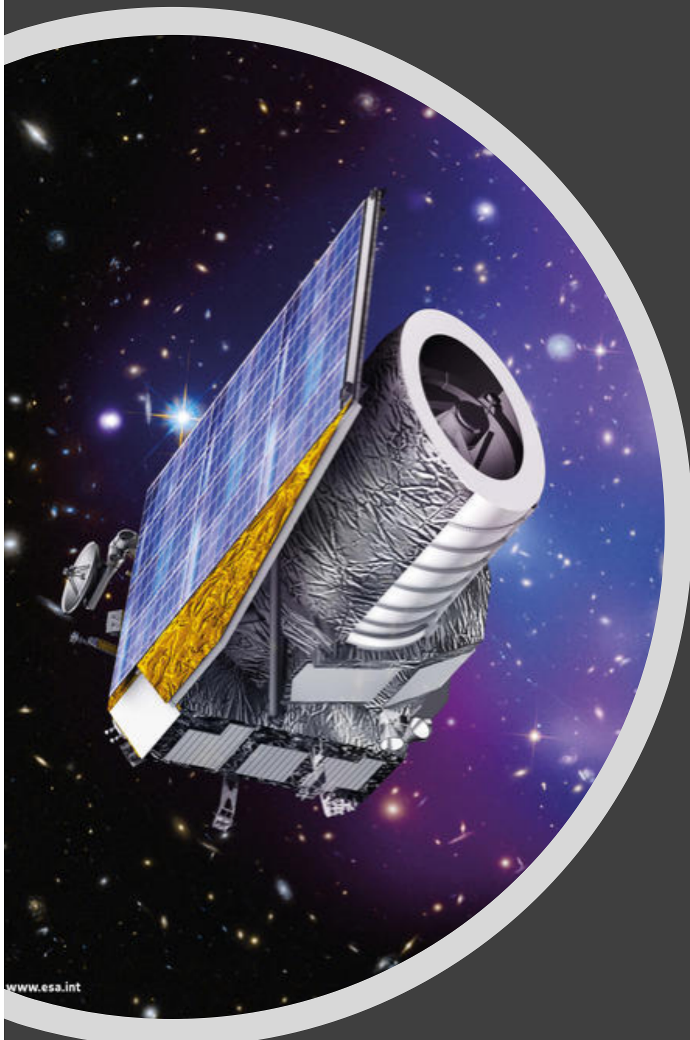
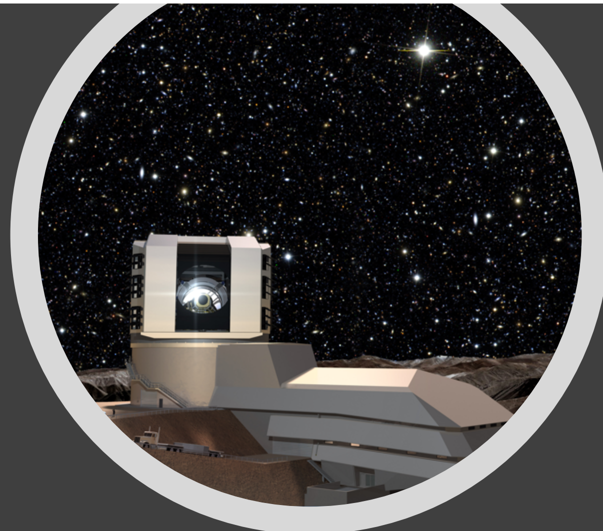
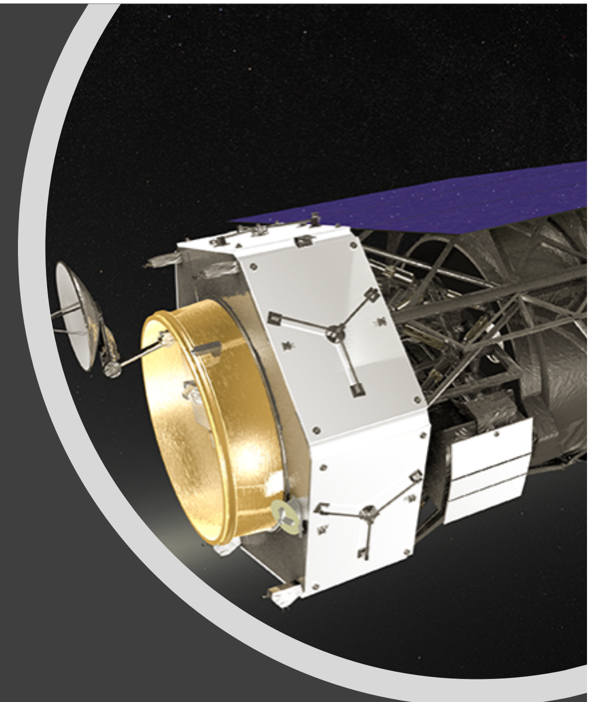
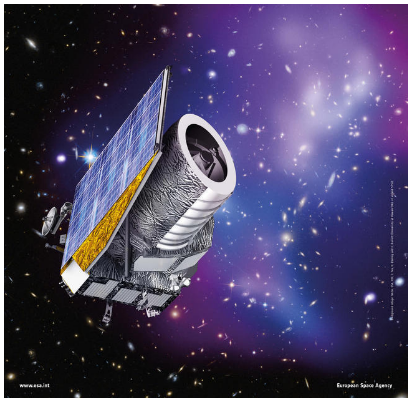




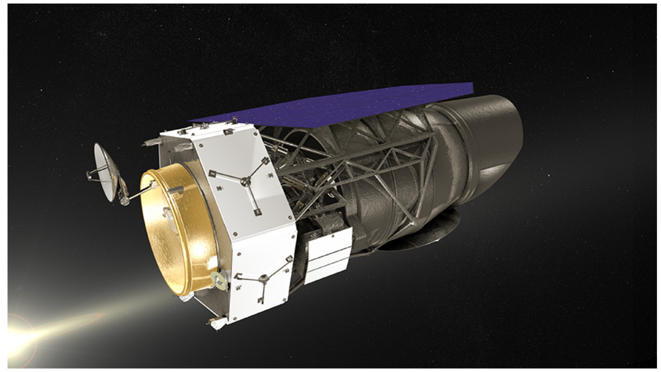


IPAC/Caltech led Tri-Agency Working Group

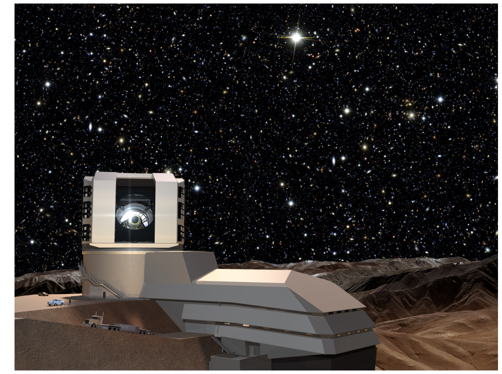
Euclid-LSST-WFIRST Joint Survey Processing



