

Long-term management of 1000s of All-Sky reference data sets using the HiPS network

ADASS – October 2016 - Trieste

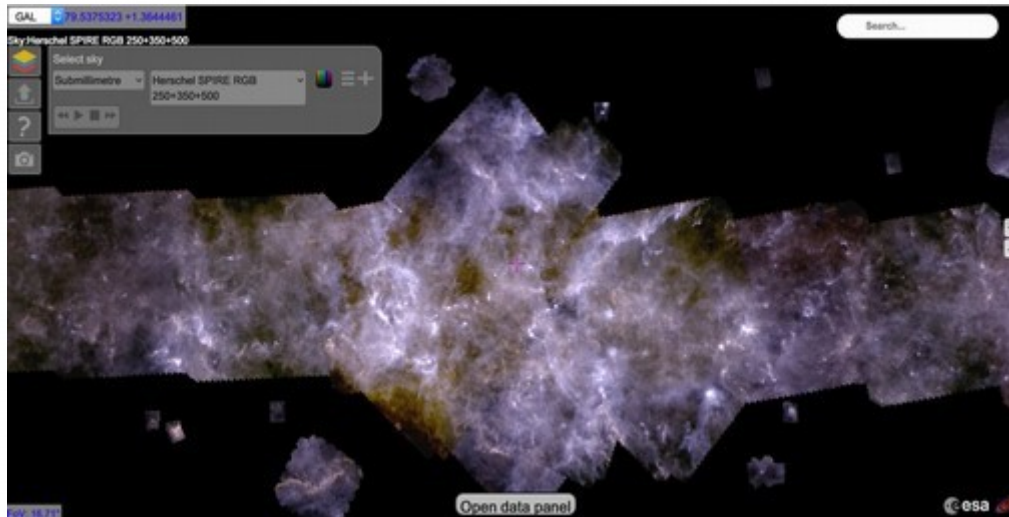
P.Fernique, T.Boch, A. Oberto,
M. Allen, D. Durand, K. Ebisawa,
B. Merin, J. Salgado



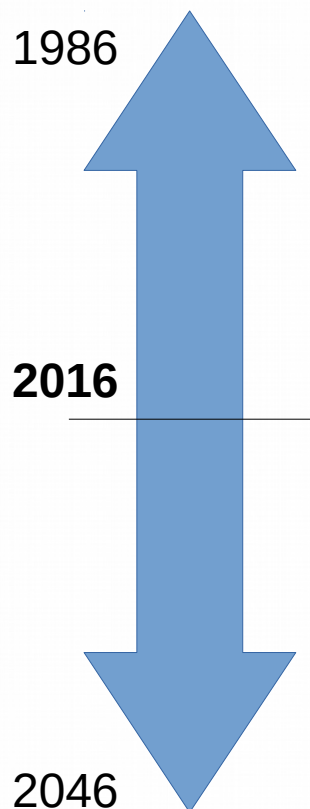
□ What's the plan ?

1) What does “**long term**” mean ?

2) Why **HiPS** could be **a good candidate** for long term management ?



□ Long term ?



- Which 30 years old files are you able to load today on your laptop ?

*FM2 ? TGA ? DOC ? GIF ? COD ? Fits ? JPEG ?
ASCII ? JPEG2000 ? HDF4 ? PDF ? PCT ? CR1 ?
PDS ? XML ? PPM, CDR ? WKS ? ...*

- Which existing files will be still **loadable/usable in 30 years** ?

□ The long term problems

- 1) Not just a data storage issue
=> The goal: **use the data** as long as possible
- 2) In fact, a couple of issues: **Data + Metadata**
- 3) Observation: “***More the data+metadata are easy to read/understand, greater is their life time***” (ex: ASCII tables)
- 4) A basic metric: **How hard is to (re)write the data reader ?**
(indication: how depth is the documentation?).

□ Long term “enablers”

- 1) **Open solution** => proprietary solution can always disappear
- 2) **Basic & Standard coding** => avoid unusable or complex bit codings, dedicated compression and/or encryption
- 3) **Meta data easily readable, easy to associate to the data** => no binary, avoid subtle data model
- 4) **Avoid dependencies** to a medium, a DB, or other specific tools
- 5) **Extensible** => avoid package barriers such as data size or pixel depth limit (will force to a future format evolution)

□ Two opposit approaches

1) **Describe the data as they have been produced** (ex: PDS)

- => The best for scientific analysis,
- => Easier for data producers,
- => Difficult for the data readers (have to support all collections of data+metadata)

2) **Translate the data in a common basic package** (ex: FITS)

- => Impact on future scientific analysis,
- => Difficult for the producers (translation required),
- => Easier for the readers (only one package to support)

□ Illustration of these approaches

1) ...the data as they have been produced...(ex:PDS)

...The collection consists of Mariner 10 images ... file consists of 22 blocks containing 31,944 bytes per block. Each block is composed of 33 logical records of 968 bytes each. The first logical record of the first block contains a single set of VICAR label information. The tapes were created on an IBM system, so the VICAR Standard Format labels are written in EBCDIC. ...

