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]	245645	73.479308
1	[Java, ByteCode, Member, Method]	13197	3.947593
2	[Java, ByteCode, Parameter]	13016	3.893451
3	[Git, Commit]	12051	3.604792
4	[File, Git]	10677	3.193790
5	[Java, ByteCode, ParameterizedType, Bound]	7172	2.145346
6	[Java, ByteCode, Bound]	7113	2.127698
7	[Java, ByteCode, Member, Field]	3503	1.047846
8	[Java, ByteCode, Bound, WildcardType]	2908	0.869864
9	[Java, Value, ByteCode, Annotation]	2822	0.844139
10	[Xml, Element]	2144	0.641331
11	[Java, ByteCode, Member, Constructor, Method]	2044	0.611418
12	[Xml, Text]	1436	0.429548
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.327844
14	[Java, ByteCode, Member, Method, Lambda]	938	0.280582
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.269215
16	[Type, File, Java, ByteCode, Class]	861	0.257549
17	[Type, File, Java, ByteCode, JavaType]	748	0.223748
18	[Java, Value, ByteCode, Primitive]	650	0.194433
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.170204
20	[Json, Key]	560	0.167512
21	[Value, Json, Scalar]	544	0.162726
22	[Type, File, Java, ByteCode, ExternalType]	396	0.118455
23	[Author, Git, Person]	293	0.087645
