# The WFCAM/UKIDSS data archive : problems and opportunities

Andy Lawrence[a], Nigel Hambly[a], Bob Mann[a]
Mike Irwin[b], Richard McMahon[b], Jim Lewis[b]
Andy Adamson[c]

[a] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland, UK
[b] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 OHA, England, UK
[c] Joint Astronomy Centre, 660 North A'ohoku Place, University Park, Hilo, Hawaii 96720, USA

## ABSTRACT

The UKIRT Wide Field Camera (WFCAM) is an IR mosaic camera that represents an enormous leap in deep IR survey capability. It will be used as both an open time facility, and to perform a public IR Deep Sky Survey (the UKIDSS project), starting in early 2004. Here we present current plans for the data archive system, which will be provided as a standard service for all UK WFCAM data whether private or public survey data. The data rate is an order of magnitude larger than any previous survey experiment. WFCAM is therefore a crucial stepping stone between current day surveys such as SuperCOSMOS, APM and SDSS, and future facilities such as VISTA and the LSST. Pipeline processing presents a technical challenge, but the strongest challenges come in operation and curation of such a pipeline and of the rapidly accumulating database. For the public archive, there is little technical challenge in simply storing the data, and the real challenge comes in the rapidly increasing expectations of the user community for the kind of on-line services available with the archive. We describe three levels of archive service and the challenges they present, and discuss the hardware and software solutions we are likely to deploy.

**Keywords:** Surveys, databases, mosaic cameras, infra-red, pipeline, archive.

## 1. INTRODUCTION

WFCAM is a wide field IR array camera being built for the United Kingdom Infra Red Telescope (UKIRT) in Hawaii. It will be used partly for open competitive time, but mostly to undertake a large public survey programme known as UKIDSS (UKIRT Infrared Deep Sky Survey). A pipeline reduction system is under construction, which will provide standardised data products, to open-time users as well as the UKIDSS team. These standardised products will be ingested into a public on-line science archive with sophisticated query and analysis facilities. In organisational terms, there are three separate, though closely interlinked projects : (i) WFCAM construction, (ii) UKIDSS survey design and implementation, and (iii) pipeline and archive construction. This paper concentrates on the nature of the WFCAM archive, but first summarises the camera construction and public survey projects.

## 2. THE UKIRT WIDE FIELD CAMERA (WFCAM)

WFCAM is under construction at the United Kingdom Astronomy Technology Centre (UKATC) in Edinburgh, and is scheduled to be delivered to Hawaii in September 2003. The project scientist is Mark Casali, and the Project Manager is Dave Lunney. The work is undertaken in collaboration with the Joint Astronomy Centre (JAC), who operate UKIRT, and in part through an MOU with Japan, who are providing some of the detector arrays. WFCAM should start regular science operations at the beginning of 2004. The project website is at `http://www.roe.ac.uk/atc/projects/wfcam`, and a summary of current status is in a recent SPIE paper.[1]

WFCAM has the widest field of view of any IR camera so far attempted, and being placed on a 4m class telescope, is fast enough to make to large deep IR surveys that are a scientific match to optical surveys such

