Taming Bloated Software Dependencies

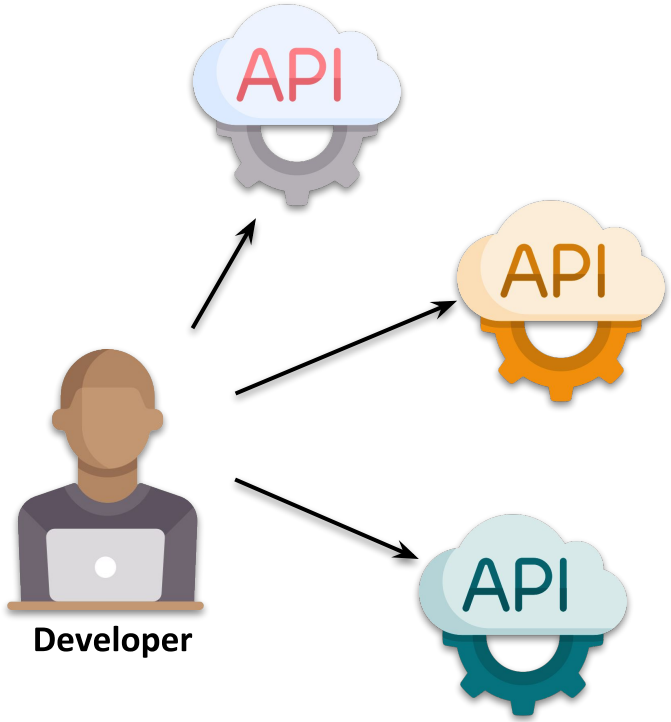
Benoit Baudry, César Soto Valero
baudry@kth.se, cesarsv@kth.se