2) ...the data in a common basic package...(ex:FITS)

*...FITS...consisted of **a binary array** preceded by **an ASCII text header** with information describing the organization and contents of the array ...*

□ Consequences...


PDS: ~2 readers,
1 lib, only unix

✖ *reply to post by BuzzDengue*

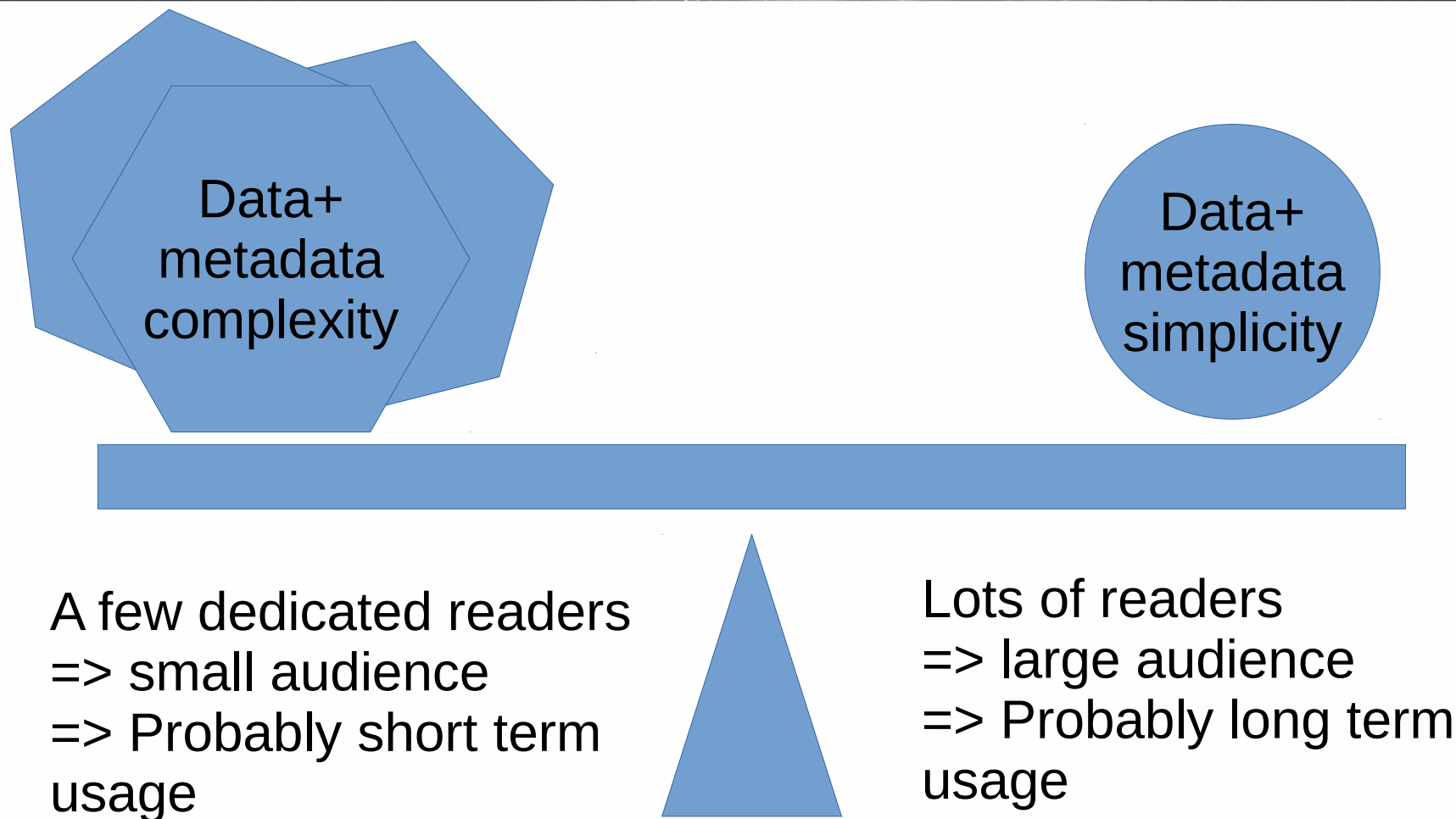
I couldn't find any program besides NASA View and the ISIS program (ISIS is a suite made up of dozens of programs) that can read PDS version 3 images. 😞

30 years
later...

