58984	5.899025
4	DELETES	36739	3.674289
5	INVOKES	35573	3.557677
6	COMMITTED	23764	2.376652
7	DEPENDS_ON	21873	2.187532
8	OF_TYPE	21347	2.134926
9	DECLARES	20683	2.068519
10	OF_RAW_TYPE	17077	1.707881
11	HAS_PARENT	14448	1.444953
12	HAS	14094	1.409549
13	RETURNS	12547	1.254833
14	RENAMES	10165	1.016608
15	READS	9078	0.907896
16	HAS_ACTUAL_TYPE_ARGUMENT	8276	0.827688
17	HAS_NEW_NAME	6190	0.619066
18	OF_GENERIC_TYPE	5902	0.590263
19	RESOLVES_TO	5225	0.522555
20	SIMILAR	3968	0.396842
21	CONTAINS	3805	0.380540
22	WRITES	3803	0.380340
23	RETURNS_GENERIC	3543	0.354338
24	ANNOTATED_BY	2809	0.280930
25	REQUIRES	2173	0.217323
26	HAS_FIRST_CHILD	2144	0.214423
27	HAS_LAST_CHILD	2144	0.214423
28	HAS_ELEMENT	2126	0.212623
29	HAS_UPPER_BOUND	1521	0.152116

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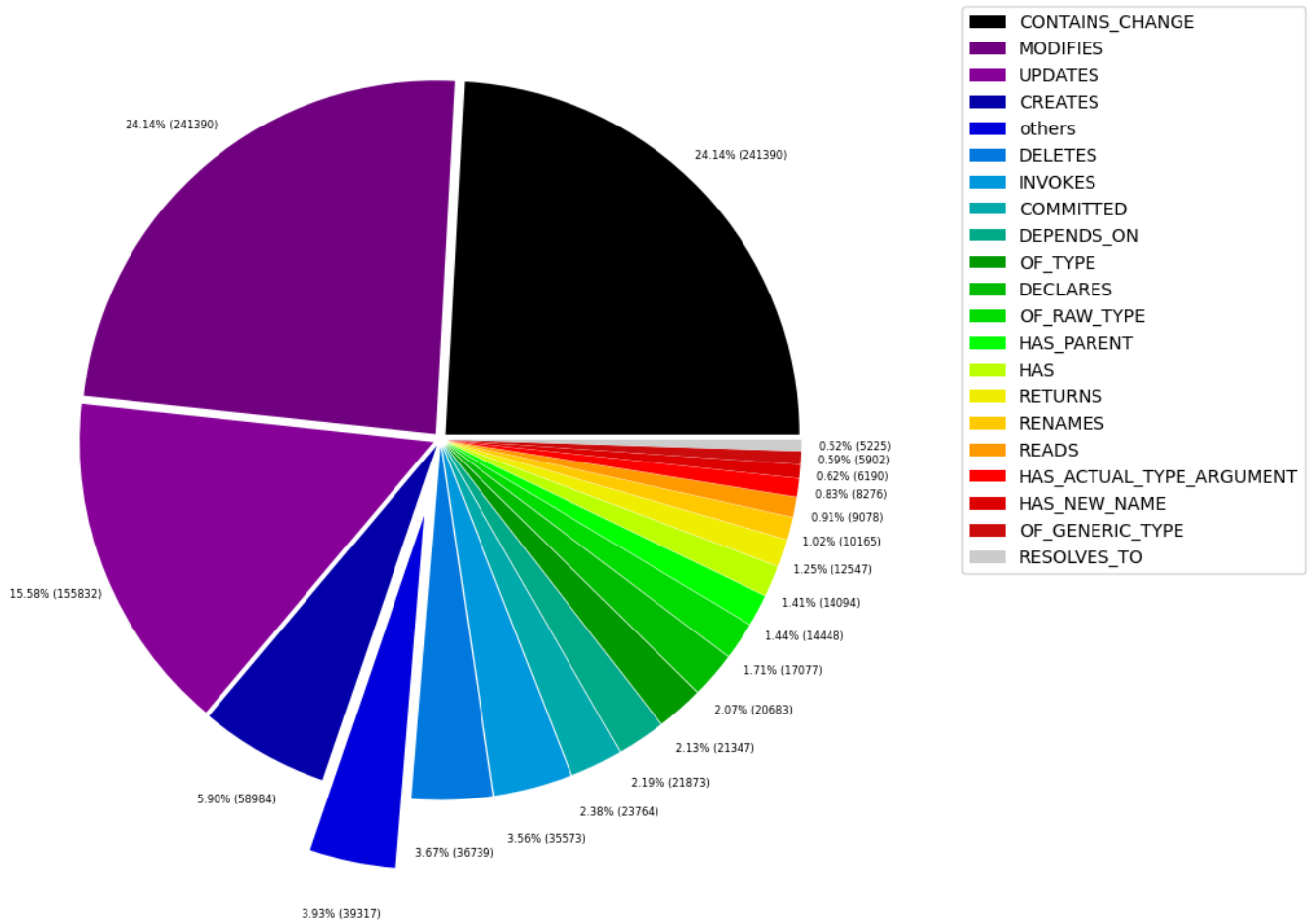
Relationship types (more than 0.5% overall)