CASTOR
Software Research Centre

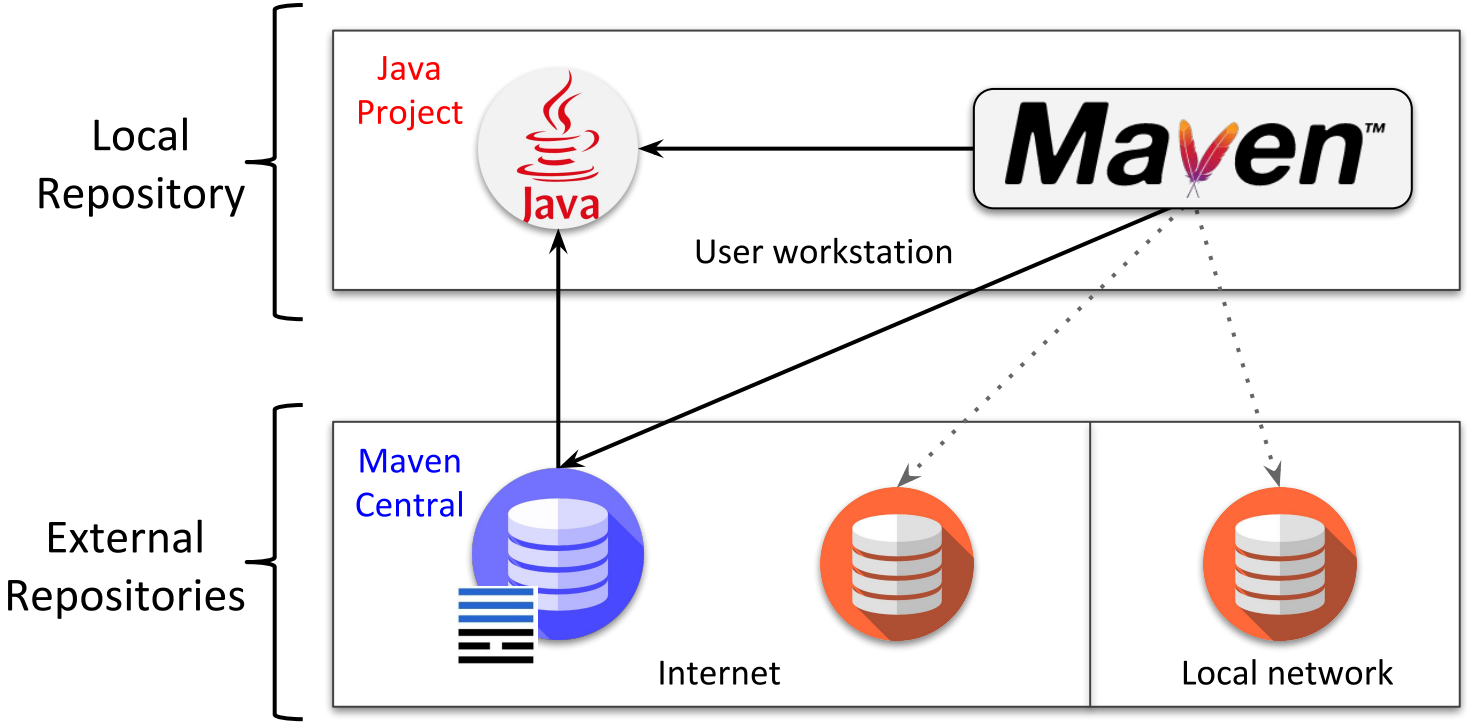
WASP
WALLENBERG AI,
AUTONOMOUS SYSTEMS
AND SOFTWARE PROGRAM

APIs: the backbone of software development

- Facilitate reusability
- Boost productivity
- Increase software quality
- Prevent dependency monoculture
- Increase fault tolerance



Software dependency management with Maven



Example dependency usage: jxls-poi

Dependency declaration

```
<dependency>
    <groupId>org.jxls</groupId>
    <artifactId>jxls-poi</artifactId>
    <version>1.0.15</version>
</dependency>
```

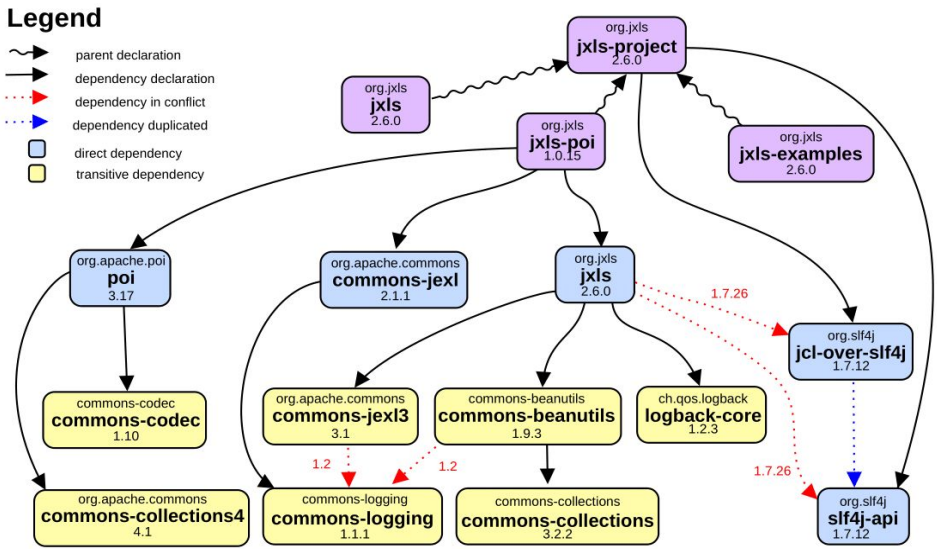
Dependency usage

```
import org.jxls.common.Context;
import org.jxls.util.JxlsHelper;

. . .

try(InputStream is = ObjectCollectionDemo.class.
getResourceAsStream("object_collection_template.xls")) {
   try (OutputStream os = new FileOutputStream(
            "target/object_collection_output.xls")) {
        Context context = new Context();
        context.putVar("objectModels", objectModels);
        JxlsHelper.getInstance()
              .processTemplate(is, os, context);
   }
}
```

