



D3.4.3: Implementierung von Langzeitarchivierungsdiensten Teil 1: Jhove2 FITS Format-Modul

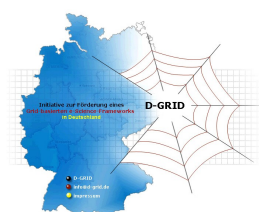
Version – 14.04.2011

Arbeitspaket 3

Verantwortlicher Partner - AWI

WissGrid

Grid für die Wissenschaft



Bundesministerium
für Bildung
und Forschung

WissGrid- Langzeitarchivierungsdienste, Teil 1: Jhove2 FITS Format-Modul

Projekt: **WissGrid**

Teil des D-Grid Verbundes und der deutschen e-Science Initiative

BMBF Förderkennzeichen: 01|G09005A-G (Verbundprojekt)

Laufzeit: Mai 2009 - April 2012

Dokumentstatus: stabile Zwischenversion

Verfügbarkeit: öffentlich

Autoren: José Mejía, Bernadette Fritzsch (AWI)

Revisionsverlauf

Datum	Autor	Kommentare
14.04.2011	J. Mejía, B.Fritzsch	Erster Entwurf

Inhaltsverzeichnis

1 Einleitung.....	4
2 Installation und Testing.....	5
3 Aktueller Stand.....	7
4 Wartung und Support.....	8

1 Einleitung

Für den Ingest von Daten in die Langzeitarchivierung ist die Formatvalidierung und Metadatenextraktion einer der ersten Schritte. Das in der Community der Bibliotheken entwickelte und mittlerweile weit verbreitete Framework JHOVE2¹ kann auch für die in WissGrid betrachteten Forschungsdaten als Grundlage genommen werden. Für die in der Klimawissenschaft gebräuchlichen Formate NetCDF und GRIB und das dort verbreitete Metadatenprofil auf Basis von ISO 19 115/139 wurde das Formatmodul JANEMEs² (J-A NetCDF Metadata Extractor) entwickelt, das konzeptionell bereits auf eine leichte Übertragbarkeit auf andere Communities (neue Formate, andere Metadatenprofile) hin angelegt wurde. Als ein mögliches Beispiel wird hier das Flexible Image Transport System (FITS) Format³, das wohl weitestverbreitete digitale Datenformat in der Astronomie, das 1981 standardisiert wurde und derzeit in der Version 3.0 verwendet wird.

Ähnlich wie NetCDF ist auch FITS ein selbstbeschreibendes Format in dem Sinne, dass zu den eigentlichen Bilddaten auch immer gleich die Metadaten im File abgelegt werden. Eine FITS Datei besteht aus mehreren Header/Datablocks — den sogenannten HDUs. Am AWI wurde auf der Grundlage der bisherigen Arbeiten ein Tool zur Metadatenextraktion aus FITS Daten implementiert. Das FITS-Modul arbeitet iterativ über alle HDUs mittels der *nom.tam.fits* Bibliothek version 1.04 und erkennt dabei den Typ und die Struktur jedes Blocks. Es wird eine Liste von Attributen und Werten pro Block erstellt. Die extrahierten Metadaten können dann in XML, Text oder JSON Format ausgegeben werden.

Dieses Dokument dient als Einleitung in das Jhove2 FITS Modul als Software Entwicklung im Rahmen von WissGrid Arbeitspaket 3. Die ausführliche Dokumentation dieses Moduls ist nach den in der dortigen Community vereinbarten Standards erstellt worden. Die aktuell gehaltene Modulspezifikation *specFITS.pdf* ist Bestandteil des Installationspakets und entspricht dem offiziellen Modul-Dokumentationsformat. In diesem Dokument wird daher für weitergehende Informationen auf diese Spezifikation verwiesen.