Euclid
1.2m telescope
ESA-led with NASA NIR detectors
2022+, Soyuz launch to L2
15000 deg^2 in VIS, y, J, H



WFIRST
2.4m telescope
NASA-led, Ball etc.
2026+,
2200 deg^2 in W,R,z,y,J,H,F814

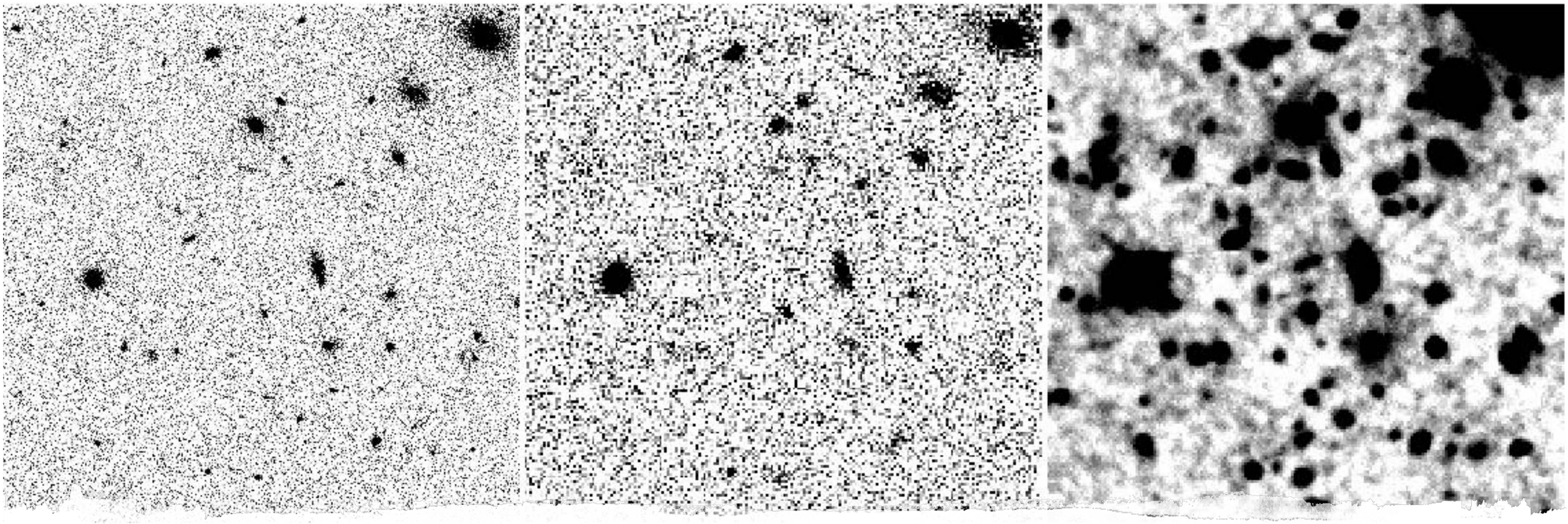


LSST
8.4m telescope in Chile
NSF-led
2022+,
20000 deg^2 in u,g,r,i,z,y

Joint Survey Processing >> Sum of Parts

Paradigm (Whole>Σparts)

- Well-documented set of standardized joint products
- Clear division of tasks between project, joint-processing and community and avoiding duplication
- Enables new multi-wavelength science while doing “old-science” better, reduces systematics
- Analysis and simulations are run on single copies of data located at the data processing institutions rather than moving (vast amounts) data between multiple institutions
- Provides a way for brain-power of projects to support community analysis – crucial when data-taking and processing is out of hands of end-user

