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]	258356	74.269468
1	[Java, ByteCode, Member, Method]	13219	3.800059
2	[Java, ByteCode, Parameter]	13029	3.745440
3	[Git, Commit]	12698	3.650288
4	[File, Git]	10809	3.107258
5	[Java, ByteCode, ParameterizedType, Bound]	7176	2.062881
6	[Java, ByteCode, Bound]	7118	2.046208
7	[Java, ByteCode, Member, Field]	3504	1.007293
8	[Java, ByteCode, Bound, WildcardType]	2908	0.835961
9	[Java, Value, ByteCode, Annotation]	2822	0.811239
10	[Xml, Element]	2144	0.616335
11	[Java, ByteCode, Member, Constructor, Method]	2040	0.586438
12	[Xml, Text]	1436	0.412806
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.315067
14	[Java, ByteCode, Member, Method, Lambda]	942	0.270796
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.258435
16	[Type, File, Java, ByteCode, Class]	864	0.248374
17	[Type, File, Java, ByteCode, JavaType]	748	0.215027
18	[Java, Value, ByteCode, Primitive]	652	0.187430
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.163570
20	[Json, Key]	560	0.160983
21	[Value, Json, Scalar]	544	0.156383
22	[Type, File, Java, ByteCode, ExternalType]	396	0.113838
23	[Author, Git, Person]	298	0.085666
24	[Value, Array]	267	0.076754
25	[Committer, Git, Person]	252	0.072442