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

Pie chart values:
- 24.14% (241390)
- 24.14% (241390)
- 15.58% (155832)
- 5.90% (58984)
- 3.93% (39317)
- 3.67% (36739)
- 3.56% (35573)
- 2.38% (23764)
- 2.19% (21873)
- 2.13% (21347)
- 2.07% (20683)
- 1.71% (17077)
- 1.44% (14448)
- 1.41% (14094)
- 1.25% (12547)
- 1.02% (10165)
- 0.91% (9078)
- 0.83% (8276)
- 0.62% (6190)
- 0.59% (5902)
- 0.52% (5225)

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.
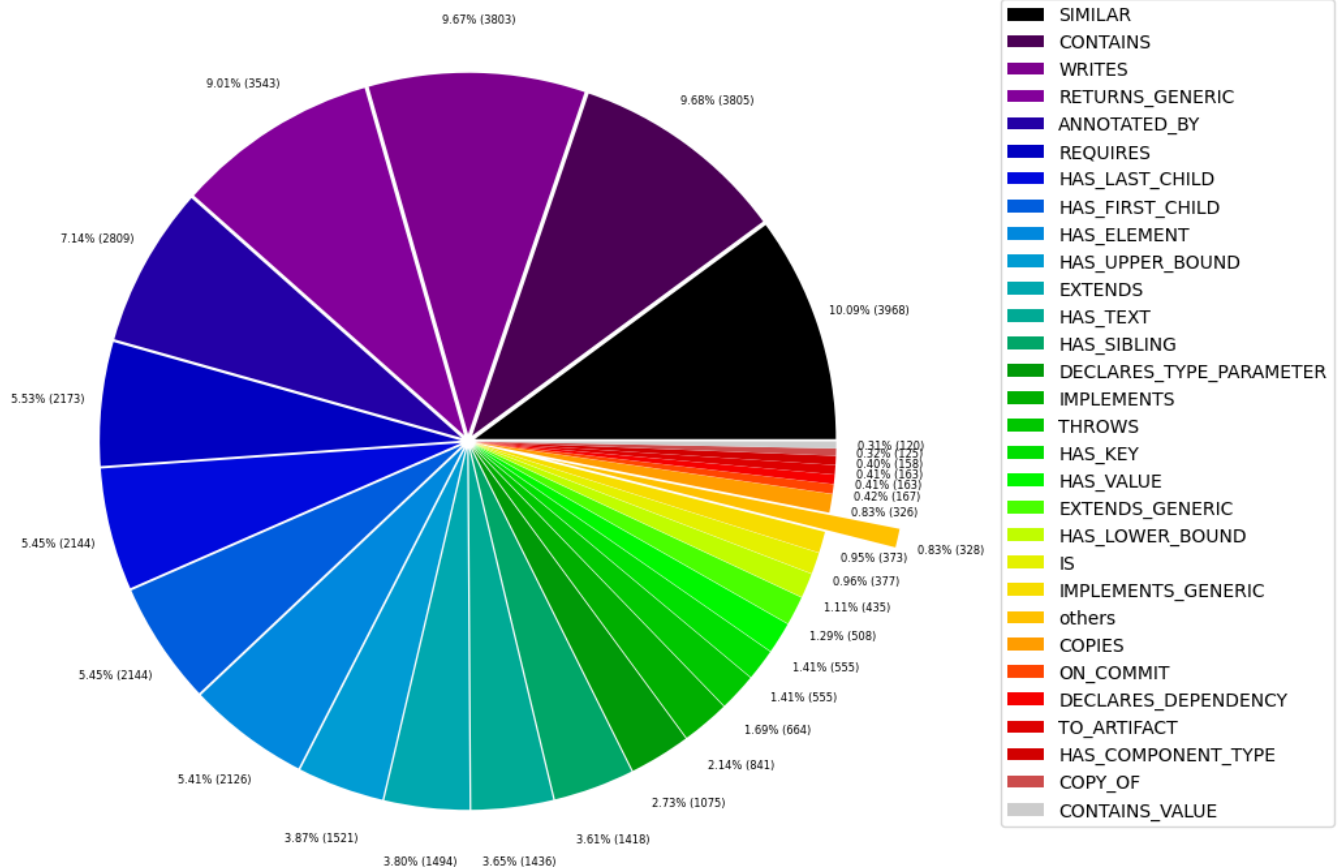
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000100
1	THROWS_GENERIC	5	0.000500
2	EXCLUDES	5	0.000500
3	HAS_ROOT_ELEMENT	9	0.000900
4	DESCRIBES	9	0.000900
5	HAS_GOAL	16	0.001600
6	HAS_EXECUTION	16	0.001600
7	OF_NAMESPACE	18	0.001800
8	HAS_ATTRIBUTE	18	0.001800
9	INCLUDES_CONCEPT	19	0.001900
10	USES_PLUGIN	21	0.002100
11	REQUIRES_TYPE_PARAMETER	21	0.002100
12	IS_ARTIFACT	21	0.002100
13	HAS_CONFIGURATION	21	0.002100
14	REQUIRES_CONCEPT	28	0.002800
15	HAS_HEAD	28	0.002800
16	DECLARES_NAMESPACE	36	0.003600
17	HAS_DEFAULT	36	0.003600
18	CONTAINS_VALUE	120	0.012001
19	COPY_OF	125	0.012501
20	HAS_COMPONENT_TYPE	158	0.015802
21	TO_ARTIFACT	163	0.016302
22	DECLARES_DEPENDENCY	163	0.016302
23	ON_COMMIT	167	0.016702
24	COPIES	326	0.032603
25	IMPLEMENTS_GENERIC	373	0.037304
26	IS	377	0.037704
27	HAS_LOWER_BOUND	435	0.043505
28	EXTENDS_GENERIC	508	0.050805
29	HAS_VALUE	555	0.055506

