Large Synoptic Survey Telescope (LSST)

LDM-503-1 (WISE Data Loaded in PDAC) Test Report

Gregory P. Dubois-Felsmann

DMTR-52

Latest Revision: 2018-04-23

D R A F T

Abstract

This is the test report for LDM-503-1 (WISE Data Loaded in PDAC), an LSST DM level 2 milestone pertaining to the LSST Science Platform, with tests performed according to LSP-00, Portal and API Aspect Deployment of a Wide-Area Dataset.

Change Record

Version	Date	Description	Owner name
1	2018-04-22	Initial version.	Dubois-Felsmann

Document curator: Gregory P. Dubois-Felsmann

Document source location: `https://github.com/lsst-dm/DMTR-52`

Version from source repository: ad5f92c

Contents

LDM-503-1 (WISE Data Loaded in PDAC) Test Report

1   Introduction

The LDM-503-1 milestone calls for the principal WISE primary mission and NEOWISE first-year catalogs to be loaded into the LSP Integration Environment, also known as the Prototype Data Access Center. The LSP-00 test specification, defined in LDM-540, was used to verify that the data were loaded and that the data service provided meets a subset of requirements for functionality of the API Aspect and the Portal Aspect of the Science Platform.

The LSP-00-25 and LSP-00-30 test cases are not included in this report, as they address image data functionality and the WISE data load to PDAC in the end did not include the image data. The decision to avoid loading the image data was taken when the time required to load the WISE catalog data turned out to be larger than originally expected, and also required unanticipated hardware upgrades. In addition to the tiem required for Loading the WISE image data, providing a useful image cutout service would have required developing an LSST-stack (`afw`) camera model for WISE, and this was judged to be an effort with very limited benefit to DM development. The key usefulness of the WISE data in the DM context is as a scientifically realistic scaling test for Qserv, requiring the handling of tens of billions of rows. These test cases can be performed against the SDSS data previously loaded in the PDAC, and will be executed and documented separately.[1]

The WISE catalog data to be loaded include the following tables:

- AllWISE Source Catalog: this catalog is comparable to the projected LSST `Object` table, arising from source detection performed on the coaddition of all the data from the pre-hibernation year of WISE operations, including the primary 4-band data as well as shorter periods of 3-band (W1,W2,W3) and 2-band (W1,W2) data taken after the exhaustion of the stored cryogen.

- AllWISE Multi-Epoch Photometry (MEP) Table: roughly comparable to LSST `ForcedSource`, arising from forced photometry on every single-epoch image, based on the sources in

---

[1] As a user convenience, the Portal Aspect in PDAC does provide some limited access to WISE image data directly from IRSA, but as it does not use LSST interfaces to do so, it is not included in this test.

the AllWISE Source Catalog. Differs from `ForcedSource` in that WISE takes all four band images at once, so every row in this table contains data from all four bands.

- Single-Epoch Photometry tables from four eras of the mission, comparable to LSST `Source`. Similarly to the MEP Table, the WISE imaging model produces measurements in each available band in each row. As the W4 and then W3 bands were shut down as the spacecraft warmed up, their columns were dropped from the table, leading to an evolution of the schema with time. The four eras and the associated official table names are:

  - All-Sky Single Exposure (L1b) Source Table, from the primary cold mission, with all four bands operating;

  - 3-Band Cryo Single Exposure (L1b) Source Table, from the brief period when the spacecraft was still cold enough to collect W3, but not W4, data;

  - Post-Cryo Single Exposure (L1b) Source Table, containing only W1 and W2 data, from the period following the primary mission when the cryogen was fully exhausted but operations were continued, after which the spacecraft was put into hibernation until 2013;

  - NEOWISE-R Year 1 Single Exposure Source Table, from the first year of operations after the reactivation of the spacecraft, and containing only W1 and W2 data.

The LSP-00 tests verified that these tables were present on the PDAC and exercised a variety of operations against them. The rest of this report covers the eight test cases into which this work was divided.

2 LSP-00-00: Verification of the presence of the expected WISE data

The tests were performed primarily from an Apple laptop computer running OS X 10.11.6, wirelessly connected to the Internet and using the NCSA VPN at `vpn.illinois.edu`. API Aspect queries were performed using a combination of the Python `requests` package and the `curl` command. Notebooks were executed using Jupyter versions up through 5.4.1 and Python version 3.6.1. The Portal was accessed using Safari 11.0.2.