FITS :
hundreds of
readers + libs,
multiple OS

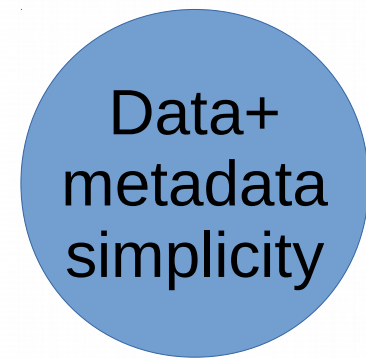
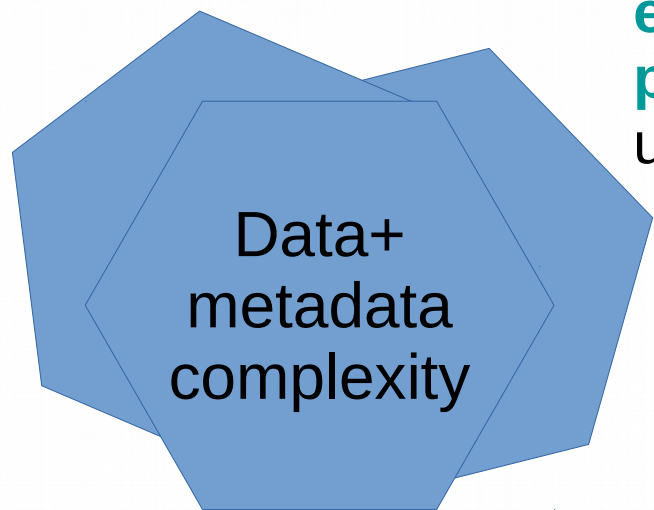
AristoArt	Windows 2000 through Windows X	139 Faces for version 1.0 (upgrade from version 4.0 - 10 years)	FITS 16, FITS, DPGC and others	"A complete software for image processing, photometry, astrometry, camera control and image stacking for digital and film images."	astroart.sourceforge.net
ARIS	Windows V, VISTA, XP, 2008, 64	Browsers version of AristoArt software above	FITS 16, 16A CCD-DEP file		aristofits.sourceforge.net
AS	Windows Vista/XP / Mac OS X  Microsoft (Private) Linux, Linux PPC, Solaris, SGI, Alpha EV6, IOP CTS	free to ASP check the FAQ	FITS	The Chandra X-ray Observatory Education site provides a connection between FITS software, Chandra data sets, and much-to-be-hoped-for use in teaching with Chandra data. You can download PDF for use interactively with Chandra data or the general one with any FITS data from any source.	cosmicbridge.ucr.edu
EELabLibraries	MacOS X for Windows OXP, Vista or Windows 7, 48 bit or 64 bit Min OS X 10.4 or later	free	FITS FIB	"To create color image using raw observations from a range of telescopes, including the NASA/ESA Hubble Space Telescope, NASA's Spitzer Space Telescope, ESA's Very Large Telescope and ESO's SMART-Survey Telescope"	eelsat.com
FITSViewer	Windows 2.0/MS/ME/SP/Vista Min: Unix/Linux	free	FITS		hottelstarcas.org
FITS_Frames	Min: OS X 10.4 or later, 64 bit architecture	\$4.99 available from the Apple app store	FITS		see also ChandraTools/FITS-Frames
FITS-GIMP	32 and 64 bit Windows, Vista, 7 and 8	ACT2007		"Displays FITS images, whenever image properties or thumbnails are shown, and allows full-sized image to be loaded into one of the latest Windows applications, like ImageJ, Adobe Photoshop, Corel Paint Shop Pro, Windows Photo Viewer, Mozilla Firefox and Microsoft Office components like Outlook, Word, Excel, PowerPoint, etc."	
FITSImageV1	Windows, Photoshop plugins for 13 through old	ACT2008		"Int., raw, and output B-, R-, and G-band images and 3-color colour FITS images"	
F.I.G.	Min: OS-X, (most Windows) and UNIX architectures	free	FITS file editor	"graphical program for viewing and editing FITS format images or tables"	fritzsky.caibn.ac.cn/en
ICCD	free	free	FITS, FITS, DPG, PGC, SHIP PCB	"converts the Starlight Express cosmic ray spots FITS files to the macintosh"	net2556.de
ImvizToolBox	Windows	free	FITS, DPG, FIT, SHIP, and Camera Rawpack	"A lightweight image viewer for FITS and other formats"	