Dependency tree



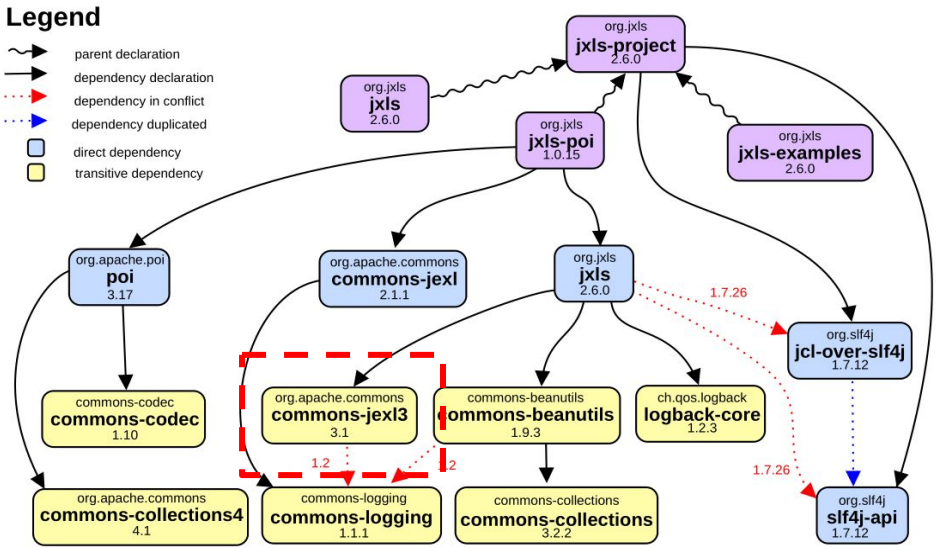
Bloated dependencies

Bytecode

```
Constant pool:
   #1 = Methodref          #9.#29         // org/jxls/common/Context."<init>":()V
   #2 = Fieldref           #3.#30         // org/jxls/transform/poi/PoiContext.varMap:Ljava/util/Map;
   #3 = Class              #31            // org/jxls/transform/poi/PoiContext
   #4 = String             #32            // util
   #5 = Class              #33            // org/jxls/transform/poi/PoiUtil
   #6 = Methodref          #5.#29         // org/jxls/transform/poi/PoiUtil."<init>":()V
   #7 = InterfaceMethodref #34.#35        // Java/util/Map.put:(Ljava/lang/Object;Ljava/lang/Object;)Ljava/lang/Object;
   #8 = Methodref          #9.#36         // org/jxls/common/Context."<init>":(Ljava/util/Map;)V
   #9 = Class              #37            // org/jxls/common/Context
...


  public org.jxls.transform.poi.PoiContext(java.util.Map<java.lang.String, java.lang.Object>);
                descriptor: (Ljava/util/Map;)V
                flags: (0x0001) ACC_PUBLIC
                Code:
                stack=4, locals=2, args_size=2
                0: aload_0
                1: aload_1
                2: invokespecial #8          // Method org/jxls/common/Context."<init>":(Ljava/util/Map;)V
                5: aload_0
                6: getfield      #2          // Field varMap:Ljava/util/Map;
                9: ldc           #4          // String util
                11: new          #5          // class org/jxls/transform/poi/PoiUtil
  ...
```

