



# Decoupling the Archive

Kimberly DuPrie, Lisa Gardner, Michael Gough, Richard C. Kidwell Jr.

kduprie@stsci.edu



ST&I | SPACE TELESCOPE  
SCIENCE INSTITUTE

The JWST archive will store numerous metadata for the various files that it contains: currently a single FITS file can have up to 250 different metadata fields in the archive, most of which map to keywords in the primary header or extension headers. Our goal is to allow for changes in the fields stored in the database without having to change the ingest code. This poster describes the method we use to decouple the archive from the ingest process.

## Ingest Manifest

When the pipeline creates a file that it would like to archive it will create an ingest manifest. This XML file contains basic information about the file such as name, location, type, etc. If it is a FITS file it will also contain keyword information scraped from the primary and extension headers and may optionally contain data from ASCII table extensions. It then posts this manifest to the ingest code via a RESTful call.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<IngestManifest ManifestVersion="0.12" ClientId="dpDev01" Mission="JWST"
xsi:noNamespaceSchemaLocation="http://jwst.stsci.edu/IngestManifest.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <File>
    <ProcessStep>Store+Catalog</ProcessStep>
    <Location>/my/directory</Location>
    <FileName>jw86500030001_uncal.fits</FileName>
    <FileType>Science FGS Level 1b</FileType>
    <Metadata>
      <HeaderExtension extName="PRIMARY" hduNum="0">
        <Keyword name="TIMESYS" value="UTC"/>
        <Keyword name="FOCUSPOS" value="0.0"/>
      </HeaderExtension>
      <HeaderExtension extName="SCI" hduNum="1">
        <Keyword name="XTENSION" value="IMAGE"/>
        <Keyword name="NAXIS1" value="2048"/>
      </HeaderExtension>
    </Metadata>
  </File>
</IngestManifest>
```

## KeywordFieldMap

The database contains the KeywordFieldMap view which maps FITS keywords to a column in a table for a particular FileType.

KeywordFieldMap

fileType	keyword	tablename	fieldname
Science FGS Level 1b	TIMESYS	ScienceCommon	timesys
Science FGS Level 1b	INSTRUME	ScienceCommon	instrument
Science FGS Level 1b	FOCUSPOS	FgsScience	focuspos
Science FGS Level 1b	NUMDTHPT	FgsScience	numdthpt

ScienceCommon

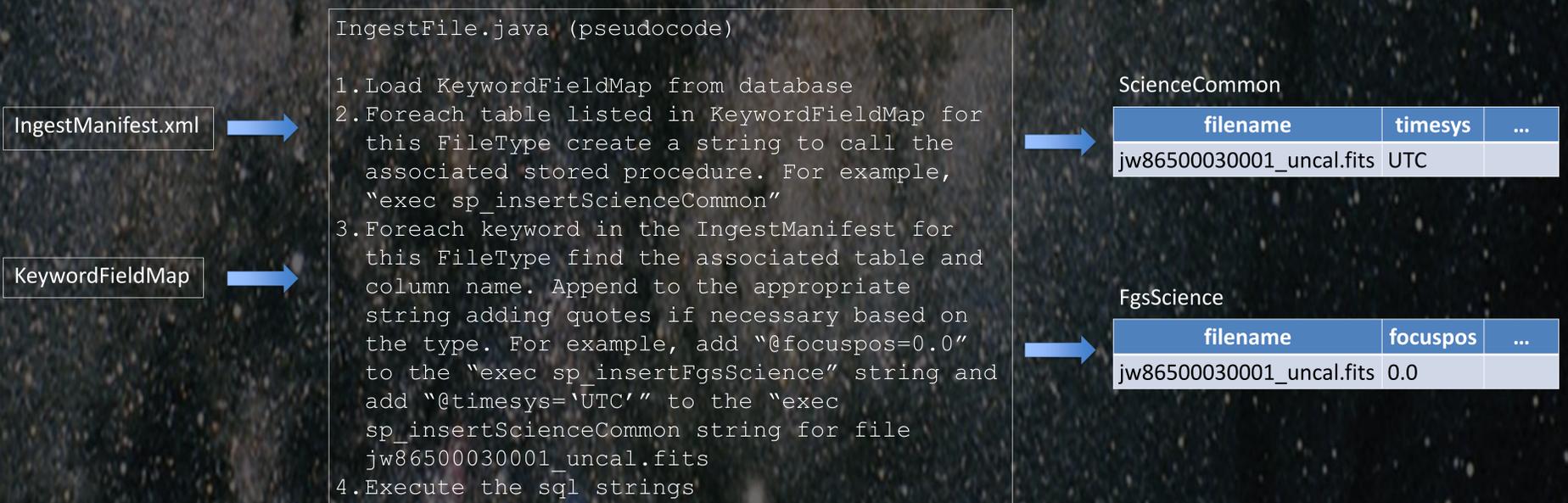
filename	timesys	instrument	...
jw86500030001_uncal.fits	UTC		

FgsScience

filename	focuspos	numdthpt	...
jw86500030001_uncal.fits	0.0		

## Ingest Code

When the ingest code runs it gets all the information it needs from the Ingest Manifest and the KeywordFieldMap view. It calls one or more stored procedures based on the tablename(s) in KeywordFieldMap for this FileType, passing in parameters and values as specified in the Ingest Manifest. If the Ingest Manifest contains a keyword that is not listed in KeywordFieldMap for this FileType then the keyword is simply skipped.



If a new keyword is added to a FITS file then it needs to be added to the Ingest Manifest, KeywordFieldMap and the appropriate table in the database. No change to the Ingest Code is needed. Disclaimer: We are using Hibernate so we also need to update the code that maps to the database table, but this code is separate from the ingest code. If we were not using an ORM then it would not be necessary to update this additional code.