24	[Value, Array]	267	0.079867
25	[Committer, Git, Person]	248	0.074184
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.070594
27	[Value, Property]	202	0.060424
28	[Type, File, Java, ByteCode, Interface]	187	0.055937
29	[Java, Value, ByteCode, Class]	172	0.051450

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', 'ParameterizedType', 'Bound']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
- ['Xml', 'Element']
- ['Java', 'ByteCode', 'Member', 'Constructor', 'Method']

73.48% (245645)
0.61% (2044)
0.64% (2144)
0.84% (2822)
0.87% (2908)
1.05% (3503)
2.13% (7113)
2.15% (7172)
3.19% (10677)
3.59% (12013)
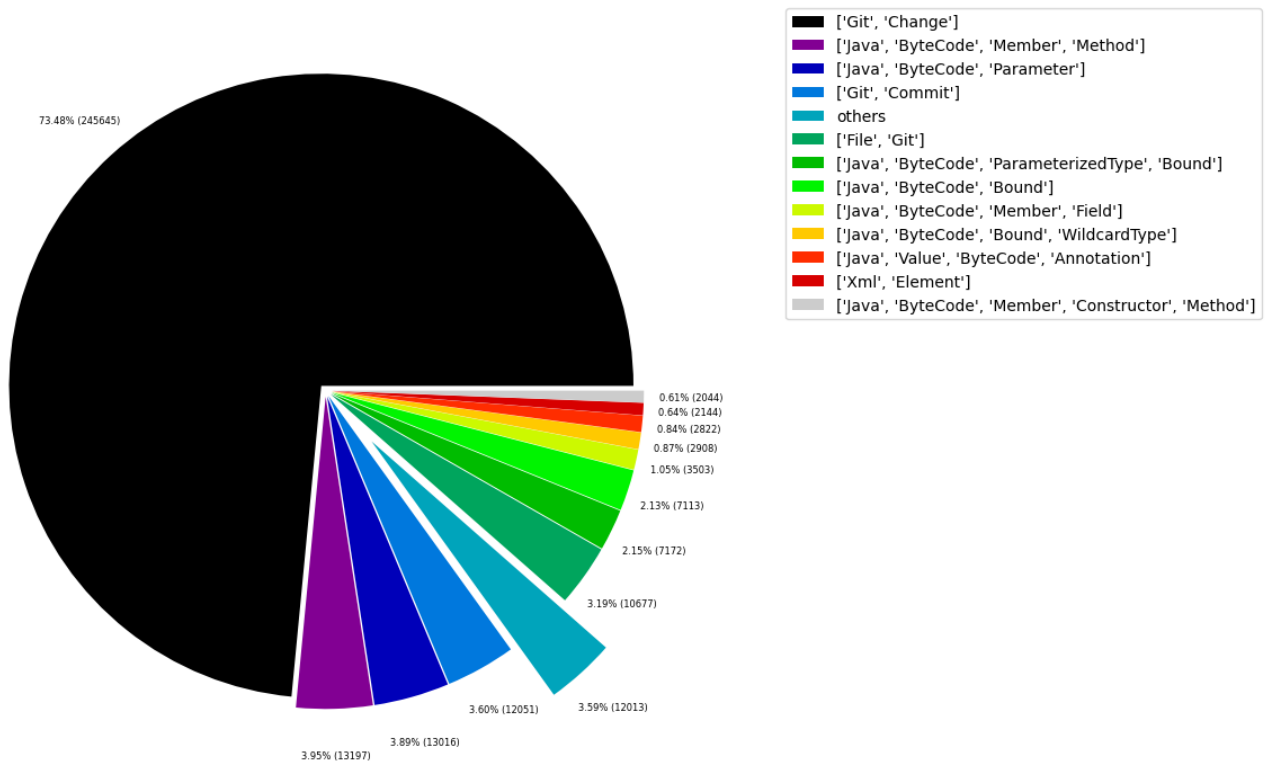3.60% (12051)
3.89% (13016)
3.95% (13197)

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	0.000299
1	[Repository, File, Git]	1	0.000299
2	[File, Json]	2	0.000598
3	[File]	3	0.000897
4	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001197
5	[Maven, Exclusion]	5	0.001496
6	[Value, Json, Array]	6	0.001795
7	[File, Maven, Xml, Pom, Document]	9	0.002692
8	[Type, File, Java, ByteCode, Void]	9	0.002692
9	[Java, ManifestSection]	9	0.002692
10	[File, Java, Manifest]	9	0.002692
