Workflow & Characterisation

André Schaaff and François Bonnarel

Acknowledgement

- Collegial work
  - T. Boch, F. Bonnarel, B. Gassmann, M. Louys, C. Pestel and A. Schaaff
  - Trainees : G. Mantelet and O. Benjelloun (~ 1 year FTE)
  - Discussions in the frame of the VO France Workflow WG

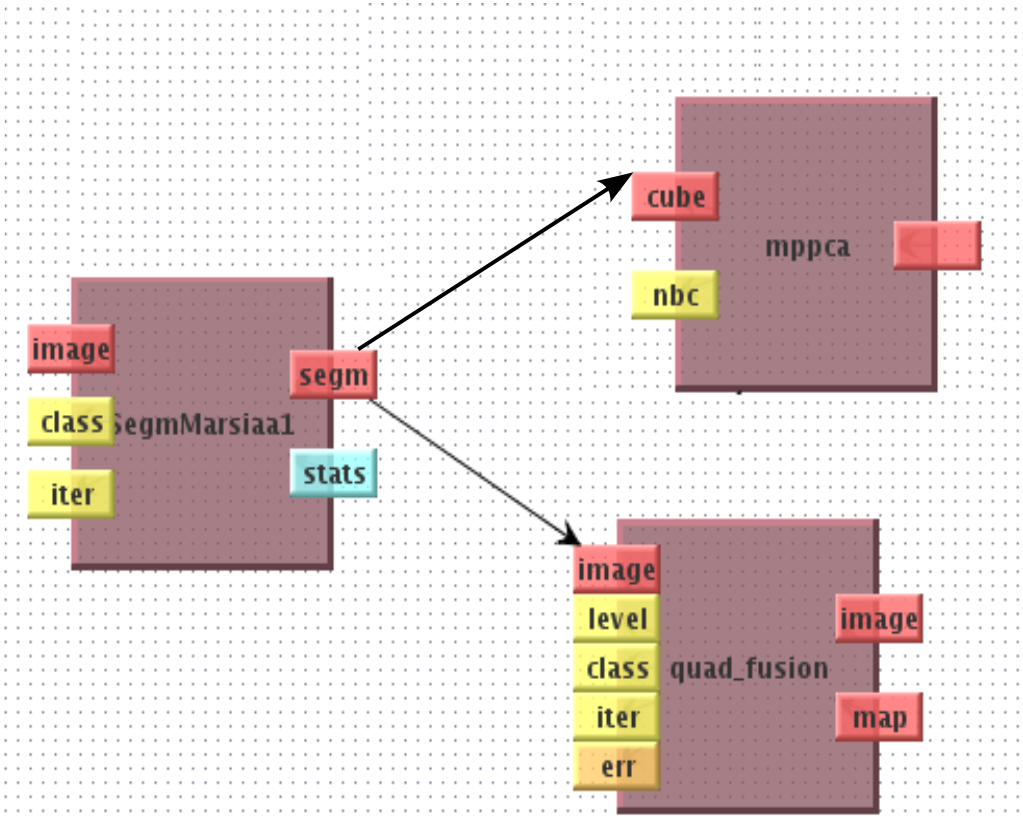
A workflow

- **Workflow : a sequence of tasks executed within a controlled context (taking into account VO standards)**
- **A Workflow involves resources like clusters, grids, databases, ...**
  - **A workflow process is dependant from the tasks it runs and of their “good” execution**
  - **How to optimize the execution of a workflow ?**
  - **...**

Optimization

- Checking of a workflow before and during its execution
  - (Test bed : AÏDA 2.0 – Workflow tool using UWS)

- Benefits
  - A checking is done on the client side before the submission to the engine
  - If validation fails : minimize resources use
  - Time gain for the user
  - ...

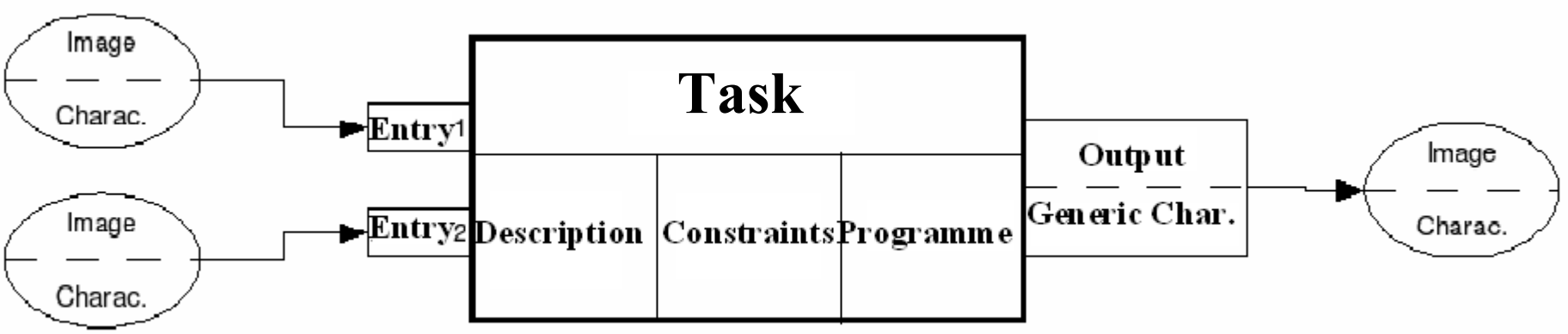


What/How to check

- **Checking of the inputs/outputs**
  - **Verify the types of the linked I/O → Not enough**
  - **Check more than the type**
    - Try to do it for tasks using FITS files as entries and use the Characterization standard
      - FITS file + its characterization file
    - A constraints file at the task level
    - A constraints checker at the workflow tool level
  - **Do this checking also during the execution**
    - Need a characterization file generation for FITS files resulting from the execution (→ FITS file + Characterization file for the next task)