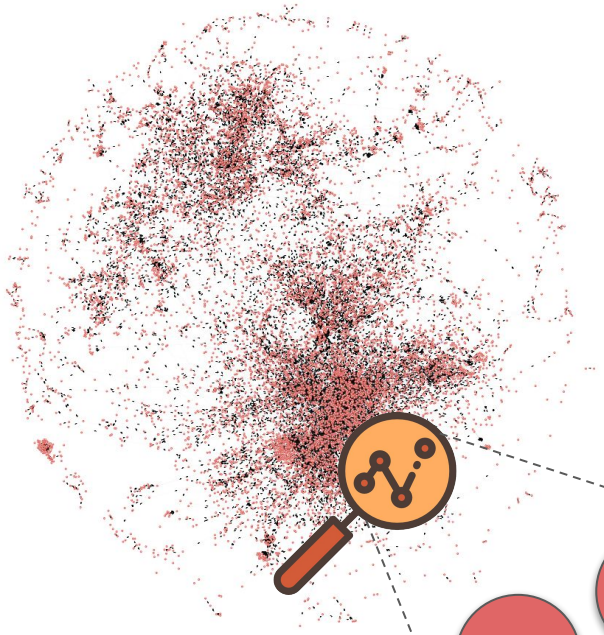
Dependency tree

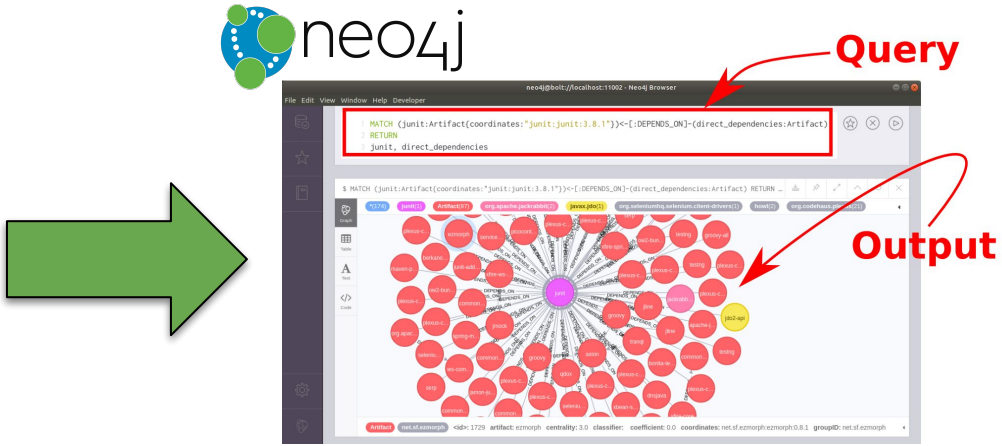


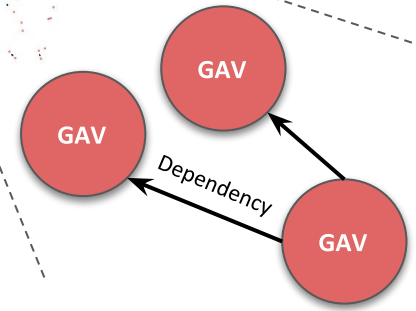
jdbl-pom-maven-plugin

- The analysis of dependencies is based on static analysis.
- The tool reports on dependencies that are:
  - Used and Declared
  - Used and Undeclared
  - Unused and Declared
  - Unused and Undeclared
- The tool produces a debloated pom file

The Maven Dependency Graph



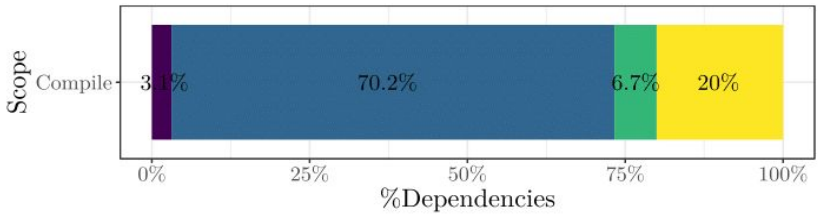


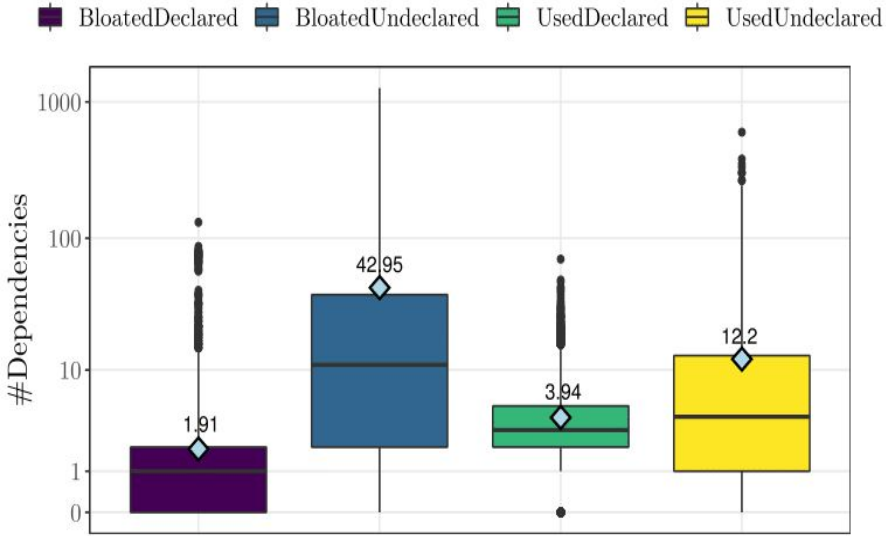


- ~2.4M artifacts
- ~223K libraries
- ~9M direct dependencies

Tool available: https://github.com/diverse-project/maven-miner
Data available: https://zenodo.org/record/1489120