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002692
12	[File, Java, ServiceLoader]	10	0.002991
13	[File, Java, Properties]	12	0.003590
14	[Maven, ExecutionGoal]	16	0.004786
15	[Maven, PluginExecution]	16	0.004786
16	[Xml, Attribute]	18	0.005384
17	[jQAssistant, Rule, Concept]	19	0.005683
18	[Maven, Plugin]	21	0.006282
19	[Maven, Configuration]	21	0.006282
20	[Type, File, Java, ByteCode, Enum]	28	0.008376
21	[Git, Branch]	29	0.008675
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008675
23	[Xml, Namespace]	36	0.010769
24	[Type, File, Java, ByteCode, Annotation]	43	0.012863
25	[File, Directory]	44	0.013162
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025426
27	[Artifact, Maven]	92	0.027520
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.027819
29	[Value, ManifestEntry]	120	0.035895

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', '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', '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']
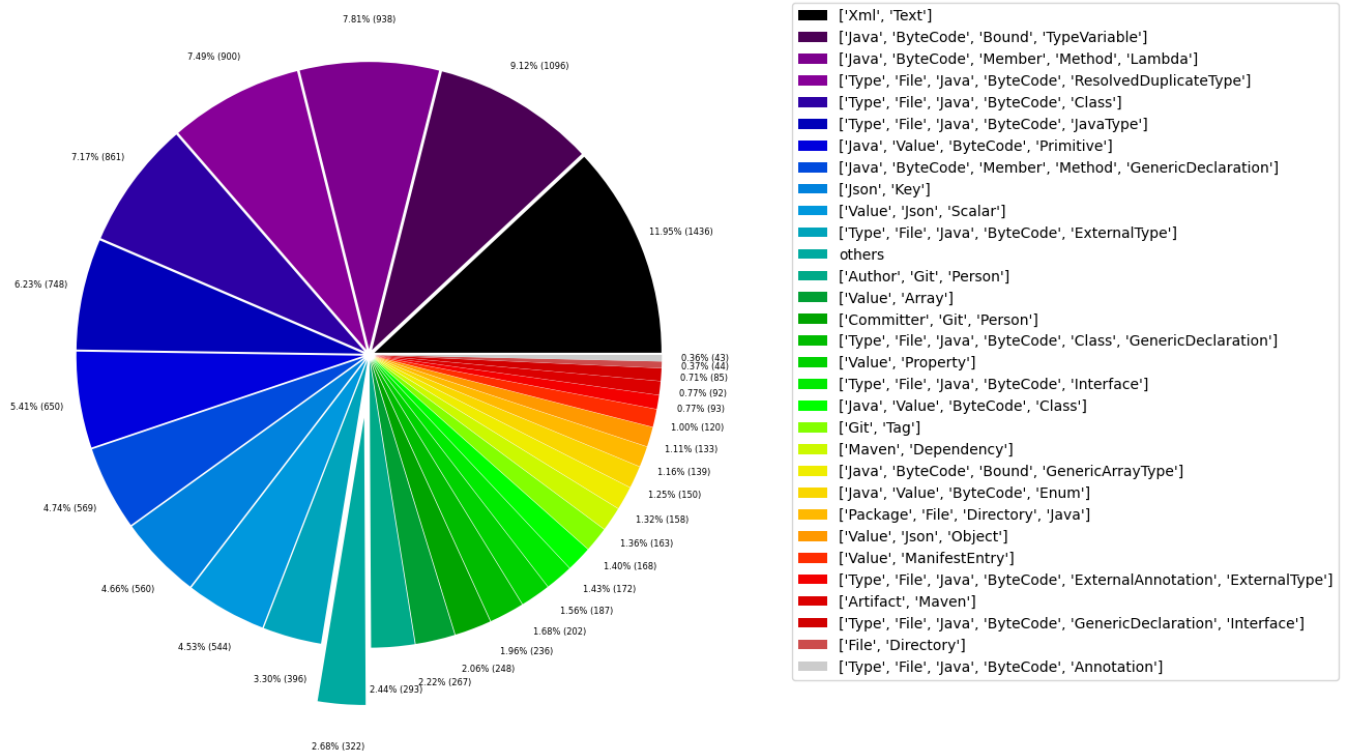- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

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.