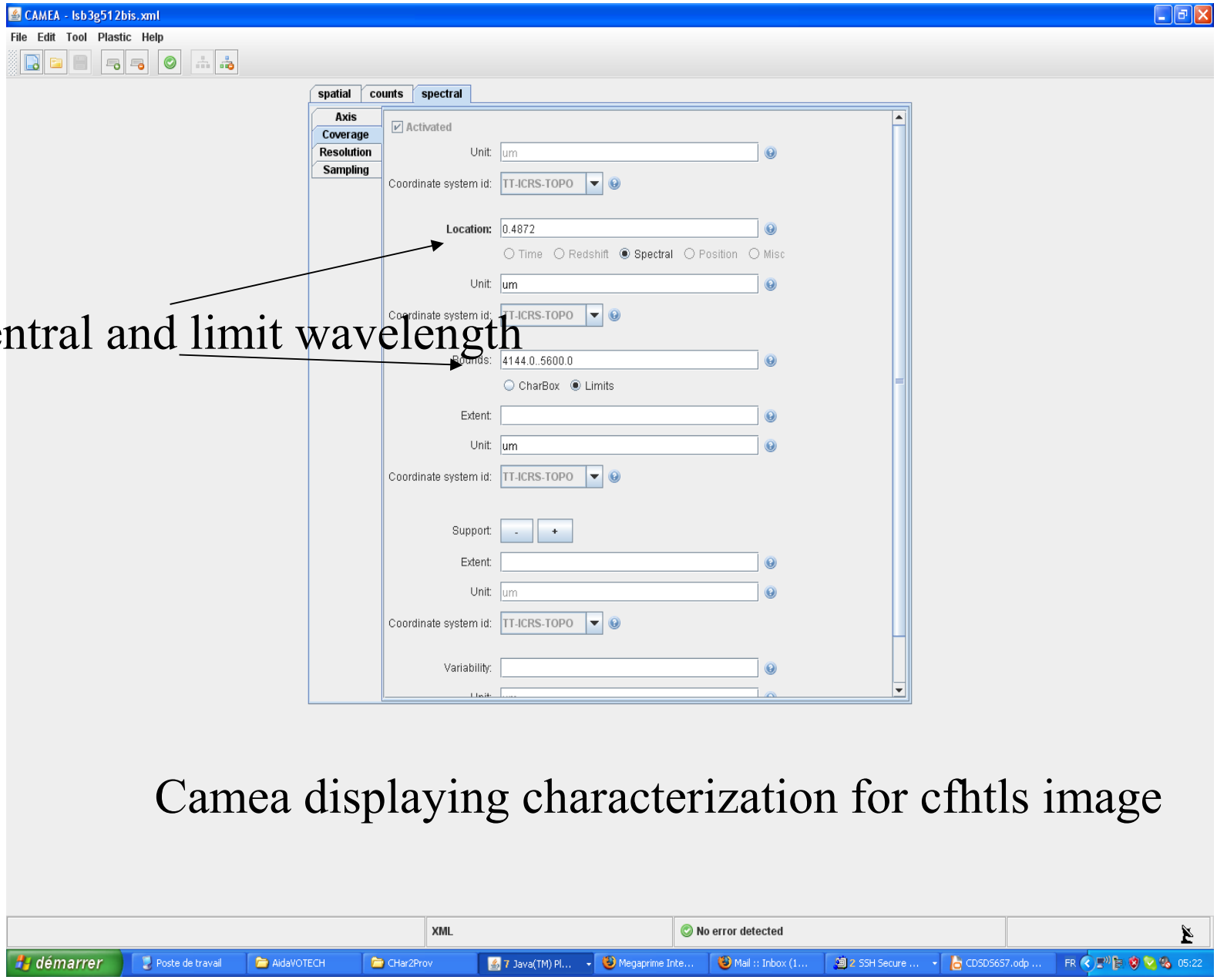
At the task level

- **Before the execution**
  - Constraints definition for the inputs
  - Validation step checks the entries



- **During the execution**
  - A characterization file should be generated for the outputs and checked with next task constraints and so on.

