VISTA UK Data Flow System

Review Panel Report

November 28, 2006

TABLE OF CONTENTS

1 Introduction

The review of the UK VISTA Data Flow System (VDFS) took place at the Institute of Astronomy in Cambridge on October 24-25, 2006. Review panel members were Gavin Dalton, Matt Jarvis, Tim Naylor, Robert Nichol and Luc Simard (Chair). CASU and WFAU staff were present to give presentations and answer questions from the review panel. VDFS Project Manager Jim Emerson and PPARC observer Colin Vincent were also in attendance.

VDFS documents were submitted for review on September 30th ahead of the face-to-face meeting to give time for panel members to read the documents and submit review items that warranted further discussion. Most of the first day was devoted to presentations and general questions while the second day focussed on specific review items.

The review panel would like to congratulate the VDFS team for all their hard work that led to this very successful review. They have designed a system that ranks amongst the best in the world even surpassing in some aspects larger efforts such as the Sloan Digital Sky Survey. We all look forward to the first VISTA data!

This report is organized as follows. Section 2 lists the documents submitted for review, Section 3 discusses general issues and recommendations, Section 4 goes over the Terms of Reference, and Section 5 focusses on review items.

2 Documents Reviewed

The list of documents submitted to the review panel is given in Table 1 below.

Document No.	Title
VIS-SPE-IOA-00009-0001	UK User Requirements
VIS-SPE-IOA-20000-0002	The VIRCAM Calibration Plan
VIS-SPE-IOA-20000-0010	Data Reduction Library Design
VIS-SPE-IOA-20000-0017	UK Pipeline General Design Overview
VDF-TRE-IOA-00011-00001	On the Calibration of WFCAM Data from 2MASS
VDF-TRE-IOA-00011-0002	Web-based pipeline and survey progress monitoring
VDF-WFA-VSA-002	Science Archive Science Requirements Analysis
VDF-WFA-VSA-003	Management and Planning
VDF-WFA-VSA-004	Pipeline / Science Archive Interface Control
VDF-WFA-VSA-006	Science Archive Hardware/OS/DBMS Design
VDF-WFA-VSA-007	Science Archive Database Design
VDF-WFA-VSA-008	Science Archive User Interface
VDF-WFA-VSA-009	Science Archive Software Architecture Design
VDF-WFA-VSA-010	Science Archive Integration into the VO

Table 1. List of review documents.

3 General Issues

3.1 Source Code Availability

Astronomers do not like to use “black boxes”, and the Panel shares their aversion. CASU gives stable and robust software modules to WFAU. The same modules should be made public in a form (i.e., source code) that allows people to examine them. This would have multiple advantages: (1) CASU software would gain “mindshare” in the community, (2) it would force good documentation and (3) it would leverage the community’s help in stamping out bugs. A big CASU concern (shared by the Panel) is that community requests for support will overwhelm CASU staff. This situation should obviously be avoided by including a very strong disclaimer to the effect that *no* support will be offered with the software.

Recommendation: Source code of software used to produce data products distributed to users should be released. A strong disclaimer clearly stating that no support will be provided should be included.

3.2 Relevance of VDFS Products to UK Users

The Panel notes that, for WFCAM, some teams are not using pipeline-processed stacks and catalogues. The Panel also notes that two VISTA survey teams plan to do their own stacking and object detection. The Panel has some concerns regarding what level of VDFS processing will be relevant in five years if user feedback is not folded back into the pipeline right now. At least one more iteration on requirements is needed.

Recommendation: The Project should make sure that VDFS running on WFCAM data can reproduce the catalogues produced by the various UKIDSS surveys so that once VISTA data have had their instrumental signatures removed, VDFS will produce the correct (and desired) data products on a timescale that retains a UK advantage.

3.3 Data Releases

There is potential for different releases from the same data as plans currently stand: VDFS, ESO and survey PI. ESO releases will essentially be the Survey PI releases.

All survey PIs have chosen to use VDFS.

Recommendation: Iterations with users should strive towards the goal of making VDFS and PI products the same! In the event that different data products will be produced, they should all be published next to each other through the same VSA interface with clear documentation/explanation to avoid user confusion.

3.4 Survey PI Expectations

Regarding VDF-WFA-VSA-002, page 26: How many expectations can be realistically met? What is the cost of a given expectation?

