WFCAM SCIENCE ARCHIVE SOFTWARE

Nigel Hambly, Ian Bond & Bob Mann

Wide Field Astronomy Unit (WFAU), Institute for Astronomy, University of Edinburgh

October 8, 2002

Contents

1 INTRODUCTION

1.1 Background

WFCAM (see http://www.roe.ac.uk/atc/projects/wfcam/index.html) will enable the next generation wide–angle sky survey to be undertaken in the UK. It follows on from the hugely successful UK Schmidt photographic surveys of the last decades of the twentieth century, the major difference between the old and the new being the data rates and volumes that will be produced. WFCAM employs 4 2k$\times$2k Rockwell devices and has an instantaneous field–of–view of 0.21 square degrees. WFCAM is expected to be on–telescope for the greater fraction of all available UKIRT time, and will have average/peak data rates of 100/230 Gbytes per night. It will commence operations in the final quarter of 2003.

There is, of course, a clear need for 4m survey facilities in the era of 8m–class telescopes; the relative performance of WFCAM (as measured by its 'grasp', or information gathering product AΩ) shows (see, for example, the original VISTA science case, available from http://www.vista.ac.uk/) that it is amongst the world's leading IR survey instruments, even when including other non–dedicated survey facilities such as VLT–IRMOS. The combined science case (for complete details, follow the URL http://www.ukidss.org/sciencecase/sciencecase.html) proposed by the UKIDSS consortium for WFCAM, for example, details a programme that is unrivalled in terms of depth, field of view and therefore survey volume. UKIDSS proposes a nested series of surveys ranging from the Large Area Survey ('LAS', 4000 sq. deg. to K=18.4), the Galactic Plane Survey ('GPS', 1800 sq. deg. to K=19), the Galactic Clusters Survey ('GCS', 1600 sq. deg. to K=18.7), the Deep Extragalactic Survey ('DXS', 35 sq. deg. to K=21) to the Ultra–Deep Survey ('UDS', 0.8 sq. de.g. to K=23). The image data alone for these amounts to $\sim$ 50 Tbytes of data, while the object catalogue and ancillary information are likely to be many Tbytes in size.

1.2 The need for science archives.

The question naturally arises as to how science exploitation of such large datasets will be undertaken. Data volumes will simply be too large for users to download and keep their own copies. Raw data processing is likely to be complicated, while calibration procedures will evolve as cameras are better characterised and more calibration data are obtained. Reprocessing of substantial amounts of pixel data may be necessary in the light of improved algorithms or for specific 'non–standard' science goals. Once data are reduced using standardised pipeline procedures, the establishment of a centralised 'science archive' offers the greatest potential for full science exploitation (see the paper presented by Lawrence et al. at the 2002 SPIE meeting in Kona, Hawaii; available online at http://www.roe.ac.uk/~nch/wfcam/misc). Again, calibration procedures can be more easily developed and applied in a controlled manner to data in a central repository – it makes sense to solve data–specific reduction and calibration problems once, yielding an optimal solution. Early community access to well calibrated data will facilitate timely science exploitation. A well constructed science archive will enhance greatly the scope of research that can be done with the survey data; in fact, many science applications will only be feasible via a sophisticated science archive. For example, much of the science that will be done with the UKIDSS LAS will rely on complementary data from the SDSS and other non–IR wavelength surveys. Given the volume of all of these datasets, some thought needs to go into the design of the archive to enable full exploitation.

In outline, what should the science archive provide? As an absolute bare minimum, a simple web interface to pixels and object catalogues along the lines of WFAU's SuperCOSMOS Sky Survey (SSS; see http://www-wfau.roe.ac.uk/sss) provides community access to the data. However, full science exploitation requires much more. A well–designed archive should:

- enable control of releases/updates/calibration and software versioning;
- include well defined and documented science products (e.g. housekeeping 'meta' data, pixels, object catalogues, merged catalogues);

- enable generalised browsing/searching/querying of observations database;
- enable generalised browsing/searching of object catalogue data;
- enable cross–referencing with existing surveys (e.g. for WFCAM, optical catalogues like the legacy SSS and SDSS).