Results: large-scale study

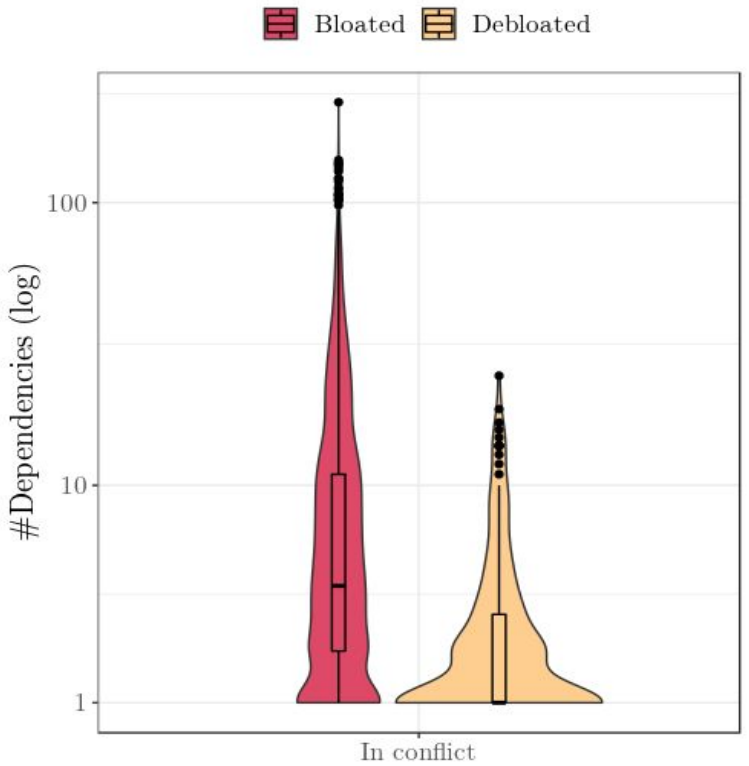




- 73% of dependencies in our dataset of Maven Central artifacts are bloated
- The average number of bloated dependencies per project is 45

Results: large-scale study

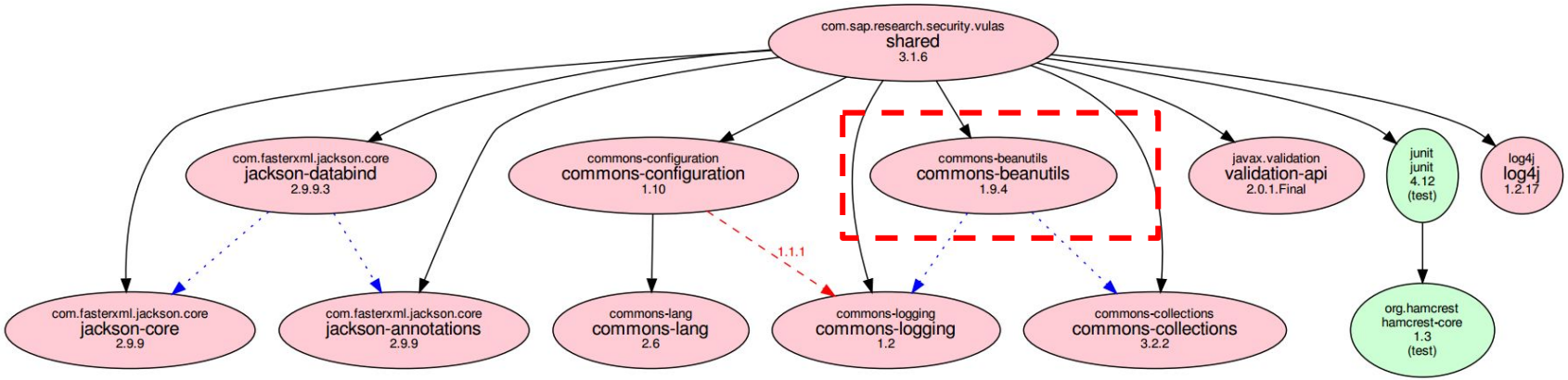
- 98.3% of the total number of dependency conflicts are superfluous, i.e. they are caused by bloated dependencies.
- By removing bloated dependencies, the average number of conflicts per project drops from 7.6 to 2.5