To ensure the UK community makes the most of the VISTA data, we strongly recommend that one additional iteration be done with the survey PIs and the user community. The wish list should include level of effort estimates to help with prioritization.

Recommendation: Survey PIs should be invited to join VDUC. This will ensure that VDFS products will fulfill the needs of its core user communities.

3.5 Funding

The six-month bridging request is clearly needed. This extra time is a great opportunity to close the loop on requirements with survey PIs.

Recommendation: In a choice between old requirements and survey PI requirements, we strongly urge the VDUC to give precedence to the PI requirements. To avoid a major wholesale modification of requirements, the survey PI wish list should include level of effort estimates and costs to ensure that this re-prioritization does not negatively impact the schedule and resources of the VDFS project.

4 Terms of Reference

1. Review the design and plans for deliverables in the UK of the VISTA Data Flow System (VDFS) project, specifically the pipeline data reduction, data archiving and curation and the delivery of data products to the UK user community.

The VDFS project team has done a great job at hitting big problems head on first. They have built a state-of-the-art processing/archiving system that is better than the SDSS system in many respects.

2. Assess the degree to which the requirements and goals, as stated in the UK User Requirements Document, are planned to be met, also taking into account the views of the VDFS Users Committee (VDUC).

Views from VDUC are to be collated later, but the Panel could not identify any major failures to meet requirements. However, the Panel was concerned that despite the requirements for WFCAM having been met, survey teams were still doing some of their own data processing. Hence, some of our recommendations in Section 3.

3. Assess the degree to which the deliverables are compatible with the Virtual Observatory and Astrogrid.

The panel views the proximity of the Astrogrid unit to WFAU as an ideal situation to ensure that VDFS deliverables will be (and will remain) compatible with the Virtual Observatory and Astrogrid even as these data exchange/distribution protocols evolve rapidly.

4. Assess the degree to which the design is efficient in terms of implementation and operation, taking into account the need also to handle WFCAM data and to deliver a VISTA data reduction library to ESO.

There seems to be sufficient headroom in implementation and operations of the VDFS to meet other commitments such as ESO software deliverables at the same time.

There is some concern regarding the proliferation of complex queries with time. These queries could eventually overload the system as the database collections expand.

Risk register for both CASU and WFAU should include the risk of a lack of person power in the face of mounting data processing/archiving pressures.

We commend the VDFS team for achieving significant savings by implementing WFCAM and VISTA jointly.

5. Comment on the feasibility of implementing the design within the budgetary and time constraints.

Limited information was provided to the review panel here. Core functionalities can be delivered in time for VISTA operations. “Headroom” should be carefully monitored to avoid the unpleasant consequences of things going wrong.

5 Review Items

Sixty-seven review items were sent to the VDFS team by Panel members. The list of all these review items is given in Table 2. Each review item included a recommended action from the relevant panel member. Section 5.1 includes individual review items with conclusions from the Panel.

Review Item ID	Document	Description
DALTON01	VDF-SPE-IOA-00009-0001	Focal plane layout
DALTON02	VDF-SPE-IOA-00009-0001	Mini filters
SIMARD10	Pipeline and archive docs	User requirements for pipeline and archive
DALTON03	VIS-SPE-IOA-20000-0017	Defocussed images
NAYLOR2	UK VISTA User reqn	Algorithm documentation
NAYLOR3		Access to precision hierarchy
NAYLOR4		Photometry in highly reddened fields
SIMARD9	VIS-SPE-IOA-20000-0017	Only calibrated science products will be released
JARVIS1		Raw data
NAYLOR1		Assessing astrometric accuracy
NICHOLA	VIS-SPE-IOA-20000-0017	Photometric accuracy
DALTON04	VIS-SPE-IOA-20000-0017	Clipped RMS
DALTON05	VIS-SPE-IOA-20000-0017	8% PS completeness
JARVIS3		Overall accuracy of photometry for crowded field and close pairs/blended objects
DALTON06	VIS-SPE-IOA-20000-0017	Federation of object catalogues
DALTON07	VIS-SPE-IOA-20000-0017	Flat fielding
DALTON08	VIS-SPE-IOA-20000-0017	Fringing
SIMARD11	VIS-SPE-IOA-20000-0017, VDF-WFA-VSA-003	Facilities: Physical space, cooling and power
SIMARD12	VIS-SPE-IOA-20000-0017	Data availability
JARVIS11	VIS-SPE-IOA-20000-0017	Human resources for CASU
NICHOLC	VIS-SPE-IOA-20000-0017	Risk register
NICHOLD	VIS-SPE-IOA-20000-0017	Risk register (same as NICHOLC)
DALTON09	VIS-SPE-IOA-20000-0017	Circular reference
DALTON10	VIS-SPE-IOA-20000-0017	Schedule
NICHOLB	VIS-SPE-IOA-20000-0017	CASU & WFAU interface
NAYLOR5	Pipeline processing	Aperture photometry
JARVIS2		Number of standard star fields
DALTON11	VIS-SPE-IOA-20000-0002	Reset frames
DALTON12	VIS-SPE-IOA-20000-0002	2MASS calibration
DALTON13	VIS-SPE-IOA-20000-0002	Astrometric calibration
DALTON14	VIS-SPE-IOA-20000-0002	Quality control
DALTON15	VIS-SPE-IOA-20000-0002	QC zero points
SIMARD2	VIS-SPE-IOA-20000-0002	Missing template for off-source observations
SIMARD1	VIS-SPE-IOA-20000-0002	Version system for acquisition, calibration and observation templates

