The MARC Project

A joint pilot project between STScI & AAS Journals

Scientists need to be able to get at the data that result in science both to *reproduce* results and produce *new* results

"non-reproducible single occurrences are of no significance to science"

–Karl Popper, *The Logic of Scientific Discovery, 1934*

While it is possible to track down data from a paper, *it's often unpleasantly hard*

Table 1. Sample of PNe with *HST* WFPC2 or WFC3 Hα and [O III] observations

PN G	Common Name	Hα Exp. Time (s)	[O III] Exp. Time (s)	Proposal ID	PN G	Common Name	Hα Exp. Time (s)	[O III] Exp. Time (s)	Proposal ID
000.3+12.2	IC 4634	1000	1000	6856	084.2+01.0	K 4-55	2460	2440	11956
001.2+02.1	Hen 2-262	280	280	9356	084.9−03.4	NGC 7027^c	500	100	11122
001.7−04.4	H 1-55	200	280	9356	089.8−05.1	IC 5117	240	320	8307
002.3−03.4	H 2-37	280	280	9356	096.4+29.9	NGC 6543	800	1600	5403
002.4+05.8	NGC 6369	640	640	9582	106.5−17.6	NGC 7662	200	500	6117, 6943, 8390
002.7−04.8	M 1-42	900	1800	11185	111.8−02.8	Hb 12	1600	1600	11093
002.9−03.9	H 2-39	280	280	9356	138.8+02.8	IC 289	2000	2000	11956
003.5−04.6	NGC 6565	160	320	11122	144.1+06.1	NGC 1501	1600	2000	11956
003.6+03.1	M 2-14	280	280	9356	189.1+19.8	NGC 2371-72	1600	1600	11093
003.8+05.3	H 2-15	280	280	9356	197.8+17.3	NGC 2392	400	400	8499
003.9−03.1	KFL 7	280	280	9356	215.2−24.2	IC 418	888	360	6353, 7501
004.0−03.0	M 2-29	200	160	9356	231.8+04.1	NGC 2438	2080	2080	11827
004.1−03.8	KFL 11	280	280	9356	215.6+03.6	NGC 2346	200	120	7129
004.8−22.7	Hen 2-436	200	160	9356	234.8+02.4	NGC 2440	1600	1600	11090
004.8+02.0	H 2-25	400	400	9356	249.0+06.9	SaSt 1-1	200	280	8332
005.2−18.6	StWr 2-21	280	280	9356	261.0+32.0	NGC 3242	100	1200	6117, 7501, 8773
006.1+08.3	M 1-20	200	160	9356	261.9+08.5	NGC 2818	1600	2000	11956
006.3+04.4	H 2-18	280	280	9356	272.1+12.3	NGC 3132	400	1200	6221, 8390
006.4+02.0	M 1-31	780	160	9356	285.6−02.7	Hen 2-47	1600	1600	11090
006.8−19.8	Wray 16-423	200	160	9356	285.7−14.9	IC 2448	200	320	11122
006.8+04.1	M 3-15	200	160	9356	294.6+04.7	NGC 3918	140	320	11122
