Internal Dependencies

References

- Analyze java package metrics in a graph database
- Calculate metrics
- Neo4j Python Driver

Artifacts

List the artifacts this notebook is based on. Different sorting variations help finding artifacts by their features and support larger code bases where the list of all artifacts gets too long.

Only the top 30 entries are shown. The whole table can be found in the following CSV report: `List_all_existing_artifacts`

Table 1a - Top 30 artifacts with the highest package count

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1b - Top 30 artifacts with the highest type count

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1c - Top 30 artifacts with the highest number of incoming dependencies

The following table lists the top 30 artifacts that are used the most by other artifacts (highest count of incoming dependencies, highest in-degree).

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1d - Top 30 artifacts with the highest number of outgoing dependencies

The following table lists the top 30 artifacts that are depending on the highest number of other artifacts (highest count of outgoing dependencies, highest out-degree).

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1e - Top 30 artifacts with the lowest package count

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1f - Top 30 artifacts with the lowest type count

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1g - Top 30 artifacts with the lowest number of incoming dependencies

The following table lists the top 30 artifacts that are used the least by other artifacts (lowest count of incoming dependencies, lowest in-degree).

artifactName	packages	types	incomingDependencies	outgoingDependencies

Table 1h - Top 30 artifacts with the lowest number of outgoing dependencies

The following table lists the top 30 artifacts that are depending on the lowest number of other artifacts (lowest count of outgoing dependencies, lowest out-degree).

artifactName	packages	types	incomingDependencies	outgoingDependencies

Cyclic Dependencies

Cyclic dependencies occur when one package uses a class of another package and vice versa. These dependencies can lead to problems when one of these packages needs to be changed.

Table 2a - Cyclic Dependencies Overview

Show the top 40 cyclic dependencies sorted by the most promising to resolve first. This is done by calculating the number of forward dependencies (first cycle participant to second cycle participant) in relation to backward dependencies (second cycle participant back to first cycle participant). The higher this rate (approaching 1), the easier it should be to resolve the cycle by focussing on the few backward dependencies.

Only the top 40 entries are shown. The whole table can be found in the following CSV report: `Cyclic_Dependencies`

Columns:

- *artifactName* identifies the artifact of the first participant of the cycle
- *packageName* identifies the package of the first participant of the cycle
- *dependentArtifactName* identifies the artifact of the second participant of the cycle
- *dependentPackageName* identifies the package of the second participant of the cycle
- *forwardToBackwardBalance* is between 0 and 1. High for many forward and few backward dependencies.
- *numberForward* contains the number of dependencies from the first participant of the cycle to the second one
- *numberBackward* contains the number of dependencies from the second participant of the cycle back to the first one
- *someForwardDependencies* lists some forward dependencies in the text format "type1 -> type2"
- *backwardDependencies* lists the backward dependencies in the format "type1 <- type2" that are recommended to get resolved

artifactName	packageName	dependentArtifactName	dependentPackageName	forwardToBackwardBalance	numberForward	numberBackward	someForwardDepend

Table 2b - Cyclic Dependencies Break Down

Lists packages with cyclic dependencies with every dependency in a separate row sorted by the most promising dependency first.

Only the top 40 entries are shown. The whole table can be found in the following CSV report: `Cyclic_Dependencies_Breakdown`

Columns in addition to Table 2a:

- *dependency* shows the cycle dependency in the text format "type1 -> type2" (forward) or "type2<- type1" (backward)

artifactName	packageName	dependentArtifactName	dependentPackageName	dependency	forwardToBackwardBalance	numberForward	numberBackward

Table 2c - Cyclic Dependencies Break Down - Backward Dependencies Only

Lists packages with cyclic dependencies with every dependency in a separate row sorted by the most promising dependency first. This table only contains the backward dependencies from the second participant of the cycle back to the first one that are the most promising to resolve.

Only the top 40 entries are shown. The whole table can be found in the following CSV report: `Cyclic_Dependencies_Breakdown_BackwardOnly`

artifactName	packageName	dependentArtifactName	dependentPackageName	dependency	forwardToBackwardBalance	numberForward	numberBackward

Interface Segregation Candidates

Well known from Design Principles and Design Patterns by Robert C. Martin, the *Interface Segregation Principle* suggests that software components should have narrow, focused interfaces rather than large, general-purpose ones. The goal is to minimize the dependencies between components and increase modularity, flexibility, and maintainability.

Smaller, focused and purpose-driven interfaces

- make it easier to modify individual components without affecting the rest of the system.
- make it clearer which client is affected by which change.
- don’t force their clients to depend on methods they don’t need.
- reduce the scope of changes since a change to one component doesn’t affect others.
- lead to a more loosely coupled architecture that is easier to understand and maintain.

Reference: Analyze java package metrics in a graph database

How to apply the results

If just one method of a type is used, especially in many places, then the result of this method can be used to call e.g. a method or constuct an object instead of using the whole object and then just calling that single method.

If there are a couple of methods that are used for a distinct purpose, those could be factored out into a separate interface. The original type can extended/implement the new interface so that there are no breaking changes. Then all the callers, that use only this group of methods, can be changed to the new interface.

Table 4 - Top 40 most used combinations of methods

The following table shows the top 40 most used combinations of methods of larger types that might benefit from applying the *Interface Segregation Principle*. The whole table can be found in the CSV report `Candidates_for_Interface_Segregation`.

fullDependentTypeName	declaredMethods	calledMethodNames	calledMethods	callerTypes

Package Usage

Table 5 - Types that are used by multiple packages

This table shows the top 40 packages that are used by the highest number of different packages. The whole table can be found in the CSV report

`List_types_that_are_used_by_many_different_packages`.

fullQualifiedDependentTypeName	dependentTypeName	dependentTypeLabels	numberOfUsingPackages

Table 6 - Packages that are used by multiple artifacts

This table shows the top 30 artifacts that only use a few (compared to all existing) packages of another artifact. The whole table can be found in the CSV report `ArtifactPackageUsage`.

artifactName	dependentArtifactName	dependentPackages	dependentArtifactPackages	packageUsagePercentage	dependentFullQualifiedPackageNames	dependent

Table 7 - Types that are used by multiple artifacts

This table shows the top 30 types that only use a few (compared to all existing) types of another artifact. The whole table can be found in the CSV report `ClassesPerPackageUsageAcrossArtifacts`.

artifactName	dependentArtifactName	packageName	dependentPackage.fqn	dependentTypes	dependentPackageTypes	typeUsagePercentage	dependentTypeName

Table 8 - Duplicate package names across artifacts

This table shows the top 30 duplicate package names across artifacts. They are ordered by the number of duplicates descending.

This might lead to confusion, makes importing more error prone and might even lead to duplicate classes where only one of them will be loaded by the class loader. If a package is named the same way in two or more artifacts this even allows another artifact to access package protected classes, methods or members which might not be intended.

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

packageName	duplicates	artifactNames

Table 9 - Annotated elements

This table shows 30 most used Java Annotations including some examples where they are used.

annotationName	languageElement	numberOfAnnotatedElements	examples