Review Item ID	Document	Description
SIMARD5	VIS-SPE-IOA-20000-0010	Asynchronous processing operations
SIMARD3	VIS-SPE-IOA-20000-0010	Pixel-to-pixel correlated noise in confidence maps
SIMARD4	VIS-SPE-IOA-20000-0010	Object detection, deblending and classification
SIMARD6	VIS-SPE-IOA-20000-0010	Header curation mechanism
SIMARD7	VIS-SPE-IOA-20000-0010	Dummy data products
SIMARD8	VIS-SPE-IOA-20000-0010	Iterative improvement of confidence maps in VIRCAM image stacking?
SIMARD13	VDF-WSA-VSA-002	Intermediate data products
SIMARD14	VDF-WSA-VSA-002	Observing conditions
SIMARD15	VDF-WSA-VSA-002, VDF-WFA-VSA-009	Astrometric and photometric re-calibrations of archived data
SIMARD16	VDF-WSA-VSA-002	Mapping archive content
SIMARD17	VDF=WSA-VSA-002	Additional requirements from VISTA Public Surveys
JARVIS4		Incorporation of VST data
JARVIS6	VDF-WFA-VSA-002	Data releases
JARVIS7	VDF-WFA-VSA-002	Requirements over and above UKIDSS
JARVIS8	VDF-WFA-VSA-002	Requirements over and above standard VISTA archive
SIMARD19	VDF-WSA-VSA-003	Funding for VDFS development
JARVIS5	VDF-WFA-VSA-003	Funding beyond current cut-off date
SIMARD18	VDF-WSA-VSA-003	System administration staff
SIMARD20	VDF-WSA-VSA-003	Risk register for science archive
SIMARD21	VDF-WSA-VSA-004	Data compression factors
SIMARD22	VDF-WSA-VSA-004	Network transfer failures
NICHOL8	VDF-WFA-VSA-006	Running the pixel server
NICHOL1	VDF-WFA-VSA-007	Database architecture: RDBMS alternatives
NICHOL2	VDF-WFA-VSA-007	Table indexing
NICHOL3	VDF-WFA-VSA-007	Table indexing (same as NICHOL2)
NICHOL4	VDF-WFA-VSA-007, VDF-WFA-VSA-008	Database tutorials and examples
NICHOL5	VDF-WFA-VSA-007/008	CasJobs?
NICHOL6	VDF-WFA-VSA-007	Value-added catalogues?
NICHOL7	VDF-WFA-VSA-007	Mask and LSS informations?
NICHOL10	VDF-WFA-VSA-007	Database mirrors
JARVIS9	VDF-WFA-VSA-008	Image cut-outs
JARVIS10	VDF-WFA-VSA-008	Catalogue and image data
NICHOL9	VDF-WFA-VSA-010	Virtual observatory

Table 2. List of review items

5.1 Individual Review Items

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON01
Issue (short description):	Focal plane layout
Document(s):	VDF-SPE-IOA-00009-0001
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Figure 3: The detector numbering is now out of step with the camera software convention which has 1-16 reversed.