007.5+04.3	Th 4-1	280	280	9356	305.1+01.4	Hen 2-90	2325	1210	8345, 9102
008.2+06.8	Hen 2-260	200	460	9356	307.5−04.9	MyCn 18	600	1400	6221
008.6−02.6	MaC 1-11	280	280	9356	309.1−04.3	NGC 5315	1600	1600	11090
009.3+05.7	Hen 3-1475	830	800	7285	312.3+10.5	NGC 5307	1600	1600	11090
010.0+00.7	NGC 6537	1240	1000	6502	319.6+15.7	IC 4406^d	540	600	8726, 9314
010.8+18.0	M 2-9	1240	1000	6502	324.0+03.5	PM 1-89	4900	2900	5404, 5864
010.8−01.8	NGC 6578	160	320	11122	327.8+10.8	NGC 5882	140	380	11122
019.4−05.3	M 1-61	240	320	8307	331.1−05.7	PC 11	200	280	8332
025.3+40.8	IC 4593	1600	1600	11093	331.3−12.1	Hen 3-1333	240	368	6039, 8390
025.8−17.9	NGC 6818	520	1300	6792, 7501, 8773	331.7−01.0	Mz 3^e	1260	1160	6856, 9050
027.6+04.2	M 2-43	520	1800	8307	341.8+05.4	NGC 6153	1000	1200	8594
034.6+11.8	NGC 6572	180	840	7501, 9839	349.5+01.0	NGC 6302^d	2100	2220	11504
036.1−57.1	NGC 7293	1800	1800	5977	351.1+04.8	M 1-19	160	160	9356
037.7−34.5	NGC 7009	400	320	8114	351.9−01.9	Wray 16-286	200	280	9356
037.8−06.3	NGC 6790	160	200	8307	352.6+03.0	H 1-8	200	280	9356
043.1+37.7	NGC 6210	320	320	6792	353.5−05.0	JaFu 2^f	3600	2000	6780
054.1−12.1	NGC 6891	1280	320	11122	354.5+03.3	Th 3-4	280	280	9356
054.2−03.4	Necklace Nebula^a	2000	2000	12675	354.9+03.5	Th 3-6	280	400	9356
057.9−01.5	Hen 2-447	520	1800	8307	355.4−02.4	M 3-14	200	160	9356
060.1−07.7	NGC 6886	1120	1020	7501, 8345, 8773	355.9+03.6	H 1-9	280	280	9356
060.8−03.6	NGC 6853	2000	1000	8726	356.1−03.3	H 2-26	280	280	9356
063.1+13.9	NGC 6720	480	720	7632, 8726	356.5−03.6	H 2-27	360	400	9356
064.1+04.3	M 1-92	680	2080	6533	356.9+04.4	M 3-38	280	280	9356
064.7+05.0	BD+30°3639	484	900	8116, 8390	357.1−04.7	H 1-43	200	280	9356
065.0−27.3	Ps 1^b	11420	1040	6751	357.2+02.0	H 2-13	280	280	9356
071.6−02.3	M 3-35	520	1000	8307	358.5−04.2	H 1-46	160	160	9356
073.0−02.4	K 3-76	6	18	6943	358.5+02.9	Wray 16-282	280	280	9356
074.5+02.1	NGC 6881	280	320	8307	358.9+03.4	H 1-19	200	280	9356
082.1+07.0	NGC 6884	1100	560	8345, 8390	359.2+04.7	Th 3-14	280	400	9356
082.5+11.3	NGC 6833	40	3	6943, 6353	359.3−00.9	Hb 5	1300	1000	6502
083.5+12.7	NGC 6826	100	100	6117					