- **Characterization file in CAMEA**
  - **Developed in the frame of VOTECH**



Central and limit wavelength

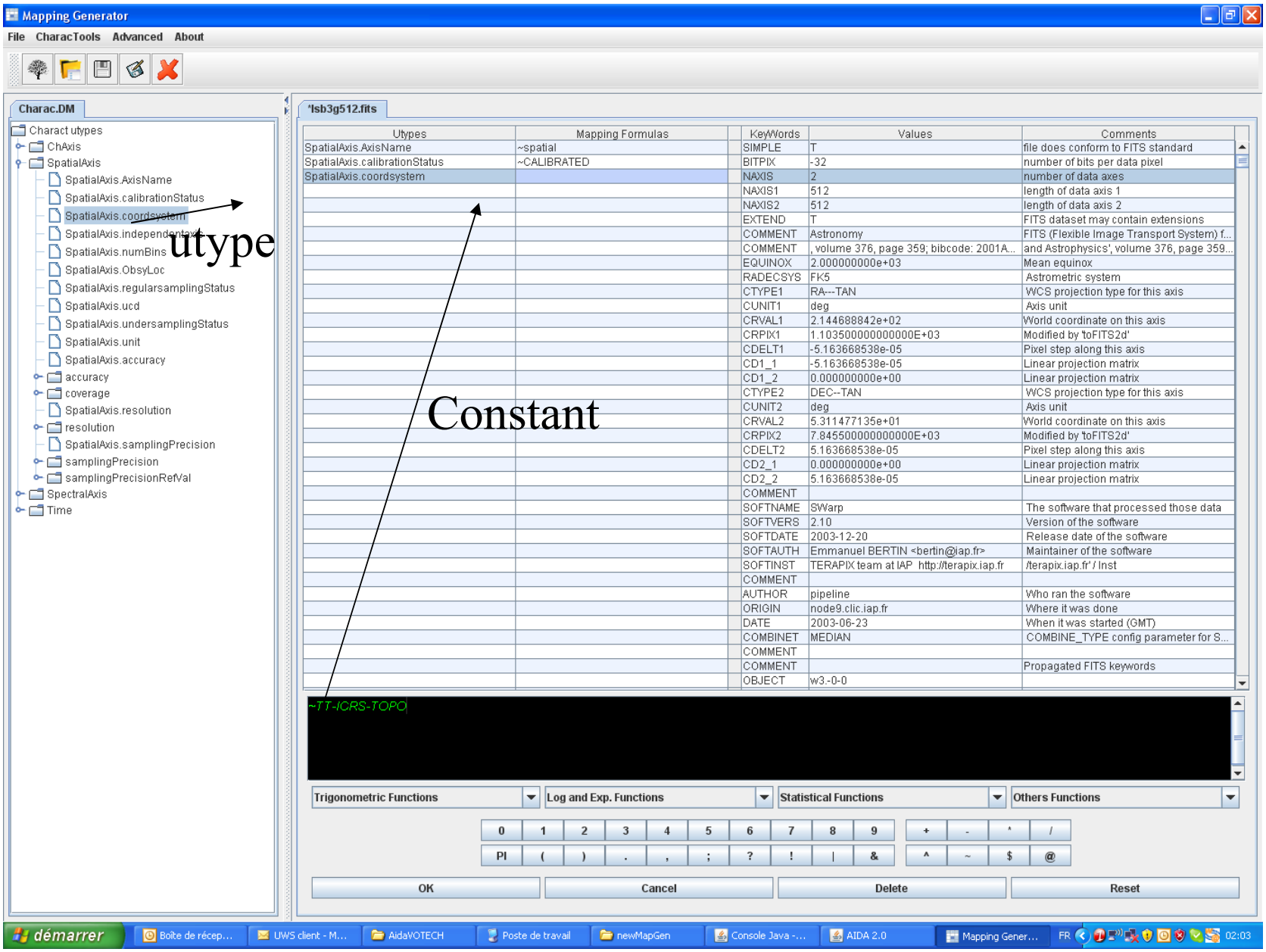
Camea displaying characterization for cfhtls image

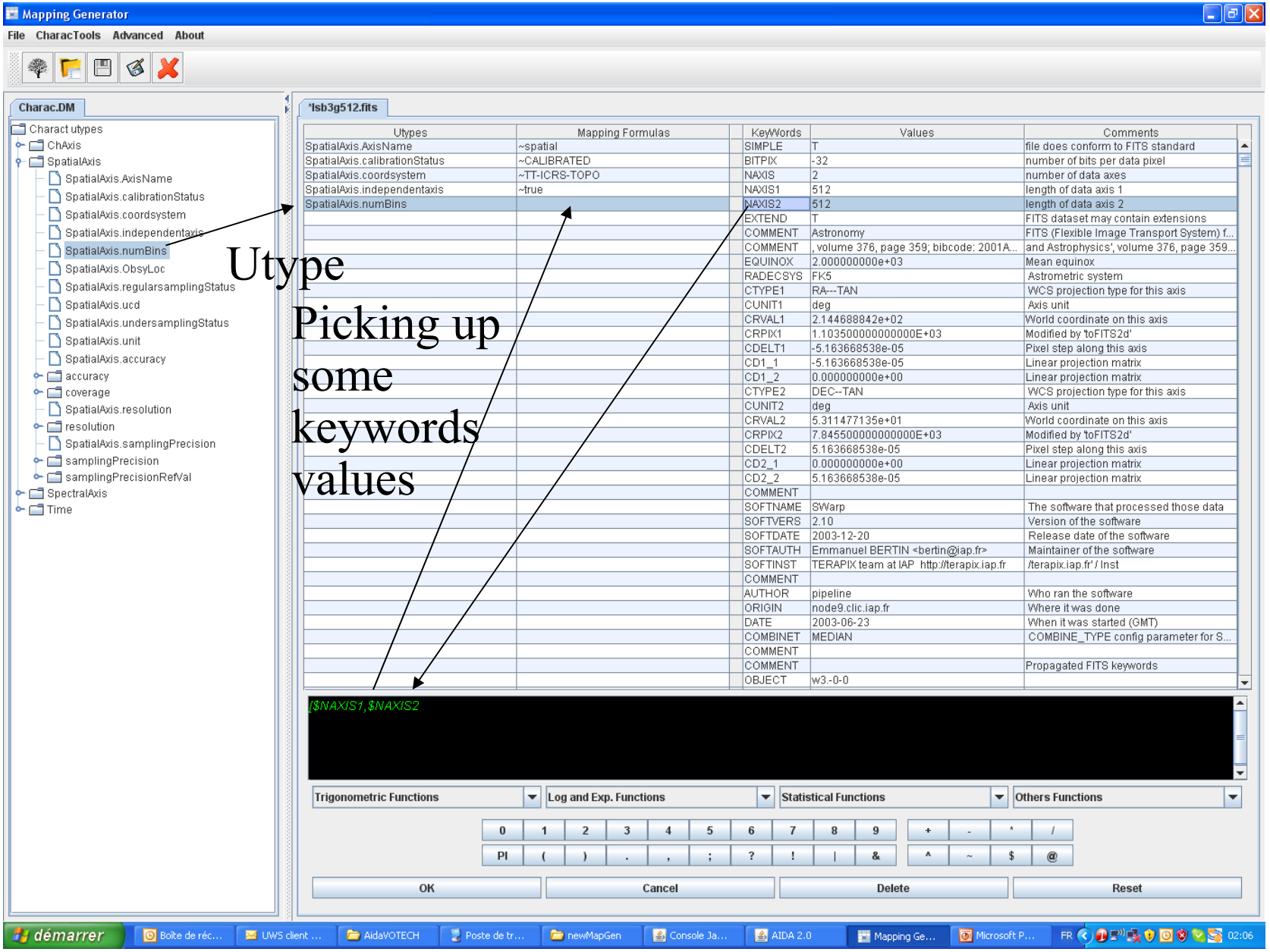
- **Constraint example for the Segmentation task**
  - **Used for the checking**