Action recommended by reviewer:

Update Figure

Conclusion of Review Panel:

Figure will be updated.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON02
Issue (short description):	Mini filters
Document(s):	VDF-SPE-IOA-00009-0001
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 3.4, paragraph 3. There is no plan in place currently to implement these mini photometric filters

Action recommended by reviewer:

Conclusion of Review Panel:

Update document.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD10
Issue (short description):	User Requirements for pipeline and archive
Document(s):	CASU and WFAU pipeline/archive documents
Reviewer:	Luc Simard

Reviewer's description of issue:

Reading through the VDFS documentation, it is clear that it will be very hard to meet some user requirements in early VISTA operations (or ever because they may be unrealistic.). How many times have these requirements been iterated between the VDFS project and the users before being used as a basis to define the proposed scope of the VDFS project? For example, Section 2 of document VIS-SPE-IOA-20000-0017 lists CASU responses to the UK user requirements. Have these responses been presented back to the users?

Action recommended by reviewer:

A description of the "requirement gathering process" (through VDUC?) would be helpful at the review.

Conclusion of Review Panel:

Iteration with user communities is needed to refine user requirements.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON03
Issue (short description):	Defocussed images
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 2.3: Pipeline should handle defocused objects: As far as I know there is no way of specifying a defocus within the observing templates, and no corresponding FITS keyword to identify such a frame to the DFS…

Action recommended by reviewer:

This requirement should either be addressed or dropped, but not left hanging.

Conclusion of Review Panel:

Requirement should be removed through proper change control.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NAYLOR2
Issue (short description):	Algorithm Documentation
Document(s):	UK Vista User Requirements
Reviewer:	Tim Naylor

Reviewer's description of issue:

There is a requirement that the algorithms be documented well enough so that the results could be reproduced by a sufficiently determined astronomer. How will this requirement be met?

Action recommended by reviewer:

I suspect that much of the documentation for this exists, but it must be presented in a cohesive way.

Conclusion of Review Panel:

This is a very important issue. No astronomer wants to use "black boxes". Proper documentation, technical papers and unit tests for each module should be made public. Furthermore, source code should be publicly released with a *very strong* disclaimer that the code is provided as is and that *no support* will be provided by CASU.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NAYLOR3
Issue (short description):	Access to precision hierarchy
Document(s):	
Reviewer:	Tim Naylor

Reviewer's description of issue:

There will be a hierarchy of photometric and astrometric precision, depending on whether two stars being compared were observed on the same detector, paw-print or night. How will the user know which is the case?

Action recommended by reviewer:

Time stamps may help with some of this, but not all.

Conclusion of Review Panel:

This is essentially a question of provenance of measurements. There needs to be a way to give users access to tables of photometric solutions. Use cases for this capability should be developed before deciding whether effort should be spent implementing this capability in the archive.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NAYLOR4
Issue (short description):	Photometry in highly reddened fields
Document(s):	
Reviewer:	Tim Naylor

Reviewer's description of issue:

Photometry in highly reddened fields is clearly an issue. I also note that for the WFCAM data the fit to the standards is limited to stars with J-K<1. Both these may be significant issues for work with reddened objects. How are we to ensure a meaningful red calibration?

Action recommended by reviewer:

Do we need another iteration with the users do work out what the most useful definition of the red magnitudes is, driven by science cases?

Conclusion of Review Panel:

This issue is in two parts: zeropoints and transformation to other photometric systems. The problem is that no photometric system includes object like T dwarfs. Color terms may be significant and must be firmed up so that science like isochrone fitting is possible.

J-K < 1 transforms very well in UK faint standards.

Solution: Create a consistent internal system and let users do their own system transformations. Pipeline should not try to provide transformations for all objects in all external photometric systems. A transformation to 2MASS has already been published.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD9
Issue (short description):	Only calibrated science products will be released
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 1 states that only calibrated science products will be released. This is the sensible thing to do once VISTA operations have reached its "cruising altitude". However, it is likely that some user communities will want to reduce some raw data themselves before they are ready to start trusting the "official" calibrated products. This will be especially likely at the beginning of survey operations (it was for the CFHT Legacy Survey!). How will this be handled at the pipeline/archive level if needed?