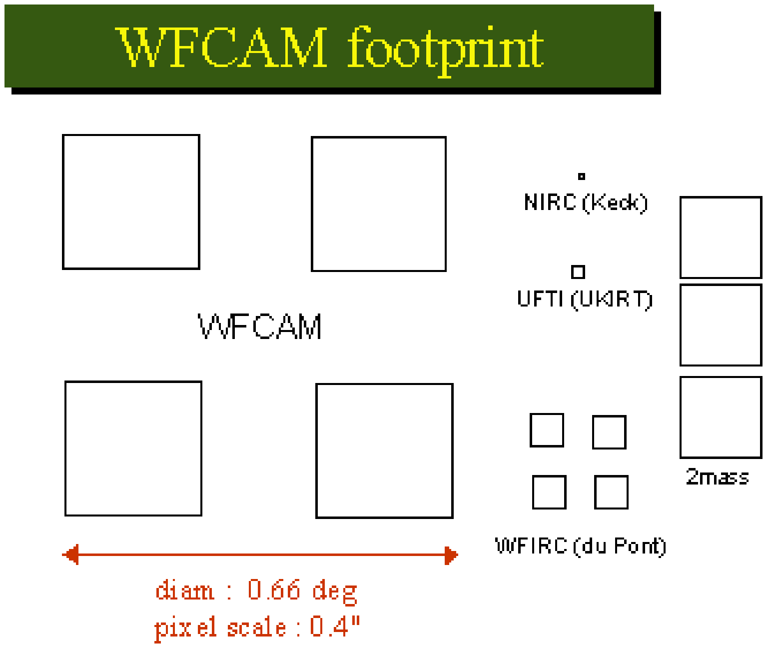


**Figure 1**: The WFCAM footprint on the sky compared to other IR cameras.

as the Sloan Digital Sky Survey (SDSS). It is based on a novel optical design using a quasi-Schmidt camera placed forward of Cassegrain focus that gives a 1° diameter field of view. It will be used in conjunction with UKIRT's tip-tilt secondary, which primarily corrects for wind-shake and dome seeing and is expected to deliver median seeing of 0.4″ across the whole of this field of view. The camera employs four Rockwell PACE $2048^2$ HgCdTe arrays operating across the JHK spectral region, with 0.4″ pixels, and placed at 90% spacing, giving an instantaneous coverage of 0.19 sq.deg. Figure 1 shows the footprint on the sky compared with other IR cameras, which spells out what a radical advance WFCAM is, especially bearing in mind that the 2MASS detectors are on a telescope with an order of magnitude less collecting area than UKIRT.

The pixel size of WFCAM is somewhat undersampled with respect to the expected median seeing, but micro-stepping, as employed successfully with 2MASS, can produce well-sampled images. It will be possible to dither/mosaic WFCAM in a variety of ways, but a standard filled-in "tile" can be made by making 2×2 micro-steps and 2×2 macro-steps, resulting in a final image with 0.2″pixels and diameter 0.87°, including small overlap regions that are observed more than once. In the K-band, exposures will not be more than $\sim$ 10 sec in order to avoid sky saturation, and this is likely to be taken as a standard stare time. Very little time is lost between micro-steps, but the macro-steps use telescope movement and involve a finite time to settle. With a standard 10 sec exposure, we expect the net observing efficiency to be $\sim$ 65%. The complete filled-in tile with sixteen separate exposures therefore takes approximately four minutes. Given the well known UKIRT performance and the expected throughput of WFCAM, then for point sources and an assumed seeing of 0.4″, and to a sensitivity of $5\sigma$, we expect to reach a depth of K=18.4 in such a four minute observation - fifty times deeper than the 2MASS survey. In a ten hour night, and allowing two hours for calibration, one could survey 115 square degrees.

## 3. THE WFCAM SCIENCE PROGRAMME

From 2004 onwards, the major instruments on UKIRT will be WFCAM, UIST, and Michelle. WFCAM will take the majority of the UK-owned time (two thirds), scheduled in advance in large blocks, in order to enable major surveys to be carried out. Of that WFCAM time, approximately 80% will be used for the UKIDSS survey (see below) and the remaining 20% will be open to competitive allocation through the PATT system.

The UKIRT Infrared Deep Sky Survey (UKIDSS) is the second major organised activity connected with WFCAM, along with the camera project itself, and the pipeline/archive project. The aim of UKIDSS is to undertake a substantial public survey programme over seven years, designed and implemented by a large consortium of astronomers at many different institutions. The UKIDSS web page is at `http://www.ukidss.org`. The science goals are summarised in two recent papers by Steve Warren.[2,3] The Principal Investigator (PI) is Andy Lawrence (Edinburgh University) and the Consortium Survey Scientist (CSS) is Steve Warren (Imperial College). Seventy astronomers are members, mostly from UK universities, plus a small number from Japan. The consortium meets regularly and every one has full voting rights on key issues, but in practice most of the work is done by a small number of working groups. As part of joining ESO, the UK has agreed that all the data from this public survey programme will be made available to astronomers from all ESO member states. We are now actively encouraging additional ESO members for both the full consortium and key working groups. The consortium has no proprietary rights. The policy specified by the UKIRT Board is that survey data will be freely available to astronomers from all ESO member states in the first instance, and to the rest of the world eighteen months after each successive release.