□ The dilemma

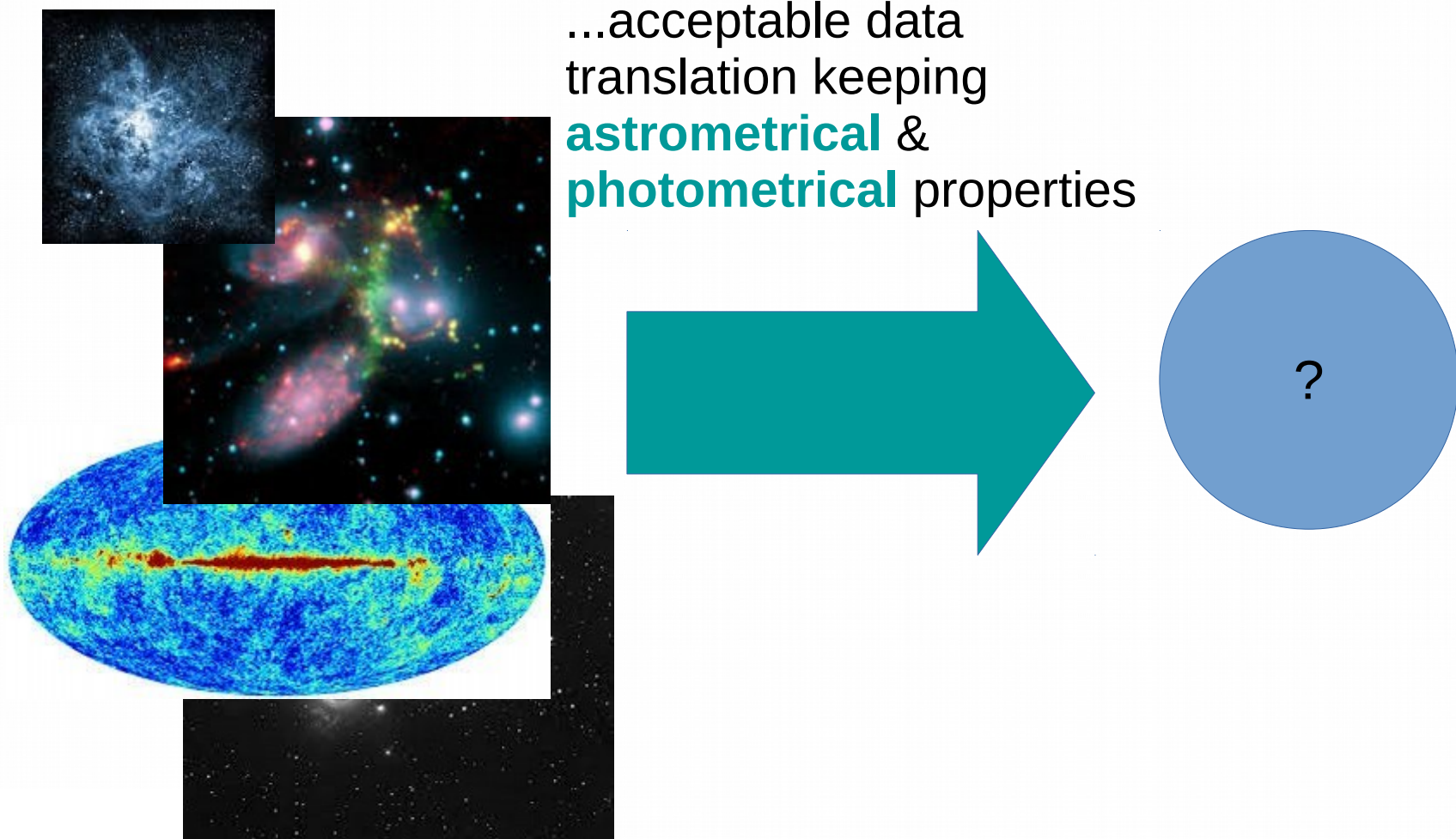


□ The method

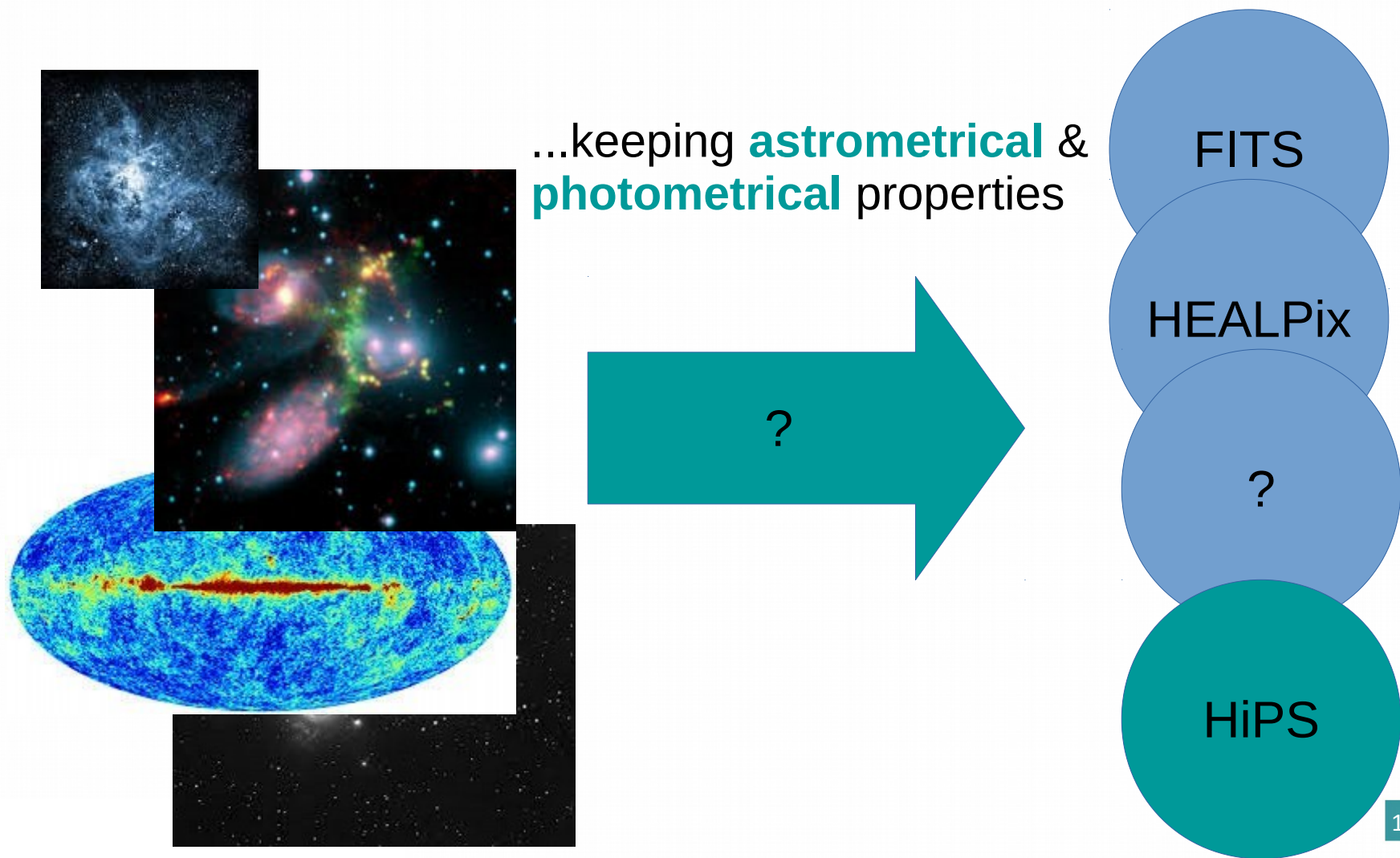
Find **an acceptable data translation keeping enough scientific properties** for future usages



□ The astronomical image context



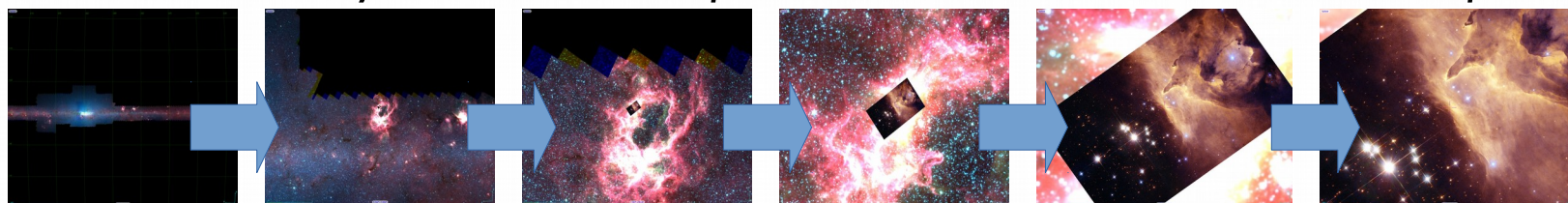
□ Which solution for astronomical images?



□ HiPS – What is it ?