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.067843
27	[Value, Property]	202	0.058069
28	[Type, File, Java, ByteCode, Interface]	188	0.054044
29	[Java, Value, ByteCode, Class]	172	0.049445

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']

74.27% (258356)
0.59% (2040)
0.62% (2144)
0.81% (2822)
0.84% (2908)
1.01% (3504)
2.05% (7118)
2.06% (7176)
3.11% (10809)
3.46% (12040)
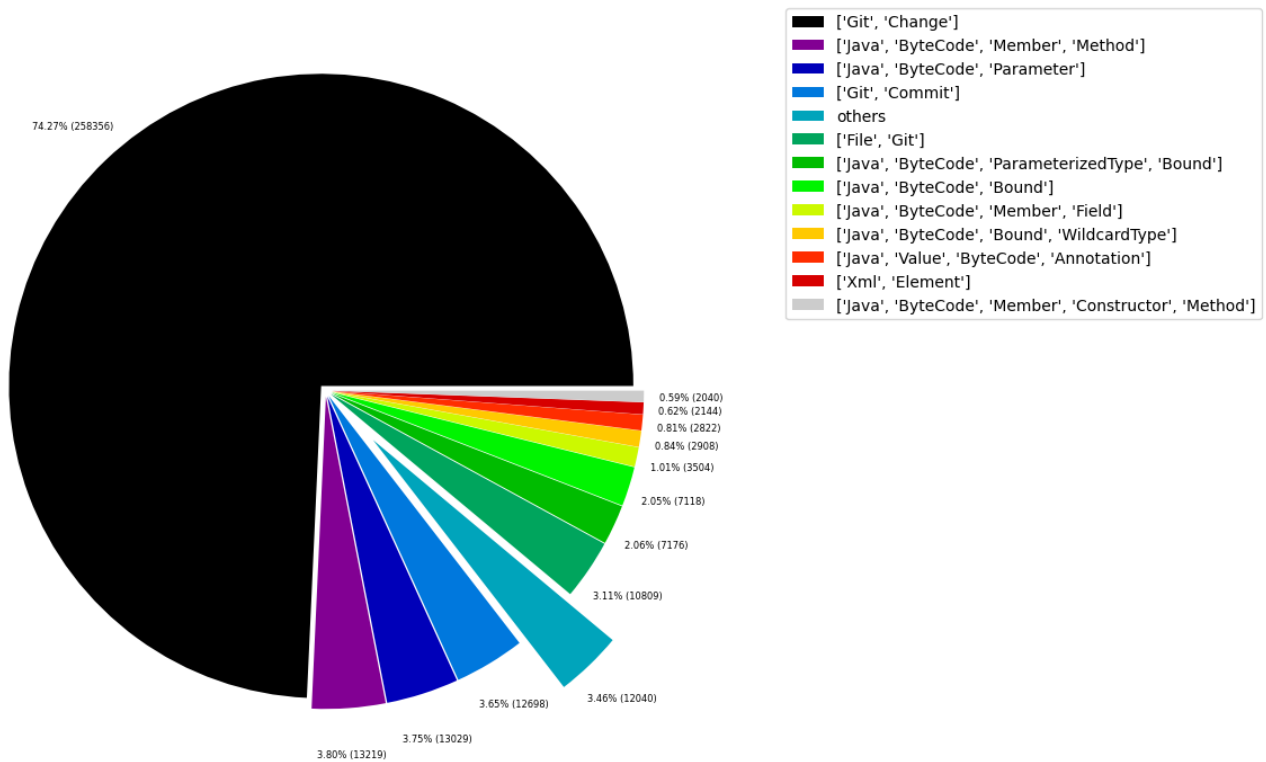3.65% (12698)
3.75% (13029)
3.80% (13219)

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	0.000287
1	[Repository, File, Git]	1	0.000287
2	[File, Json]	2	0.000575
3	[File]	3	0.000862
4	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001150
5	[Maven, Exclusion]	5	0.001437
6	[Value, Json, Array]	6	0.001725
7	[File, Maven, Xml, Pom, Document]	9	0.002587
8	[Type, File, Java, ByteCode, Void]	9	0.002587
9	[Java, ManifestSection]	9	0.002587
10	[File, Java, Manifest]	9	0.002587
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002587
12	[File, Java, ServiceLoader]	10	0.002875
13	[File, Java, Properties]	12	0.003450
14	[Maven, ExecutionGoal]	16	0.004600
15	[Maven, PluginExecution]	16	0.004600
16	[Xml, Attribute]	18	0.005174
17	[jQAssistant, Rule, Concept]	19	0.005462
18	[Maven, Plugin]	21	0.006037