Additionally, we expect to follow new developments in web server protocols (e.g. VOTable – see http://vizier.u-strasbg.fr/doc/VOTable – a proposed protocol for exchange of astronomical data embedded in XML). Clearly, such functionality adds considerable value to the survey data by creating the potential for advanced science exploitation. Furthermore, in the era of 'e–science' and 'virtual observatories' we envision yet more advanced archive functionality:

- database 'federation' with any available database, e.g. UKIDSS UDS linked with data from the Subaru Deep Field and XMM datasets;
- data analysis tools, e.g. statistical and/or cluster analysis in N–dimensional parameter space;
- advanced visualisation on remote hosts, e.g. ability to pan around a wide–angle image at full resolution in real time;
- PSF analysis and fitting in image data; PSF matching for difference imaging;
- ability for user upload of analysis code, e.g. for pixel reanalysis, upload of code for detection of low surface brightness galaxies or other large–scale features;
- advanced image classification, e.g. galaxy morphological type
- reprocessing of large fractions of a pixel dataset.

In short, the WFCAM/UKIDSS science programme comprises surveys of unprecedented scope with data volumes that are hugely increased over previous experience in the UK. Moreover, user expectation as to the kinds of science that should be possible with these surveys is growing. Archive developments for this major new project will benefit *all* UK astronomers over the next few decades and will continue the traditional leading role of the UK in survey astronomy which underpins much of the research done here. The science archive is so much more than a simple repository of data – it is the means by which users will exploit the huge WFCAM data volumes.

1.3 This document

Ultimately, successful science exploitation of WFCAM will depend on user access to the large data volumes generated by this instrument. This document has been written for the purposes of the WFCAM software review to describe WFAU's solution to this problem. The intention is to outline the science requirements being placed on the WFCAM Science Archive (WSA) as a whole; discuss in more detail those requirements pertaining to WSA software; describe the current work at WFAU relevant to the implementation of a working WSA; and finally an overview of the design of the WSA as it currently stands. The WFAU WSA homepage is at http://www.roe.ac.uk/~nch/wfcam.

2 REQUIREMENTS

The specified requirements are available online at http://www.jach.hawaii.edu/~adamson/wfarcrq.html where more details can be found. These are the external, top–level requirements placed on the science archive by the 'customers', viz. JAC and UKIDSS.

2.1 Specified requirements

The WFCAM science Archive must provide rapid and straightforward access to WFCAM data from both the UKIDSS surveys and generalised open-time usage. The WFCAM science archive should:

Top–level requirements:

T1: provide the maximum possible potential for capitalizing on the UKIDSS surveys.

Rationale: UKIDSS amounts to 80% of WFCAM time and rapid exploitation is crucial to the UK's future strategy. Since exploitation is through the science archive, the need for T1 is clear.

T2: contain data from both UKIDSS surveys and open-time usage

Rationale: WFCAM data are likely to be far larger than the average UK University site can hold. Standard pipeline products held centrally are the solution.

T3: be flexible and copes with alterations in the UKIDSS survey design over time

Rationale: UKIDSS has an initial two-year allocation followed by review. This review may result in changes in both priority and design of the surveys. It is imperative that the science archive should not be a blockage in this process.

T4: not preclude usage from the GRID and inclusion in Virtual Observatory

Rationale: obvious, and clearly relates to T1

T5: allow simple and complex queries with appropriate interfaces for both

Rationale: a single complex interface suitable for data federation would be excessive when all the user wants is a K frame around a given target.

T6: be simple to use for PR purposes

Rationale: UKIDSS is a truly world-beating undertaking. The science archive will be the obvious point from which to draw PR material and should not be a hindrance to this.

