The Importance of Mid-Range Telescope Data Archives - an Example of Ondřejov Observatory 2-meter Telescope

Petr Škoda*

Astronomical Institute of the Academy of Sciences
of the Czech Republic

Abstract

There are less then ten ground-based optical telescopes in the world with the automatic web-based data archives. Most of them are huge highly over-allocated systems which allow only short-time (several days) observation of usually very faint targets.

On the other side there is a lot of mid-range telescopes larger than 1 m (some are also used for long-term monitoring of a relatively bright objects) that do not provide other archiving than simple bulk backups on tapes without querying or retrieval capabilities.

In last years the situation is getting much worse as these smaller telescopes are being decommissioned due to economic difficulties. Thus the valuable data may be lost. For improvement of this situation a simple SQL based archives with WWW GUI built at the smaller observatories would be of high value Such archives should be set up very quickly without extensive SW development.

In our contribution we present such a data archive which is under development in Stellar Department of Astronomical Institute in Ondvrejov for easy access to electronic spectra obtained by Reticon and CCD detectors of Ondrejov observatory 2-meter telescope. The system is based on cheap RDBMS Adabas D - Linux Edition for handling the observing logs and file headers. The data are presented on Apache web server by the WDBI form interpreter which uses the Perl DBI interface for communication with database. In the near future the archive will be enhanced by the simple automatic reduction pipeline and graphical output of reduced spectra. The design goals, user requirements and necessary trade-offs are discussed as well as the possibility of future improvements.

*skoda@sunstel.asu.cas.cz

1 The Importance of Astronomical Archives

Some of principal discoveries in astronomy were based on long term observational monitoring of selected objects by simple means and using the data by someone else (e.g. formulation of Kepler laws using the observing log of Tycho de Brahe).

For present astronomical research is, however, quite common the shift of interest to the huge telescopes that can allow due to great over-allocation only short glimpses of an (usually extremely faint) target. To allow more people to use these very expensive data, the integral part of such a huge telescope (or space mission) project is a public archive with a nice, mostly WWW-based, front-end. By deep searching of many sites mentioned in AstroWeb metaindex[1] we have found only several ground base optical telescopes with at least list of observation published on Web (if you know about another, please let us know): La Palma data archive of ING telescopes[2],CFHT archive[3], Anglo-Australian Telescope archive[4], ESO NTT and VLT archive[5], UK Schmidt Telescope plate archive[6], UKIRT archive[7] and Japaneese Mitaka-Okayama-Kiso observatories archive[8].

On the other side, there is a plenty of smaller telescopes (with diameter about 1 m and larger) on various observatories all over the world - some of them are often used for long-term monitoring campaigns (e.g. to investigate Be stars, recurrent novae, cataclysmic variable or other types of relatively bright objects with unpredictable behaviour). The data secured here are mostly given to the visiting astronomer and saved somewhere on tape but no one else can see the data even after proprietary period, simply because of absence of any archiving and presentation facilities - often even the observing logs are only in hand-written form.

In recent years the situation is getting much worse as smaller telescopes are being decommissioned to save money for building the huge super-telescopes or due to common economic difficulties. We have heard rumours about financial problems of telescopes at Observatoire de Haute-Provence, telescopes at Pic du Midi or some telescopes at Kitt Peak (2m and Coude feed) that might be soon closed in case of further budget cuts. An example of a productive telescope already decommissioned due to budget restrictions is the 1.5m Coude Auxiliary Telescope (CAT) at La Silla. As such a subsidiary telescopes have some sort of public electronic archive very seldom (we have found on the Web only the list of observing logs of 182cm telescope of Asiago Observatory[9] and a static list of spectra secured by Elodie system at OHP[10]) and there are no spare money and/or and skilled database specialists for building fancy archive from the scratch, the valuable data may thus be effectively lost for next generations.