2.1   API Aspect tests

The API Aspect tests made use of the "v0" `dbserv` service at

$$\texttt{http://lsst-qserv-dax01.ncsa.illinois.edu:5000/db/v0/tap/sync.}$$

This service accepts SQL queries, forwards them to Qserv for execution, and returns the results as JSON structures. The open-source Python *Requests* package[2] was used to perform the HTTP GET and POST operations required, and to parse the returned JSON structures into Python dictionaries.

2.1.1   Table presence verification

API Aspect `dbserv` queries (`SHOW DATABASES` and `SHOW TABLES`) were used to establish the set of databases available, and within those databases, identify the WISE catalog tables of interest for the tests.

WISE catalog	PDAC database	PDAC table
AllWISE Source Catalog	`wise_00`	`allwise_p3as_psd`
AllWISE Multi-Epoch Photometry Table	`wise_00`	`allwise_p3as_mep`
All-Sky Single Exposure (L1b) Source Table	`wise_4band_00`	`allsky_4band_p1bs_psd`
3-Band Cryo Single Exposure (L1b) Source Table	`wise_3band_00`	`allsky_3band_p1bs_psd`
Post-Cryo Single Exposure (L1b) Source Table	`wise_2band_00`	`allsky_2band_p1bs_psd`
NEOWISE-R Year 1 Single Exposure Source Table	`neowiser_yr1_00`	`neowiser_yr1_p1bs_psd`

TABLE 1: WISE tables as discovered through the PDAC DAX service

2.1.2   Table Schemas

Schemas for the tables were obtained from `DESCRIBE` queries on `dbserv`. The schemas were compared programmatically against those obtained from the appropriate IRSA program interface:

---

[2] http://github.com/requests/requests

```
https://irsa.ipac.caltech.edu/cgi-bin/Gator/nph-dd?mode=xml&catalog=table.
```

The results were parsed with the Python standard library class `xml.etree.ElementTree`. Comparison of the schemas was carried out in Python, without regard to the ordering of columns. The following discrepancies were observed (details are preserved in the test notebook):

- The `dec` column in each table was renamed to `decl` to avoid a conflict with a reserved word in the Qserv implementation database.

- The Qserv implementation adds special columns `chunkId` and `subChunkId` to each partitioned table.

- The original column names in the `wise_2band_00.allsky_2band_p1bs_psd` table all appear in the PDAC version converted to uppercase.

- For some tables, The IRSA-published schemas via the program interface have additional columns compared to what is displayed in the IRSA Gator UI and, apparently, in addition to what was included in the export files used to load the PDAC Qserv. This was assessed to be irrelevant to the objectives of this test, and no serious effort was made to determine the reasons in time for this report, as this would have required IRSA staff time that was in short supply. IRSA support tickets will be filed to ensure that the discrepancies are ultimately understood and, if possible and desirable, corrected. The affected tables are:

    - `allwise_p3as_psd`: 36 columns appear in the IRSA program interface schema that do not appear in the PDAC table. Most of these are associated with a proper motion model fit.

    - `allsky_4band_p1bs_psd`: 20 columns appear in the IRSA program interface schema that do not appear in the PDAC table. 16 of these seem to be intermediate data products of the photometry pipeline, including zero-point corrected PSF and aperture magnitudes and the zero points themselves.

    - `neowiser_yr1_p1bs_psd` (compared against the IRSA program interface for the "live" table `neowiser_p1bs_psd`): 24 columns appear in the IRSA program interface schema that do not appear in the PDAC table. These again appear to be a set of intermediate data products.

2.1.3 Row counts

Row counts were obtained from `COUNT(*)` queries against all six tables. These queries require accessing all chunks and subchunks in the Qserv partitioning and therefore resulted in table scans taking several minutes of wall clock time each. The row counts and representative query durations are shown in Table 2. The pattern of query times has not been fully understood; `allwise_p3as_psd` is a notable outlier. The row counts were compared to those advertised in the WISE documentation at IRSA. They agree exactly except for the AllWISE Multi-Epoch Photometry Table, `allwise_p3as_mep`, for which 42,759,337,365 rows are reported in the IRSA documentation, a difference of 761,630,509 or 1.8%. This discrepancy has not yet been understood.