T7: allow access to survey data before all observations are complete, and must not be disrupted by regular ingest of new survey data.

Rationale: This is consistent with rapid exploitation (and was a PDR requirement). Access to interim releases of the data will clearly be needed.

T8: allow arithmetic operations and options from the advanced processing toolkit on pixel data to be specified in requests

Rationale: multicolour imaging is perfectly possible given standard pipeline output and should not require the data to be transferred to the users home site just to stitch three colours together.

T9: Must be scalable to VISTA data volumes

Rationale: WFCAM and VISTA represent the UK's deep infrared survey capability. To use the same science archive design for both will ensure maximum exploitation of both.

T10: Must be able to merge reduced frames taken in non–photometric conditions with other data from the same survey

Science archive contents and functions (minimum)

C1: Contains calibrated object catalogues resulting from the pipeline, for both UKIDSS and open-time observations

Rationale: obvious, basic science archive function.

C2: Ingests and stores pipeline output frames for later online processing, generates compressed pixel images on the fly for rapid web-based access, carries out immediate cross–referencing with existing UKIDSS survey data and produces consolidated UKIDSS catalogue in a given field

Rationale: again basic science archive functionality. Compression by at least a factor of order 10 is necessary in the early stages, whether or not the need disappears later due to enhanced network speeds. It is assumed (but not required) that this will be lossy compression.

C3: Is able to recalibrate a given field or fields in the event of revised calibration information (specifically, photometric and astrometric), and allow database queries on the recalibrated quantities

Rationale: changes in calibration information are frequently encountered in survey operations, and the science archive itself may lead to such changes.

C4: Is able to cross-calibrate photometric information using areas of overlap between processed frames, where available.

Rationale: this is not a sensible function of the CASU pipeline, which is required only to produce results on a night-by-night basis. The science archive will have all photometric information and calibrations for all superframes, and is where this should happen.

C5: Allows public access to subsets of survey data on a variety of different search criteria (specified below)

Rationale: basic science archive functionality.

C6: Allows rapid on-line cross-referencing of search results with other catalogues

Rationale: consistent with T1, this requirement is expanded on later.

C7: Allows for generation of finder charts via a web form

Rationale: simple to provide and useful when observing at a site remote from the UK

C8: Holds housekeeping information for all archived data

Detailed requirements

Consistent with the above, the Science archive:

D1: must allow searching individual (or all) UKIDSS surveys on at least the following criteria (or combination of them):
Positional rectangle RA, Declination (J2000 coordinates)
Positional rectangle Galactic Coordinates (l, b)
Circular sky patch within a specified radius from a given RA, Dec
Circular sky patch within a specified radius from a given Galactic Coordinate
Circular sky patch within a specified radius of a resolvable source name
Source colour index (e.g. YJHK magnitudes and any linear combination thereof) between any of the available survey colours for the particular survey
Source parameter ranges (using any of the available source parameters)

Rationale: These are all basic science archive functions

D2: must allow searching within open-time programme data on the same criteria (where possible), returning whatever data are available (may only be image frames)

Rationale: open-time data become UKIRT science archive data as normal, and in many instances will be as searchable as data from the UKIDSS survey.

D3: must allow similar queries to be repeated for all objects in a given user-supplied source catalogue, or those in any catalogue also stored locally to the UKIDSS science archive.

Rationale: the user may have their own catalogue generated at a different wavelength, and may wish to obtain UKIDSS catalogue or image data at this list of field centres.

D4: must allow combinations of queries on a number of different source catalogues

Rationale: this requirement anticipates compound queries such as "return all UKIDSS image data around infrared sources with axial ratios greater than 1.5 and which have no detection in the SDSS". An arbitrary number of source catalogues should be allowed, subject only to their local availability.

D5: must allow functions [methods] to be used in setting up complex queries (i.e. for colour index if not stored separately)

Rationale: for consistency with T9.