The UKIDSS project is not an end-to-end private enterprise, like SDSS. Neither is it a traditional public service programme, operated entirely by an observatory on behalf of a community. Rather, it is a kind of altruistic super-open-time proposal. The telescope and camera are public UKIRT facilities, whose construction, maintenance, and operation are the responsibility of UKATC and JAC. The UKIDSS consortium is a collection of private individuals who proposed to the UKIRT Board that a very large survey should be carried out by themselves, but on behalf of the whole community, with no proprietary ownership of the data by the consortium. The agreed responsibilities of the consortium are for scientific design and implementation of the survey, in the context of a working camera, and a working data pipeline system (see below). The UKIDSS concept was first proposed in 1998, alongside the camera project itself, with a final full scientific proposal[4] to the UKIRT Board in March 2001 requesting 1200 nights spread over seven years. The programme is approved subject to rolling review and should begin in early 2004.

The scientific goals of UKIDSS are summarised elsewhere in these proceedings,[2] but we will give a brief flavour here. It is not a single survey, but a suite of five surveys with a kind of "wedding cake structure". There are three very large shallow surveys totalling almost 7000 sq.deg. to a depth of K=18-19 : the *Large Area Survey (LAS)* covering a significant fraction of the SDSS area; the *Galactic Plane Survey (GPS)* covering much of the Milky Way visible to UKIRT at $b = \pm 5°$; and the *Galactic Cluster Survey (GCS)*, covering many key large open clusters. Next there is a smaller but deeper survey, the *Deep Extragalactic Survey (DXS)* covering 35 sq.deg. to K=21, over four selected areas with substantial multiwavelength data available. Finally there is the *Ultra Deep Survey (UDS)* covering 0.75 sq.deg. to a depth of K=23, centred on the Subaru Deep Field, which also has a wealth of very deep multiwavelength data available or imminent. Each one of these component surveys is two orders of magnitude larger than exisiting surveys to comparable depth. The prime purpose of all the surveys is to produce a legacy database that will fuel many scientific projects for years to come, but the design has been optimised with a number of specific goals in mind. These include detection of z=7 quasars and the epoch of re-ionisation; discovery of the very nearest and smallest sub-stellar objects; the shape of the initial mass function, and its environmental dependence; large scale structure and the value of the equation of state parameter $\omega$; the growth of large scale structure at z=0, z=1, and z=3; the abundance of high redshift ellipticals; and the connection between EROs, AGN, and high-z ellipticals.

## 4. THE WFCAM PIPELINE/ARCHIVE PROJECT

All UK-owned WFCAM data, whether public UKIDSS data or open-time private data, will be processed through a pipeline producing standardised data products, and ingested into an on-line science archive. The UKATC, as builders of the camera, will provide software for operation of the instrument, and extraction of data from it. Following this, the responsibility for the pipeline and archive system belongs jointly to the JAC, and the two UK rolling-grant funded wide field astronomy groups, the Cambridge Astronomy Survey Unit (CASU) of the Institute of Astronomy, Cambridge University, and the Wide Field Astronomy Unit (WFAU) of the Institute for Astronomy, Edinburgh University. In addition, the National Astronomical Observatory of Japan (NAOJ) at Hilo, Hawaii, and Mitaka, Tokyo will collaborate to develop further advanced pipeline facilities. The JAC has end-to-end responsibility. CASU is taking prime responsibility for the pipeline processing, and WFAU for the science archive, but the boundaries between these activities is blurred, and all teams are working closely together. Finally, the pipeline/archive project is formally integrated with the VISTA Data Flow System (VDFS) project, which is led by Jim Emerson at QMUL.