Guerrero+ 2013

*We searched MAST for HST WFPC2 or WFC3 coeval Hα and [O iii] images of PNe available by March 2013. This search yielded Hα and [O iii] images for **103** PNe obtained through the F656N and F502N filters, respectively*

*We present a catalogue of photometric and structural properties of **228** nuclear star clusters (NSCs) in nearby late-type disc galaxies. These new measurements are derived from a homogeneous analysis of all suitable Wide Field Planetary Camera 2 (WFPC2) images in the Hubble Space Telescope (HST) archive.*

Table 1. Main properties of the galaxy sample with measured NSC properties. (All 228 galaxies are listed in the online version of the table).

Galaxy	RA (hh:mm:ss)	Dec. (dd:mm:ss)	$m - M$ (mag)	$E(B - V)$ (mag)	B (mag)	$B - V$ (mag)	I (mag)	R_{25} (kpc)	ϵ	PA (deg)	Incl. (deg)	Type	t
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
DDO078	10:26:27.78	67:39:25.1	27.82	0.018	15.8	−	−	1.063	0.00	−	0.0	I	10.0
IC 4710	18:28:37.95	−66:58:56.1	29.75	0.079	12.51	0.57	11.19	4.494	0.15	−	34.9	Sm	8.9
NGC 1258	3:14:05.50	−21:46:27.3	32.28	0.022	13.88	−	12.35	5.870	0.26	20.5	43.7	SABc	5.7
NGC 3319	10:39:09.47	41:41:12.5	30.7	0.013	11.77	0.41	11.46	7.289	0.51	36.	62.7	SBc	5.9
NGC 5334	13:52:54.44	−1:06:52.4	32.78	0.041	12.97	−	12.19	17.729	0.28	18.2	44.8	Sc	5.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...

Notes: The values for all columns are taken from HyperLeda, except for columns 4 and 5, which are taken from NED. More specifically, the distance modulus $m - M$ in column 4 is the median value in NED. If the latter is not available, we adopt the redshift-derived distance modulus, modz, from HyperLeda.

Georgiev & Böker 2013

At Space Telescope there is 1 FTE spent trying to piece this all together; without the authors it's hard!

Assign Dataset Names to HST papers

Set status	DSN Comments: *The content cannot contain http:*	Person entering information
○ Not Started	It might be possible to track down some of the specific datasets for this paper, but it would be very time consuming with possibly little payoff. I'm tagging as impossible for now. –Sara	jilllag
○ In progress (P)		
○ Initial Entry (E)		
○ Question (Q)		
○ Best we could -Partially Impossible (B)		
⊙ Impossible (I)		
○ Complete (C)		
○ Recheck (R)		

[Update HSTDSN status] [HST] [PaperCentral]

There are 47 combinations of program ID and Instrument for 2014MNRAS.441.3570G

There are 8 poss

[AddDSN] [Mark all]

> It might be possible to track down some of the specific datasets for this paper, but it would be very time consuming with possibly little payoff. I'm tagging as impossible for now. –Sara

Mark	DataSet Name	Association Name	Target	RA	Dec	Start Time	Exposure Length	Inst Config	Filter/Grating	PI Name
☐	U9OX0301									
☐	U9OX0302									
☐	U9OX0303									
☐	U9OX0304M		SN2006MY	12 43 39.302	+16 23 36.22	2007-04-27 00:02:16	300	WFPC2	F555W	SMARTT
☐	U9OX0305M		SN2006MY	12 43 39.302	+16 23 36.22	2007-04-27 01:01:16	500	WFPC2	F814W	SMARTT
☐	U9OX0306M		SN2006MY	12 43 39.302	+16 23 36.22	2007-04-27 01:15:16	700	WFPC2	F814W	SMARTT
☐	U9OX0307M		SN2006MY	12 43 39.302	+16 23 36.22	2007-04-27 02:37:16	700	WFPC2	F450W	SMARTT
☐	U9OX0308M		SN2006MY	12 43 39.302	+16 23 36.22	2007-04-27 02:51:16	700	WFPC2	F450W	SMARTT
Mark	**DataSet Name**	**Association Name**	**Target**	**RA**	**Dec**	**Start Time**	**Exposure Length**	**Inst Config**	**Filter/Grating**	**PI Name**

[AddDSN] [Mark all] [Unmark all]

There are 21 possible datasets found for 10829 / WFPC2

[AddDSN] [Mark all] [Unmark all]

Mark	DataSet Name	Association Name	Target	RA	Dec	Start Time	Exposure Length	Inst Config	Filter/Grating	PI Name
☐	U9NP0201M		NGC6509	17 59 22.171	+06 16 58.39	2007-04-22 17:24:16	600	WFPC2	F606W	MARTINI
☐	U9NP0202M		NGC6509	17 59 22.194	+06 16 58.50	2007-04-22 17:40:16	600	WFPC2	F606W	MARTINI
☐	U9NP0203M		NGC6509	17 59 22.216	+06 16 58.61	2007-04-22 17:56:16	600	WFPC2	F606W	MARTINI
☐	U9NP0401M		NGC4519	12 33 27.424	+08 39 38.17	2007-12-03 04:18:17	600	WFPC2	F606W	MARTINI
☐	U9NP0402M		NGC4519	12 33 27.446	+08 39 38.02	2007-12-03 04:34:17	600	WFPC2	F606W	MARTINI