D6: must have a web-based user interface to the object catalogue database implemented in an SX–like manner, and pixel database access implemented with in-house software. Ideally a simple interface for very quick searching on a given object name or position would be provided separately.

Rationale: clearly a web interface is needed, at least in the short term, and the SX similarity requirement follows from the desire to maximise the power of matching with Sloan for the LAS. For very quick returns of UKIDSS data around a given object, either the basic web interface must be straightforward enough to make this very quick, or a separate form should be provided.

D7: must allow plotting of returned parameters direct from the user interface, in selected pairs (x-y plots) or histograms (in the case of a single parameter)

D8: must return pixel images, confidence maps and catalogue data (where reduction recipes generate it) to the user in (at least) gzipped FITS format

D9: must be able to return pixel data in any available passband, over a contiguous field up to [TBD but one array across seems a sensible number], together with matched object catalogue

D10: must be able to generate (on–the–fly) and return larger areas from survey data, blocked down as required

D11: must be able to generate and return stacked images using a variety of stacking algorithms, where survey data are available

D12: must be able to generate and return merged multi-colour multi-parameter catalogues with best available astrometric and photometric calibration from survey data

D13: must support database federation with other source catalogues (initially SDSS and Schmidt sky survey)

D14: must be able to generate and return meaningful optical/IR colours for all objects in the overlap with SDSS, whether or not detected in the Sloan data.

D15: must support ANDing of a given query with another, where both have already been executed

D16: must support the returning of only a subset of the entire possible array of object parameters

D17: must allow trial-and-error searches (e.g. return number of source hits rather than entire catalogue), for any valid query

D18: must allow repetition of queries using previous versions of astrometric and photometric calibration

D19: must be able to produce a finder chart for any region within which survey data exist, returning both object shapes and positions (ellipse map) and a colour pixel image if suvey data in more than one colour are available.

D20: must allow access to best or duplicate data for objects in overlapping survey data (e.g. to generate light curves and/or proper motions from duplicate observations of the same object).

D21: must allow general access to all housekeeping data – e.g. for a given survey area, what is currently available, how good it is, etc.

D22: must specify calibrated quantities through coefficient values and calibration scheme

D23: must allow a summary of data available to be generated for a given search region

Security

A1: Archived data must be accessible only by validated users

A2: Archived data must be uncorruptable by science archive users and managers

A3: Allows data protection on the basis of proprietary date (per frame)

A4: Science archive must be quickly recoverable in the event of corruption by hardware/software faults etc.

Rationale: Clear need to insure against data loss (e.g. backup on removable media and/or 100% redundant storage with data striping)

2.2 Software requirements analysis

The software requirements for the WSA implied by the above are new to WFAU – previous *ad hoc* solutions based on flat–file storage and limited interactive network access (e.g. as implemented for the SSS) are simply not scalable to WSA data volumes nor WSA user expectations. In fact, these new data volumes are new to UK astronomy generally, and this is one of the main science drivers for the AstroGrid project (http://www.astrogrid.org/; we expect AstroGrid studies to inform our development of the WSA).

The software requirements fall into the following broad categories:

2.2.1 DBMS software

Top–level requirements T1–T10 imply a flexible DBMS with distributed processing/storage for data capacity and fast response times. Requirements C6, D3, D4, D13 (explicitly) and D14 require DBMS 'federation' of source catalogues and flexibility in schema design. Requirement D6 explicitly requires a DBMS that has an implementation of extensible SQL–like querying (see later). Overall, the requirements imply a commercial DBMS along with an operating system capable of implementation on high aggregate IO hardware, probably a distributed storage system. This in turn implies some degree of software parallelisation.

2.2.2 Ingest/curation software not catered for by the DBMS

Basic ingest, indexing along with tracking housekeeping data will be provided by any commercially available DBMS. Additional software is needed, for example to reformat object catalogue tables from FITS and generally strip housekeepong data (requirements C8, D21, A1, A3). Pairing software (over and above simple DBMS 'joins' between different tables) will be needed for robust merging of datasets obtained in different colours and at different times.