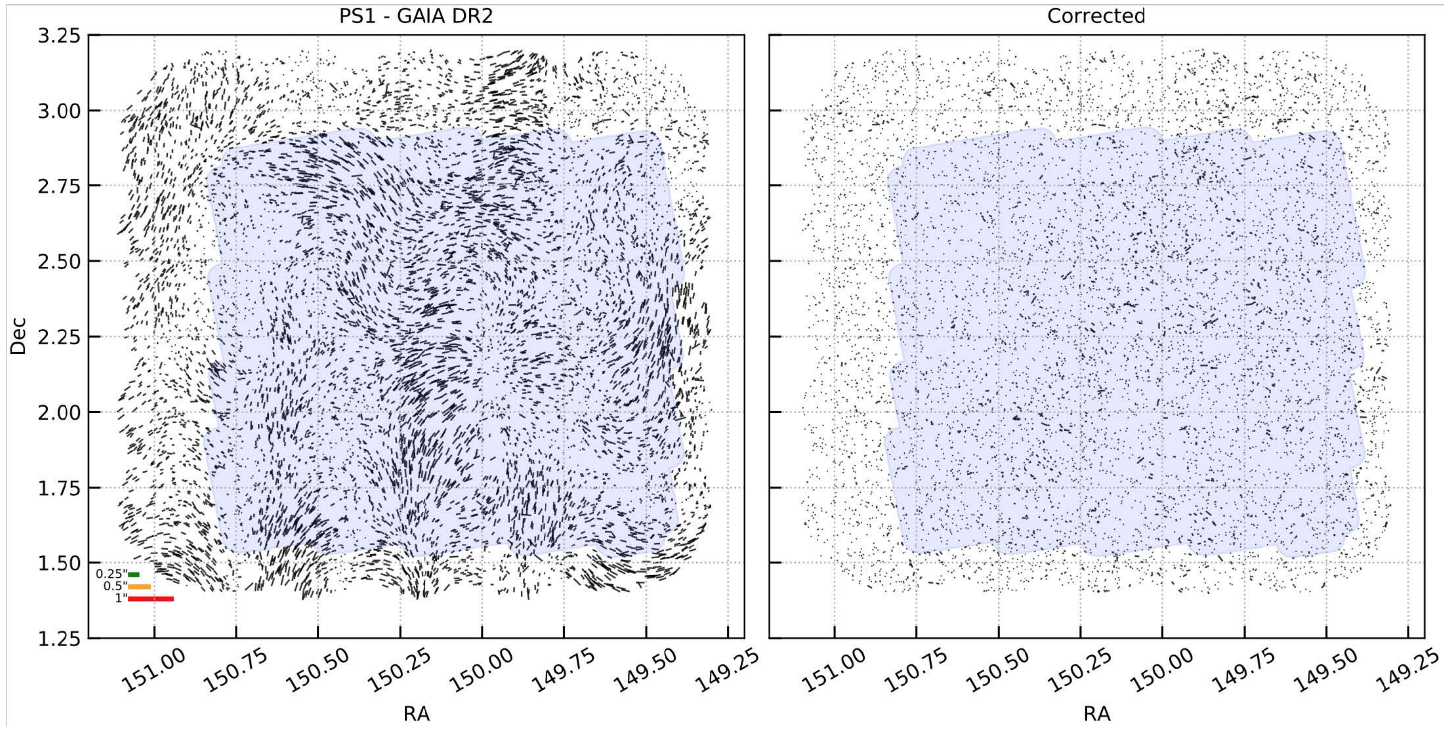


Courtesy of B. Lee (IPAC)

An illustration of the problem: simulated LSST, Euclid, and WFIRST images

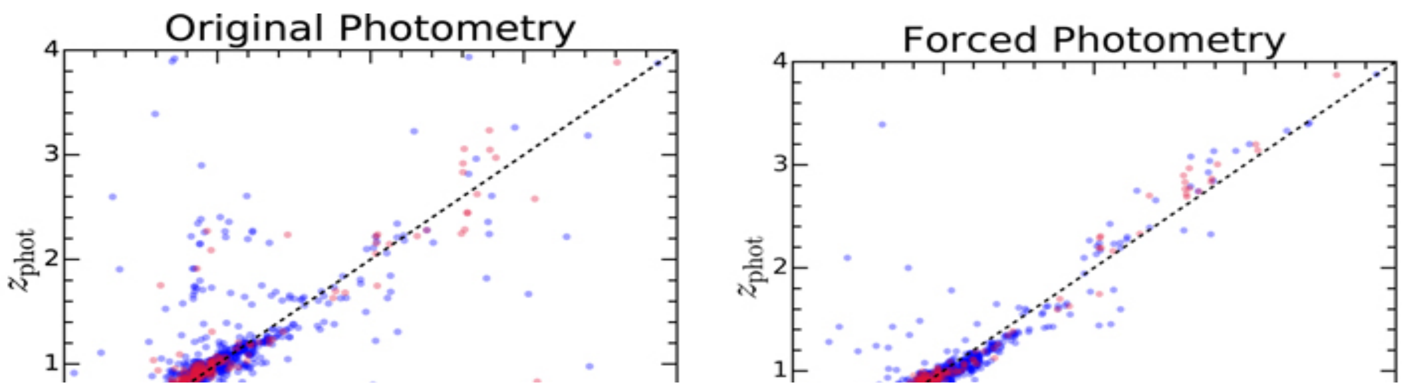
- An extragalactic field (40”X40”) seen by the different projects (WFIRST~0.1 , Euclid VIS~0.18”, Euclid NIR~0.3”, LSST~0.7” FWHM)
- Clear evidence for source blending in LSST
- Re-extract photometry from the pixel level data using space-data as priors.
- Joint processing is NOT catalog matching.

Joint pixel level processing reduces astrometric distortion, improves photometry



Courtesy of G. Brammer (DARK)

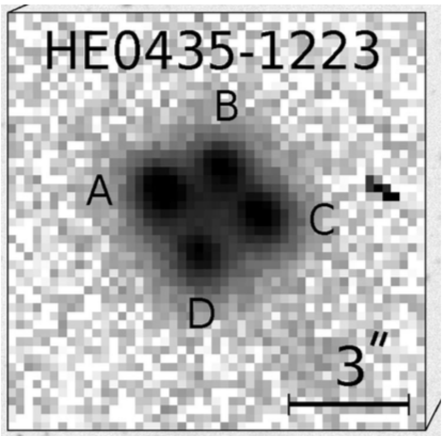
Joint pixel level processing reduces confusion, improves photometry and reduces the scatter in phot-z

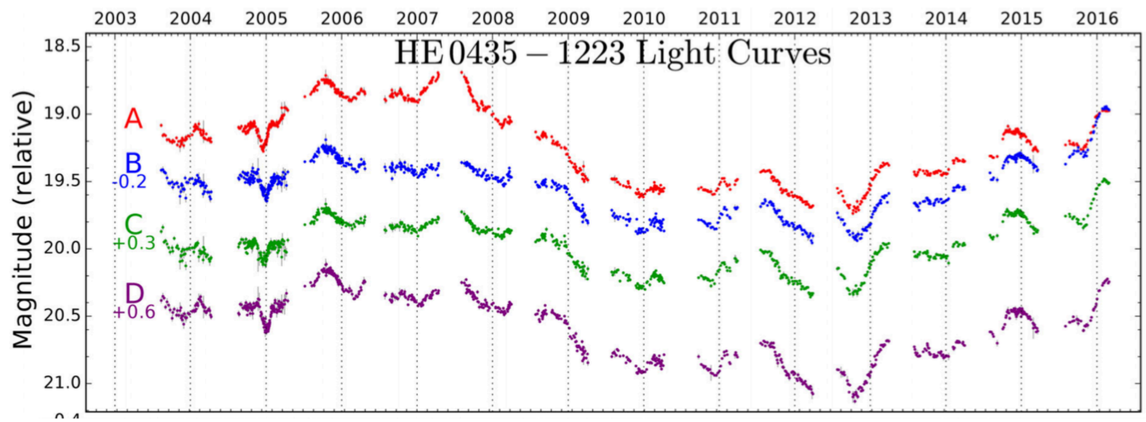


4 Science Pillars:

- Cosmology – weak lensing phot-z, strong lensing time delays, Type Ia SNe
- Reionization and galaxy evolution – morphologies, cross-calibrated line-fluxes, galaxy physical properties for LISA
- Microlensing – LSST bulge field surveys can be made 2-3 mag more sensitive through deconfusion
- Moving objects – Providing a ~10-20 year time baseline for moving objects