Example

Open-source vulnerability assessment tool license Apache 2.0

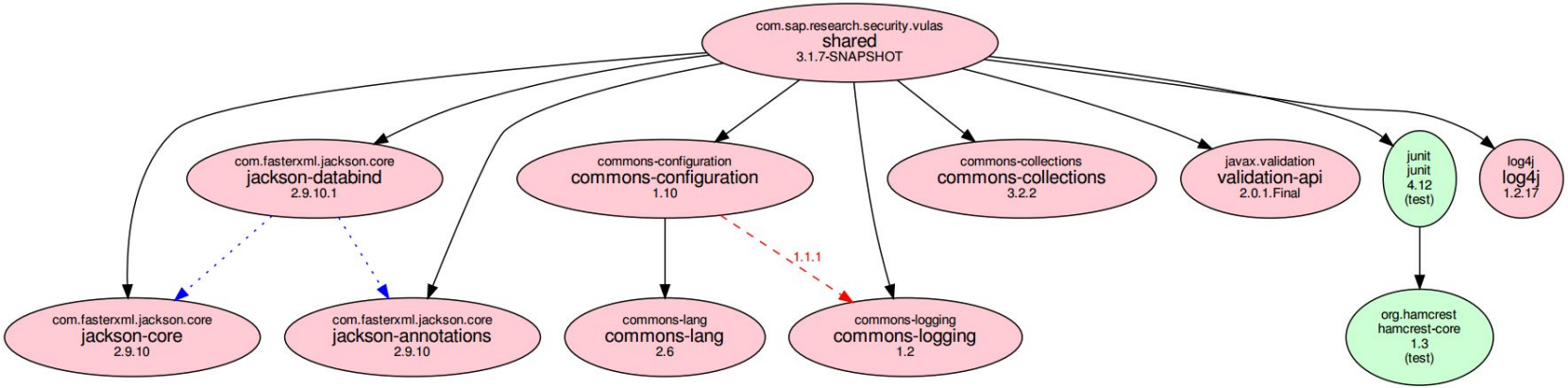
PRs welcome build passing release v3.1.6



Example

Open-source vulnerability assessment tool license Apache 2.0

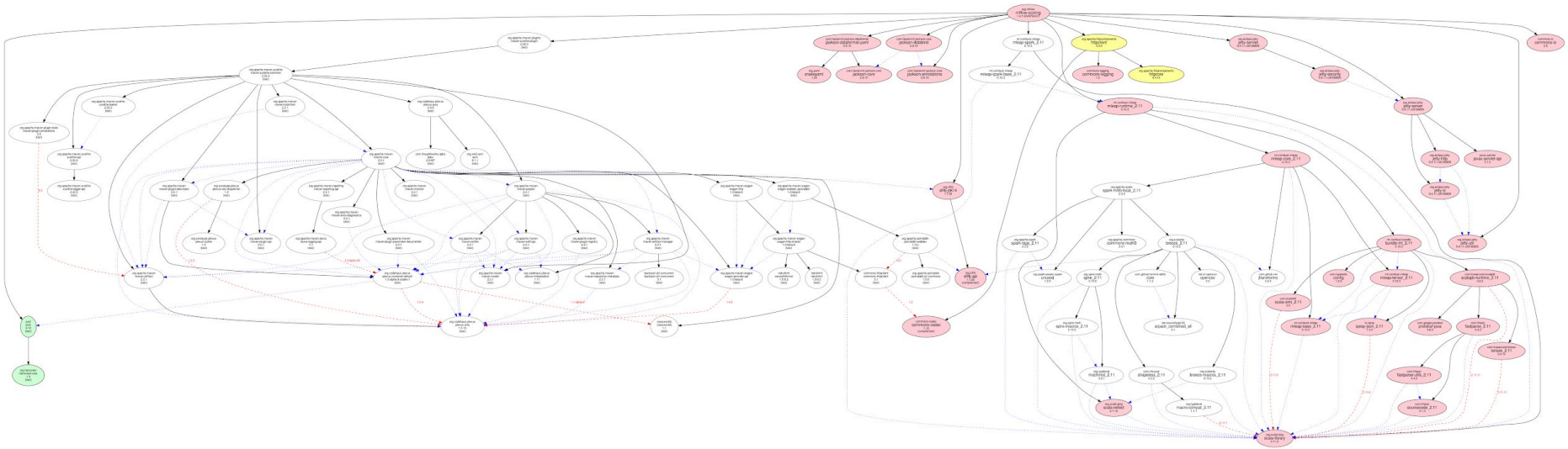
PRs welcome build passing release v3.1.6



