

PDFBox Library

# Table of Contents

How to configure/use .....	2
Classpath .....	2
Supporting encrypted PDFs.....	2
Bootstrapping .....	2
API .....	4
Known issues.....	5
Dependencies .....	5

This module (`isis-module-pdfbox`) provides a wrapper around [Apache PDFBox](#).

# How to configure/use

## Classpath

Update your classpath by adding this dependency in your dom project's `pom.xml`:

```
<dependency>
  <groupId>org.isisaddons.module.pdfbox</groupId>
  <artifactId>isis-module-pdfbox-dom</artifactId>
  <version>1.16.1</version>
</dependency>
```

Check for later releases by searching [Maven Central Repo](#).

For instructions on how to use the latest `-SNAPSHOT`, see the [contributors guide](#).

## Supporting encrypted PDFs

Optionally, encrypted PDFs can be handled by adding the [Legion of the Bouncy Castle](#) encryption libraries. These can be included in your Maven project using three or so dependencies.

At the time of writing these were:

```
<dependency>
  <groupId>org.bouncycastle</groupId>
  <artifactId>bcprov-jdk15on</artifactId>
  <version>1.54</version>
</dependency>

<dependency>
  <groupId>org.bouncycastle</groupId>
  <artifactId>bcpkix-jdk15on</artifactId>
  <version>1.54</version>
</dependency>

<dependency>
  <groupId>org.bouncycastle</groupId>
  <artifactId>bcmail-jdk15on</artifactId>
  <version>1.54</version>
</dependency>
```

See the [PDFBox documentation](#) for the most up-to-date information.

## Bootstrapping

In the `AppManifest`, update its `getModules()` method, eg:

```
@Override
public List<Class<?>> getModules() {
    return Arrays.asList(
        ...
        org.isisaddons.module.pdfbox.dom.PdfBoxModule.class,
        ...
    );
}
```

# API

This module provides a single service, called `PdfBoxService`. It has the following API:

```
@DomainService(nature = NatureOfService.DOMAIN)
public class PdfBoxService {

    public byte[] merge(byte[]... pdfByteArrays) throws IOException { } ①
    public byte[] merge(List<File> fileList) throws IOException { ... } ②
    public byte[] merge(File... files) throws IOException { ... }

    public void merge(
        OutputStream outputStream,
        InputStream... inputStreams
    ) throws IOException { } ③
}
```

- ① merge an arbitrary number of PDFs, represented as arrays of bytes, to another byte array
- ② merge a list of files (to avoid memory issues)
- ③ merge an arbitrary number of PDFs, represented as an input stream, to an output stream

# Known issues

None known at this time.

Do note however that stitching together large PDFs can be memory intensive. This in part is because Apache Isis' **Blob** datatype is just a wrapper around a **byte[]** byte array, rather than some sort of stream.

## Dependencies

Maven can report modules dependencies using:

```
mvn dependency:list -o -pl modules/lib/pdfbox/impl -D excludeTransitive=true
```

which, excluding Apache Isis itself, returns these compile/runtime dependencies:

```
org.javassist:javassist:jar:3.19.0-GA  
org.apache.pdfbox:pdfbox:jar:2.0.5
```

For further details on 3rd-party dependencies, see:

- [Apache PDFBox](#)
- [Javassist](#)
- (optionally) [Legion of the Bouncy Castle](#), to handle encrypted PDFs.