Hierarchical Progressive Survey

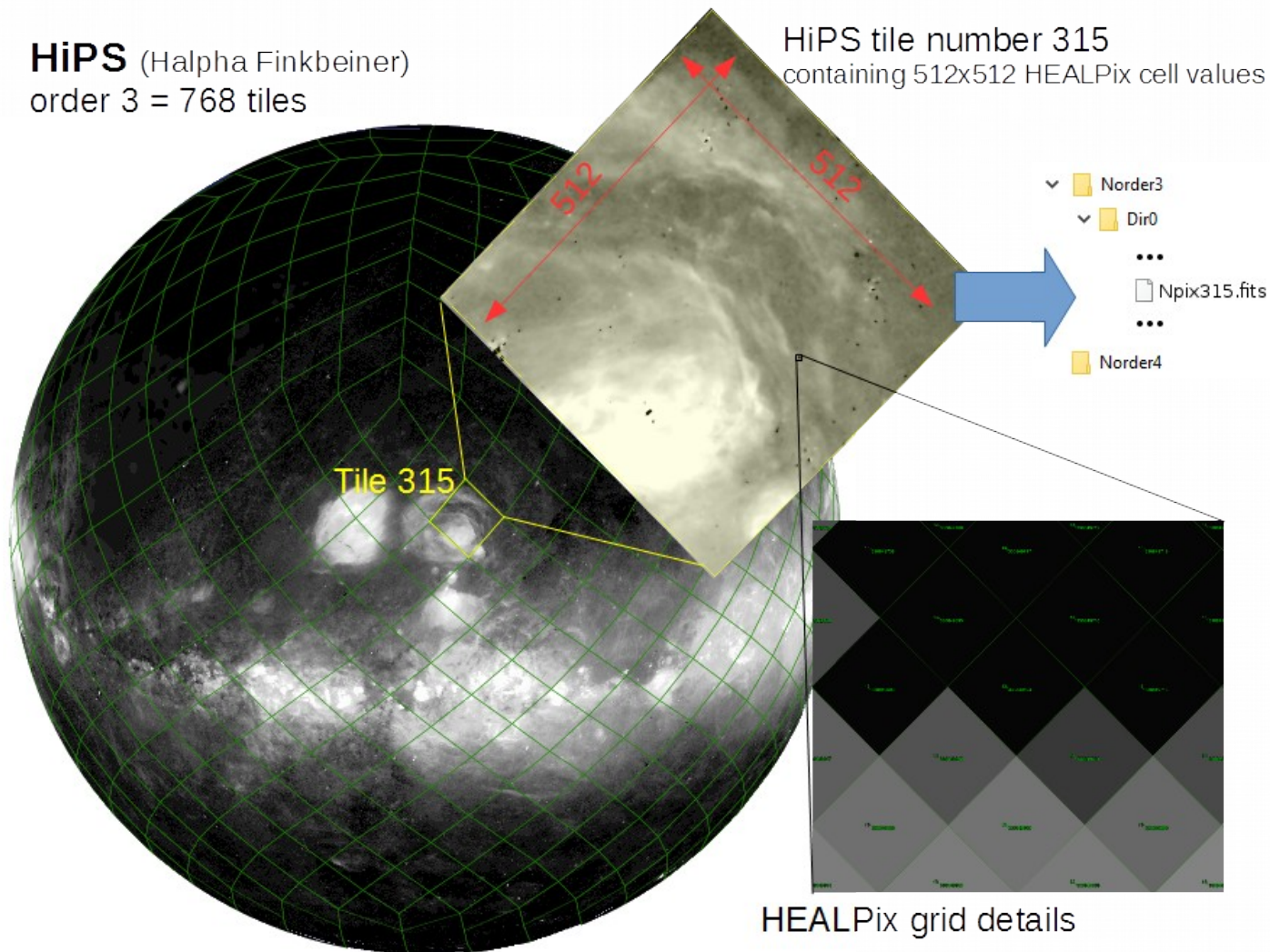
“The more you zoom in on a particular area, the more details show up”