PDAC WISE table	Rows	COUNT(*) duration
allwise_p3as_psd	747,634,026	798s
allwise_p3as_mep	41,997,706,856	1457s
allsky_4band_p1bs_psd	9,479,433,101	993s
allsky_3band_p1bs_psd	3,703,319,374	278s
allsky_2band_p1bs_psd	7,337,642,955	535s
neowiser_yr1_p1bs_psd	18,468,575,586	759s

TABLE 2: Row counts in PDAC WISE tables, with durations of the associated `COUNT(*)` queries

2.2 Portal Aspect

The Portal Aspect was accessed in order to verify the accessibility of the WISE catalog data.

This work covers step 6 of the LSP-00-00 test case procedure.

2.2.1 Step 6c

The front page of the Portal Aspect for this test was:

```
https://lsst-sui-proxy01.ncsa.illinois.edu/suit/
```

2.2.2   Step 6c

The front page presents an "LSST Data" screen which offers the catalog data available for access, associated with three "projects": SDSS, AllWISE, and the WISE and NEOWISE single-epoch photometry.

The radio buttons under the project selector show the tables available for selection. All six WISE tables covered by this test specification are available, as seen in the following screen shots.

FIGURE 1: AllWISE tables available for query in the Portal

Given a selected table, the UI presents a selector for the type of search to be performed, offering a variety of options for spatial region searches including cone, ellipse, box, and polygon, as well as a multi-object cone search and a free-form all-sky search. This UI element is shown in Figure 3 below.

2.2.3   Step 6c

The Portal also provides a UI element, shown in Figure 4 displaying the entire table schema in a scrolling region, with, among other features, the ability to apply restrictions to any column
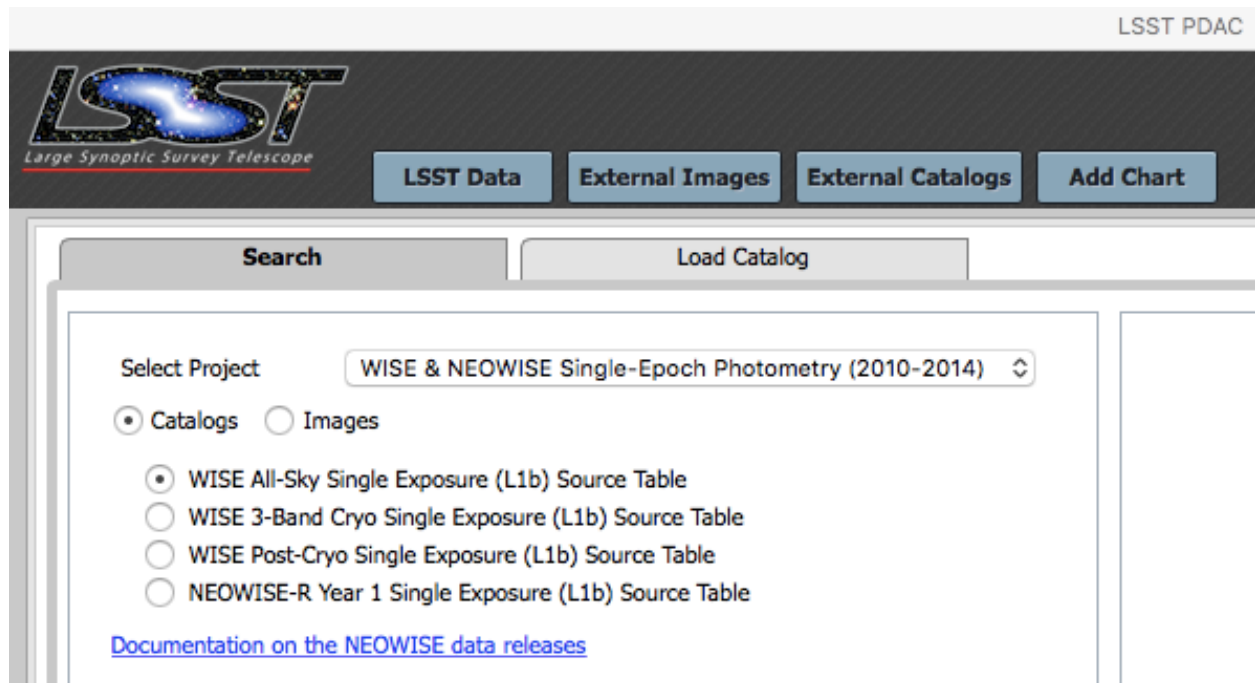
FIGURE 2: WISE and NEOWISE single-epoch photometry tables available for query in the Portal