All WFCAM observations will be in a queued mode, using UKIRT's new Observing Management Protocol (OMP) system. Furthermore, there will be very few possible observing modes. For both UKIDSS and open-time observations we will be enforcing fixed standard calibration procedures. All these simplifying factors make a standard pipeline possible for all WFCAM data. However, no standard pipeline will ever squeeze all the possible information out of the data, and there will always be users, or types of question, that require different assumptions or algorithms in the processing. We are not attempting to construct an all-purpose completely flexible pipeline toolkit, but rather a processing pipeline to produce pre-agreed standard data products. These should be good enough for most purposes, but the design is optimised to produce the survey products demanded by the UKIDSS project, with optimal stacking, mosaicing, and source extraction, and uniform astrometric and photometric calibration across survey fields. The Science Requirements Document (SRD) for the pipeline and archive is under construction as we write, and is expected to be agreed with the JAC and with the UKIDSS consortium by the end of 2002. More advanced processing, and an on-the-fly user-pipeline toolkit may also be constructed, but these are over and above the commitment to the standard processing.

The processing can be seen as divided into several stages - data acquisition, the summit pipeline, the basic pipeline, further survey-wide processing, ingestion to the archive, refinement of calibration, and serving the data to users. At the summit, data from all exposures within a single night at the same telescope pointing position (including micro-steps within that position) and using the same filter, are co-added on the spot by the *Data Acquisition System (DAS)*. Regardless of the macroscopic dither pattern through the night, data from the four arrays are kept distinct, as are data from different pointings, so that the data written to media consist of a collection of $4096^2$ co-added *frames* within one night. These frames are then the analogy of traditional plates and form the basic units of the archive from which all other data products are constructed. The main purpose of the *summit pipeline* is to generate near real-time Data Quality Control (DQC) information from the co-added frames. It will use fixed library frames for instrument signature removal (e.g. flat fielding) and will do a first cut source extraction. The reduced frames, DQC, and a statistical analysis of the source lists, will be examined in Hawaii by UKIRT and or UKIDSS staff and used to update a *survey progress database*, which then feeds back to the observing queue.

The raw data (i.e. the collection of co-added 4K frames from each night) will be sent to the UK on a daily basis for processing with the *basic pipeline* in Cambridge. The default plan is that the data will be sent on a nightly tape, but we are discussing the possibility of sending a hot-swappable hard disk drive, as used by the ESO NGAST system. The basic pipeline will use the same software as the summit pipeline, but it will process the real calibration data, and will estimate a separate PSF from each frame. The pipeline removes instrument signature, does a first cut photometric and astrometric calibration, and a default source extraction. The result is a calibrated version of the collection of 4K frames, and a separate source list for each frame.

A series of *further processing steps* is needed to make final survey products, but these steps can only be carried out as the survey data accumulates in the science archive, so they need to be especially carefully planned between the CASU and WFAU teams. The most obvious thing is optimal stacking of matching frames from different nights, and mosaicing to produce a final large pixel-map for each survey. (Once again the first aim is a standard single pixel map, but we will also develop the facility to build images of any given sky-area on the fly from the constituent frames using different sampling and stacking choices). The next step is improved PSF generation including variations over the field and within a frame stack. Next there is improved source extraction from the stacked data, including detection and parameterisation of Low Surface Brightness Objects and transient events. Then we have *pairing* of sources across catalogues in different filters to make YJHK colours, and pairing with objects in external catalogues, such as the SDSS. Finally, as large area surveys are accumulated, we will revisit astrometric and photometric calibration looking for systematic gradients and step functions, and making external checks, and eventually, over several years, deriving proper motions and variability parameters.

Finally the data is ingested into a public science archive housed in Edinburgh. This will have both interactive and batch modes, and all data in it will be well calibrated and documented. It will contain all the final survey pixel maps and paired source-lists, but will also contain all the constituent nightly 4K frames and their standard one-filter source lists, as well as a browsable database of the available frames. As a minimum, the user will be offered the ability to download arbitrary subsets of these data, including on-the-fly mosaics of small areas using specified combinations of frames. However, we expect to offer rather more functionality, as discussed in the section after next.

## 5. THE DATA VOLUME CHALLENGE

Current day sky survey databases, such as DPOSS, SuperCOSMOS, or SDSS, are all roughly 10TB in size. From 2004 onwards, we expect WFCAM to be producing 20 TB/year of archived raw data. The leap upwards is because of smaller pixel size, short exposures, and repeated coverage. From 2006, we expect VISTA to be producing perhaps 100 TB/year. A few years after that, if and when the LSST becomes a reality, it may be producing 1 PB/year of data we wish to save. WFCAM then is an important intermediate step and test case. Are there significant problems associated with this increased data volume ?