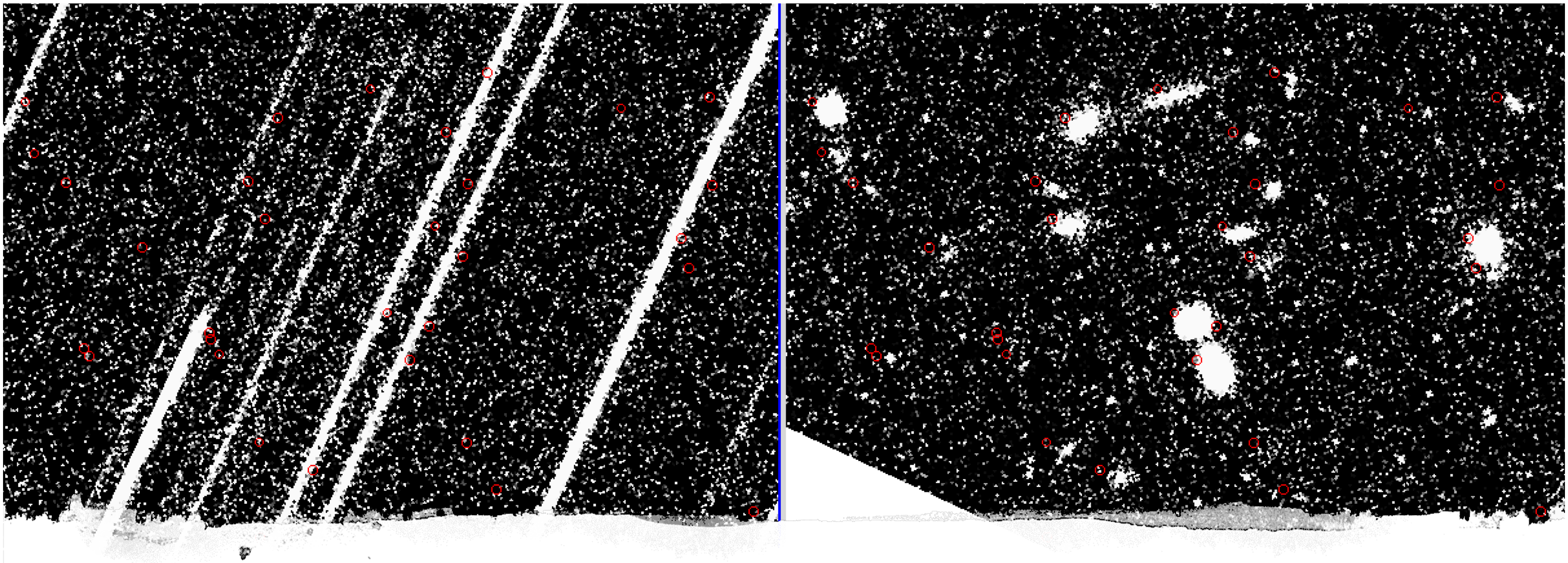
Strong-lens time delays in Euclid/LSST deep fields





H0liCows Team: Bonvin et al., Suyu et al., Treu & Marshall
Cadence of 6-10 days over many years
LSST will provide seeing limited data over similar timescales – photometry can be fit using priors

$$\Delta t_{ij} = \frac{D_{\Delta t}}{c}\left[\frac{(\boldsymbol{\theta}_i - \boldsymbol{\beta})^2}{2} - \psi(\boldsymbol{\theta}_i) - \frac{(\boldsymbol{\theta}_j - \boldsymbol{\beta})^2}{2} + \psi(\boldsymbol{\theta}_j)\right],$$



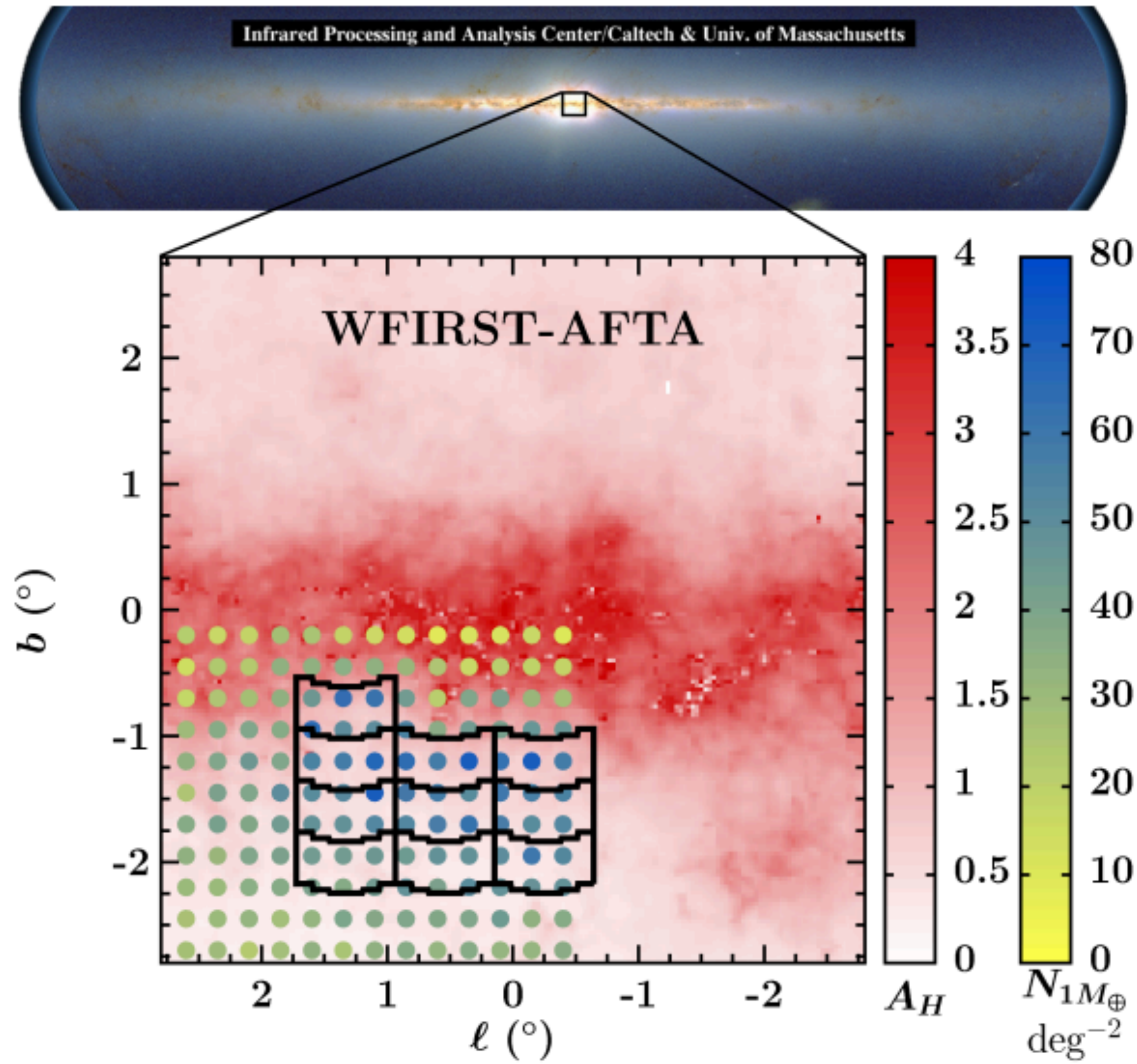
Combination of Euclid/WFIRST grism and LSST data