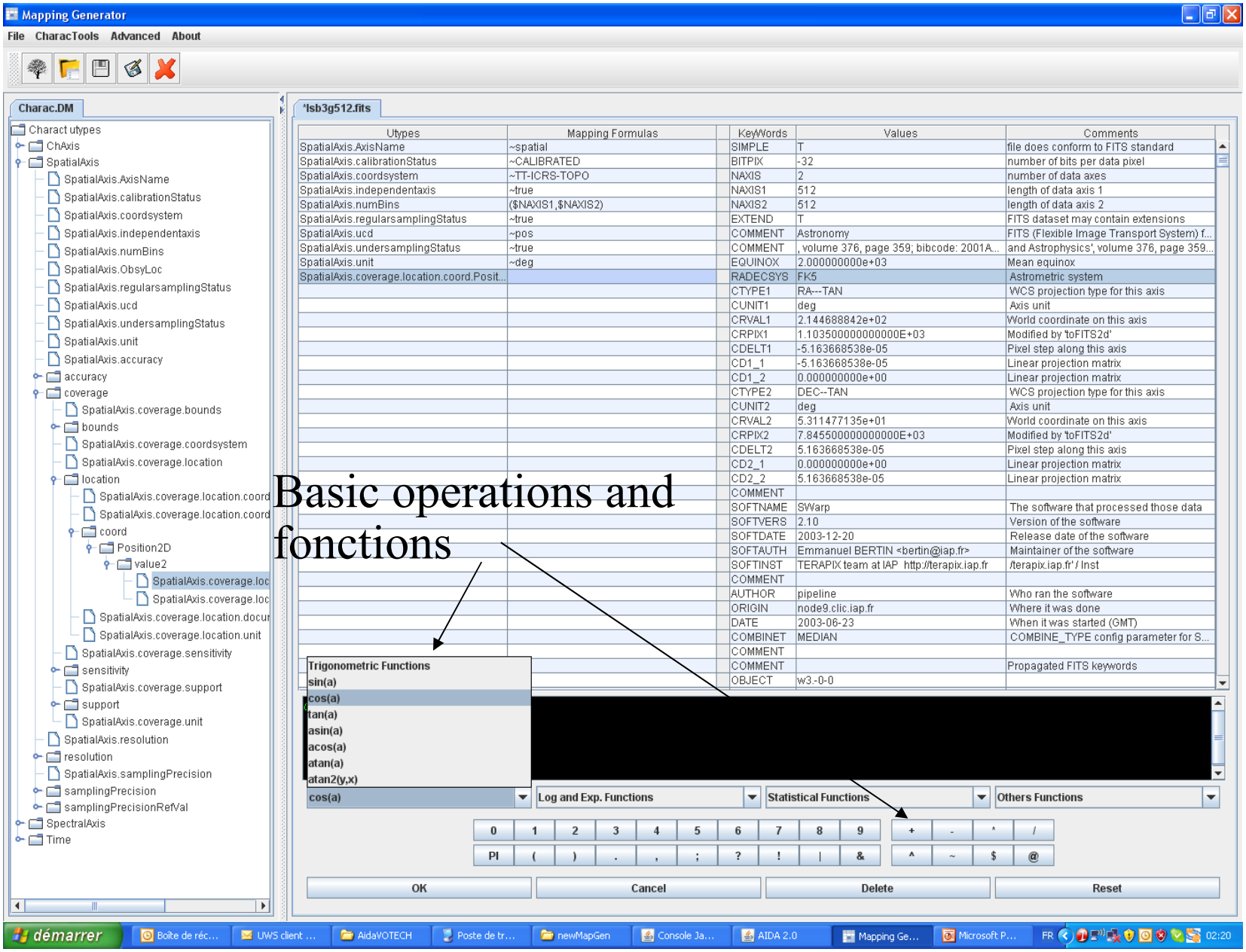
```
AxisShortcut SPATIAL: Axis[ucd="pos"]
AxisShortcut SPECTRAL: Axis[ucd="em"]
AxisShortcut FLUX: Axis[independentAxis="false"]

IF NOT(EQUAL(:SPATIAL.Coverage.location.coord_system_id)) OR NOT (EQUAL(:SPATIAL.coordsystem))
        warning ("Input images do not have identical coordinate systems")
ELSE
#0.check units on all spatial axis
 IF (EXISTS(:SPATIAL.Coverage))
        EQUAL(:SPATIAL.Coverage.location.unit) OR
        EQUAL(:SPATIAL.Coverage.unit) OR
        EQUAL(:SPATIAL.unit) )

# 1. check for same resolution or sampling among input images
IF (EXISTS(:SPATIAL.Resolution))
        NEAR(:SPATIAL.Resolution.resolutionRefVal.Resolution, 0.3)
#       NEAR(:SPATIAL.Resolution.resolutionRefVal.Resolution2.C2, 0.3)
        EQUAL(:SPATIAL.Resolution.unit) OR EQUAL(:SPATIAL.unit)
ELSE
        NEAR(:SPATIAL.SamplingPrecision.samplingPrecisionRefVal.samplingPeriod.C1, 0.3)
        NEAR(:SPATIAL.SamplingPrecision.samplingPrecisionRefVal.samplingPeriod.C2, 0.3)
        EQUAL(:SPATIAL.SamplingPrecision.unit) OR EQUAL(:SPATIAL.unit)
FI

# 2. Check for identical sizes for images
IF (EXISTS(:SPATIAL.numbins))
        EQUAL(:SPATIAL.numbins)
ELSIF (EXISTS(:SPATIAL.numbins2))
        EQUAL(:SPATIAL.numbins2.i1) AND EQUAL(:SPATIAL.numbins2.i2)
ELSIF (EXISTS(:SPATIAL.numbins3))
        EQUAL(:SPATIAL.numbins3.i1) AND EQUAL(:SPATIAL.numbins3.i2) AND
        EQUAL(:SPATIAL.numbins3.i3)
ELSE
        ERROR( "Input images do not have the same sizes" )
FI
```

