



Interoperable Time Series Access

Project Definition Document

Project ID:	1.4
Revision:	1.0
Project Lead:	John Good (IPAC)
QA&T Engineer	Dick Shaw
Document Status:	Final
Reviewed	Live Review: 27 Apr 2011 Approved: 16 Aug 2011 (RP)

1 Project Description

The time domain is often referred to as the last frontier in astrophysics. Although much is known about time-variable phenomena, from solar and stellar pulsations to gamma-ray bursts, our sampling of the data in the time domain is terribly incomplete. A number of new time domain surveys are now in progress or being planned which programs will undoubtedly give rise to discoveries of previously unknown time-variable phenomenon, and will enrich our understanding of already-known phenomenon by increasing our sample sizes by orders of magnitude.

There are already several significant repositories of time series data and much to be gained by properly interconnecting their data and services. The purpose of this project is to be a pathfinder for this kind of collaboration by connecting the data at the Time Series Center (TSC) at Harvard and the NASA/IPAC/NEoScl Star and Exoplanet Database (NStED) with services at both centers (specifically to start with the Periodogram service at NStED and the Classifier service at TSC).

This project is a little unusual in that it is not concerned with the end services *per se* but rather with the protocols and communications mechanisms needed to interconnect them. It is explicitly limited to the two named centers, though it is expected that the work will result in time series data and services everywhere becoming more interoperable. It is also not responsible for the scientific validity or speed of the end services, just ensuring that added interconnections produce the same result as would be returned if the interconnection were handled manually.

A simple example will help illustrate what is expected.

A user of NStED has used that system to identify a set of sources having particular attributes. They want to see if TSC has light curves for any of these and, if so, analyze them using the NStED Periodogram service in a mode tuned to detecting and characterizing transiting planets.

Without this work they would have to:

- Take the list of sources NStED provides and export it as a table of coordinates.
- Go to TSC and enter each coordinate. When a match, download the data file (TSC format).
- Translate the TSC data into a format NStED can understand.
- Upload the translated light curve to the NStED Periodogram, configure it for transits, and run the fit.



VIRTUAL ASTRONOMICAL OBSERVATORY

This project will allow the user to streamline and even automate this process by introducing a uniform standard for the light curve data and by taking the user out of the interconnection loop.

2 Project Interdependencies

This project is expected to depend on the following IVOA standards currently in development

- Anything relevant to a time series data model (TBD at the moment). Some of specific near-term activities (and even requirements) are only needed in lieu of having a formal time series metadata access specification and since it would be unrealistic to expect such a specification in Year 1 of this project.

Specifically, the first two requirements involve the sharing of metadata so that each center has access to information on the other's holdings. This may be fully replaced at a later date by standard (e.g. STSP) services that allow for uniform remote metadata mining.

Other dependencies

- This project is highly dependent on the two external project to update their existing services to support VO data formats and access methodologies (specifically the Simple Time Series Protocol once it is defined).

3 High-level Requirements

Req.#	Requirement statement	Verif. Method*
1.	Each of the two centers (NStED and TSC) must provide their metadata to the other in the form of a tab-delimited table containing at least three columns: the RA, Dec and a handle that can be transformed into a URL for downloading a light curve. In addition, they will provide each other with whatever formulae are needed to perform this URL transformation.	I
2.	Based on the shared metadata, each center must provide public position-searchable interfaces to the other's holdings which return lists of remote light curves. The details of where this capability will be used is a local decision and may evolve with time but most likely will be provided to the user as ancillary information wherever the local interface provides region search results.	T
3.	Each center must provide links on its interface to send its local data to the remote service (e.g. NStED will provide links on its results pages to send NStED light curves to the TSC classifier).	T
4.	Each center must provide documentation describing how to send third-party data directly to the processing services.	I
5.	When a formal IVOA specification is available for Time Series access, both the metadata sharing and time series file formats must be updated to conform.	T

*Verification Method Codes: T = Testing, D = Demonstration, I = Inspection, A = Analysis.

NOTE 1: The NStED Periodogram service already has information of this sort (<http://nsted.ipac.caltech.edu/periodogram/cgi-bin/Periodogram/nph-simpleupload>) but this will need to be revised when the data formats change.

NOTE 2: This may not happen in Year 1 but is important enough that it is appropriate to have a Year 1 requirement to do it if possible.

4 Roadmap

Information given here may be deprecated during the design phase.

This work is expected to result in the following types of products:

- Interconnected archives using new VO Time Series protocols.
- Publicly available services (Periodogram and Classifier) which can be sent standard Time Series data from any source.

The work for this project is expected to be delivered by:

- September 30, 2011