Action recommended by reviewer:

Consider providing some raw data at the beginning of VISTA operations to independent teams with demonstrated abilities in IR processing.

Ensure that all user communities are aware that there is currently no user requirement to provide raw data. See SIMARD10 for more on this.

Conclusion of Review Panel:

Panel was informed that raw data would be available through the ESO archive.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS1
Issue (short description):	Raw data
Document(s):	
Reviewer:	Matt Jarvis

Reviewer's description of issue:

Will raw data be available to the science teams conducting the first public surveys in order to cross-check the reduction pipeline?

Action recommended by reviewer:

Just a comment on this at the review is fine.

Conclusion of Review Panel:

Panel was informed that raw data would be available through the ESO archive.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NAYLOR1
Issue (short description):	Assessing astrometric accuracy
Document(s):	
Reviewer:	Tim Naylor

Reviewer's description of issue:

How will we assess whether the targets for astrometric accuracy have been met? Whilst comparison with 2MASS will allow us to see that the accuracy is better than 100mas, some of the targets are well below this.

Action recommended by reviewer:

Should be discussed at panel review.

Conclusion of Review Panel:

Stack residuals computed by the CASU pipeline (examples of which were shown by M. Irwin) provide the necessary information.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NicholA
Issue (short description):	Photometric accuracy
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Bob Nichol

Reviewer's description of issue:

In section 2.3, the goal of 1% was said to be impossible. What are the ramifications of missing that goal?

The SDSS is now claiming 1% using their apache wheel technique and ubercal. Why not use z and Y band data from DES to calibrate against and check for large-scale gradients?

Also explain the consequences of not handling the defocused bright stars?

Action recommended by reviewer:

Indication of the consequences of not reaching the goal

Conclusion of Review Panel:

1% goal may be impossible depending on issues such as focus stability and flatness of the focal plane etc. etc. In any case, software should not be the limiting factor here.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON04
Issue (short description):	Clipped RMS.
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 2.4: Clipped rms for each pixel. If this is to be ignored by VDFS there should be either a change request or a waiver request. It is insufficient to leave a hanging statement that it would be difficult to implement within the proposed infrastructure. Since the requirement exists in a document that has been signed (and therefore agreed) by the head of CASU, this needs to be addressed. The response also makes assumptions about the motivation of the requirement that go beyond that stated in the requirement itself.

Action recommended by reviewer:

Implement the agreed requirement.

Conclusion of Review Panel:

Go back to the source of this requirement and iterate. Change requirements (if needed) through proper change control.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON05
Issue (short description):	8% PS completeness
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 2.6, item 10.7: This requirement is nothing to do with LSB galaxies, but simply the completeness of FAINT galaxies. What does 10-s at 4.5x actually mean anyway

Action recommended by reviewer:

Address the requirement.

Conclusion of Review Panel:

Pipeline should not preclude users from doing their own completeness tests. Galaxy completeness tests should be designed and performed.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS3
Issue (short description):	Overall accuracy of photometry for crowded fields and close pairs/blended objects.
Document(s):	
Reviewer:	Matt Jarvis

Reviewer's description of issue:

When generating the object catalogues from individual paw prints how well are merged objects identified. In crowded fields – i.e. Galactic centre how will the source density effect the derived photometry. Are there any simulations of this?

Action recommended by reviewer:

Quantitative prediction of the accuracy of photometric calibrations for fields with increasing degree of source density.

Conclusion of Review Panel:

Approximate answer is the Rayleigh criterion. Realistic errors will be provided by the PSF fitting photometry. Outliers from the stellar locus could be used o monitor how well photometric deblending is doing. Simulations have been done at CASU, and reports are available on the CASU webpage.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON06
Issue (short description):	General comment on federation of object catalogues
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

What algorithm will be used to link objects that have variable structure as a function of waveband (e.g. faint galaxies with bright resolved star forming regions).

Action recommended by reviewer:

Discuss.

Conclusion of Review Panel:

This is documented in an archive paper in preparation. List driven photometry will be implemented.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON07
Issue (short description):	Flat Fielding
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 3.1: overview: Flat Fielding: Please clarify the use of the term ‘gain’ here. Do you mean variations in quantum efficiency, variations in total system throughput, or variations in the actual electronic gain of the system.