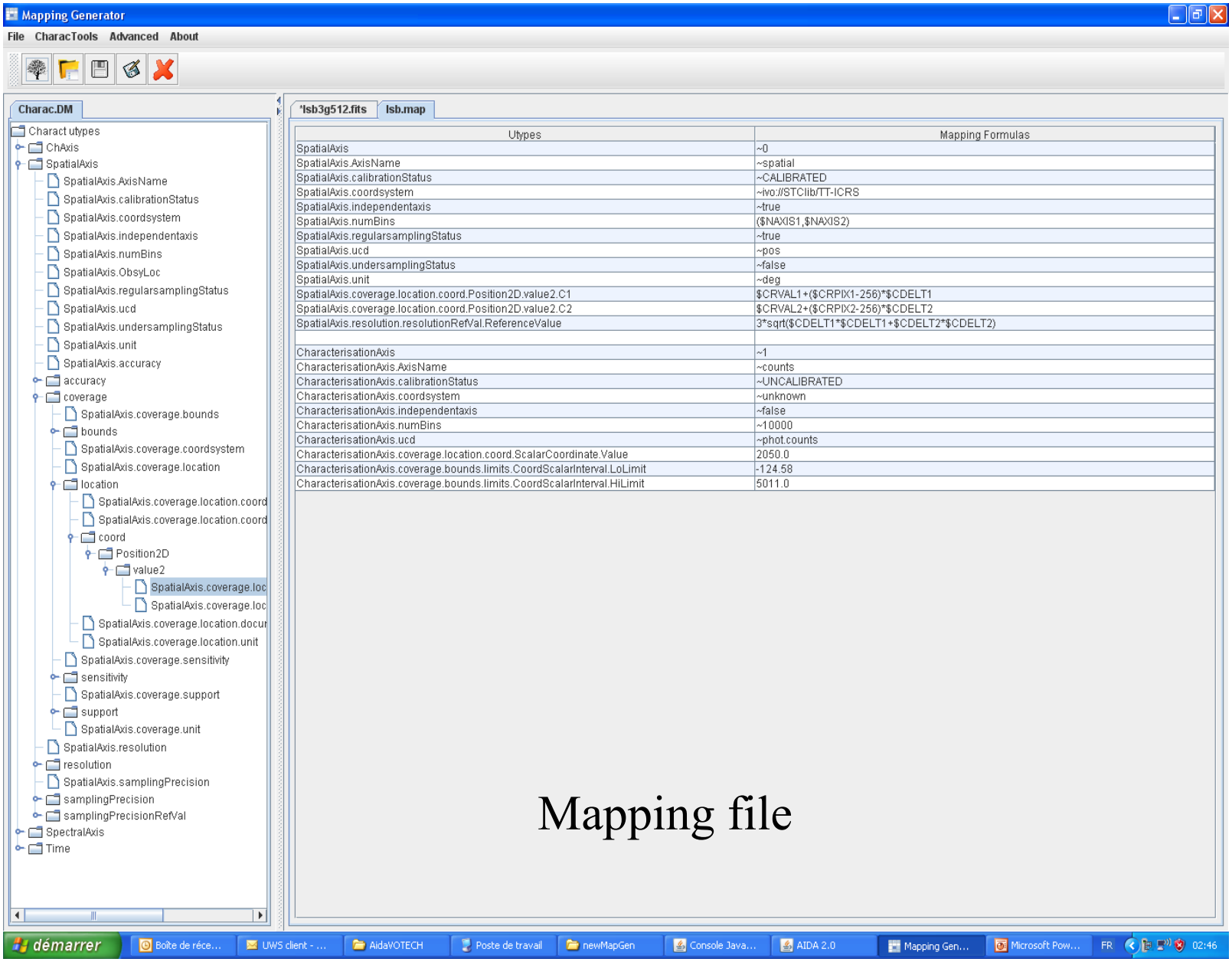
```
....
# 3. check that all input images overlap
EQUAL(:SPATIAL.Coverage.location.unit) OR EQUAL(:SPATIAL.Coverage.unit) OR EQUAL(:SPATIAL.unit)
EQUAL(:SPATIAL.coordsystem)
EQUAL(:SPATIAL.Coverage.location.coord)

# 4. Check for Observable range
EQUAL(1[]:FLUX.coverage.bounds.unit) OR EQUAL(:SPATIAL.Coverage.unit) OR EQUAL(:SPATIAL.unit)
IF (1[]:FLUX.bounds.limitHi - 1[]:FLUX.bounds.limitLo >= 100)
        WARNING("(Observables: max-min <100) Normalisation required before processing ")
FI

# 5.Check for spectral bounds overlap between input images
IF (1[1]:SPECTRAL.bounds.limitHi >1[2]:SPECTRAL.bounds.limitLo ) OR
 (1[2]:SPECTRAL.bounds.limitHi >1[1]:SPECTRAL.bounds.limitLo )
        WARNING("(Waveband overlap between input images : check for correlation ", Normalisation required before processing ")
FI

#EQUAL(:FLUX.ucd)
#1[]:FLUX.bounds.extent < 100

# MASK: Verifier que le mask est donne en entree:
IF (EXISTS(2))
        WARNING("Coucou !!! La deuxieme entre (le masque) est fourni !")
        # 1. ObservableAxis: min=0, max=1:
        2:FLUX.ucd = "meta.code.class"
        2:FLUX.unit = "unitless"
        2:FLUX.coverage.numbins = 2 AND 2:FLUX.bounds.limitLo = 0 AND 2:FLUX.bounds.limitHi >= 0

        # 2. SpatialAxis: numbins = numbins INPUT:
        # line 8 !
FI
```