- Euclid will currently extract spectra for all sources in an input catalog
- Serendipitous sources which are in an LSST-only catalog will not necessarily be extracted e.g. Ly-a from Lyman-break dropouts
- Will need on-the-fly deblending in blue grism data

Without joint processing, you'll be out of luck

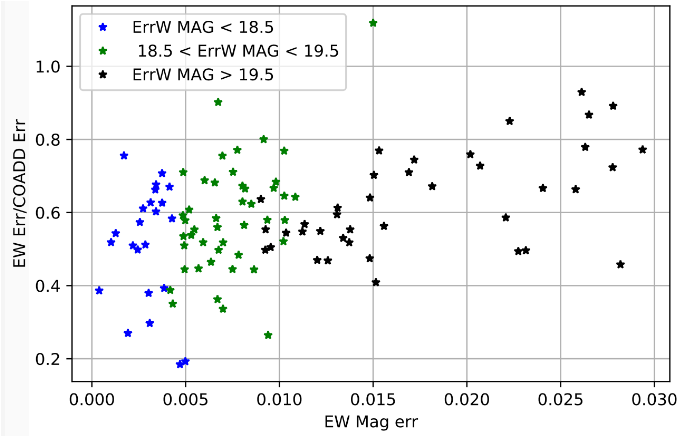
Microlensing

- LSST will likely undertake a microlensing survey (Rachel Street, Rosanne di Stefano)
- Both towards the bulge and out of plane. Bulge will be strongly affected by confusion, requiring WFIRST or a dedicated Euclid survey to provide priors.
- High spatial resolution extinction map will allow luminosities and spectral types of lensing systems
- Prior based deblending allows pushing detection thresholds for planets by ~2 mag i.e. super-Earths



Key volume driver question: single epoch or coadded image for seeing-limited data?

- For stars, current indications are photometry on single epoch images is needed:
  - We are getting 20% improvement in photometry to x2 at the bright end (see figure on right)
  - DAOPHOT does not handle PSFs which have diffraction spikes at different PAs with variable seeing, very well
  - For galaxies, optimal LSST coadds may result in photometric scatter than is 10% worse than combined single epoch photometry.

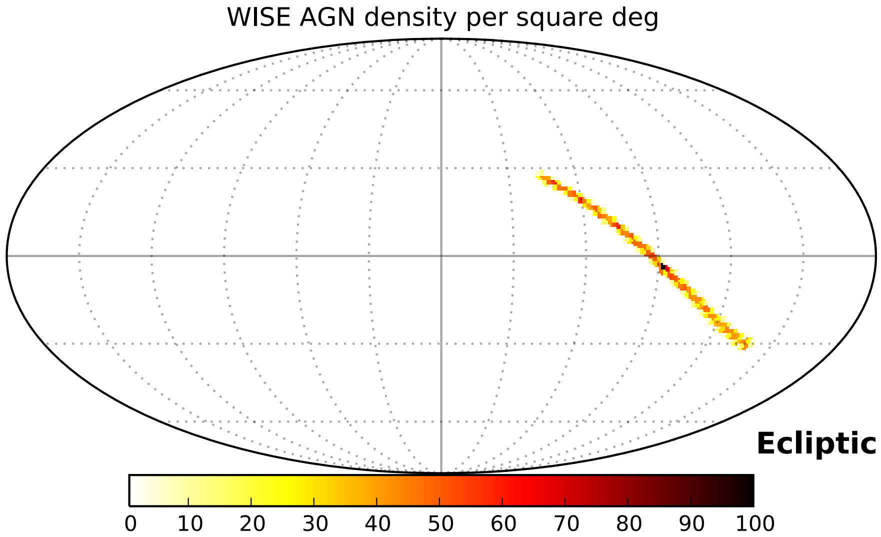


- At any rate single epoch frames need to be stored to handle variable objects e.g. Type Ia SN, strongly lensed QSOs, AGN etc.
- So we will be working with ~100 PB of image level data from the three projects.

Agile Extra Galactic Science: AGN variability

- W1- W2 >0.8 selects both type I & II AGNs
- ~14,000 AGN over Stripe82
- ~52 AGNs per sq. degrees
- Advantage of Infrared