We need to carefully distinguish different kinds of data. A single $2048^2$ image with 4 bytes per pixel takes up 16MB. We will be producing four of these simultaneously every 10 seconds, and may also be employing 1 second non-destructive readouts in order to have useful information on 2MASS stars in the field which might saturate in 10 seconds. The DAS therefore needs to cope with a maximum rate of 12MB/sec. But not all of this is output by the DAS or saved. The peak rate for saved data is given by assuming a collection of four 4K frames each of which is an interleaved set of 2×2 micro-stepped 2K frames, with individual 10 second exposures and an observing efficiency of 0.65. If such a shallow survey mode, with no co-adds, is carried on continously for a 14 hour winter night, we get a *peak recorded data rate* of 230 GB/night. A considerable part of the science programme will involve such shallow surveying, but on the other hand much will involve co-adding at a single sky position in any one night. If we take the planned UKIDSS programme as representative of all WFCAM use, and take an average night as 10 useable hours, then we get an *average recorded data rate* of 100 GB/night. Given the expected useage of WFCAM, the *raw frame data* then grows at $\sim$ 20TB/year. Following processing, there will be calibrated versions of each frame, plus housekeeping information, confidence arrays, and object catalogues, which may in total add up to 30 TB/year for the *calibrated frame data*. After mosaicing and stacking, the final derived *UKIDSS survey maps* are $\sim$ 50 TB in total. The density of objects will vary greatly from one location to another, but the final *UKIDSS survey object catalogues* are expected to be about 5 TB in full, although much smaller stripped down catalogues with only a few key parameters will also be available. Table 1 summarises these data rates and volumes.

**Table 1**: Estimated WFCAM data rates/sizes

| | |
|---|---|
| Peak data flow at instrument | 12 MB/s |
| Peak recorded data rate | 230 GB/nt |
| Average recorded data rate | 100 GB/nt |
| Raw frame data accumulation | 20 TB/yr |
| Calibrated frame data and ancillary data accumulation | 30 TB/ yr |
| 2010 archive : frame data | 350 TB |
| 2010 archive : stacked survey maps | 50 TB |
| 2010 archive : survey object catalogues | 5 TB |

Are these rates/volumes scary or not ? In terms of pipeline processing hardware, the rates are fast but not very intimidating. A small cluster will be needed to keep up. Likewise the archive storage does not present serious technological problems, but some moderate expense. Much larger cost and logistical problems are involved in the human operation and curation of the pipeline and archive, along with quality control and science calibration. Between now and 2010, this is where most of the money goes. What about the software development for the pipeline ? If we were building from scratch, or if we were building a completely flexible general purpose pipeline, this could easily eat up many millions of pounds. This is largely why our first priority is a well defined standard pipeline, although we are building this in a modular fashion to allow a cookbook approach to alternative reductions. Likewise the cost of archive software development is strongly dependent on ambition, so we are approaching this in layers (see below). It is also important that we are building on the existing experience and algorithms of CASU and WFAU. For example, the processing of Schmidt plates with SuperCOSMOS by WFAU may seem old fashioned, but actually involves a similar data rate to SDSS (20 GB/day), quality control and feedback to the observing team, removal of instrumental signature, source extraction and photometric and astrometric calibration, proper publication of experimental procedures as well as assessment of survey products,[7,?,?] and ingestion into an on-line science archive (`http://www-wfau.roe.ac.uk/sss`). Likewise, CASU as well as having similar experience with APM, has been responsible for processing and archiving data for the CCD mosaic public INT Wide Field Survey (INT-WFS)[8–10] (see `http://www.ast.cam.ac.uk/ wfcsur/index.php`), and already has a working pipeline for a smaller four-array IR camera, CIRSI.

## 6. USER EXPECTATIONS IN THE VO ERA

Storing hundreds of TB of data on spinning disk is not expected to be a major technological challenge, though it will not be a trivial expense. If Moore's law continues, storage will become easier as time goes by. The expensive part of an archive system is the human effort in operation, curation, calibration and documentation. But what about software development of the on-line archive service ? This rather depends on the level of service offered, and the expectations of users are increasing considerably as we enter the Virtual Observatory (VO) era. One can distinguish three levels of on-line archive service : data access, complex queries, and database manipulations. We expect to provide at least the first two for WFCAM, and in collaboration with AstroGrid, the UK VO project, we hope to provide service at the third level.

