Next Generation Search Interfaces

William Roby,[1] Xiuqin Wu,[1] Loi Ly,[1] and Tatiana Goldina[2]

[1]*Caltech, Mail Code 100-22, Pasadena, CA, 91125-2200, USA*

[2]*3038 Old Glenview Rd., Wilmette, IL 60091, USA*

Abstract. Astronomers are constantly looking for easier ways to access multiple data sets. While much effort is spent on VO, little thought is given to the types of User Interfaces we need to effectively search this sort of data. For instance, an astronomer might need to search Spitzer, WISE, and 2MASS catalogs and images then see the results presented together in one UI. Moving seamlessly between data sets is key to presenting integrated results. Results need to be viewed using first class, web based, integrated FITS viewers, XY Plots, and advanced table display tools. These components should be able to handle very large datasets.

To make a powerful Web based UI that can manage and present multiple searches to the user requires taking advantage of many HTML5 features. AJAX is used to start searches and present results. Push notifications (Server Sent Events) monitor background jobs. Canvas is required for advanced result displays. Lesser known CSS3 technologies makes it all flow seamlessly together.

At IPAC, we have been developing our Firefly toolkit for several years. We are now using it to solve this multiple data set, multiple queries, and integrated presentation problem to create a powerful research experience.

Firefly was created in IRSA, the NASA/IPAC Infrared Science Archive (http://irsa.ipac.caltech.edu). Firefly is the core for applications serving many project archives, including Spitzer, Planck, WISE, PTF, LSST and others. It is also used in IRSA's new Finder Chart and catalog and image displays.

1. Firefly Overview

Firefly is IRSA's web-based Front end to access science archives. It provides sophisticated data visualization capabilities not available in many web-based archives. Our efforts to create a tightly integrated and user-friendly interface has generated numerous positive feedback from scientists. Users can browse and understand large volumes of archive data in a fraction of the time it took in the past.

AJAX technology gives us an opportunity to radically add value to basic data search and retrieval. In addition to searching and downloading, scientists can study, visualize, investigate, and begin to do science while browsing the archive. Not only can these advanced capabilities be an incredible time saver, but also improve the quality of the research.

Firefly is being reused for multiple archives because we make all the components configurable using XML. This has allowed us now to implement additional archive systems at a fraction of the cost of the first system, Spitzer Heritage Archive.

Firefly is implemented as a web application. This differs from a web page. With a web page, every time the display changes a new page is loaded. With a web application, there is just one page that is completely controlled by the JavaScript much like a desktop application. The Spitzer, WISE, Planck, and PTF archives all use this approach.

Firefly is also available through an API called Firefly Tools. Firefly Tools exposes the most powerful components of Firefly in a way that can be used by any web page with no

prerequisites. It allows any web developer access to Firefly's FITS visualizers, XY Plots or Table Tool with just a very few lines of JavaScript. This allows more classic style web pages to use the most important Firefly components with little change to the page. The IRSA Catalog search and some of the smaller IRSA image search datasets have been enhance significantly by using Firefly tools without requiring a major rewrite of the system.

2.   Major Firefly Tools Components

Firefly provides an environment for writing astronomy archive web-applications. The results of searches are viewed with very advanced visualizations components. These firefly components, each provide a different way to view the data.

2.0.1.   Tabular Display.   Firefly has implemented "Excel-like"Ï tables on the webpage. In an easy, interactive way, a user can sort the results, filter the data on multiple constraints, hide or show columns or select data for download or visualization. The Firefly Tools server is optimized to show very large tables without significant performance degradation. We have tested it with over 10 million rows, 50 column tables.

2.0.2.   Fits Visualization.   Firefly provides a first-class FITS visualization on the Web without any plugins. All of the significant components you would expect to see in a basic desktop FITS application are available with data that Firefly displays. The FITS visualizer allows for plotting anything that the table shows that has a Lon and Lat. Therefore, users can overlay multiple catalogs over their FITS image and interact with it.

2.0.3.   2D Line Graphs.   Firefly shows 2D line graphs interactively so that a user can read the data point values as he moves his mouse around or zooms in to investigate the data at a finer level. These graphs are used for spectrum or plotting table columns. The user can specify any column that the tables are showing. The user can also define equations from the rows for the XY Plot, such as (row1*row2/tan(row3)

2.0.4.   Data Model

Each of these visualizers work with the same data model. Therefore, any change made in any of the visualizers is reflected on all the others. If the user filters his data from the tables, then the filter view of the data is plotted on the FITS viewer and shown in the XY Plots. If a single point is highlighted in the FITS viewer then the same point is highlighted in the XY Plots and the row is highlighted in the tables. If there are Multiple fits viewers then the change is reflected in each one.

This gives the user significant power. The user can sort and filter his data, then see it in multiple types of interactive views. The data model will transparently decimate the data so that the browser does not get overwhelmed by large volumes of data. Drawing speeds will stay fairly consistent no matter the data size. See Figure 1.

