MNRAS **000**, 1–2 (2017) Preprint 22 February 2017 Compiled using MNRAS LATEX style file v3.0

Preliminary Data Access Center : User Report

Krzysztof Suberlak [1]★ Željko Ivezić, [1]
[1] *Department of Astronomy, University of Washington, Seattle, WA, United States*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
A report on user experience of the Preliminary Data Access Center (PDAC). Employing the SDSS and GAIA datasets we test the quality and ease of access to the data. PDAC will pave the way to the Science User Interface and Tools (SUIT). We employ both in-detail study of individual objects, and a statistical study of an ensemble of objects. We evaluate user-friendliness of the current interface, and make recommendations for its future improvements.

1 INTRODUCTION

This is a document to report on the user experience testing of the Preliminary Data Access Center. The Large Scale Synoptic Telescope (LSST) will produce a big volume of data. Such unprecedented data stream poses new challenges to provide an easy access for users, in such a way that they can quickly find what they need, and thus be able to focus on the science goal that they would like to achieve. The detail description of such online user-interface called Science User Interface and Tools is outlined in documents LDM-130 (SUIT requirements) and LDM-492 (SUIT Vision). An idea of having an interface to the data is not new : there exists Aladin, SDSS CAS jobs, IPAC IRSA, Mikulsky NASA Archive, NED, and many other archives. These allow a user to query for data (either via SQL query, or interface), returning the data table. Some user interfaces (eg. IRSA) have some rudimentary plotting capabilities. There have been ideas of a new interface, that would not only eg. plot the lightcurve and display the spectrum, but also allow the user to run some machine learning algorithms, or simple models that can help narrow down the query, or obtain science results in the browser. Namely, Victor Pankratius, from MIT, in his talk "Computer-Aided Discovery: Towards Scientific Insight Generation with Machine Support" outlined the idea of an ipython notebook - access to data, which lives in the cloud, is allocated some CPU share and memory, and allows one to upload / download the data and run the model in real time, which is especially helpful to geoscientists doing fieldwork, where new data acquisition conditions their next step.

Indeed, astronomers may find that quick look into the data, finding eg. all stars that exhibit RR Lyr variability and have been observed in a certain region of the sky, is very helpful.

Here we outline the user experience of PDAC (see PDAC technical description on [1]

Currently, the PDAC v1, under tab 'LSST Data' in the upper-left corner of the interface (see Fig. 1) includes the Summer 2013 DM-stack reprocessed SDSS Stripe 82 data, hosted at the NCSA on the LSST prototype ("integration cluster") hardware, in Qserv [Gregory Dubois-Felsmann, priv.comm. 02-20-2017, slack]. The reprocessing included:

- coadding the data from all epochs in each of the ugriz SDSS filters. Measurements on coadds (per object) are available as `RunDeepSource` table, accessible via Catalogs –> 'DeepSource' . The single-band coadded images with MariaDB metadata are available as `DeepCoadd` table, accessibla via Images –> 'DeepCoadd' .
- using i-band detections to seed forced photometry on all epochs in all bands. The results of photometry are available as `RunDeepForcedSource` table, accessible via Catalogs –> 'Deep Forced Source' .
- For reference , the individual calibrated single epoch images are available as `Science_Ccd_Exposure` table, accessible via Images –> 'Science CCD Exposure'

Details of the S82 LSST reprocessing can be found in the PDAC document https://confluence.lsstcorp.org/display/DM/Properties+of+the+2013+SDSS+Stripe+82+reprocessing.

PDAC v1 under tab 'External Catalogs' also provides access to all NASA/IPAC Infrared Science Archive(IRSA) publicly accessible catalogs, including GAIA, WISE, etc. (see Fig. 2). These are stored at Infrared Processing and Analysis Center (IPAC) http://www.ipac.caltech.edu/project/lsst.

2 METHODS

We perform single-object tests and statistical tests on an ensemble of objects .

First, we study in detail a particular source - we consider examples of variable objects, confirmed by previous studies (eg. RR Lyrae from Sesar+2010, Table 1). We download these from the S82 dataset on PDAC, run Lomb-Scargle periodogram to find period, and plot the phased lightcurve.