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:
- SIMILAR
- CONTAINS
- WRITES
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_LAST_CHILD
- HAS_FIRST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- THROWS
- HAS_KEY
- HAS_VALUE
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- others
- COPIES
- ON_COMMIT
- DECLARES_DEPENDENCY
- TO_ARTIFACT
- HAS_COMPONENT_TYPE
- COPY_OF
- CONTAINS_VALUE

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]	241390	11882	
1	[Git, Change]	MODIFIES	[File, Git]	241390	241390	
2	[Git, Change]	UPDATES	[File, Git]	155832	241390	
3	[Git, Change]	CREATES	[File, Git]	58984	241390	
4	[Git, Change]	DELETES	[File, Git]	36739	241390	
5	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	21529	13077	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14439	11882	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11882	290	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11882	244	
9	[Git, Change]	RENAMES	[File, Git]	10165	241390	
10	[Java, ByteCode, Method, Member]	HAS	[Java, ByteCode, Parameter]	8300	13077	
11	[Java, ByteCode, Method, Member]	READS	[Java, ByteCode, Member, Field]	8101	13077	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6360	13000	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6190	10651	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3569	7112	
15	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3043	7166	
16	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2906	7166	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2655	13000	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7166	
19	[Java, ByteCode, Constructor, Method, Member]	WRITES	[Java, ByteCode, Member, Field]	2426	2038	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...	2340	2821	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2111	2038	
23	[Java, ByteCode, Method, Member]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2088	13077	
24	[Java, ByteCode, Constructor, Method, Member]	HAS	[Java, ByteCode, Parameter]	2023	2038	
25	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13000	
26	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2008	13077	
27	[Java, ByteCode, Method, Member, Lambda]	INVOKES	[Java, ByteCode, Method, Member]	1882	936	
28	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1802	7166	
29	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	1796	2038	

Graph Density

```
total_number_of_nodes (vertices): 329794
total_number_of_relationships (edges): 999894
-> total directed graph density: 9.193264996024815e-06
-> total directed graph density in percent: 0.00091932649960248315
```