2.2.3 Analysis software

Pixel manipulation software is needed (requirement C2). Generalised pixel manipulation software to stack and/or mosaic images on–the–fly is required for user requests of image data at arbitrary positions and of arbitrary size (requirements D9–D11 and D19). Pixel manipulation/analysis algorithms from the basic pipeline (CASU) must be used where appropriate.

2.2.4 User/manager/curator interface software

Flexible querying (requirements T4, T5, C5, D5, D6, D15, D16, D21) implies use of a Structured Query Language (SQL) interface to the DBMS. Moreover, requirements T4, T8, D7 and D17 imply more enhanced functionality than is catered for in basic SQL. In general, preliminary WSA 'use cases' imply interaction with and manipulation of data from queries with a typical science use case sometimes requiring an iterative series of basic queries to achieve its goal. Software is required for remote client GUIs; web service software is required for communication between remote host and database server.

3 OVERVIEW

3.1 Current status at WFAU

WFAU has a proven track record in survey astronomy. We have been routinely handling survey data at rates of several tens of Gbytes per day for some time now, and have much experience in delivering survey products to the international astronomical community – see, for example, details of the SuperCOSMOS Sky Survey (URL given previously). WFAU has formally undertaken science archiving for *all* WFCAM data (not just UKIDSS data).

During the past 3 yr period, WFAU has been instigating R&D projects and establishing collaborative links to address large database problems. At the basic level, we have gained significant experience in the production, curation and user support of the SSS archive (which is ~ 2 Tbyte in volume). With an eye on more advanced database technology and the future requirements for large astronomical databases in the UK we have been following the development of the archive system for the multi–terabyte SDSS (see http://www.sdss.org/). This has produced a solution to the generic large astronomical database problem, and our proposed solution to the WSA development problem is based on this since we believe the SDSS system is unlikely to be bettered using available resources. WFAU has established links with the Johns Hopkins University group responsible for these developments, and has established a UK mirror for the SDSS–EDR. We are undertaking implementation of the SSS archive in SQLServer, and are collaborating with Jim Gray of Microsoft Research in the ingestion of this terabyte dataset to study the scaling issues. We intend to build on this, and our experience in creating the SSS and other similar products, to design and implement the WSA.

On the hardware side, we have employed monolithic RAID technology as a solution to mass storage and fast random access for the SSS. We have established an 8–node Beowulf cluster for our SDSS–EDR mirror, and have also acquired a 12–node rack–mounted Beowulf for experimentation with parallelisation techniques (both Beowulfs were funded through University and JREI sources). We have begun an analysis of the hardware requirements for the WFCAM science archive as part of an end–to–end data flow analysis and have made preliminary approaches to vendors, e.g. Eclipse Computing (providers of bespoke hardware solutions like our monolithic RAID system and Beowulf clusters) and Sun Microsystems. Finally, the Blackford Hill site is now connected to SuperJanet via two 1 Gbit s^{-1} network links as part of a SRIF upgrade award.

Funding for WFAU is mainly via PPARC Rolling Grant. We are currently half way through the 4 year award period 2000–2004, and have applied for funding for the overlapping period 2002–2006. WSA development forms a major part of the work for this period (the relevant section of the RG application, on which this document is based in part, is online at http://www.roe.ac.uk/˜nch/wfcam).

3.2 How the requirements will be met

An outline plan for the WSA development project is presented in our RG funding application, which details a set of work packages, including software design, that are currently being worked on. This work is in its early stages – we are currently working on the formal science requirements and resulting top–level UML design (http://www.roe.ac.uk/˜nch/wfcam). All further development work, including software work packages, will follow on from this. The next Section gives a broad outline of the WSA as we currently envision it, with the emphasis on softare solutions. We also discuss the feasibility of our proposed design.