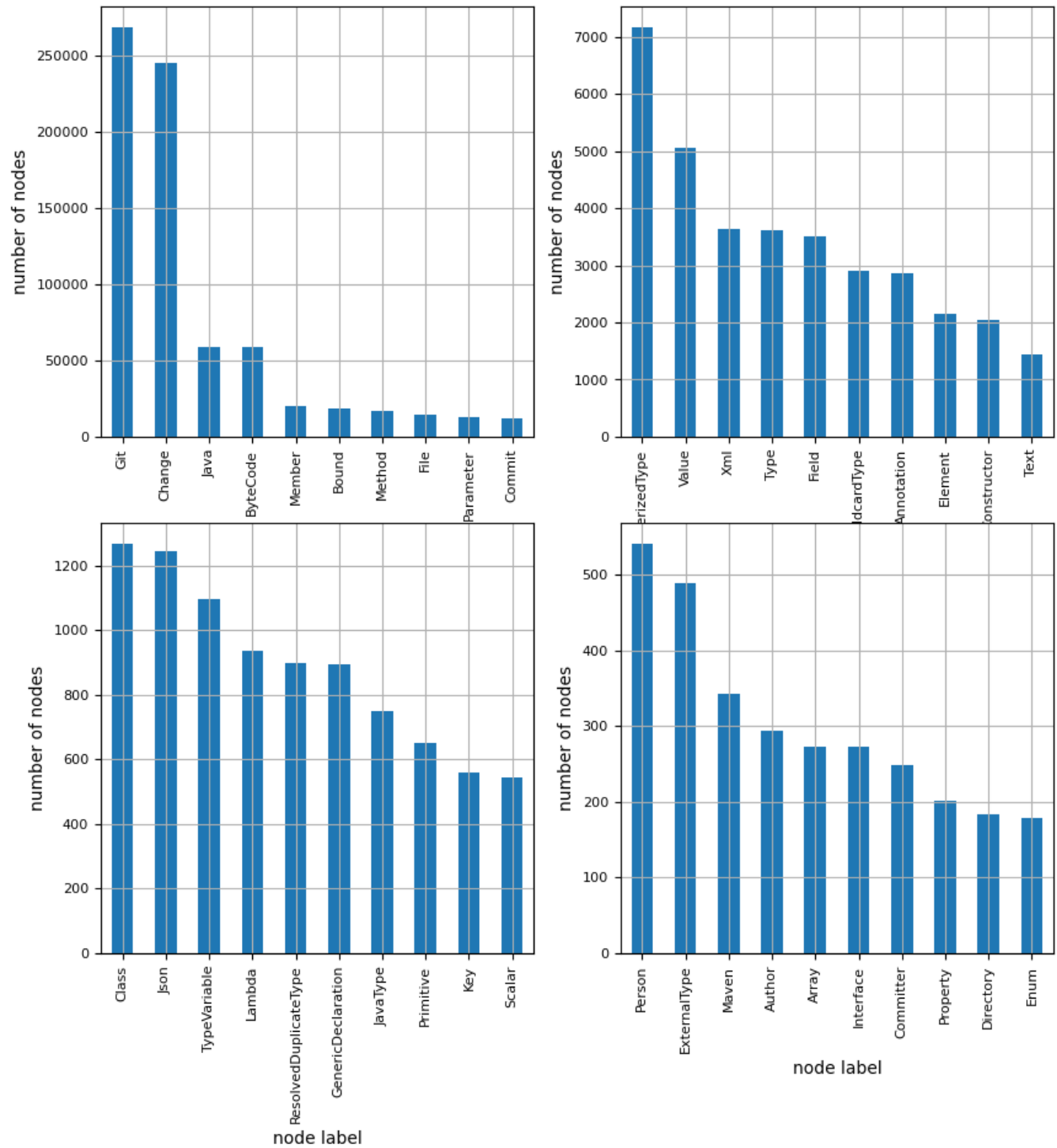
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	269112	80.498946
1	Change	245645	73.479308
2	Java	59315	17.742780
3	ByteCode	59127	17.686544
4	Member	20255	6.058838
5	Bound	18447	5.518015
6	Method	16752	5.010993
7	File	14530	4.346330
8	Parameter	13016	3.893451
9	Commit	12051	3.604792
10	ParameterizedType	7172	2.145346
11	Value	5066	1.515383
12	Xml	3643	1.089723
13	Type	3615	1.081348
14	Field	3503	1.047846
15	WildcardType	2908	0.869864
16	Annotation	2865	0.857002
17	Element	2144	0.641331
18	Constructor	2048	0.612614
19	Text	1436	0.429548
20	Class	1269	0.379593
21	Json	1245	0.372414
22	TypeVariable	1096	0.327844
23	Lambda	938	0.280582
24	ResolvedDuplicateType	900	0.269215
25	GenericDeclaration	894	0.267420
26	JavaType	748	0.223748
27	Primitive	650	0.194433
28	Key	560	0.167512
29	Scalar	544	0.162726
30	Person	541	0.161828
31	ExternalType	489	0.146274
32	Maven	343	0.102601
33	Author	293	0.087645
34	Array	273	0.081662
35	Interface	272	0.081363
36	Committer	248	0.074184
37	Property	202	0.060424
38	Directory	183	0.054740
39	Enum	178	0.053245

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: 1037158