Action recommended by reviewer:

Clarify

Conclusion of Review Panel:

The word “gain” should be replaced by “throughput” here.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON08
Issue (short description):	Fringing
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 3.2.1.2: What exactly is meant by the use of the phrase 'it appears' here… I thought Newton had proved this one beyond all reasonable doubt?

Action recommended by reviewer:

?

Conclusion of Review Panel:

None.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD11
Issue (short description):	Facilities: physical space, cooling and power
Document(s):	VIS-SPE-IOA-20000-0017, VDF-WFA-VSA-003
Reviewer:	Luc Simard

Reviewer's description of issue:

The status of the facilities that will be hosting VDFS is unclear. In VIS-SPE-IOA-20000-0017, section 4.5, p. 19 reads "Final choices of actual hardware will be based on a total systems approach, taking into account not only simple processing per pound, but physical footprint (real estate), power requirements (including cooling), reliability and serviceability" and VDF-WFA-VSA-003, section 4.2, p. 9 discusses hardware acquisition. Are server rooms with adequate power and cooling already reserved for VDFS? Where are their costs budgeted?

Action recommended by reviewer:

If rooms are not secured yet, this should obviously be resolved as soon as possible. Building new server rooms requires very long lead times.

Conclusion of Review Panel:

Required facilities are already in place.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD12
Issue (short description):	Data Availability

Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 4.5.3 on p. 22 is entitled “Off-line archive”, and this is confusing. The WFAU VSA is definitely not off-line. In fact, I am wondering how many on-line copies of the released data products will be available at any given time. There should be at least two copies so that VDFS can transparently deliver a requested data product to a user even if one copy is unavailable due to maintenance or a transient failure. Given that users will be requesting products through WFAU, will the WFAU system have direct access to the CASU on-line copy in case of a failure at WFAU? What mechanism will be used?

Action recommended by reviewer:

Two copies of all released data products should be on-line at all times.

Conclusion of Review Panel:

Catalogues have two on-line copies, but flat pixel files do not due to cost. Requirements are 3% downtime during normal working hours and 6% outside these hours. Based on WFCAM statistics, it looks like uptime requirements are being met.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS11
Issue (short description):	Human resources for CASU
Document(s):	VIS-SPE-IOA-200000-0017
Reviewer:	Matt Jarvis

Reviewer's description of issue:

It is quite worrying that the continuation of the pipeline is essentially dependent on renewal of the CASU rolling grant, particularly when nobody is quite sure what is going to happen when the LFRC takes over from PPARC. This is obviously crucial and what would happen if only 2 FTEs were funded rather than 3 etc?

Action recommended by reviewer:

What will be the impact of staff shortages? Discuss at review.

Conclusion of Review Panel:

If required FTE’s are not allocated, then the same level of WFCAM service will not be possible.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOLC
Issue (short description):	Risk Register
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Bob Nichol

Reviewer's description of issue:

I found this a little flippant and offhand. I also did not understand the scoring scheme. The lost of key staff does seem a large risk and the mitigation strategy seemed weak. It is hard to simply replace key staff as by definition they are "key" staff! Also getting leaving staff to document before they leave is hard. Also we can expect bug fixes all the time – look at the SDSS!

More detail on this please

Action recommended by reviewer:

Discussion of UKIDSS/WFCAM completenesses

Conclusion of Review Panel:

CASU has sufficient overlap such the loss of one person would not be fatal. CASU has had historically low turn-over rates.

Systematic problem is that long-term projects such as VISTA are funded through a series of short-term grants.

It takes one year at WFAU to replace a key staff member. One must pay attention to University administration rules that may cause staff problems. It is important to ensure that people see a career path.

Reprocessing is the bad news item. As a mitigation plan, more people will have to be hired if reprocessing falls behind.

Scoring system should be explained.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOLD
Issue (short description):	Risk Register
Document(s):	VIS-SPE-IOA-20000-0017

Reviewer:	Bob Nichol

Reviewer's description of issue:

I found this a little flippant and offhand. I also did not understand the scoring scheme. The lost of key staff does seem a large risk and the mitigation strategy seemed weak. It is hard to simply replace key staff as by definition they are "key" staff! Also getting leaving staff to document before they leave is hard. Also we can expect bug fixes all the time – look at the SDSS!