4 DESIGN

4.1 DBMS

Originally we envisioned using Objectivity as the low–level DBMS in the WSA, echoing the implementation of the SDSS–EDR. However, due to some performance issues experienced by the SDSS science archive team at JHU, we are following the SDSS group and investigating Microsoft 'SQLServer' (a relational DBMS). We currently envision that the WSA design will therefore be based on SQLServer, but we note that we are not inflexible as to the final DBMS choice – we also intend to investigate the IBM's 'DB2' (again, a relational DBMS).

Feasibility of DBMS choice will come down to performance of the implementation of the SSS archive, which is similar in size to the projected size of the WSA. This work is being undertaken at the time of writing. From the software point of view, we do not anticipate any implementation problems for Tbyte–sized archives.

4.2 Ingest/curation software

In the same way as we have been writing code for data I/O and curation of the SSS, we will write similar codes for the WSA. These will be based on Fortran/C/C++ modules using standard subroutine libraries (CFITSIO, SLALIB etc).

4.3 Analysis software

A major consideration here is in balancing currently available astronomical data analysis packages with in house software components. Given the large volume of data that will be involved, large scale image and data analysis software will be written using a low level language. C++ is the language choice because it is object oriented and allows encodeing of very complex data structures.

Software modules include:

- Utilities for basic image manipulation operations such as pixel rebinning, resampling to new coordinate systems, etc. Image classes will be developed in C++ with appropriate I/O interfaces (based on CFITSIO in the case of FITS images).
- Source extraction, measurement, and classification tools. Existing software packages include DOPHOT, DAOPHOT, and SExtractor. It would probably also be necessary to develop bespoke software to provide additional functionality – an example here would be in the construction of source catalogues merged across the different passbands.
- Astrometric and photometric calibration tools. Probably requires a combination of existing and bespoke software.
- Difference imaging tools. These would be useful for constructing emission line images derived from subtracting a broad band (e.g. K) image from a narrow band image (e.g. H2). Also useful for variability studies. We have software that can generate difference images by optimally matching the seeing on the two images involved (see http://www-wfau.roe.ac.uk/sss/halpha).
- We also plan to implement software from drawing upon our astroinformatics research. Software has been developed that allows the identification of artifacts such as satellite trails and diffraction halos in CCD images (Storkey et al. 2002, in preparation)
- Image combining software. Tools for combining CCD frames taken in overlapping regions of the sky to produce single layer image mosaics and multiple layered images derived from stacking frames. This would involve development of special classes to deal with mosaic cameras such as WFCAM.

As far as on–the–fly applications are concerned, the feasibility of manipulation of large images is questionable when they reach Gbytes in size. All of the formal science requirements will be subject to the limitations of implementation constraints and network bandwidth. However, we suggest that no user would need or expect rapid response to use cases requiring manipulation of large imaging datasets.

4.4 User interface software

We are required to provide remote user access to the archive on a number of levels ranging simple form–based GUI through to advanced data mining, finally enabling the user to carry out remote data exploration, analysis, and visualisation. To enable this functionality, we envisage the following options:

- Form interfaces loadable in a Web browser. These could be developed using CGI scripting, but PHP is an interesting new technology that will also be considered.
- Remote client GUI based tools for advanced SQL querying of the archive. Probably implemented as a browser applet written in Java.
- Web services can provide users with remote data analysis capability by means of remote procedure calls embedded in SOAP encode envelopes and transported by http.

Web services implemented via SOAP can easily communicate across platforms. Web services also allow the users to make their own custom client interfaces in their own platform of choice which does not necessarily have to be on the same platform as the service. Web services will also enable the archive to be accessed as an AstroGrid resource.

Currently, we expect the WSA end–user interactive interface to be implemented using a SkyServer–like system, i.e. having the SQL capability of SQLServer along with a remote client GUI query tool (a browser applet probably implemented in Java). Web services will be implemented as detailed above.