- **Mapping Generator**
  - **Developed in the frame of VOTECH**


Mapping Generator
File CharacTools Advanced About
Charac.DM
Charact utypes
ChAxis
SpatialAxis
SpatialAxis.AxisName
SpatialAxis.calibrationStatus
SpatialAxis.coordsystem
SpatialAxis.independantaxis
SpatialAxis.numBins
SpatialAxis.ObsyLoc
SpatialAxis.regularsamplingStatus
SpatialAxis.ucd
SpatialAxis.undersamplingStatus
SpatialAxis.unit
SpatialAxis.accuracy
accuracy
coverage
SpatialAxis.resolution
resolution
SpatialAxis.samplingPrecision
samplingPrecision
samplingPrecisionRefVal
SpectralAxis
Time
utype
Constant
'lsb3g512.fits
Utypes | Mapping Formulas | KeyWords | Values | Comments
SpatialAxis.AxisName | ~spatial | SIMPLE | T | file does conform to FITS standard
SpatialAxis.calibrationStatus | ~CALIBRATED | BITPIX | -32 | number of bits per data pixel
SpatialAxis.coordsystem | | NAXIS | 2 | number of data axes
NAXIS1 | 512 | length of data axis 1
NAXIS2 | 512 | length of data axis 2
EXTEND | T | FITS dataset may contain extensions
COMMENT | Astronomy | FITS (Flexible Image Transport System) f...
COMMENT | , volume 376, page 359; bibcode: 2001A... | and Astrophysics', volume 376, page 359...
EQUINOX | 2.000000000e+03 | Mean equinox
RADECSYS | FK5 | Astrometric system
CTYPE1 | RA---TAN | WCS projection type for this axis
CUNIT1 | deg | Axis unit
CRVAL1 | 2.144688842e+02 | World coordinate on this axis
CRPIX1 | 1.1035000000000000E+03 | Modified by 'toFITS2d'
CDELT1 | -5.163668538e-05 | Pixel step along this axis
CD1_1 | -5.163668538e-05 | Linear projection matrix
CD1_2 | 0.000000000e+00 | Linear projection matrix
CTYPE2 | DEC--TAN | WCS projection type for this axis
CUNIT2 | deg | Axis unit
CRVAL2 | 5.311477135e+01 | World coordinate on this axis
CRPIX2 | 7.8455000000000000E+03 | Modified by 'toFITS2d'
CDELT2 | 5.163668538e-05 | Pixel step along this axis
CD2_1 | 0.000000000e+00 | Linear projection matrix
CD2_2 | 5.163668538e-05 | Linear projection matrix
COMMENT
SOFTNAME | SWarp | The software that processed those data
SOFTVERS | 2.10 | Version of the software
SOFTDATE | 2003-12-20 | Release date of the software
SOFTAUTH | Emmanuel BERTIN <bertin@iap.fr> | Maintainer of the software
SOFTINST | TERAPIX team at IAP http://terapix.iap.fr | /terapix.iap.fr' / Inst
COMMENT
AUTHOR | pipeline | Who ran the software
ORIGIN | node9.clic.iap.fr | Where it was done
DATE | 2003-06-23 | When it was started (GMT)
COMBINET | MEDIAN | COMBINE_TYPE config parameter for S...
COMMENT
COMMENT | | Propagated FITS keywords
OBJECT | w3.-0-0
~TT-ICRS-TOPO
Trigonometric Functions
Log and Exp. Functions
Statistical Functions
Others Functions
0 1 2 3 4 5 6 7 8 9 + - * /
PI ( ) . , ; ? ! | & ^ ~ $ @
OK
Cancel
Delete
Reset
démarrer
Boîte de récep...
UWS client - M...
AidaVOTECH
Poste de travail
newMapGen
Console Java -...
AIDA 2.0
Mapping Gener...
FR 02:03