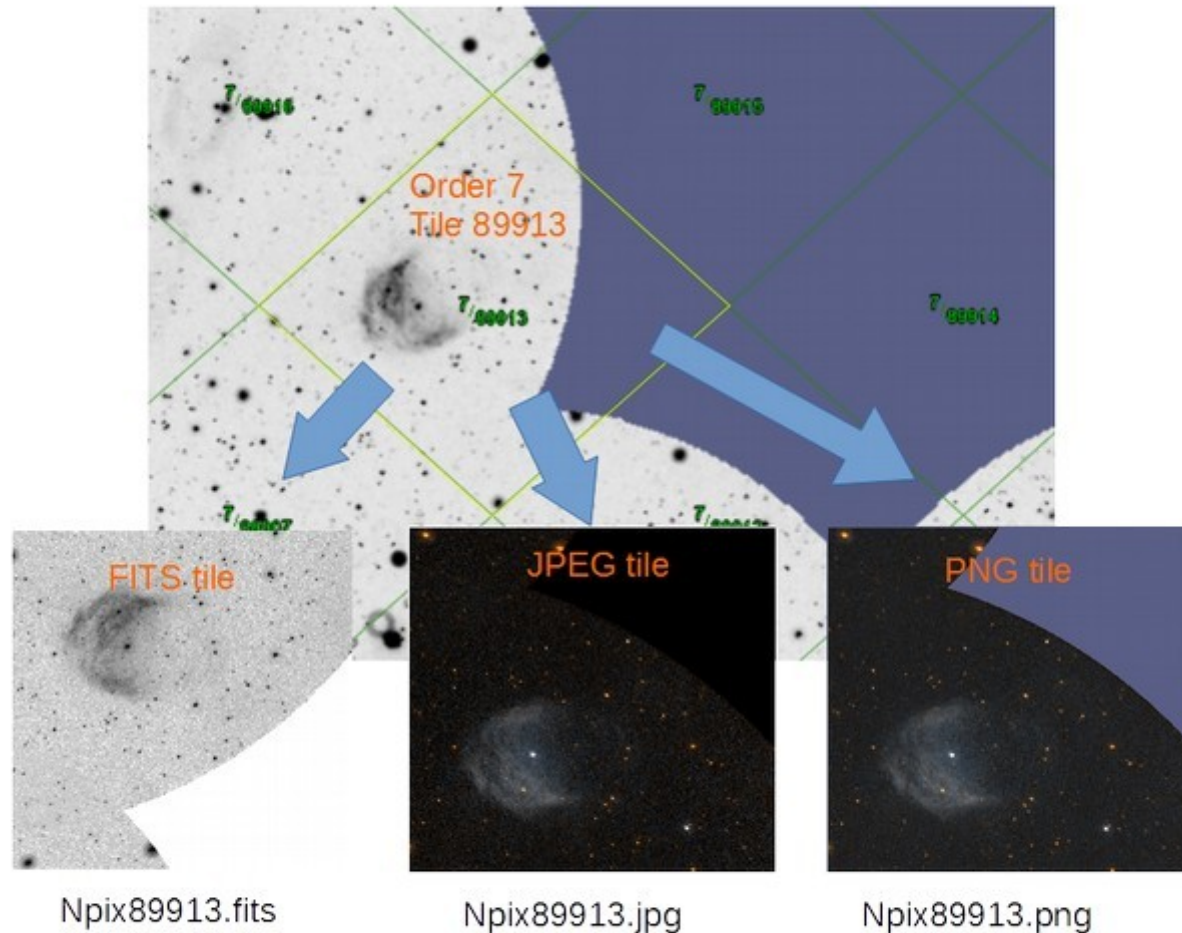
- Conserves **scientific data properties** (pixel localization & value)
- Based on the **HEALPix geometry**
- Dedicated to **any kind of image collections** (surveys, pointed observations, and more...)
- Stored as a **basic hierarchy of regular files**

- HiPS = Mosaic of HEALPix (spherical) tiles...
... stored in a hierarchy of files

HiPS (Halpfa Finkbeiner)
order 3 = 768 tiles



- Each tile/file can be coded in FITS, JPEG or PNG



□ State of art

(November 2016)