Feasibility of such a solution is currently being assessed via our implementations of the SSS and mirror of the SDSS.

A APPENDICES

A.1 Acronyms and URLs

AstroGrid – the UK's astronomy data–grid project, which will form the country's contribution to the global VO
URL: http://www.astrogrid.org/

CASU – Cambridge Astronomical Survey Unit (operated within the Institute of Astronomy at the University of Cambridge, funded by PPARC Rolling Grant).
URL: http://www.ast.cam.ac.uk/~mike/casu.

CFITSIO – C/C++ subroutine library for low–level access to FITS data.
URL: http://heasarc.gsfc.nasa.gov/fitsio

DAOphot – community–standard software suite for multiple, simultaneous profile fitting photometry and astrometry for point sources in crowded fields.

DB2 – commercial relational DBMS from IBM

DXS – Deep eXtragalactic Survey component of UKIDSS

DBMS – Data-Base Management System

EDR – Early Data Release of the SDSS

GPS – Galactic Plane Survey component of UKIDSS

GUI – Graphical User Interface

IRMOS – InfraRed Multi–Object Spectrograph for the VLT

JAC – Joint Astronomy Centre at Hawai'i

JREI – Joint Research Equipment Initiative (set up by the DTI, HEFC and the UK Research Councils to provide strategic investment in key scientific infrastructure for research)

LAS – Large Area Survey component of UKIDSS

PHP – Preprocessor Hypertext Protocol

PPARC – Particle Physics and Astronomy Research Council

PSF – Point Spread Function

RAID – Redundant Array of Interchangeable Disks (robust mass storage with data striping, and hot–swappable disks).

SDSS – Sloan Digital Sky Survey
URL: http://www.sdss.org/

SExtractor – community–standard pixel analysis software suite for detection and parameterisation of sources in an image.

SOAP – Simple Object Access Protocol

SQL – Structured Query Language

SQLServer – Commercial relational DBMS product from Microsoft

SSS – SuperCOSMOS Sky Survey (WFAU's flagship legacy digitised optical sky survey based on Schmidt photographic plate material; the SSS database is ~ 2 Tbyte in size making it one of the largest object catalogues currently publicly available).
URL: http://www-wfau.roe.ac.uk/sss

UDS – Ultra–Deep Survey component of UKIDSS

UKIDSS – UKIRT Infrared Deep Sky Survey
URL: http://www.ukidss.org/

UKIRT – United Kingdom Infrared Telescope

VISTA – Visible and Infrared Survey Telescope for Astronomy (the UK and ESO's new 4m survey telescope to be located in Chile; this will have an infrared camera larger than WFCAM but will not start operating until 2006/7).
URL: http://www.vista.ac.uk/

VOTable – Vitrual Observatory Table; a proposed transfer protocol for astronomical data in the era of datagrids, e–science and datamining.
URL: http://vizier.u-strasbg.fr/doc/VOTable

VLT – Very Large Telescope; and array of 4 8m telescopes operated in Chile by ESO.

VO – Virtual Observatory: new concept in e–science where data– and compute–grids are linked to form a data mining resource for research in astronomy

WFAU – Wide Field Astronomy Unit (operated within the Institute for Astronomy at the University of Edinburgh, funded in the main by PPARC Rolling Grant).
URL: http://www.roe.ac.uk/wfau

WFCAM – Wide Field CAMera; the new infrared imager for UKIRT
URL: http://www.roe.ac.uk/atc/projects/wfcam/index.html

WSA – WFCAM Science Archive; the point of access for the astronomy community to obtain WFCAM data and undertake astronomical reasearch.
URL: http://www.roe.ac.uk/~nch/wfcam

XML – eXtensible Markup Language

XMM – Xray Multiple–Mirror telescope (spacecraft Xray observatory)