for application in the query to be performed. It was not a goal of this test to specifically validate the correct display of the schema, but some spot checks were performed and the results were as expected. For instance, the schema display was observed to follow the expected pattern of whether W3- and W4-related were available in the selected table(s). (The nature of the scrolling region provided for this selector inhibits easy capture of the entire content. It would be useful for the final Portal to provide a UI element facilitating download of the entire schema.) The schema anomalies reported in section 2.1.2 above were also observed here, e.g., the uppercase conversion of the original column names in the `allsky_2band_p1bs_psd` column.

Note that units and descriptions are not provided in the UI, though the layout anticipates them; this is a consequence of the absence of a full `metaserv` implementation at the time of the tests.

The test specification calls for doing spot queries around (ra=0, dec=0); it turns out to be more useful to do them away from ra=0 so that in the default plots in the Portal plotting tool the points are not split across values just below 360 and just above 0. (The values can be rescaled to (-180,180), of course, but this requires an extra step.)
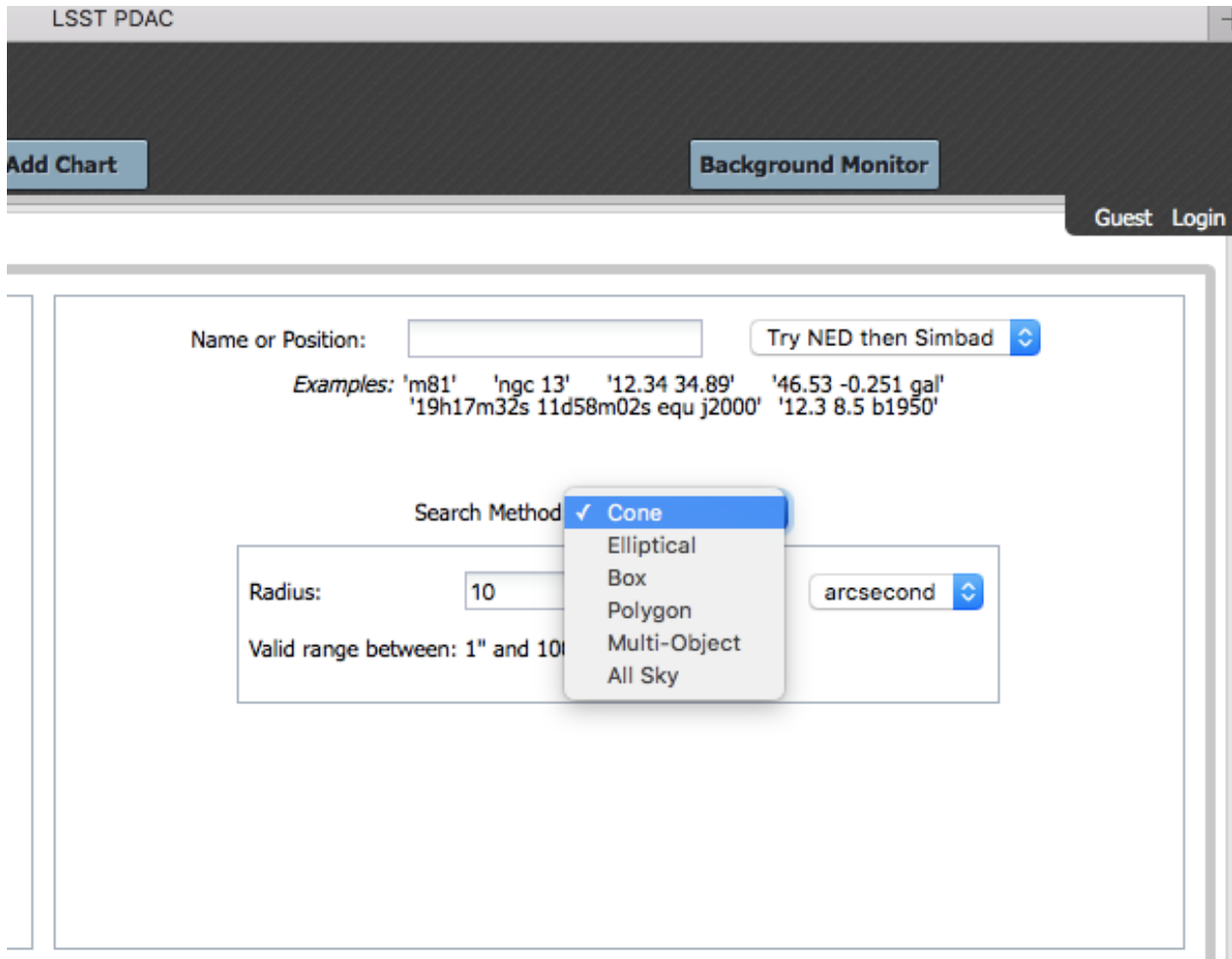
FIGURE 3: Search methods available for Portal queries on the WISE tables