¹ <http://jhove2.org/>

² <http://aforge.awi.de/gf/project/jhove2/frs/>

³ Wells, D. C.; Greisen, E. W.; Harten, R. H.: "FITS - a Flexible Image Transport System". *Astronomy and Astrophysics Supp. Ser.* 44: 363–370. June 1981. <http://adsabs.harvard.edu/abs/1981A&AS...44..363W>

2 Installation und Testing

Die Installation Jhove2s mit dem FITS Modul erfolgt nach dem in der Jhove2 offiziellen Anleitung⁴ erklärten Vorgehen und besteht aus den folgenden Schritten:

1. Java 6 Einrichten

- Überprüfung der aktuellen Java-Installation

Befehl: `java -version`

- Installation oder Aktualisierung Javas

Zum Download der Java-Software für Windows, Linux oder Solaris besuchen Sie <http://java.sun.com/javase/downloads/>. Diese Seite bietet Ihnen die Möglichkeit zum Download des Java SE Development Kit (JDK) oder der Java Runtime Environment (JRE). Die JRE reicht aus, um JHOVE2 benutzen zu können. Nur wenn eigene Modifikationen in Jhove2s oder in dem FITS-Modul vorgenommen werden sollen, laden Sie zusätzlich noch die grundlegenden Entwickler-Tools (JDK) herunter.

2. FITS Modul Archiv herunterladen und extrahieren

- FITS ZIP-Archiv ist auf <http://aforge.awi.de/gf/project/jhove2fits/frs/> zum freien Download verfügbar.
- Verschiedene ZIP-Dienstprogramme sind auf den meisten Plattformen vorinstalliert. Das folgende Beispiel extrahiert eine JHOVE2 ZIP-Archiv in das aktuelle Verzeichnis: `unzip jhove2.zip`

3. Anpassung des JHOVE2 Skriptes, wenn nötig

- `jhove2.bat` für MS-Windows
- `jhove2.sh` für unix-ähnliche Systeme

4. Anpassung der Konfigurationsdateien im `config` Verzeichnis, wenn nötig

- Spring XML Configuration Files
- Displayer Property Files
- Java Persistence Files

⁴ Siehe http://bitbucket.org/jhove2/main/wiki/documents/JHOVE2-Users-Guide_20110222.pdf

WissGrid- Langzeitarchivierungsdienste, Teil 1: Jhove2 FITS Format-Modul

Für einen ersten Test wird im Installationspaket die Beispiel-Datei *rosat_pspc_rdf2_3_basic.fits* bereitgestellt. Die Aufrufe

```
./jhove2.sh -o output.xml src/test/resources/examples/fits/rosat_pspc_rdf2_3_basic.fits bzw.
```

```
./jhove2.sh -d XML -o output.xml src/test/resources/examples/fits/rosat_pspc_rdf2_3_basic.fits bzw.
```

```
./jhove2.sh -d JSON -o output.xml src/test/resources/examples/fits/rosat_pspc_rdf2_3_basic.fits
```

liefern dann die darin enthaltenen Metadateninformation in den Outputformaten *Text*, *XML* oder *JSON*.

Das Installationspaket enthält weitere Beispiele für FITS Dateien unter *src/test/resources/examples/fits* sowie die Modulspezifikation *specFITS.pdf*, die eine komplette Übersicht der Benutzung und technischen Details des hier beschriebenen Moduls bietet.

3 Aktueller Stand

Das FITS Modul kann unter <http://aforge.awi.de/gf/project/jhove2fits/frs/> heruntergeladen werden. Unsere Software ist ein unter der BSD-Lizenz freigegebenes Jhove2-Formatmodul.

4 Wartung und Support

Das Modulcode wird von José Mejía Villar, M.Sc., vom Alfred-Wegener-Institut für Polar- und Meeresforschung in Bremerhaven, Deutschland gepflegt.

Der Entwickler kann für Supportanfragen unter <mailto:Jose.Mejia@awi.de> erreicht werden. Hier erhalten Sie fachliche Unterstützung in allen Fragen rund um das Jhove2 FITS-Modul und zur Fehlerbehebung und können uns Feedback auf Englisch, Spanisch oder Deutsch geben.