Mapping Generator
File CharacTools Advanced About
Charac.DM
Charact utypes
ChAxis
SpatialAxis
SpatialAxis.AxisName
SpatialAxis.calibrationStatus
SpatialAxis.coordsystem
SpatialAxis.independentaxis
SpatialAxis.numBins
SpatialAxis.ObsyLoc
SpatialAxis.regularsamplingStatus
SpatialAxis.ucd
SpatialAxis.undersamplingStatus
SpatialAxis.unit
SpatialAxis.accuracy
accuracy
coverage
SpatialAxis.resolution
resolution
SpatialAxis.samplingPrecision
samplingPrecision
samplingPrecisionRefVal
SpectralAxis
Time
'lsb3g512.fits
Utypes
Mapping Formulas
KeyWords
Values
Comments
SpatialAxis.AxisName ~spatial
SpatialAxis.calibrationStatus ~CALIBRATED
SpatialAxis.coordsystem ~TT-ICRS-TOPO
SpatialAxis.independentaxis ~true
SpatialAxis.numBins
SIMPLE T file does conform to FITS standard
BITPIX -32 number of bits per data pixel
NAXIS 2 number of data axes
NAXIS1 512 length of data axis 1
NAXIS2 512 length of data axis 2
EXTEND T FITS dataset may contain extensions
COMMENT Astronomy FITS (Flexible Image Transport System) f...
COMMENT , volume 376, page 359; bibcode: 2001A... and Astrophysics', volume 376, page 359...
EQUINOX 2.000000000e+03 Mean equinox
RADECSYS FK5 Astrometric system
CTYPE1 RA---TAN WCS projection type for this axis
CUNIT1 deg Axis unit
CRVAL1 2.144688842e+02 World coordinate on this axis
CRPIX1 1.103500000000000E+03 Modified by 'toFITS2d'
CDELT1 -5.163668538e-05 Pixel step along this axis
CD1_1 -5.163668538e-05 Linear projection matrix
CD1_2 0.000000000e+00 Linear projection matrix
CTYPE2 DEC--TAN WCS projection type for this axis
CUNIT2 deg Axis unit
CRVAL2 5.311477135e+01 World coordinate on this axis
CRPIX2 7.845500000000000E+03 Modified by 'toFITS2d'
CDELT2 5.163668538e-05 Pixel step along this axis
CD2_1 0.000000000e+00 Linear projection matrix
CD2_2 5.163668538e-05 Linear projection matrix
COMMENT
SOFTNAME SWarp The software that processed those data
SOFTVERS 2.10 Version of the software
SOFTDATE 2003-12-20 Release date of the software
SOFTAUTH Emmanuel BERTIN <bertin@iap.fr> Maintainer of the software
SOFTINST TERAPIX team at IAP http://terapix.iap.fr /terapix.iap.fr' / Inst
COMMENT
AUTHOR pipeline Who ran the software
ORIGIN node9.clic.iap.fr Where it was done
DATE 2003-06-23 When it was started (GMT)
COMBINET MEDIAN COMBINE_TYPE config parameter for S...
COMMENT
COMMENT Propagated FITS keywords
OBJECT w3.-0-0
Utype Picking up some keywords values
[$NAXIS1,$NAXIS2
Trigonometric Functions
Log and Exp. Functions
Statistical Functions
Others Functions
0 1 2 3 4 5 6 7 8 9 + - * /
PI ( ) . , ; ? ! | & ^ ~ $ @
OK
Cancel
Delete
Reset
démarrer
Boîte de réc...
UWS client ...
AidaVOTECH
Poste de tr...
newMapGen
Console Ja...
AIDA 2.0
Mapping Ge...
Microsoft P...
FR
02:06


Mapping Generator
File CharacTools Advanced About
Charac.DM
Charact utypes
ChAxis
SpatialAxis
SpatialAxis.AxisName
SpatialAxis.calibrationStatus
SpatialAxis.coordsystem
SpatialAxis.independentaxis
SpatialAxis.numBins
SpatialAxis.ObsyLoc
SpatialAxis.regularsamplingStatus
SpatialAxis.ucd
SpatialAxis.undersamplingStatus
SpatialAxis.unit
SpatialAxis.accuracy
accuracy
coverage
SpatialAxis.coverage.bounds
bounds
SpatialAxis.coverage.coordsystem
SpatialAxis.coverage.location
location
coord
Position2D
value2
SpatialAxis.coverage.sensitivity
sensitivity
SpatialAxis.coverage.support
support
SpatialAxis.coverage.unit
SpatialAxis.resolution
resolution
SpatialAxis.samplingPrecision
samplingPrecision
samplingPrecisionRefVal
SpectralAxis
Time
*lsb3g512.fits
Utypes
Mapping Formulas
KeyWords
Values
Comments
SpatialAxis.AxisName ~spatial SIMPLE T file does conform to FITS standard
SpatialAxis.calibrationStatus ~CALIBRATED BITPIX -32 number of bits per data pixel
SpatialAxis.coordsystem ~TT-ICRS-TOPO NAXIS 2 number of data axes
SpatialAxis.independentaxis ~true NAXIS1 512 length of data axis 1
SpatialAxis.numBins ($NAXIS1,$NAXIS2) NAXIS2 512 length of data axis 2
SpatialAxis.regularsamplingStatus ~true EXTEND T FITS dataset may contain extensions
SpatialAxis.ucd ~pos COMMENT Astronomy FITS (Flexible Image Transport System) f...
SpatialAxis.undersamplingStatus ~true COMMENT , volume 376, page 359; bibcode: 2001A... and Astrophysics', volume 376, page 359...
SpatialAxis.unit ~deg EQUINOX 2.000000000e+03 Mean equinox
SpatialAxis.coverage.location.coord.Posit... RADECSYS FK5 Astrometric system
CTYPE1 RA---TAN WCS projection type for this axis
CUNIT1 deg Axis unit
CRVAL1 2.144688842e+02 World coordinate on this axis
CRPIX1 1.1035000000000000E+03 Modified by 'toFITS2d'
CDELT1 -5.163668538e-05 Pixel step along this axis
CD1_1 -5.163668538e-05 Linear projection matrix
CD1_2 0.000000000e+00 Linear projection matrix
CTYPE2 DEC--TAN WCS projection type for this axis
CUNIT2 deg Axis unit
CRVAL2 5.311477135e+01 World coordinate on this axis
CRPIX2 7.8455000000000000E+03 Modified by 'toFITS2d'
CDELT2 5.163668538e-05 Pixel step along this axis
CD2_1 0.000000000e+00 Linear projection matrix
CD2_2 5.163668538e-05 Linear projection matrix
COMMENT
SOFTNAME SWarp The software that processed those data
SOFTVERS 2.10 Version of the software
SOFTDATE 2003-12-20 Release date of the software
SOFTAUTH Emmanuel BERTIN <bertin@iap.fr> Maintainer of the software
SOFTINST TERAPIX team at IAP http://terapix.iap.fr /terapix.iap.fr' / Inst
COMMENT
AUTHOR pipeline Who ran the software
ORIGIN node9.clic.iap.fr Where it was done
DATE 2003-06-23 When it was started (GMT)
COMBINET MEDIAN COMBINE_TYPE config parameter for S...
COMMENT
COMMENT Propagated FITS keywords
OBJECT w3.-0-0
Basic operations and fonctions
Trigonometric Functions
sin(a)
cos(a)
tan(a)
asin(a)
acos(a)
atan(a)
atan2(y,x)
cos(a)
Log and Exp. Functions
Statistical Functions
Others Functions
0 1 2 3 4 5 6 7 8 9 + - * /
PI ( ) . , ; ? ! | & ^ ~ $ @
OK
Cancel
Delete
Reset
démarrer
Boîte de réc...
UWS client ...
AidaVOTECH
Poste de tr...
newMapGen
Console Ja...
AIDA 2.0
Mapping Ge...
Microsoft P...
FR
02:20