FIGURE 4: Schema browser and column-restriction API in the Portal, for WISE 3-band table

100-arcsecond cone searches were successfully performed at (ra=1, dec=0) for all six tables. The `allsky_3band_p1bs_psd` search returned no rows, corresponding to this sky location not being included in the brief period of 3-band observation. Screen shots are provided for `allwise_p3as_psd` and `neowiser_yr1_p1bs_psd`, and tabular results for all five non-empty queries are stored in the DMTR-52 repository. Note the much larger number of sources for the NEOWISE-R single-epoch table query.
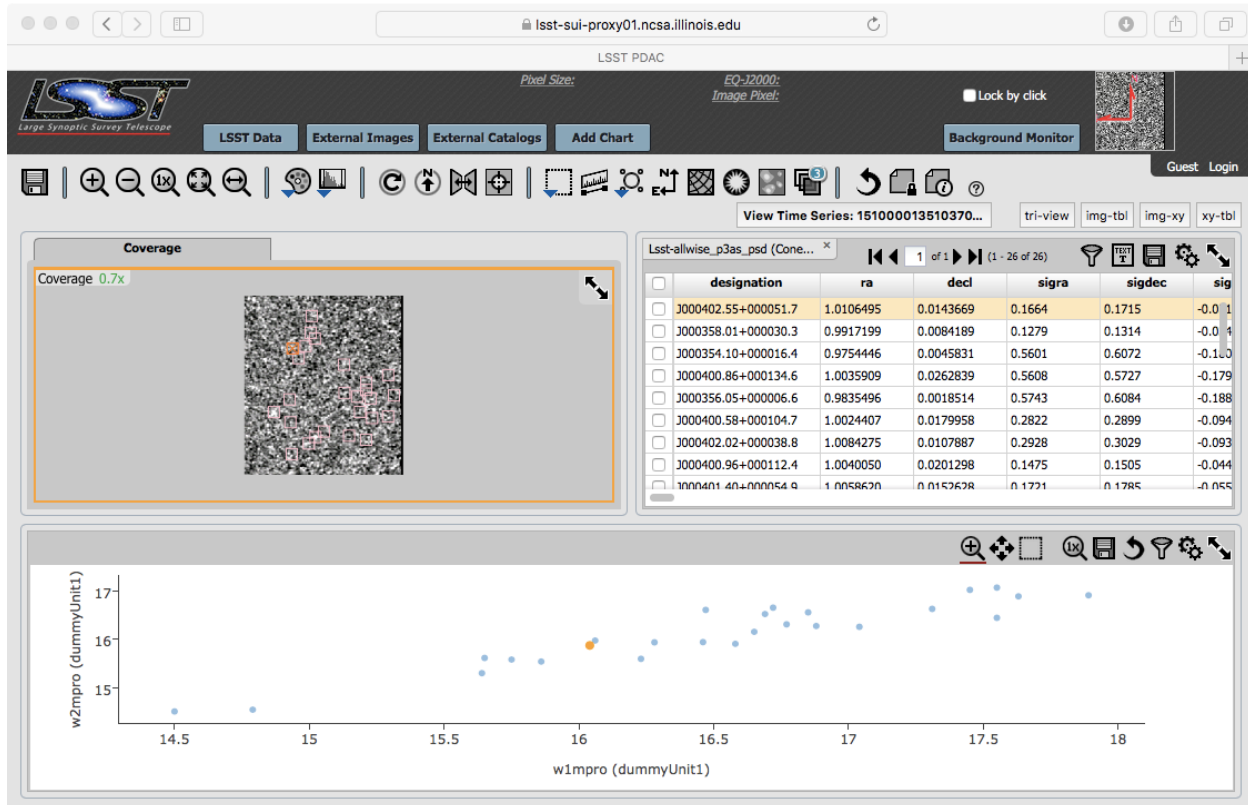