3.   Advance Search User Interface

Because Firefly serves such advanced visualization tools, it brings us to the next step of how to create the more integrated user experience as he searches for data. We are considering the best way for the user to access the data he needs through one tool. As we move more towards an environment that is informed by VO protocols, we want to create a UI that allows that sort of flexibility.
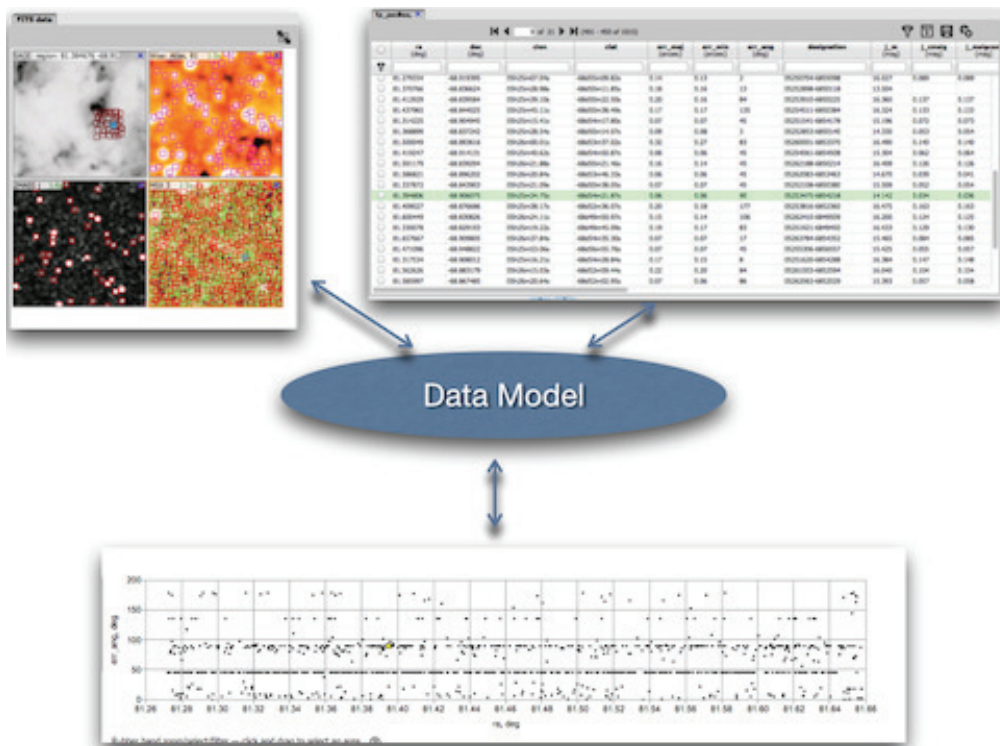
Figure 1. The FITS viewer, XY Plot, and Table all use the same data model. If data in anyone changes then it is reflected in all three. Data might change by highlighting, selecting, sorting, or filtering.

Search of data archives can be divided into two basic type of searches. The first is the inventory search. The user does not know exactly the dataset he wants to search, he is simply asking, "What is available?" There has already been a lot of effort put into this type of search.

The second type is the one we want to address. This is when the user is interested in some popular dataset, such as an all sky, survey or a space mission such as Spitzer or Hubble. In this case, the user is asking, "What does my favorite data set have in the area I am interested in? Our advance search UI attempts to implement this use case.

The typical UI for this type of search in the past has been to have a search form and a results area. The search form will produce one result view. The results of searches cannot be mixed. When the user want more data he must start over. See Figure 2.

The key to this type of searching is flexibility. We want to to be able to initiate searches and then the see the results presented together. Consider the following user case: I want to search 2Mass and WISE catalogs and see the results on WISE, Spitzer, DSS images.

More generally, a user will want to retrieve any of the catalogs or images dataset available and see them in one result viewer. The user can add or remove results at will but can see any combination together. In addition to this, all previous searches and results are remembered in a workspace. The user can go the workspace at anytime and retrieve something he did in the past.

Therefore, this web application must have three views. The Query form that can specify any dataset. The Results viewer that can show many datasets at one time. And finally the workspace that holds all previous searches and uploaded data.

Each of these views are always available. The user can move between each at will. This workflow allows the user to search and view exactly what he wants from one web application See Figure 3.

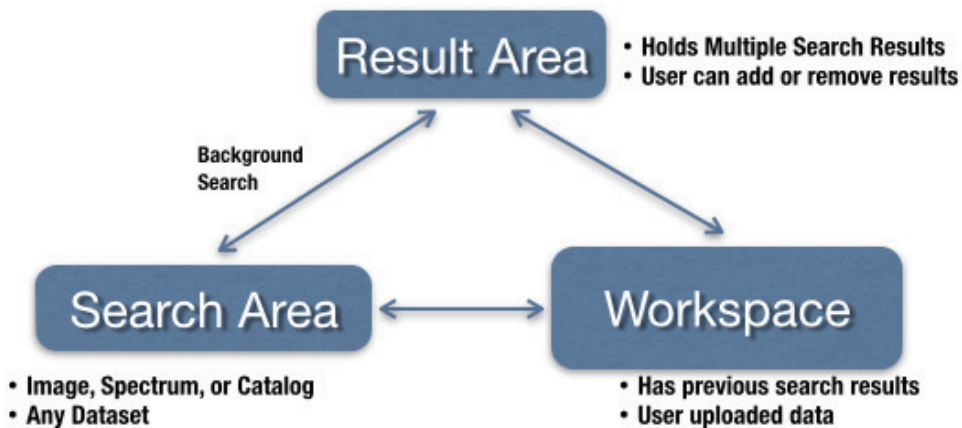Figure 2.   Classic web search flow.

Figure 3.   Any section of the search interface is accessible from the other with out destroying the view.

3.0.5.   Status

We are currently well into the process of creating this new approach to searching and viewing results. However, there is still some work to do. We will be rolling out a partial implementation from IRSA in the next few months. We will use also use this design as the starting point for our LSST work.