Places similar to STScI have DOI'ed their own content

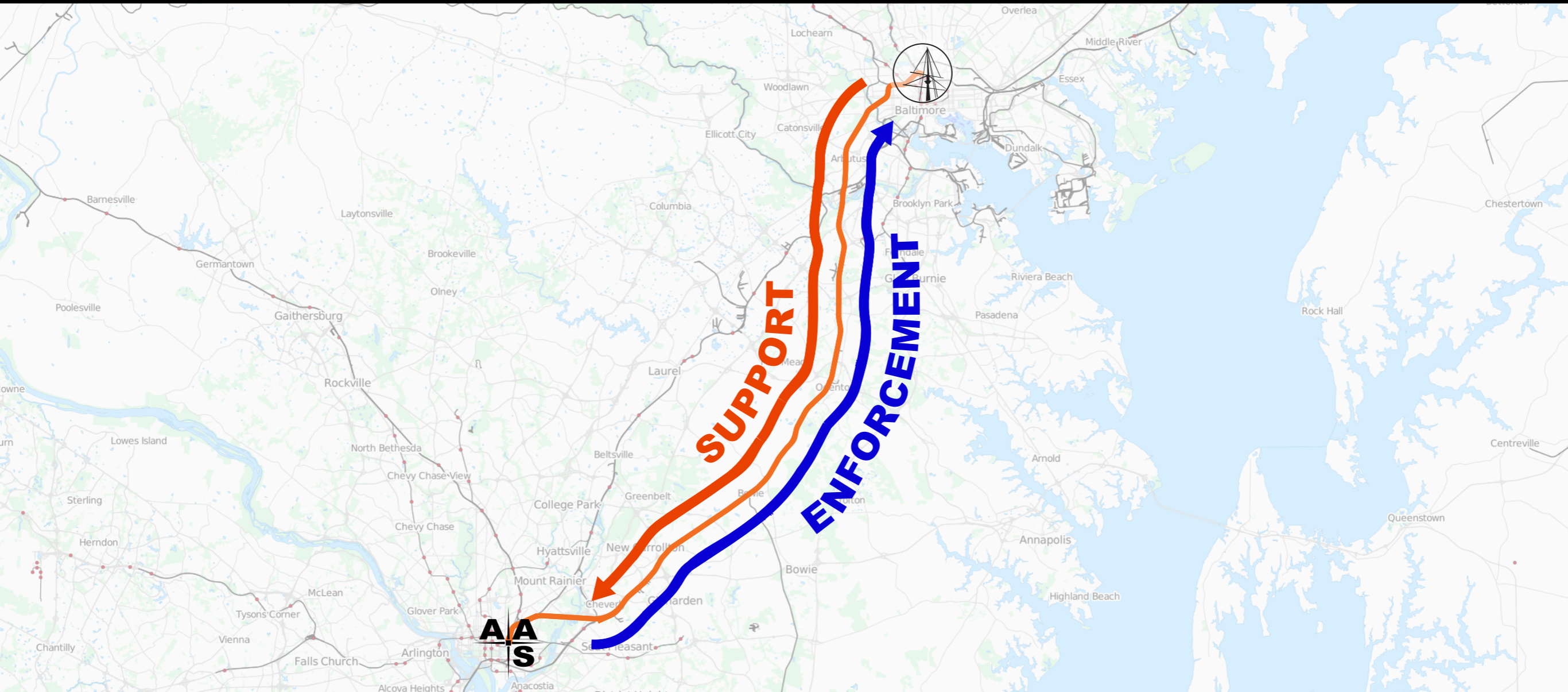# The Collective Action Problem:

People are Lazy*

[citation needed]