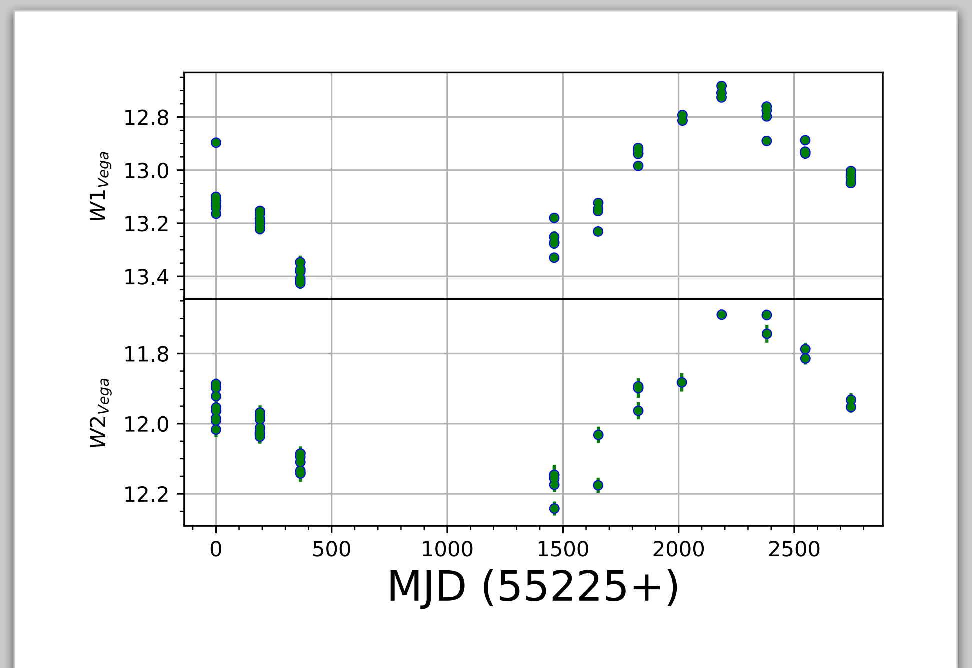
  Easier to separate stars and AGNs

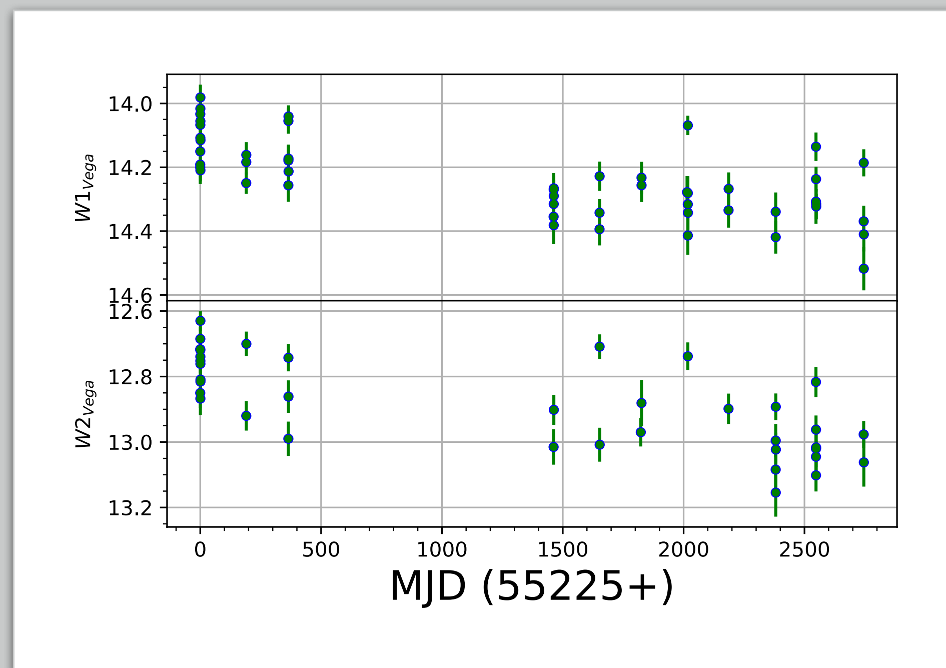


Stern et al.2012

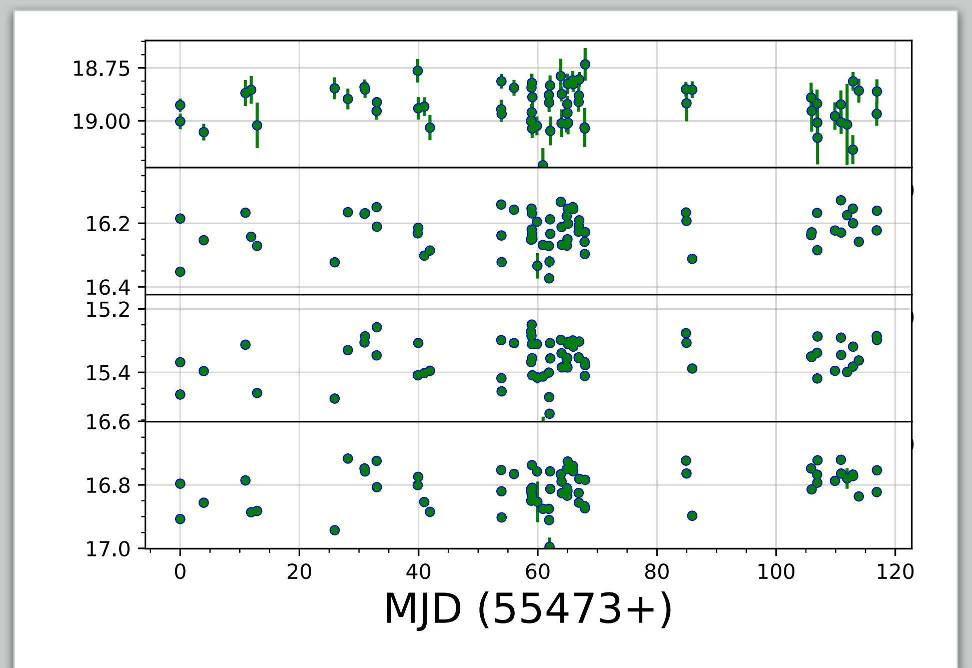
WISE promising!

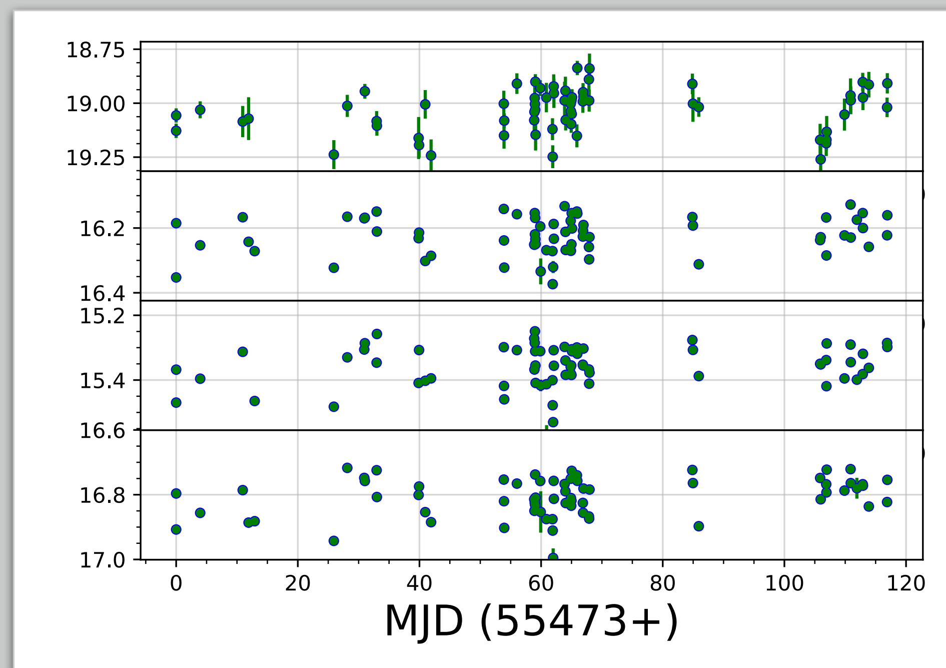
First sensitive full-sky survey for both type 1 and type 2 luminous AGN.