Straightforward *data access* is the standard today. One can offer either distinct data subsets such as plates or frames, chosen from a browsable list, or pixel maps and catalogues over user-defined areas, created from some survey database. Typically this is done through a web-browser and CGI script interface, with the user seeing a JPEG image, and then offered download of a FITS file. For such data subset access, serving a large volume doesn't make much difference, as long as the data is sensibly indexed. Access for arbitrary small datasets normally takes only a few seconds. Download of larger datasets to users is limited by network bandwidth, not by service at the data centre. The most common technical solution is storage on a RAID disk array speaking to a web server. The data can be stored in and interrogated through a Data Base Management System (DBMS), but actually flat files and home grown software are often used and work very fast. (Such systems however don't transfer well to the *next* database.)

With the second level of service, *complex queries*, one can construct questions in SQL or a similar language, along the lines of "give me a list of objects redder than X in this area of sky, with measurable proper motion, that have such-and-such quality flag better than Y (unless R-mag is brighter than Z, in which case accept anything), and that were found on a Tuesday". This kind of service is becoming slowly more common now. An example is the SDDS science archive, used either with the user-installed SDSS qT query tool, or through a web browser interface (see *http://archive.stcsci.edu/sdss*). Another example is the 2dF Galaxy Redshift Survey web page run from Mount Stromlo (`http://www.mso.anu.edu.au/2dfGRS`). This requires a proper DBMS of some kind (Objectivity for SDSS, soon changing to MS SQL server, and miniSQL for the 2DFGRS) and some sort of user-tool interfacing between the astronomer and the DBMS. To be able to answer arbitrary complex queries on very large multi-parameter datasets efficiently needs intelligent indexing and cacheing. Even then, there will always be occasions when very large numbers of table rows need to be searched through sequentially and only brute force will do. Most such requirements are CPU-disk I/O limited, and sometimes seek-time limited. Even at 100 MB/s, a 10 TB database takes over a day to search. The growingly accepted solution is to have a cluster with many CPUs searching in parallel. Such a *search engine* can provide the catalogue storage and the CPU power at the same time.

The leading edge now, which can be seen as Level-2a, is in *federated queries*, i.e. the ability to make joint queries of arbitrary databases distributed round the world - e.g. “give me all the objects in the UKIDSS LAS survey which were not seen in the SDSS but do have an X-ray ID in either a Chandra or an XMM observation, and check the list against the ESO VLT observing log”. This is the key problem being tackled by the various VO projects worldwide, and involves standardisation of data and metadata, but also use of new standardised “web service” data exchange methods with XML formatted data, SOAP message wrappers, and service description with WSDL. It also needs some kind of astronomical registry service, and a standardised method of ‘single sign on” using some kind of digital certificate rather than a multiplicity of passwords. The prospects of the necessary technological solutions being in place by the end of WFCAM’s first year of operation are good, so we anticipate participating in these kind of federated queries, and indeed will work closely with the AstroGrid project in particular to make the WFCAM archive and associated services VO-ready.

The third level of service involves *large database manipulations*. What we have in mind is things like calculation of correlation functions, cluster analysis in N-D parameter space, making statistical digests so that one can find objects $5\sigma$ outside the main clump, visualisation and exploration of multi-faceted datasets, and so on. Today, such data-intensive calculations are the province of specialised “power users” on their own machines, but we expect that such calculations will increasingly be provided as a standard, fast, service at the data centre, and that it will become common to do exploratory analysis this way as well as rigorous calculations. Furthermore, because of the increasingly large archive volume and network limitations, it will be more practical to use a service provided by a data centre than to download huge amounts of data and hack your own code. Such data-intensive calculations usually need $\mathrm{N}^2$ or $\mathrm{N}\log N$ algorithms, so PC farms, with slow interconnect between nodes, are too slow, and one needs a proper multi-processor SMP machine. In other words, to offer this kind of service, then as well as a PC farm search engine, one needs an expensive *analysis engine*, and facility-class data analysis software to go with it. This is a major challenge but one we hope to work towards in combination with AstroGrid and other VO projects.