The way out of this sad perspectives may be the building of simple public data archives based on small SQL databases that will allow at least querying the observing log to get information whether the the particular object was observed, in which instrumental setup and where to obtain the datafile (e.g. contact of data owner during and/or after the proprietary period, the internal file ID, tape or optical media name etc.). The more sophisticated engines may be able to handle the requests automatically and prepare the data on some staging area for FTP transfer by an external user. He may be asked to register prior to archive usage obtaining

[1] `http://cdsweb.u-strasbg.fr/astroweb.html`
[2] `http://archive.ast.cam.ac.uk`
[3] `http://www.cfht.hawaii.edu/ObsInfo/Archive`
[4] `http://site.aao.gov.au/arc-bin/wdb/aat_database/observation_log/make`
[5] `http://archive.eso.org/eso/archive.html`
[6] `http://www.aao.gov.au/local/www/apmcatbin/ukscat.html`
[7] `http://archive.ast.cam.ac.uk/ukirt_arch/`
[8] `http://moka.nao.ac.jp/index.html`
[9] http://leda.pd.astro.it/asiago/tel182/archive/log182.html
[10] `http://www.obs-hp.fr/ vin/ar1996.html`

the password for further access to raw data. The handling of passwords of external users as well as their e-mail notification when the data are available may be quite simple once the database system has already been set. Problematic situation exists in some observatories that do not have stated fixed rules for releasing the data to the public or even have no fixed proprietary period. Some principal investigators (especially those involved in research of long-term stellar variability) do not want to release into public use the observations more than five or even ten years old. The setup of the public archive is then the good reason to establish the institutional data politics.

Such a well established data archive may also help in saving of valuable data from closed telescopes or abandoned reserach programs by adopting them and making accesible for interested public (e.g. students and post-docs).

2 The Outline of Archive Setup

The automatic data archive with user friendly interface consists of several basic parts:

Relational database management system (RDBMS) with SQL query language. There are available expensive RDBMS with huge HW demands requiring complex administration like Oracle, Sybase, DB2, Interbase etc. as well as free of charge shareware or even GNU licensed small systems like MySQL, mSQL or PostgreSQL. Very interesting are the Linux ports of the commercial full-fledged system in its smaller Personal Edition version - like Sybase ASE or the ADABAS-D PE.

Database importing tool - it might be a bunch of scripts in shell, sed, awk or even compiled languages the task of which is the stripping off the (usually FITS) exposure headers, distinguishing the separate header items (according to keywords in FITS), making necessary value conversions and finally the importing of the selected values into database tables. At the start of the building of the archive the very handy may become the tool that can directly import values from older (non-FITS) logs in form of ASCII tables.

Web presentation tool - It is the subsystem allowing the inclusion of database item values into the HTML form and preparation and submission of SQL queries according to form input values. Most of the commercial RDBMS have such a proprietary tool (it is a quite common setup for electronic bussines). The most popular free solution is the scripting language PHP/FI and newer PHP3[11] implemented on almost all operating systems for a plenty of RDBMS backends. Very important is also its ability to compile it as Apache server module. The less pleasant is the necessity of programming the WWW user interface from the scratch. The quickest way of setting working WWW database interface is the prefabricated form interpreter WDB[12] originally written for ESO data archive by Bo Frese Rasmussen and enhanced under name WDBI[13] for table insertion, deletion and updating by Jeff Rowe. Although the visual design capabilities are somehow restricted the ease of use and extensibility (by writing additional PERL procedures for special data handling) is marvelous.

Auxiliary tools - for handling of database of registered users and sending them e-mail notification, quick-look data visualization or even automatic reduction

[11] `http://www.php.net`
[12] `http://archive.eso.org/wdb/html/`
[13] http://www.wdbi.net

pipeline.Using the Netscape plugins and MIME type capability list the browser can call e.g. IRAF `ximtool` or `xgterm` to work with FITS files just by clicking of file pointer in SQL query results table.

3 Ondřejov Observatory 2-m Telescope

The Ondřejov observatory, located 35 km (20 miles) south-east of Prague in the altitude of 534 m was founded in 1898 by J.J. Frič, the manufacturer from Prague, as a private observatory and in 1928 donated to the Czechoslovak state for the needs of the Charles University. In 1954, the Observatory merged with the Czechoslovak Academy of Sciences and was transformed into the Astronomical Institute of the Academy. In 1992, the Astronomical Institute became a part of the Czech Academy of Sciences.

The 2-m telescope built by Carl Zeiss, Jena, was put into operation in 1967 as a Czechoslovak national instrument and since the beginning it was devoted to high dispersion spectroscopy (17, 8.5 and 4 Å /mm) of early-type (mainly Be) stars in its 64-m coudé focus. Its plate archive contains almost 6000 spectral plates exposed until the beginning of year 1993. Scientists from Ondřejov stellar department came in late 60-th with the hypothesis of binary origin of the Be phenomenon and since then there had been a number of new spectroscopic binaries discovered using Ondřejov spectra.