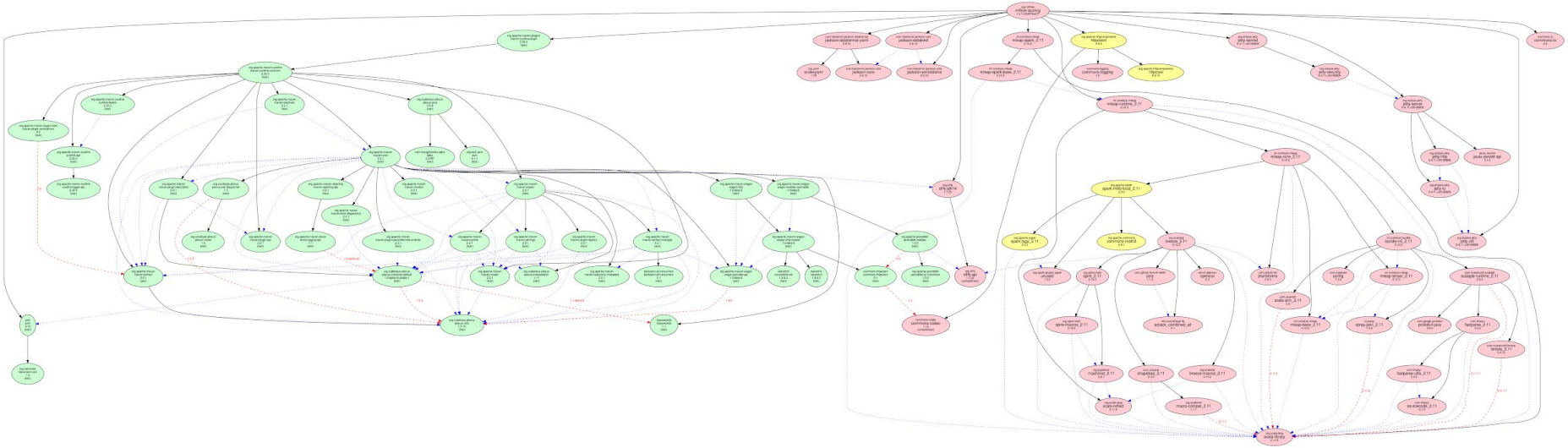
henrikplate commented 9 days ago • edited Contributor Author

It definitly is useful. In fact, I was able to remove the direct dependency `commons-beanutils` :)

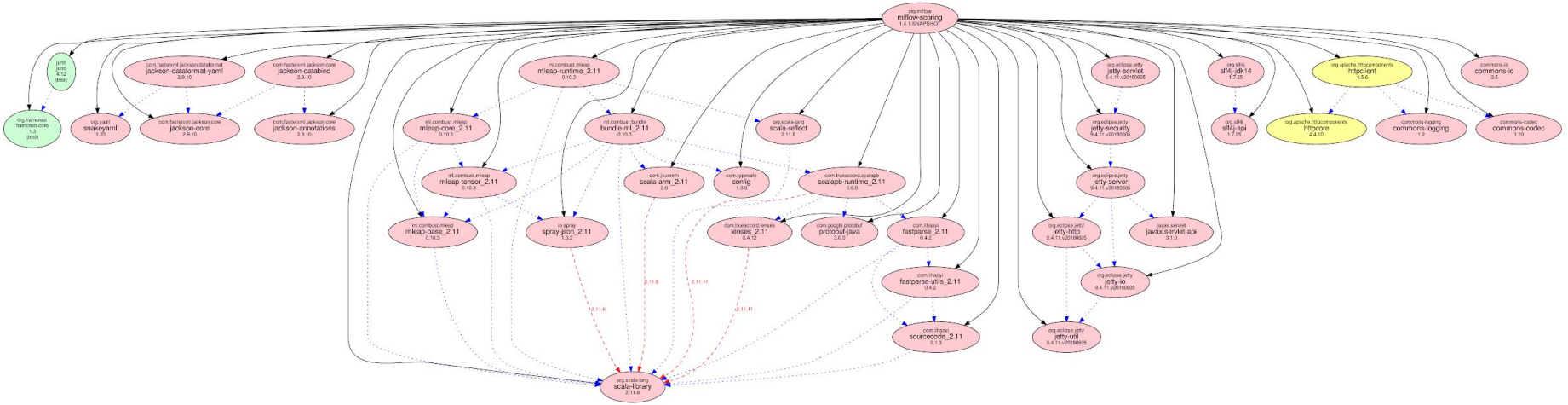
Example: **MLflow** original dependency tree



Example: **MLflow** original dependency tree



Example: **MLflow** debloated



Try it yourself!

https://github.com/castor-software/royal-debloat