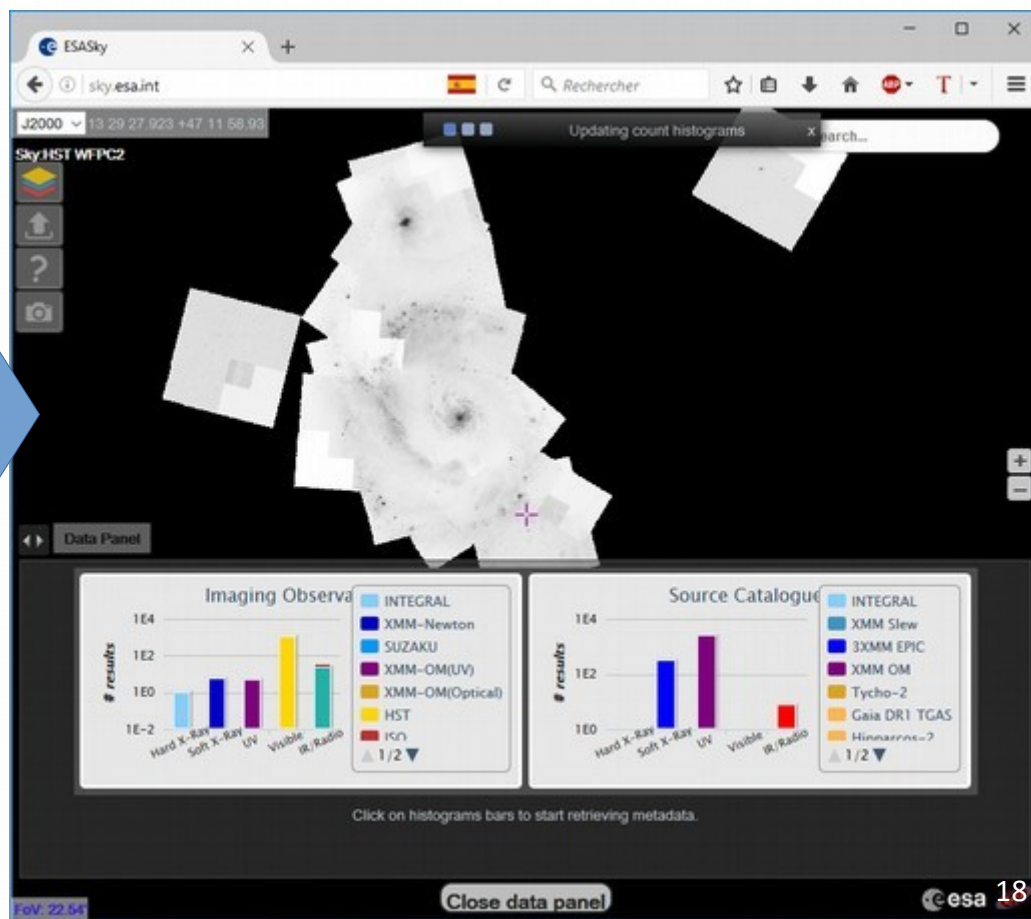
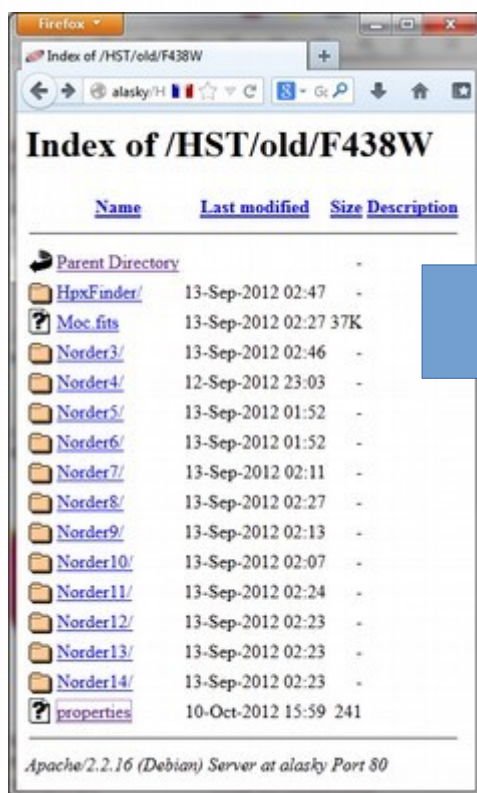
- **300+ HiPS** for **100TB** data
- **HiPS clients :**
 - Aladin Desktop (CDS), Aladin Lite (CDS), MIZAR (CNES)
 - + in dev: STScI portal (NASA), openWWT (Microsoft), proto (China), ...
 - + Aladin Lite implementation: ESAsky (ESAC), JUDO2 (JAXA), Skymap (LIGO/virgo) ...
 - + Aladin Lite web page inclusion: Simbad, Vizier, GLIMPSE360, CADE, ADS allsky, CASSIS, Akari-Viewer, VistaOrion, AstroDEEP, CDS portal, Gamma Sky, MOPRA, ...
 - + Aladin Desktop usage “diversion”: Arches walker
- **300 000+ HiPS tiles requested / day**

□ State of art (November 2016)

- **I2+** HiPS servers
 - CDS, SSC-XMM, IAS, IRAP/CADE, IPAC, ADS, ESAC, JAXA, AMIGA, Spanish-VO, Vista-Orion, TGSSADR...
- HiPS network infrastructure for registration & publication (see IVOA standard)
- **IVOA standard** (WD stage)
 - <http://www.ivoa.net/documents/HIPS/20160623>
- HiPS generator → Aladin/Hipsgen (perf: 10h/1Tpix),
- Paper → 2015A&A...578A.114F

□ Is HiPS a good candidate for long term usage?

- Today...



□ HiPS characteristics

- ✓ Astrometry: HEALPix based (400 μ arcsec)
- ✓ Photometry: FITS tiles HEALPix based
- ✓ Coding: basic files (FITS | JPEG | PNG, ASCII)
- ✓ Metadata: simple representation (key=value)
- ✓ Big data ready: hierarchical (\Rightarrow no limit)
- ✓ Open: IVOA standard
- ✓ Dependencies: HEALPix lib

□ The gamble...

- Tomorrow !