The MARC Project

connecting the AAS Journals to MAST (*and beyond…*)
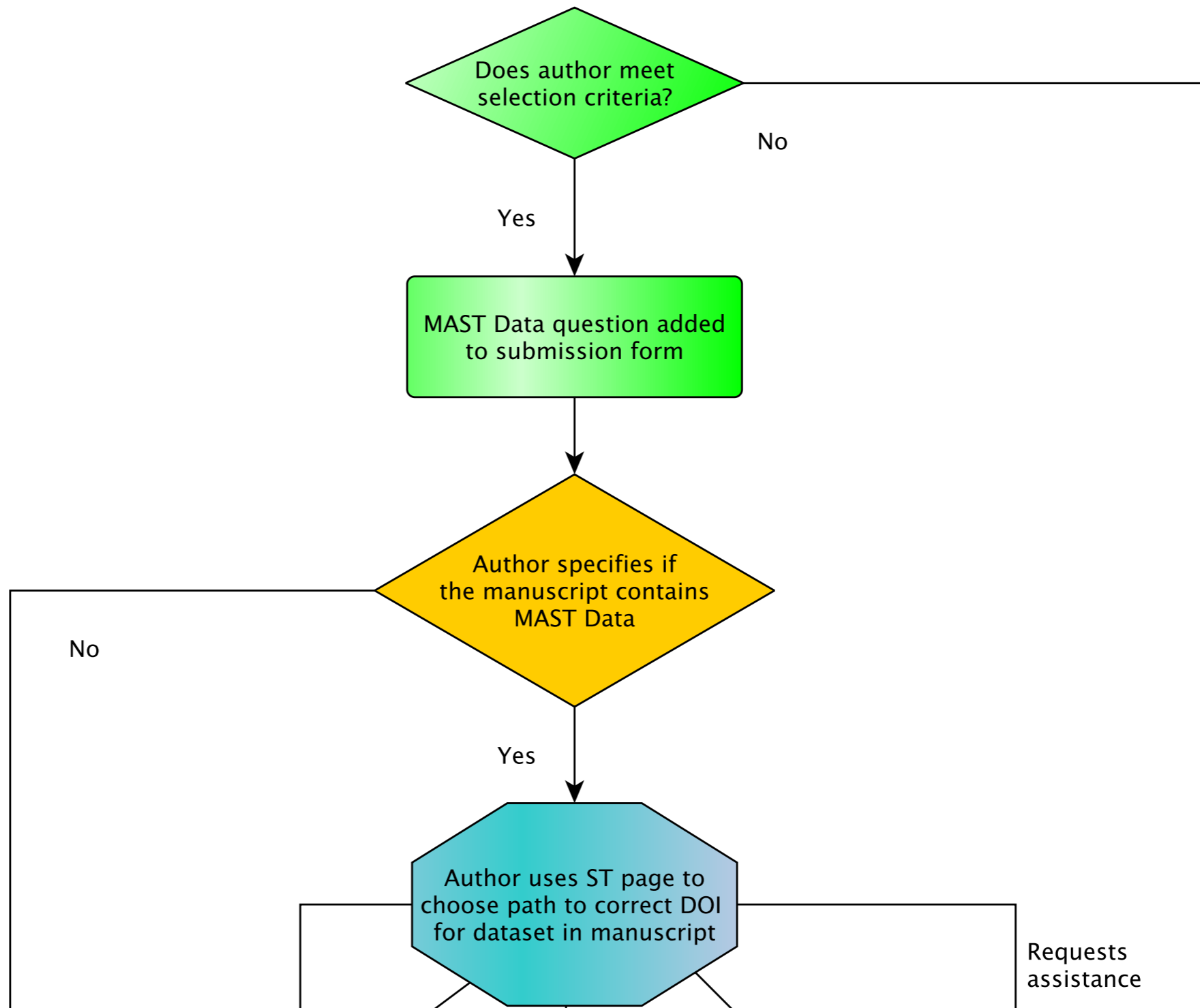
We've constructed a workflow for Authors, ApJ/AJ/ApJL/ApJS, and MAST

Author	Journal	MAST

Does author meet selection criteria?