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	245645	23.684434
1	MODIFIES	245645	23.684434
2	UPDATES	159018	15.332090
3	CREATES	59721	5.758139
4	DELETES	37196	3.586339
5	INVOKES	35676	3.439784
6	COMMITTED	24102	2.323850
7	DEPENDS_ON	21887	2.110286
8	OF_TYPE	21373	2.060727
9	DECLARES	20696	1.995453
10	OF_RAW_TYPE	17084	1.647194
11	HAS_PARENT	14652	1.412707
12	HAS	14110	1.360448
13	RETURNS	12562	1.211194
14	HAS_COMMIT	12051	1.161925
15	HAS_FILE	10677	1.029448
16	RENAMES	10290	0.992134
17	READS	9097	0.877108
18	HAS_ACTUAL_TYPE_ARGUMENT	8284	0.798721
19	HAS_NEW_NAME	6199	0.597691
20	OF_GENERIC_TYPE	5903	0.569151
21	RESOLVES_TO	5203	0.501659
22	SIMILAR	3968	0.382584
23	WRITES	3812	0.367543
24	CONTAINS	3806	0.366964
25	RETURNS_GENERIC	3543	0.341607
26	ANNOTATED_BY	2810	0.270933
27	REQUIRES	2175	0.209708
28	HAS_FIRST_CHILD	2144	0.206719
29	HAS_LAST_CHILD	2144	0.206719

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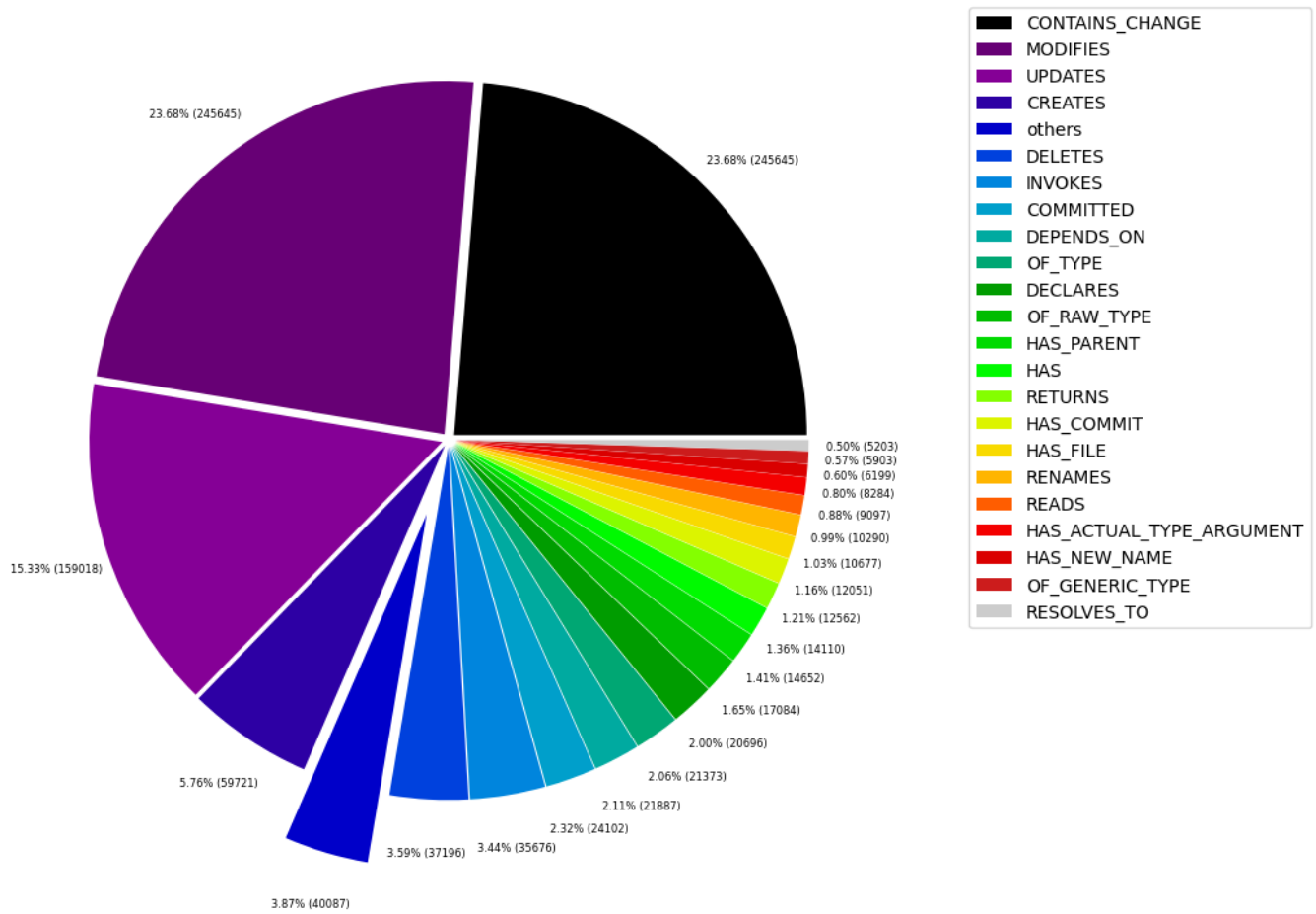
Relationship types (more than 0.5% overall)

■	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
■	RESOLVES_TO

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.