19	[Maven, Configuration]	21	0.006037
20	[Type, File, Java, ByteCode, Enum]	28	0.008049
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008337
22	[Xml, Namespace]	36	0.010349
23	[Git, Branch]	37	0.010636
24	[Type, File, Java, ByteCode, Annotation]	43	0.012361
25	[File, Directory]	44	0.012649
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024435
27	[Artifact, Maven]	92	0.026447
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.026735
29	[Value, ManifestEntry]	120	0.034496

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']
- ['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', 'ExternalAnnotation', 'ExternalType']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
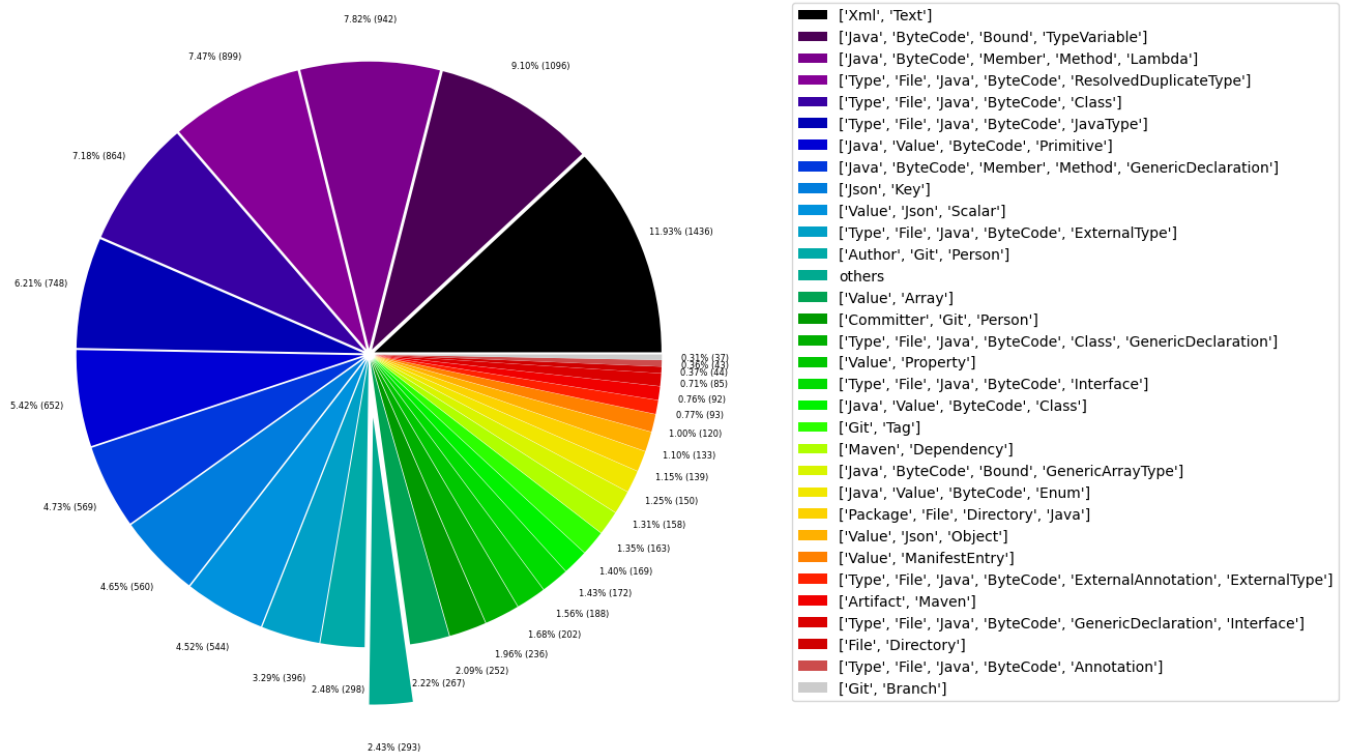- ['Git', 'Branch']

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.
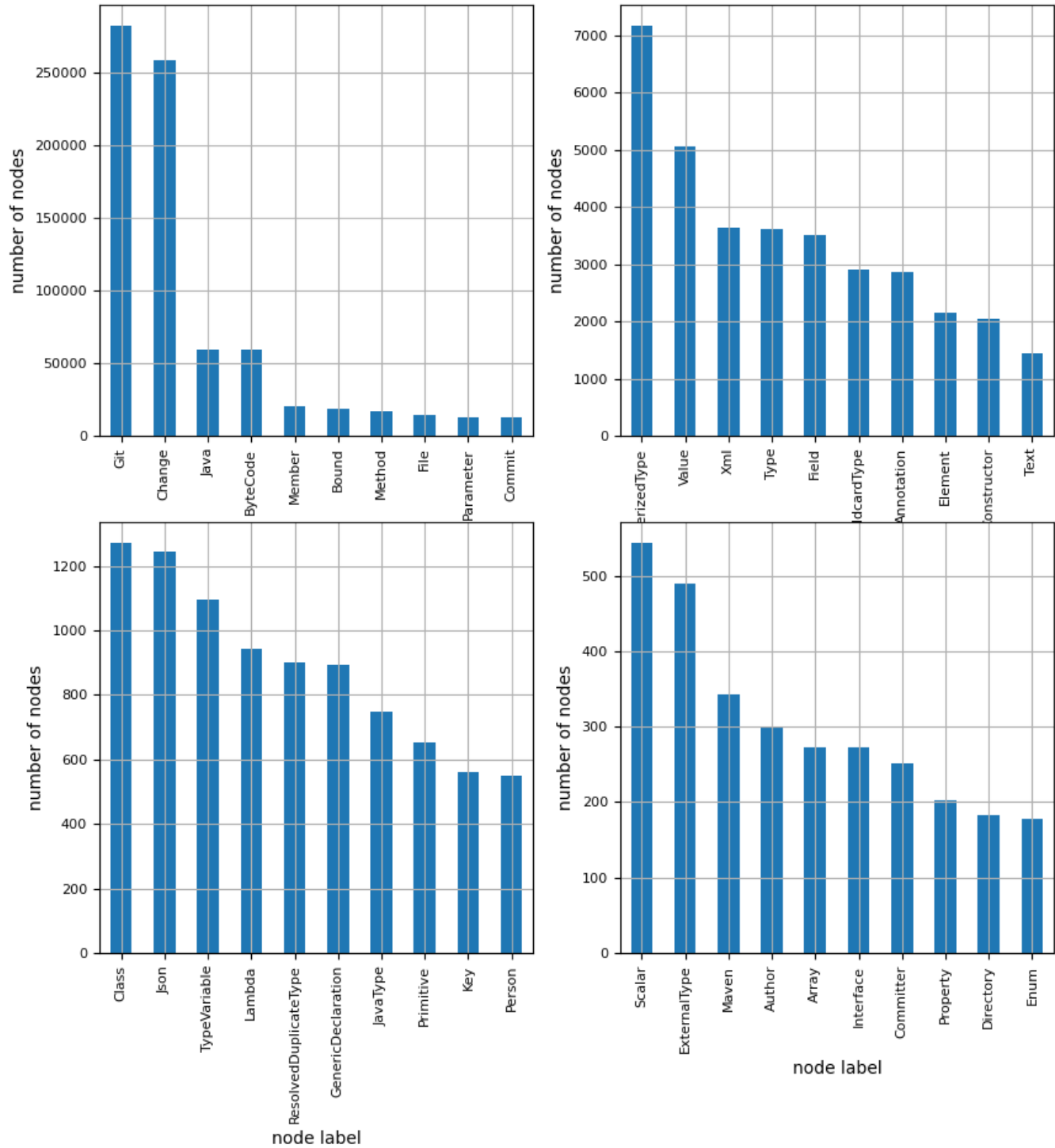
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	282620	81.244628
1	Change	258356	74.269468
2	Java	59365	17.065626
3	ByteCode	59177	17.011582
4	Member	20278	5.829306
5	Bound	18456	5.305537
6	Method	16774	4.822013
7	File	14665	4.215740
8	Parameter	13029	3.745440
9	Commit	12698	3.650288
10	ParameterizedType	7176	2.062881
11	Value	5068	1.456895
12	Xml	3643	1.047251
13	Type	3618	1.040065
14	Field	3504	1.007293