Mapping Generator
File CharacTools Advanced About
Charac.DM
Charact utypes
ChAxis
SpatialAxis
SpatialAxis.AxisName
SpatialAxis.calibrationStatus
SpatialAxis.coordsystem
SpatialAxis.independentaxis
SpatialAxis.numBins
SpatialAxis.ObsyLoc
SpatialAxis.regularsamplingStatus
SpatialAxis.ucd
SpatialAxis.undersamplingStatus
SpatialAxis.unit
SpatialAxis.accuracy
accuracy
coverage
SpatialAxis.coverage.bounds
bounds
SpatialAxis.coverage.coordsystem
SpatialAxis.coverage.location
location
SpatialAxis.coverage.location.coord
SpatialAxis.coverage.location.coord
coord
Position2D
value2
SpatialAxis.coverage.loc
SpatialAxis.coverage.loc
SpatialAxis.coverage.location.docur
SpatialAxis.coverage.location.unit
SpatialAxis.coverage.sensitivity
sensitivity
SpatialAxis.coverage.support
support
SpatialAxis.coverage.unit
SpatialAxis.resolution
resolution
SpatialAxis.samplingPrecision
samplingPrecision
samplingPrecisionRefVal
SpectralAxis
Time
'lsb3g512.fits
lsb.map
Utypes
Mapping Formulas
SpatialAxis ~0
SpatialAxis.AxisName ~spatial
SpatialAxis.calibrationStatus ~CALIBRATED
SpatialAxis.coordsystem ~ivo://STClib/TT-ICRS
SpatialAxis.independentaxis ~true
SpatialAxis.numBins ($NAXIS1,$NAXIS2)
SpatialAxis.regularsamplingStatus ~true
SpatialAxis.ucd ~pos
SpatialAxis.undersamplingStatus ~false
SpatialAxis.unit ~deg
SpatialAxis.coverage.location.coord.Position2D.value2.C1 $CRVAL1+($CRPIX1-256)*$CDELT1
SpatialAxis.coverage.location.coord.Position2D.value2.C2 $CRVAL2+($CRPIX2-256)*$CDELT2
SpatialAxis.resolution.resolutionRefVal.ReferenceValue 3*sqrt($CDELT1*$CDELT1+$CDELT2*$CDELT2)
CharacterisationAxis ~1
CharacterisationAxis.AxisName ~counts
CharacterisationAxis.calibrationStatus ~UNCALIBRATED
CharacterisationAxis.coordsystem ~unknown
CharacterisationAxis.independentaxis ~false
CharacterisationAxis.numBins ~10000
CharacterisationAxis.ucd ~phot.counts
CharacterisationAxis.coverage.location.coord.ScalarCoordinate.Value 2050.0
CharacterisationAxis.coverage.bounds.limits.CoordScalarInterval.LoLimit -124.58
CharacterisationAxis.coverage.bounds.limits.CoordScalarInterval.HiLimit 5011.0
Mapping file
démarrer
Boite de réce...
UWS client - ...
AidaVOTECH
Poste de travail
newMapGen
Console Java...
AIDA 2.0
Mapping Gen...
Microsoft Pow...
FR
02:46

We need a mapping (FITS → utypes)

- **003.7858-39.2202.fits + MappingSpecificAxis.map ----> 003.7858-39.2202.uty**

```
%CharacterisationAxis 1
%SpatialAxis.AxisName spatial
%SpatialAxis.independentaxis TRUE
%SpatialAxis.calibrationStatus CALIBRATED
%SpatialAxis.samplingPrecision.samplingPrecisionRefVal.samlingPeriod -0.000277777784317036
-0.000277777784317036
%SpatialAxis.coverage.bounds.limits.Coord2VecInterval.LoLimit2Vec 3.872320772806-39.08143766442968
%SpatialAxis.unit deg
%SpatialAxis.undersamplingStatus FALSE
%SpatialAxis.coordsystem FK5
%SpatialAxis.accuracy.statError.ErrorRefval.ErrorRefValue Unknown
%SpatialAxis.resolution.resolutionRefVal Unknown
%SpatialAxis.ucd pos
%SpatialAxis.numBins 512 1024
%SpatialAxis.regularsamplingStatus TRUE
%SpatialAxis.coverage.bounds.limits.Coord2VecInterval.HiLimit2Vec 3.762143519194-39.36588211557032
%SpatialAxis.accuracy.sysError.ErrorRefval.ErrorRefValue Unknown
%SpatialAxis.coverage.location.coord.Position2D.Value2.C1 3.8172321
%SpatialAxis.coverage.location.coord.Position2D.Value2.C2 -39.223659890

%CharacterisationAxis 2
%TimeAxis.AxisName time
%TimeAxis.coordsystem TT-ICRS-WAVELENGTH-TOPO
%TimeAxis.undersamplingStatus TRUE
%TimeAxis.numBins 1
%TimeAxis.accuracy.satatError.ErrorRefVal.ErrorRefValue Unknown
%TimeAxis.resolution.resolution.resolutionRefVal Unknown
...
```

- **Characterization library (B. GASSMANN) is used to convert this format to an XML file**

- **Demo**

Summary

- **Done**
  - **Definition of workflow use cases with Characterized image entries**
  - **Definition of a constraint language and integration in a Workflow test bed (AÏDA 2.0), definition of constraint files for the use cases**
- **Ongoing work**
  - **Increase the validation scope for less human interaction**
    - During the execution : finalize the Characterization file generation for the FITS outputs
    - Before the execution : study how to define a “virtual” Characterization file for an output before the execution...