Second, we query the S82 database against a small

[1] https://confluence.lsstcorp.org/display/DM/Guide+to+PDAC+version+1

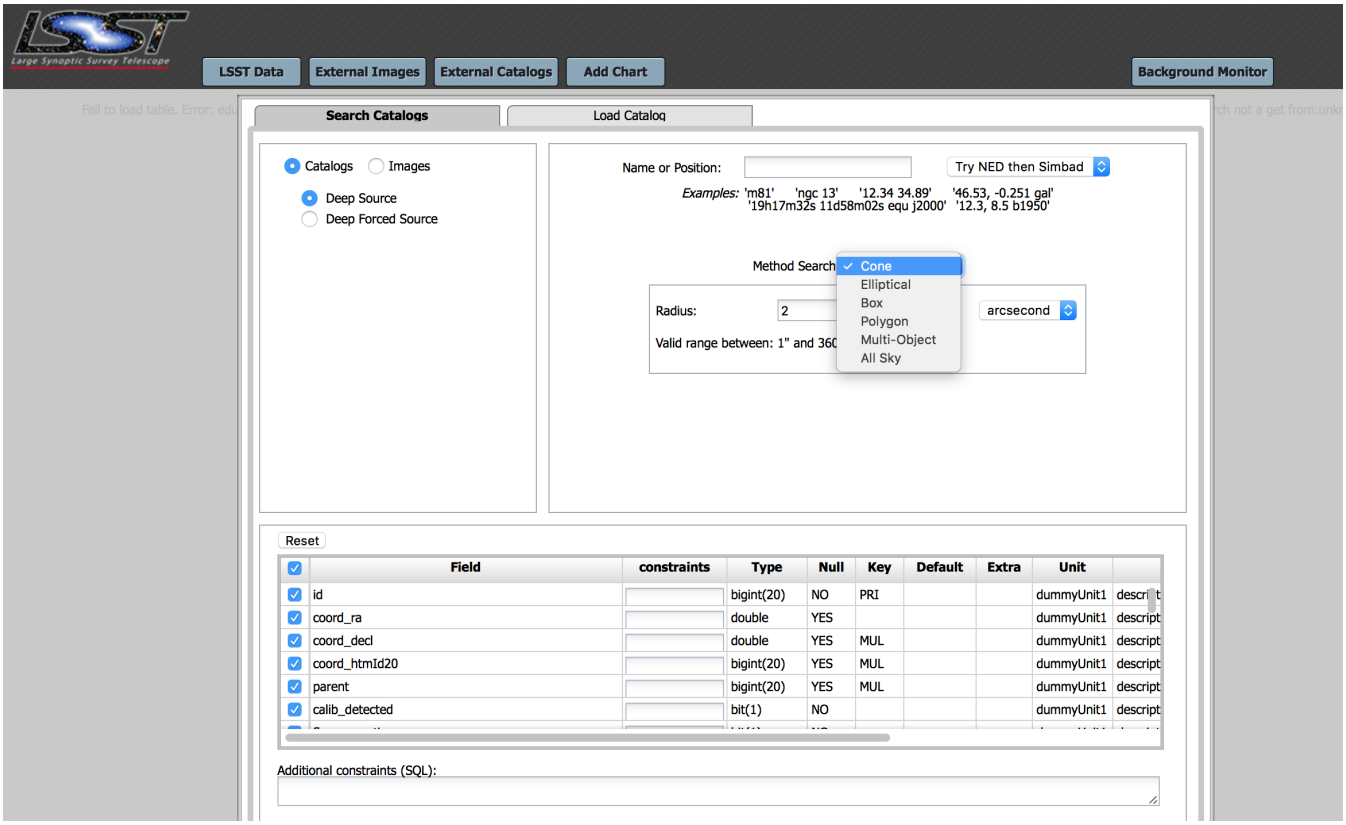

Figure 1. The main user interface of PDAC ver. 1

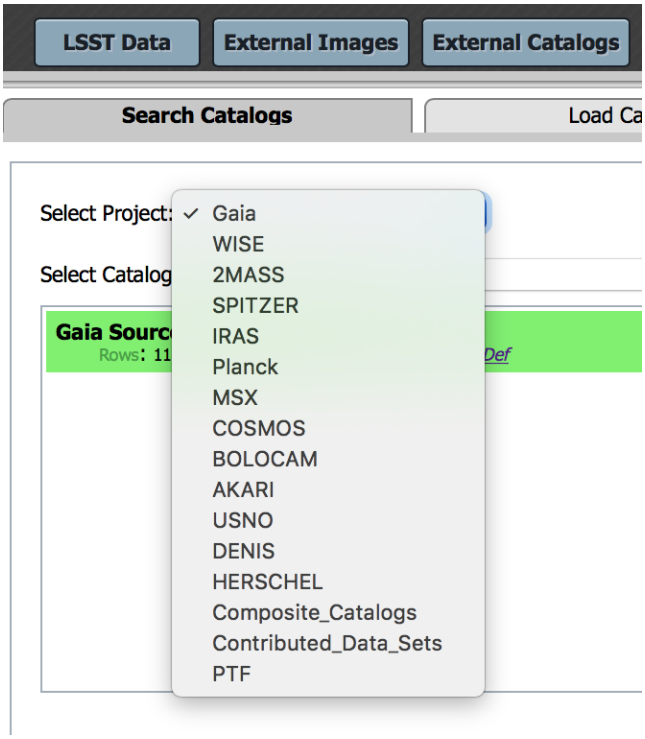

Figure 2. IPAC- hosted catalogs , accessible via IRSA.

subset of a given S82 patch (few degrees), downloading lightcurves for ~ 100000 objects in that area of the sky. We plot color-color diagrams, as in Sesar+2007, Fig.3 ,4.

3 RESULTS

4 CONCLUSIONS

ACKNOWLEDGEMENTS

Thank you !

APPENDIX A: SOME EXTRA MATERIAL

If you want to present additional material which would interrupt the flow of the main paper, it can be placed in an Appendix which appears after the list of references.

This paper has been typeset from a TEX/LATEX file prepared by the author.