No

Yes

MAST Data question added to submission form

Author specifies if the manuscript contains MAST Data

No

Yes

Author uses ST page to choose path to correct DOI for dataset in manuscript

Requests assistance

We've constructed a workflow for Authors, ApJ/AJ/ApJL/ApJS, and MAST

| Author | Journal | MAST |

Manuscript Information | **Files** | Validate | Submit

| Upload Files | Replace Files | File Type | File Description | File Order | Dataset Deposition | Dataset Upload |

MAST * ⓘ

Please enter all details for datasets uploaded to publicly accessible repositories in the fields below.

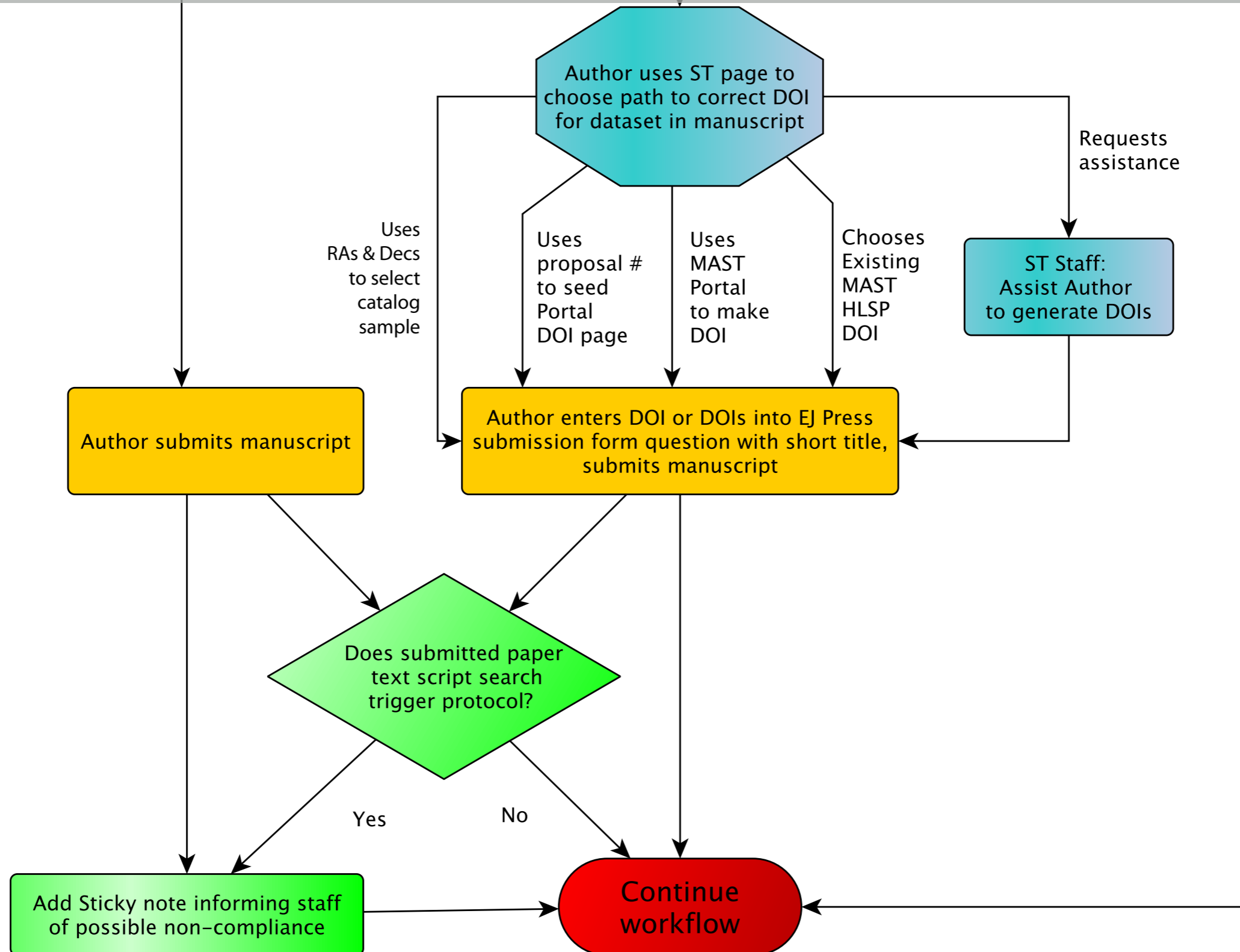For short title, please enter a tag for the data set (e.g. 'WFC3 images')