At the end of the year 1992 were the photographic plates replaced by Reticon AF1872 detector with use of which almost 3650 stellar spectra were secured, giving with calibrations the total of 15000 frames up to date. In addition to normal observation with Reticon (still in use) the department had been ocassionaly testing a CCD camera for last three years. Due to the quite poor quality of its LORAL 2688x512 CCD detector the stellar department had decided to buy a new CCD camera with ISA 2000x800 chip. It should be put into operation at the half of next year.

The observing conditions in the low altitude and in the vicinity of Prague (strongly shining at the north-west horizon) as well as the local weather (about 130 partly clear nights in a year with average seeing about 2–4 arcsec) are incomparable with the world largest telescopes in high altitudes but the advantage of 2-m telescope is the ability of continuous long-term monitoring (as the weather permits) on a smaller number of stars. The observing program contains more than 200 stars, but only about 30 of them are monitored in a systematic manner (e.g checking the behaviour in particularly interesting phases). There are time series of some stars larger than 200 spectra over the span of 3-5 years.

4 Reticon Data Archive Requirements and Design Goals

From the specific assignation of the telescope as given above also follows the main goals of an archive. At present the most important for the research are the Reticon electronic spectra. As they are quite simple to handle and there exists programs for their reduction, graphic display and listing of exposures of a particular star widely used in stellar department (written in Turbo Pascal under DOS), it was a challenge to show to stellar astronomers the advantages of a database for querying the relations between exposures (e.g a task "find all stars with exposure time longer than 5 hours and ADU value less than 100 – is the star really as faint or is there mistake in header ?).

The most important task was the selection of RDBMS system. As we didn't have any experience with SQL and databases at all, we have refused the cheap public domain systems like MySQL or PostgreSQL (partly also due to the lack of some SQL features like HAVING clause or nested queries) and from the commercial world we have chosen the cheap ADABAS-D for Linux Entry Edition[14] (for less than 300 USD).

There exists even a free Personal Edition version (with restriction of the database size and number of users) which is a part of SuSE[15] Linux distribution. The database is a modern robust system with the features like transactions, triggers and even with syntax emulation of Oracle and ANSI SQL standards. It is well supported by Perl DBD: and DBI: modules as well as by PHP3.0.

The online tutorials and a quite detailed HTML documentation was very helpful at the setup phase as well as the FAQ with detailed examples. Adabas-D has also a nice table importer (can import from fixed format by stating the range of columns with particular data or by stating the given separator between values. There is also a nice graphical Tk based query and administration tool and the SQL console tools that my be started from UNIX shell or various other shells like Tcl/Tk, Perl, Java ... The distribution also contains ODBC drivers. and the commercial version contains a dynamic HTML generating tool called WebDB with example forms.

After testing all the available web presentation tools (WebDB, PHP3.0 and WDBI) we have chosen the WDBI. The setup is quite smooth on RedHat5.2 distribution and Adabas 10 (the necessary DBI and DBD:Adabas[16] modules work quite well), on RedHat 6.0 and Adabas 11.0 its a quite tricky (but achievable) so we can't recommend such a setup. As the database front end should be chosen `sqlserver_dbi.pl` modul. The form generation script `mkfdf` must be rewritten for Adabas but you can write it directly from the scratch using the example forms.

The more detailed information can be given by author of this poster and should available soon on his experimental WWW server[17]. The present state of the archive can be seen on its home page[18]. But the location may soon change.

The Reticon log file contains for each observation an 80 byte long ASCII row with the values at fixed positions. So it is quite simple to create one table (called RETLOG) and import all rows by the Adabas importer called `xdbload`.

As there is only very small number of target stars and their names are written in the unified way by data acquisition SW (choosing them from the observing program catalog) we may provide the user with the scroll list of all stars instead of some sort of name resolver like SIMBAD or NED. The same can be applied to the number of spectral filters and dichroic mirrors or the type of image slicer . All this can be handled automatically by the WDBI with the help of its `makelist` atribute.

The number of exposures is quite small so the database will work nicely even on small computers, and in case of database failure the data may be quickly reloaded from original ASCII tables. After some playing with the `html` atribute quite interesting query form may be produced. See the examples of Reticon Data Archive in presented figures.