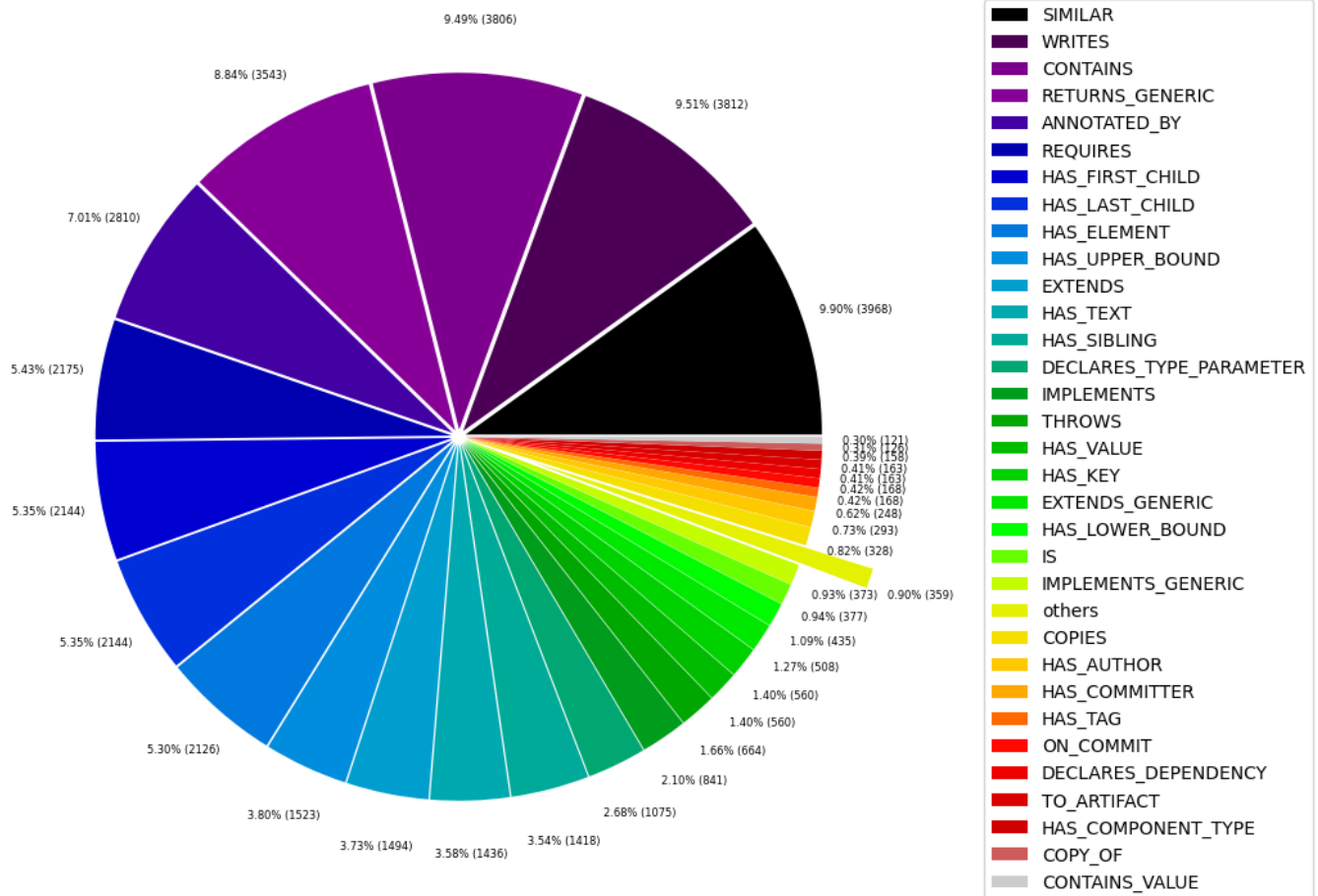
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000096
1	THROWS_GENERIC	5	0.000482
2	EXCLUDES	5	0.000482
3	HAS_ROOT_ELEMENT	9	0.000868
4	DESCRIBES	9	0.000868
5	HAS_GOAL	16	0.001543
6	HAS_EXECUTION	16	0.001543
7	OF_NAMESPACE	18	0.001736
8	HAS_ATTRIBUTE	18	0.001736
9	INCLUDES_CONCEPT	19	0.001832
10	USES_PLUGIN	21	0.002025
11	REQUIRES_TYPE_PARAMETER	21	0.002025
12	IS_ARTIFACT	21	0.002025
13	HAS_CONFIGURATION	21	0.002025
14	REQUIRES_CONCEPT	28	0.002700
15	HAS_BRANCH	29	0.002796
16	HAS_HEAD	30	0.002893
17	DECLARES_NAMESPACE	36	0.003471
18	HAS_DEFAULT	36	0.003471
19	CONTAINS_VALUE	121	0.011666
20	COPY_OF	126	0.012149
21	HAS_COMPONENT_TYPE	158	0.015234
22	TO_ARTIFACT	163	0.015716
23	DECLARES_DEPENDENCY	163	0.015716
24	ON_COMMIT	168	0.016198
25	HAS_TAG	168	0.016198
26	HAS_COMMITTER	248	0.023911
27	HAS_AUTHOR	293	0.028250
28	COPIES	328	0.031625
29	IMPLEMENTS_GENERIC	373	0.035964

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
- WRITES
- CONTAINS
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_FIRST_CHILD
- HAS_LAST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- THROWS
- HAS_VALUE
- HAS_KEY
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- others
- COPIES
- HAS_AUTHOR
- HAS_COMMITTER
- HAS_TAG
- 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]	245645	12051	
1	[Git, Change]	MODIFIES	[File, Git]	245645	245645	
2	[Git, Change]	UPDATES	[File, Git]	159018	245645	
3	[Git, Change]	CREATES	[File, Git]	59721	245645	
4	[Git, Change]	DELETES	[File, Git]	37196	245645	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21552	13071	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14643	12051	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12051	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12051	293	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12051	248	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10677	1	
11	[Git, Change]	RENAMES	[File, Git]	10290	245645	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8320	13071	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8121	13071	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6368	13016	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6199	10677	
16	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7113	
17	[Java, ByteCode, ParameterizedType, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3049	7172	
18	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7172	
19	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, ParameterizedType, Bound]	2655	13016	
20	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7172	
21	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2432	2044	
22	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2341	2822	
23	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
24	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2113	2044	
25	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2088	13071	
26	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2029	13071	
27	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2023	2044	
28	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13016	
29	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1879	938	

Graph Density

```
total_number_of_nodes (vertices): 334305
total_number_of_relationships (edges): 1037158
-> total directed graph density: 9.280267108198784e-06
-> total directed graph density in percent: 0.0009280267108198784
```