FIGURE 5: Query at ra=1, dec=0 on the AllWISE Source Catalog

3   LSP-00-05: Demonstration of low-volume and/or indexed queries against the WISE data via API

4   LSP-00-10: Demonstration of table-scan queries against the WISE data via API
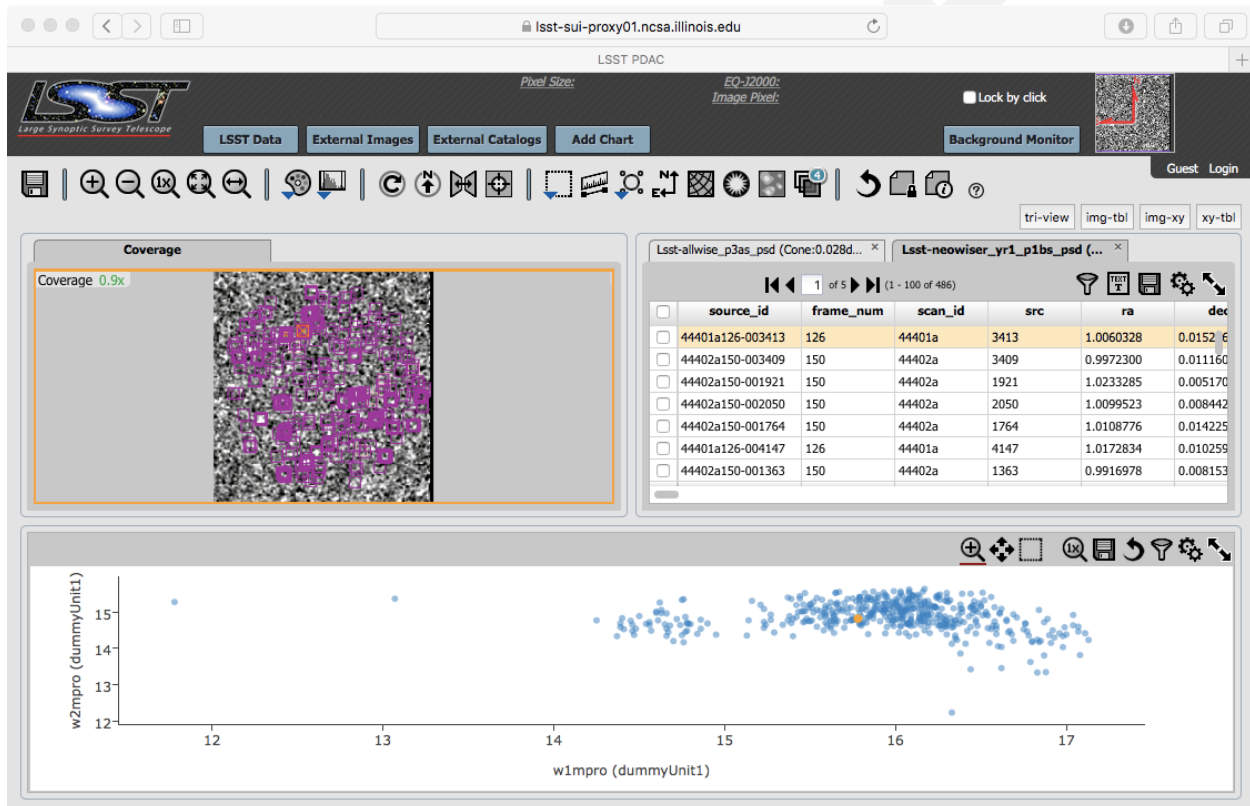
FIGURE 6: Query at ra=1, dec=0 on the NEOWISE-R Year 1 Single-Epoch Source Catalog

5   LSP-00-15: Execution of basic catalog queries in the Portal

6   LSP-00-20: Operation of the UI for interaction with tabular data results

7   LSP-00-35: Linkage of catalog query results to related catalog data