15	WildcardType	2908	0.835961
16	Annotation	2865	0.823600
17	Element	2144	0.616335
18	Constructor	2044	0.587588
19	Text	1436	0.412806
20	Class	1272	0.365661
21	Json	1245	0.357900
22	TypeVariable	1096	0.315067
23	Lambda	942	0.270796
24	ResolvedDuplicateType	899	0.258435
25	GenericDeclaration	894	0.256998
26	JavaType	748	0.215027
27	Primitive	652	0.187430
28	Key	560	0.160983
29	Person	550	0.158108
30	Scalar	544	0.156383
31	ExternalType	489	0.140573
32	Maven	343	0.098602
33	Author	298	0.085666
34	Array	273	0.078479
35	Interface	273	0.078479
36	Committer	252	0.072442
37	Property	202	0.058069
38	Directory	183	0.052607
39	Enum	178	0.051170

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: 1079659
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	258356	23.929407
1	MODIFIES	258356	23.929407
2	UPDATES	169779	15.725243
3	CREATES	61161	5.664844
4	DELETES	37961	3.516018
5	INVOKES	35731	3.309471
6	COMMITTED	25396	2.352224
7	DEPENDS_ON	21931	2.031290
8	OF_TYPE	21393	1.981459
9	DECLARES	20723	1.919402
10	OF_RAW_TYPE	17093	1.583185
11	HAS_PARENT	15377	1.424246
12	HAS	14125	1.308283
13	HAS_COMMIT	12698	1.176112
14	RETURNS	12577	1.164905
15	HAS_FILE	10809	1.001149
16	RENAMES	10545	0.976697
17	READS	9161	0.848509