## 7. DESIGN CHOICES

We are in the middle of a design process, but expect to freeze implementation decisions quite soon. A Science Requirements Document, specifying both data product definition and archive functionality, is being developed with JAC and the UKIDSS consortium and should be agreed by the end of 2002. We have been experimenting with a variety of DBMS packages, and will make a final choice in March 2003. Our working assumption is that we will follow developments by the SDSS team in order to save development effort and also to simplify UKIDSS - SDSS cross matching. This will probably mean using MS SQL Server and an updated equivalent of the SDSS query tool. We expect to do further development in parallelisation of SQL server, and an improved user interface. On the hardware side, we will be making choices by December 2002. It is likely that we will implement storage and search engine simultaneously with a cluster of roughly 50 nodes, in combination with a single large RAID array. (We may also have backup nearline storage such as DVD towers or a tape robot). In addition we expect that our analysis engine will be a share of a high powered multi-processor SMP machine, through a collaboration with the UK National e-Science Centre (NeSC), which is a joint venture of Edinburgh and Glasgow Universities. This variety of hardware will by default hang off our Local Area Network, but we have also been in discussion with some vendors over the possibility of a more integrated approach using a dedicated fibre-channel SAN. As well as the obvious cost issue, a key driver in our final hardware decisions is the requirement to find a solution that is (affordably) scaleable to VISTA and beyond.

## 8. CONCLUSIONS

WFCAM represents a leap forward technically, scientifically, and in terms of data volume. We intend to produce standardised data products from pipeline processing of all WFCAM data, including making reliable final survey products for the UKIDSS project. All data will be made available through an on-line science archive, and we expect this to be the main route through which science will be done with WFCAM. Pipeline processing and archive provision present only modest technological problems even for such a large database, but require significant human resource in operation, maintenance, curation, calibration, documentation and so on. The software development problem is as large as one cares to make it, especially for on-line archive service. The simplest levels are relatively easy to implement, but more and more sophisticated kinds of query, federated query, and data analysis services are being expected by users. Our aim is therefore to build improved service incrementally, but also to work explicitly in conjunction with funded VO projects.

## REFERENCES


1. D.Henry, et al, *Design status of WFCAM : a Widefield camera for the UK infrared Telescope*, **Proc. SPIE, 4841,** in press, 2002.
2. S.Warren, et al, *Scientific goals of the UKIDSS project*, **Proc. SPIE, 4836,** in press, 2002.
3. S.Warren, *The UKIRT Infrared Deep Sky Survey becomes an ESO public survey*, **ESO Messenger, 108**, 31-34, 2002.
4. A.Lawrence, S.Warren et al, *A proposal for a large public survey programme using the UKIRT Wide Field Camera*, paper to UKIRT Board March 2001. Publicly available at `http://www.ukidss.org`
5. N.Hambly et al, *The SuperCOSMOS Sky Survey : I. Introduction and description (User Guide)*, *Mon.Not.R.astr.Soc.,* **326,** 1279-1294, 2001.
6. N.Hambly et al, *The SuperCOSMOS Sky Survey : II. Image characterisation, parameterisation. classification and photometry*, *Mon.Not.R.astr.Soc.,* **326,** 1295-1314, 2001.
7. N.Hambly et al, *The SuperCOSMOS Sky Survey : III. Astrometry*, *Mon.Not.R.astr.Soc.,* **326,** 1315-1327, 2001.
8. R.G.McMahon, N.A.Walton, M.J.Irwin, J.R.Lewis, P.S.Bunclark, and D.H.Jones, *The INT Wide Field Imaging Survey (WFS)*, **New Astronomy Reviews, 47**, 97, 2001.
9. M.J.Irwin and J.R.Lewis, *INT WFS pipeline processing*, **New Astronomy Reviews, 45**, 105, 2001.
10. M.J.Irwin *Detectors and Data Analysis Techniques for Wide Field Optical Imaging*, in 7th Canary Islands Winter School, *ed J.M.Espinosa*, 1997.