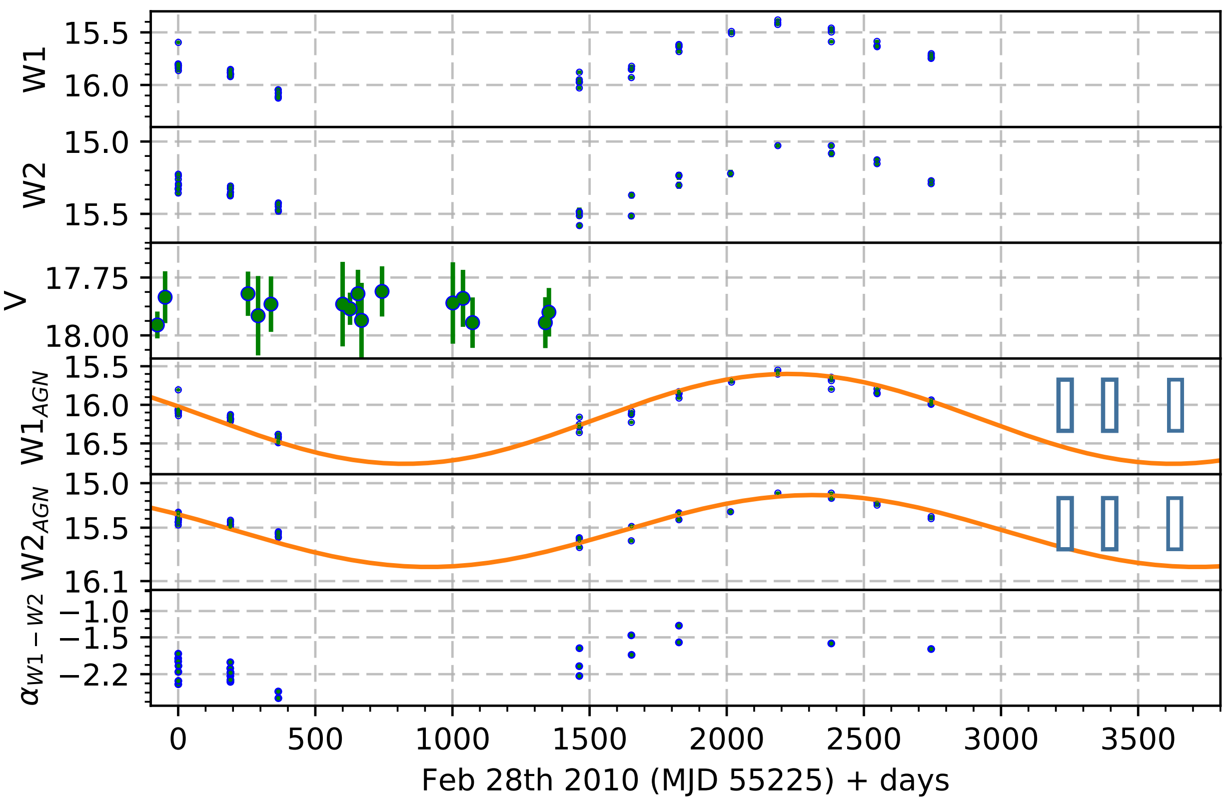
AGN variability: WISE (MIR)