DOI *	Short Title *	
12345	Sample Data Set	Clear
54321	Sample Data Set	

Additional Data Files

Back | Save and Exit | Next

We've constructed a workflow for Authors, ApJ/AJ/ApJL/ApJS, and MAST

Author **Journal** **MAST**

Yes

Author uses ST page to choose path to correct DOI for dataset in manuscript

Requests assistance

Uses RAs & Decs to select catalog sample

Uses proposal # to seed Portal DOI page

Uses MAST Portal to make DOI

Chooses Existing MAST HLSP DOI

ST Staff: Assist Author to generate DOIs

Author submits manuscript

Author enters DOI or DOIs into EJ Press submission form question with short title, submits manuscript

Does submitted paper text script search trigger protocol?

Yes   No

Add Sticky note informing staff of possible non-compliance

Continue workflow

To save catalog samples people report SQL queries which are both incomplete and non-archival

```
SELECT
   round(p.ra,6) as ra, round(p.dec,6) as dec,
   p.run, p.camcol, p.field,              --- comments are preceded by ---
   round(p.extinction_r,3) as rExtSFD, --- r band extinction from SFD
   round(p.modelMag_u,3) as uRaw,    --- N.B. ISM-uncorrected model mags
   round(p.modelMag_g,3) as gRaw,    --- rounding up
   round(p.modelMag_r,3) as rRaw,
   round(p.modelMag_i,3) as iRaw,
   round(p.modelMag_z,3) as zRaw,
   round(p.modelMagErr_u,3) as uErr,
   round(p.modelMagErr_g,3) as gErr,
   round(p.modelMagErr_r,3) as rErr,
   round(p.modelMagErr_i,3) as iErr,
   round(p.modelMagErr_z,3) as zErr,
   (case when (p.flags & '16') = 0 then 1 else 0 end) as ISOLATED,
   ISNULL(round(t.pmL,3), -9999) as pmL, --- proper motion data are set to
   ISNULL(round(t.pmB,3), -9999) as pmB, --- -9999 if non-existent (NULL)
   ISNULL(round(t.pmRaErr,3), -9999) as pmErr  --- if pmErr < 0 no pm data
INTO mydb.dustSample
FROM phototag p LEFT OUTER JOIN propermotions t ON
   (p.objID = t.objID and t.match = 1 and t.sigra < 350 and t.sigdec < 350)
         --- quality cut on pm
WHERE
   p.type = 6 and                    --- select unresolved sources
   (p.flags & '4295229440') = 0 and --- '4295229440' is code for no
                                 --- DEBLENDED_AS_MOVING or SATURATED objects
   p.mode = 1  --- PRIMARY objects only, which implies
             --- !BRIGHT && (!BLENDED || NODEBLEND || nchild == 0)]
   p.modelMag_r < 21  --- adopted faint limit
--- the end of query
```