In the near future we plan to add more columns of general interest computed from the present ones with Adabas internal PLSQL language (e.g. JD, HJD and heliocentric correction) and then we would like to prepare a simple processing pipeline. It should choose according to JD difference the nearest calibration with the same spectrograph setup (flat, offset, arcs) and then run the reduction producing the output picture directly on screen. The more complicated task will be the setup of

[14] `html://www.adabas.com/`
[15] `http://www.sus.com`
[16] `http://www.symbolstone.org/technology/perl/DBI/index.html`
[17] `http://syrma.asu.cas.cz/wdbi/adabas`
[18] `http://syrma.asu.cas.cz/index.html`

Reticon Data Archive of Ondrejov Observatory 2m Telescope

Please enter qualifiers in the fields below and press the '**Search**' button.

Search Reset Help Home

Return a maximum of 20 records per page from this query.

Report Full Detail
Report Number of Matches Only

Target ALL	**Exposure Type** ALL	**Observers**

Observation Date *DD.MM.YYYY* >04.11.1995	**UT Start** *HH:MM:SS*	**Exposure ID**

Exposure Time	**Count Value**	**ADU**

Slit/IS Type ALL	**Spectral Filter** ALL	**Dichroic Mirror** ALL

Grating ID ALL	**Grating Angle** ° ´ >30

Use full-page output.

Figure 1: Example of the 2m archive query form

2m-Telescope Reticon Archive Results

Search Criteria :

Target %
Observation Date > 19951104
Grating Angle ° ′ > 30

Matches: 80. 1 - 20 are shown on this page. Click on the button at the bottom of the page for more records.

Query Home

Target	Exposure Type	Observers	Observation Date	UT Start	Exposure ID	Exposure Time	Count Value	ADU	Slit/IS Type	Spectral Filter	Dichroic Mirror	Grating ID	Grating Angle °	′
offset	O		29.12.1995	03:30:21	9693	0	0	369	IS-RED	1	B	3	33	15
flat field	F		29.12.1995	03:32:23	9694	5	1.5	2831	IS-RED	1	B	3	33	15
ThAr15mA	C		29.12.1995	03:33:47	9695	2	0	3991	IS-RED	1	B	3	33	15
eta UMa	S	KubTl	29.12.1995	03:35:04	9696	430	0.8	768	IS-RED	1	B	3	33	15
ThAr15mA	C		29.12.1995	03:45:01	9697	3	0	4000	IS-RED	1	B	3	33	15
flat field	F		29.12.1995	03:45:25	9698	5	1.4	2575	IS-RED	1	B	3	33	15
offset	O		29.12.1995	03:48:40	9699	0	0	369	IS-RED	1	B	3	39	20
flat field	F		29.12.1995	03:52:50	9700	165	50	1558	IS-RED	1	B	3	39	20
ThAr15mA	C		29.12.1995	03:56:09	9701	7	0	3767	IS-RED	1	B	3	39	20
eta UMa	S	KubTl	29.12.1995	03:56:45	9702	939	2	116	IS-RED	1	B	3	39	20
ThAr15mA	C		29.12.1995	04:13:26	9703	8	0	3768	IS-RED	1	B	3	39	20
flat field	F		29.12.1995	04:13:57	9704	97	30	926	IS-RED	1	B	3	39	20

Figure 2: Example of the 2m archive results form

an archive of exposures taken by the new CCD camera (FITS headers, large disk space area needed - about 4 MB/frame).

5 Conclusions

The smaller telescopes at most observatories may have been producing quite valuable data by providing long-term coverage of a small number of brigher objects (especially variable stars). On the other side, the data are not stored in publically available archives (they are in best case only backed up to some tape or CD) and the external user (like student) cannot check whether the observation of particular object is available (and in given time).

In case of economical difficulties the instrument may be decommisioned and the data are thus effectively lost. To prevent this the quite simple data archives may be built using cheap (or free of charge) SW tools without large demand on administration and programming skills of database experts. With WDBI form interpreter even an skilled astronomer may prepare working archive of basic exposure information in several days. Such a archive may then adopt the observing logs of the decommisioned device or even is data.

Acknowledgments: We are grateful to Jim Lewis from RGO, Cambridge for giving the basic information about WDB and for showing us the detailed structure of the internals of the La Palma data archive maintained by him in RGO.

This work was partially supported with the subvention by the Ministry of Educati on, Youth and Sports of the Czech Republic (grant LB98251)