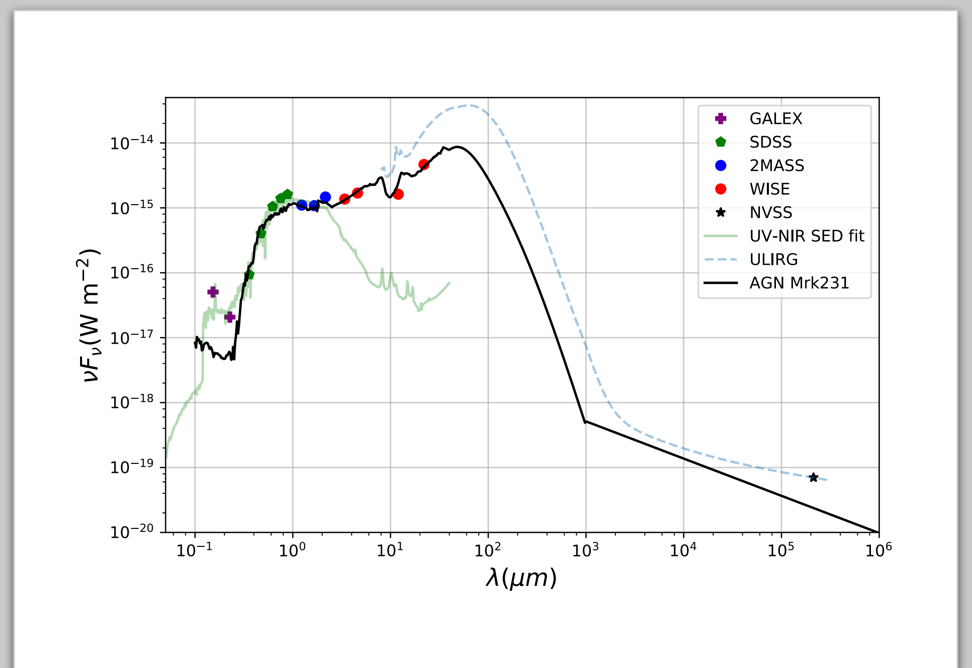


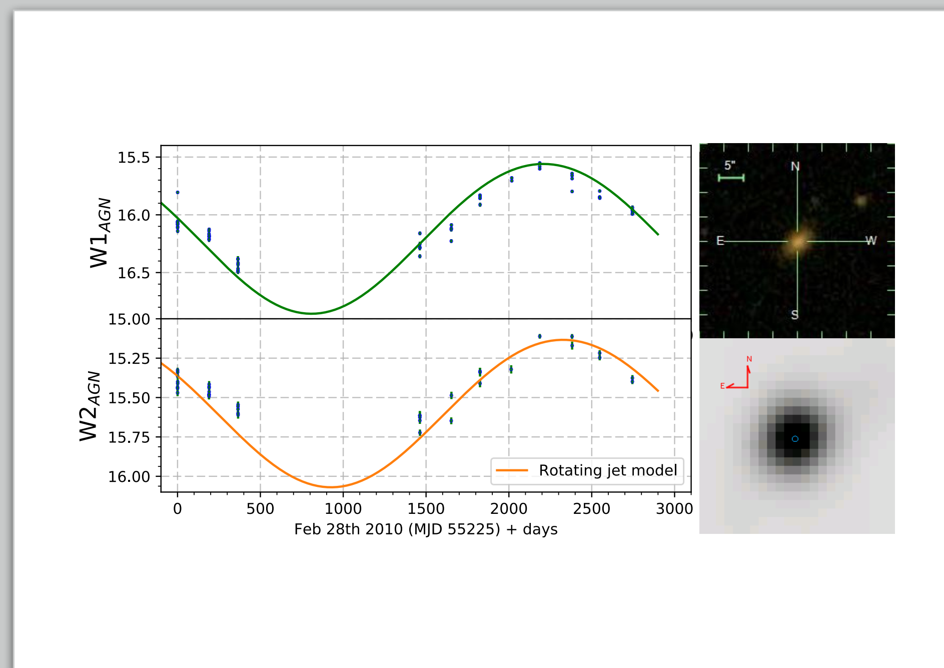


AGN variability: PTF (Optical)

Agile Extra Galactic Science: AGN variability





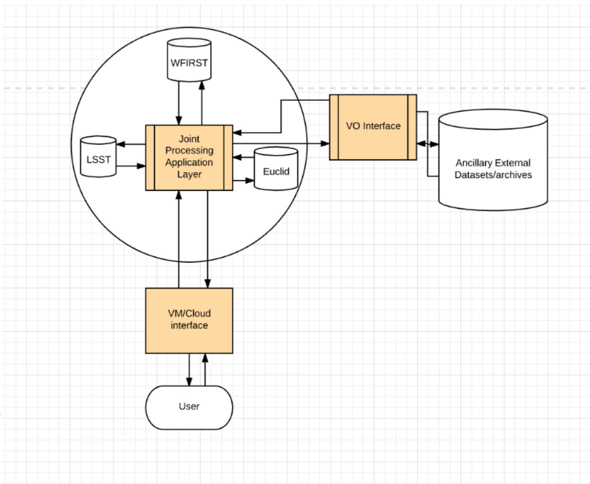


AGN variability: Ancillary datasets

Prakash et al. (in prep)

System Architecture for Implementation

- Clear need for computing to be located at data
  - Otherwise networking times could take months to get results!
- Agile, low-volume use cases will allow the setup of containers which work on data that is pushed to available computing (e.g. Open Science Grid, NASA HPC, AWS and Azure).
- Favoring a hybrid model with the data centers building up computation resources – these are crucial for batch mode jobs and fast interactive jobs.



Status: We are in Phase 2

Task	a.i	a.ii	a.iii	a.iv	b.i	b.ii	b.iii	c.i	c.ii	c.iii	c.iv	c.v
	SCIENCE				ALGORITHMS			ARCHITECTURE				
Topics	Cosmology	Agile Extragalactic Science	Galactic, Solar System, Stellar Streams, Reionization	Microlensing, time baselines for SSOs	LSST Deconfusion, optimal photometry for phot-z, cross-mission color selection, dust corrections			Computing infrastructure, networking, hardware, archives				
Lead(s)	J. Newman (UPitt.)	R. Chary (IPAC)	R. Paladini & S. Wachter (IPAC)	G. Helou (IPAC), W. Dawson (LLNL)	R. Lupton (Princeton) & H. Ferguson (STScI)	R. Lupton & H. Ferguson	P. Melchior (Princeton)	W. Dawson (LLNL)	W. Dawson	B. Rusholme (IPAC) & A. Smith (STScI)	P. Appleton (IPAC)	H. Teplitz (IPAC) & A. Smith (STScI)
Members	Momcheva, Ferguson, Schneider, Prakash, Chary, Capak	Ferguson, Momcheva, Prakash, Capak, Armus, Wood-vasey, Malhotra, McEnery	Ferguson, Momcheva, Kirkpatrick, Chary, Grillmair	van der Marel, Carey, Grillmair	Dawson, Melchior, Schneider, Schulz, B. Lee, Appleton	Dawson, Melchior, Ferguson, Schneider, B. Lee, Grillmair, Armus	Dawson, Ferguson, Koekemoer, Lupton, Schulz	Schneider, Fox, Groom, Ebert	Fox, Flynn, Ebert	Fox, Groom, Berriman	Smith, Fox, Wachter, B. Lee, Rusholme, Berriman	Fox, Wachter, Groom, Rusholme, Berriman

- Three types of tasks in Phase 2: Science requirements, Algorithms, System architecture. Based on these, effort is scoped for Phases 3 + 4 (Implementation starting ~now/FY19)
- 40 people from IPAC, STScI, LSST, JPL, GSFC etc. putting in time
- Integrate ancillary datasets like WISE, GALEX, Planck, IRAS, Pan-STARRS, and Gaia.

Timeline

- Started scoping activity July-2017, built team with representatives from the three projects as well as external people
- Science requirements finalized in Dec 2017
- Interim report submitted to NASA in early-April 2018
- Many algorithms are clearly available, all are discussed but wont be cross-evaluated until prototyping starts with real data
- Architecture is converging to a hybrid system
- Building prototype using Subaru/HSC + Hubble/ACS data in COSMOS
- Submission of final report to NASA, NSF, DOE in Mar 2019
- Presume a ~6 month review cycle
- If Phase 3 starts in Sep 2019 and lasts 3 years (FY20-22), system will be ready in Sep 2022, just in time for LSST and Euclid first "small" public data release.
- Functional interface in FY2023, just in time for Euclid and LSST data.
- Updates are synchronized a few (~3-6) months after public data releases from the projects.