❗ **Crossmatching**
❗ **Subselection**
❗ **Evolving SQL**

50 Billon Row Database (PanSTARRS)
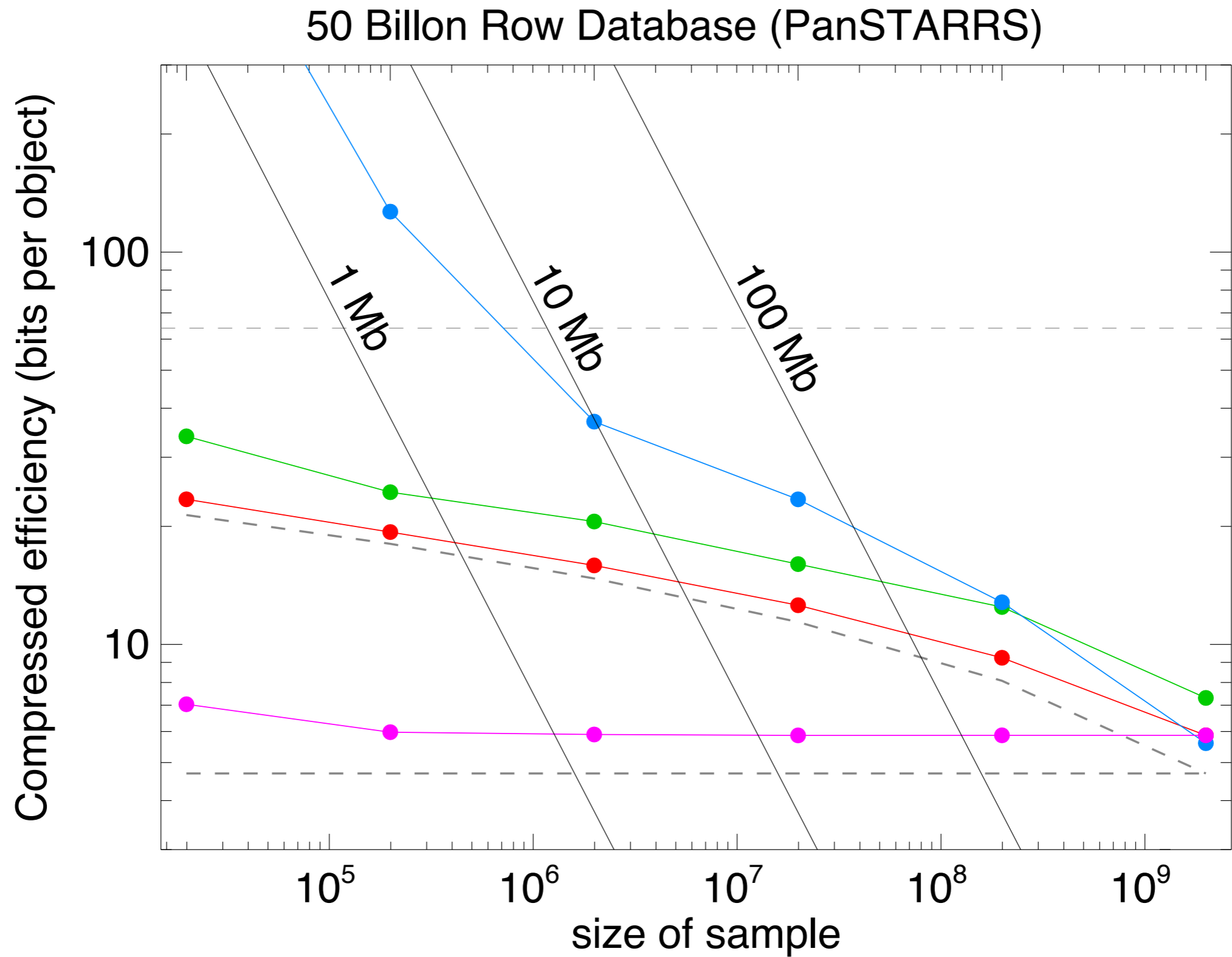
Questions

- Do you see avoidable pitfalls in roll-out to STScI Staff? The broader author community?

- What are our highest next priorities? Tightening exisiting product? Finer-grained HLSPs? Saved Catalog Samples?

- Are there other groups we should connect with to to discuss Saved Catalog Samples?

- Who should our next link be to? MNRAS?