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.767650
19	HAS_NEW_NAME	6237	0.577682
20	OF_GENERIC_TYPE	5906	0.547025
21	RESOLVES_TO	5215	0.483023
22	SIMILAR	3977	0.368357
23	WRITES	3818	0.353630
24	CONTAINS	3814	0.353260
25	RETURNS_GENERIC	3545	0.328344
26	ANNOTATED_BY	2810	0.260267
27	REQUIRES	2174	0.201360
28	HAS_FIRST_CHILD	2144	0.198581
29	HAS_LAST_CHILD	2144	0.198581

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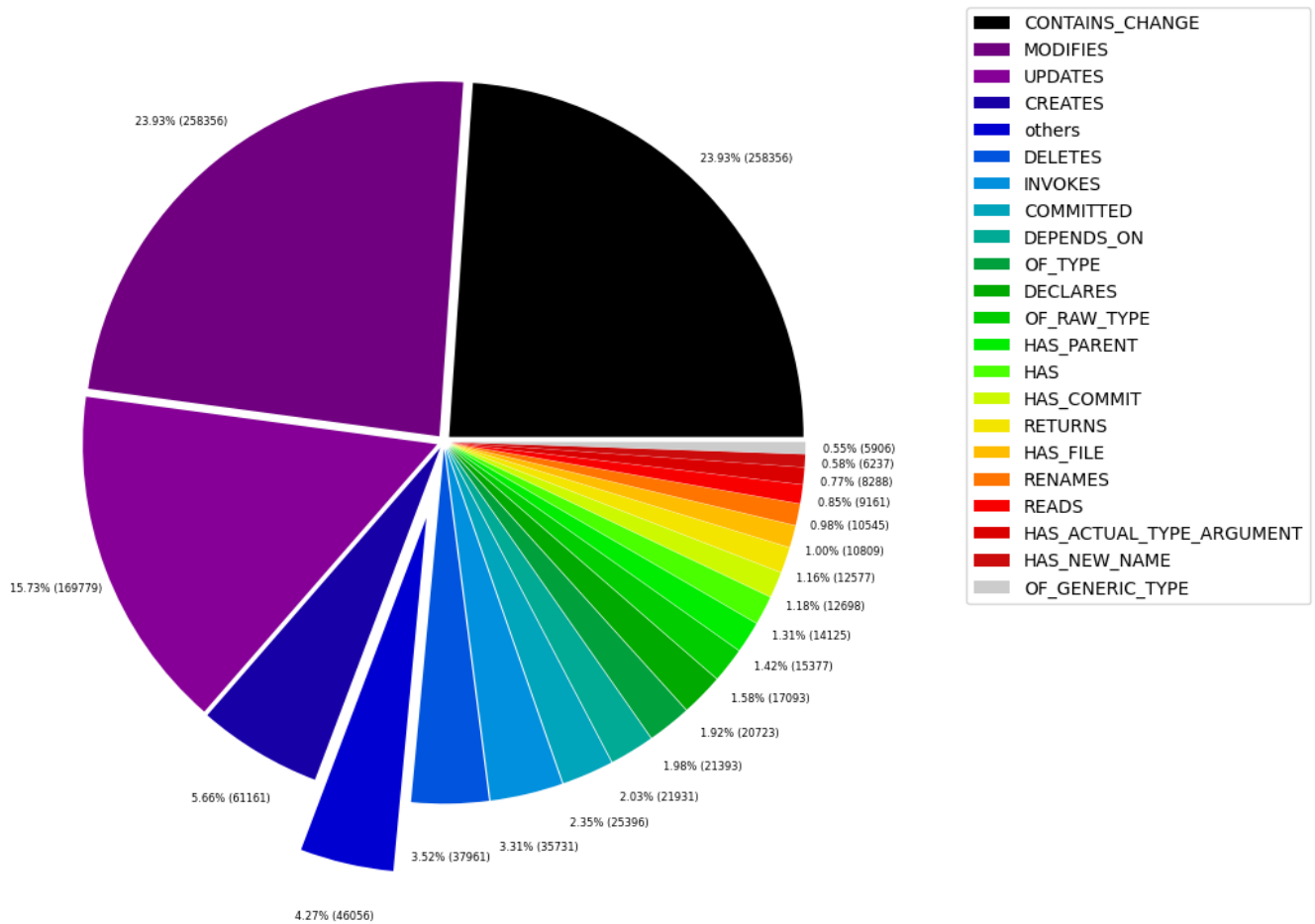
Relationship types (more than 0.5% overall)

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

Chart values:
- 23.93% (258356)
- 23.93% (258356)
- 15.73% (169779)
- 5.66% (61161)
- 4.27% (46056)
- 3.52% (37961)
- 3.31% (35731)
- 2.35% (25396)
- 2.03% (21931)
- 1.98% (21393)
- 1.92% (20723)
- 1.58% (17093)
- 1.42% (15377)
- 1.31% (14125)
- 1.18% (12698)
- 1.16% (12577)
- 1.00% (10809)
- 0.98% (10545)
- 0.85% (9161)
- 0.77% (8288)
- 0.58% (6237)
- 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.
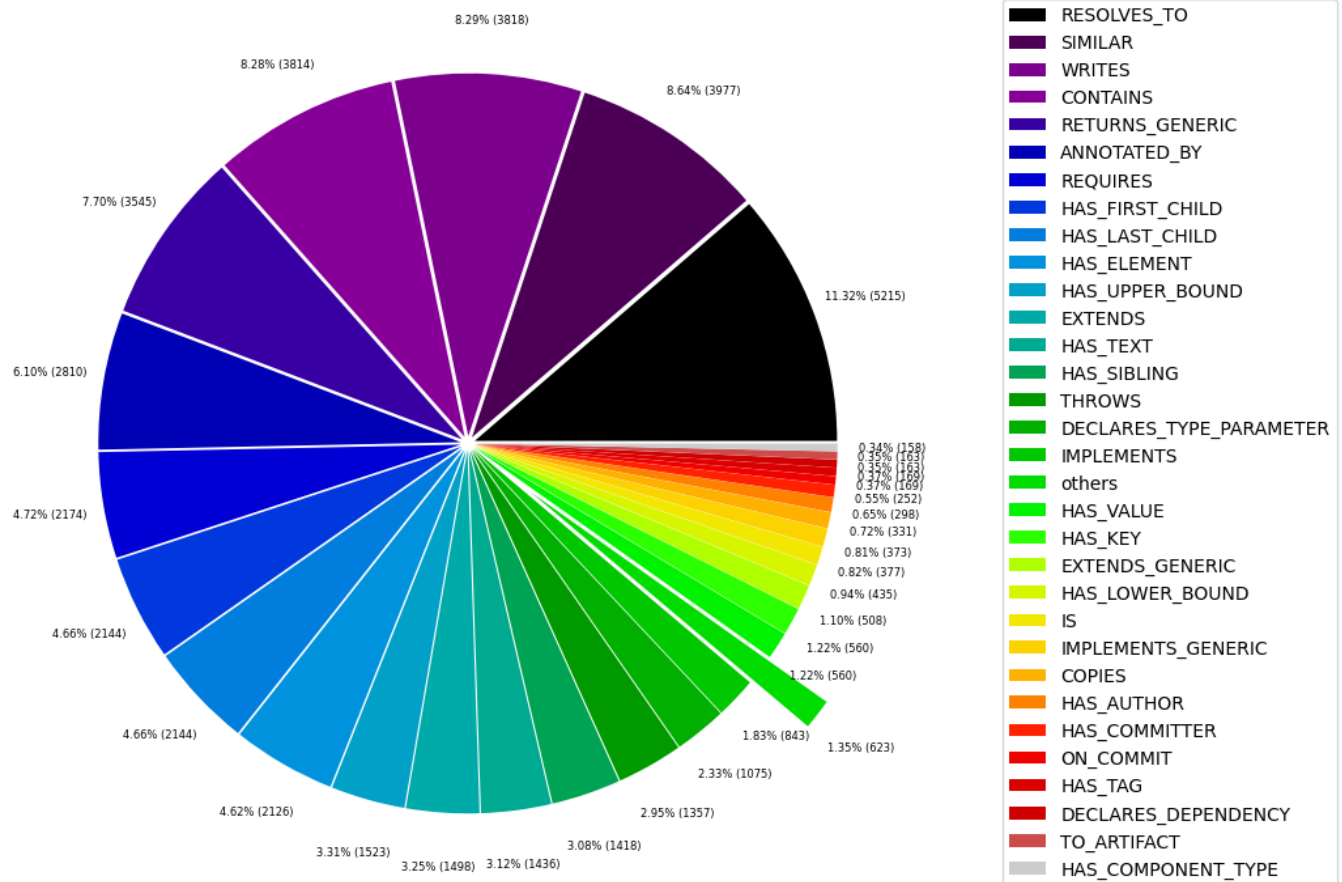
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000093
1	THROWS_GENERIC	5	0.000463
2	EXCLUDES	5	0.000463
3	DESCRIBES	9	0.000834
4	HAS_ROOT_ELEMENT	10	0.000926
5	HAS_GOAL	16	0.001482
6	HAS_EXECUTION	16	0.001482
7	OF_NAMESPACE	18	0.001667
8	HAS_ATTRIBUTE	18	0.001667
9	INCLUDES_CONCEPT	19	0.001760
10	USES_PLUGIN	21	0.001945
11	REQUIRES_TYPE_PARAMETER	21	0.001945
12	IS_ARTIFACT	21	0.001945
13	HAS_CONFIGURATION	21	0.001945
14	REQUIRES_CONCEPT	28	0.002593
15	HAS_DEFAULT	36	0.003334
16	DECLARES_NAMESPACE	36	0.003334
17	HAS_BRANCH	37	0.003427
18	HAS_HEAD	38	0.003520
19	CONTAINS_VALUE	121	0.011207
20	COPY_OF	126	0.011670
21	HAS_COMPONENT_TYPE	158	0.014634
22	TO_ARTIFACT	163	0.015097
23	DECLARES_DEPENDENCY	163	0.015097
24	ON_COMMIT	169	0.015653
25	HAS_TAG	169	0.015653
26	HAS_COMMITTER	252	0.023341
27	HAS_AUTHOR	298	0.027601
28	COPIES	331	0.030658
29	IMPLEMENTS_GENERIC	373	0.034548

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]	258356	12698	
1	[Git, Change]	MODIFIES	[File, Git]	258356	258356	
2	[Git, Change]	UPDATES	[File, Git]	169779	258356	
3	[Git, Change]	CREATES	[File, Git]	61161	258356	
4	[Git, Change]	DELETES	[File, Git]	37961	258356	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21637	13123	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15368	12698	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12698	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12698	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12698	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10809	1	
11	[Git, Change]	RENAMES	[File, Git]	10545	258356	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8358	13123	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8184	13123	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6372	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6237	10809	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5095	817	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4496	817	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7118	
19	[Java, ByteCode, ParameterizedType, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, ParameterizedType, Bound]	2657	13029	
22	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2438	2040	
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]	2188	817	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2120	2040	
28	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2091	13123	
29	[Type, File, Java, ByteCode, Class, GenericDec...]	DECLARES	[Java, ByteCode, Member, Method]	2042	227	

Graph Density

```
total_number_of_nodes (vertices): 347863
total_number_of_relationships (edges): 1079659
-> total directed graph density: 8.92218824094801e-06
-> total directed graph density in percent: 0.000892218824094801
```