More detail on this please

Action recommended by reviewer:

Discussion of UKIDSS/WFCAM completenesses

Conclusion of Review Panel:

Same as NICHOLC.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON09
Issue (short description):	Circular reference
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 3.2.2: Section 3 here should refer to section 3.1?

Action recommended by reviewer:

Correct reference

Conclusion of Review Panel:

Correct reference to be provided.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON10
Issue (short description):	Schedule
Document(s):	VIS-SPE-IOA-20000-0017

Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 6: Since this section says that schedule will be discussed at the review, but it is an issue driven review, please outline the schedule.

Action recommended by reviewer:

Outline Schedule

Conclusion of Review Panel:

Data should get to Cambridge as soon as possible. Data transfer ICD with ESO should be written and tested. It may be required at first to make extra copies of the data and ship them directly to Cambridge. LTAO tapes are preferred. The VISTA system will not be under Paranal system control at first, and some electronic data transfers might be possible then.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NicholB
Issue (short description):	CASU & WFAU interface
Document(s):	VIS-SPE-IOA-20000-0017
Reviewer:	Bob Nichol

Reviewer's description of issue:

I noticed several places where the requirements were offloaded to WFAU but there is then no subsequent discussion of these issues in the WFAU document (at least not as easily assessable as presented in this document).

This raises the managerially issue of how these two organizations work together and ensures things don't fall through the cracks?

Action recommended by reviewer:

Conclusion of Review Panel:

Regular meetings between CASU and WFAU address this issues.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NAYLOR5

Issue (short description):	Aperture photometry.
Document(s):	Pipeline processing of wide-field near-infrared data from WFCAM
Reviewer:	Tim Naylor

Reviewer's description of issue:

Whilst section 4.5 ("aperture fluxes as a panacaea" (*sic*)) makes a good case for aperture photometry, it states that it extracts most of the signal-to-noise. How well does it succeed for WFCAM, how well might its succeed for VISTA? Is the loss scientifically acceptable?

Action recommended by reviewer:

It would be good to have some simulations using the expected VISTA PSF.

Conclusion of Review Panel:

All tests have shown no improvement with PSF-fitting photometry although the errors look better. PSF photometry will be performed.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS2
Issue (short description):	Number of standard star fields
Document(s):	
Reviewer:	Matt Jarvis

Reviewer's description of issue:

Is the slewing to standard star fields every 2 hours necessary for the accuracy of the photometric calibrations? What do you lose by going to 3 standard star fields per night etc? Experience with WFCAM should make this easy to answer I think. So decreasing the number of standard star fields MAY increase the overall efficiency of VISTA, possibly dramatically as the current strategy will take ~1hour per night.

Action recommended by reviewer:

What is the ideal balance between number of standard star fields and acceptable photometric calibration accuracy?

Conclusion of Review Panel:

Resolved. First year of VISTA should have aggressive standard field observation frequency that will be re-assessed based on results.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON11
Issue (short description):	Reset Frames
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 4.2: What will be gained by measuring a 10s reset frame? It seems that the important thing here is to know the structure and possible variability of the data contained in the first read after the reset in a CDS frame, and since this read is always 1s after the reset in the current readout scheme I don't understand the meaning of a 10s reset frame.

What would perhaps me more illuminating (no pun intended) would be to map out the apparent systematic variations in 1s reset frames as a function of the background intensity (e.g. using the flat field screen or twilight). I would not be at all surprised if there were significant differences between reset frames taken with the dark filter and those taken on bright K-band sky during observations.

Action recommended by reviewer:

Clarify the nature of the effect to be corrected and the strategy for correction.

Conclusion of Review Panel:

First part is fine. Reviewer misread document.

During the daytime, reset frames at different illumination levels should be taken while pointed at dome screen. This should be a commissioning activity, and it should be added to the calibration plan. There must a corresponding observing template.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON12
Issue (short description):	2MASS calibration
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 5.2: The photometric accuracy of the 2MASS data varies considerably at the faint end (i.e. between Ks=14.5 and 15.7) where most of the objects useful to VISTA will fall. For example, a 10s exposure in Ks gives a 5-sigma point source sensitivity of 17.5 mags (Vega) in a 1.2" aperture with 0.7" FWHM. –How long do the individual exposures have to be before the calibrators are saturated out?

-Note that this being OK for WFCam does not necessarily imply it's OK for VISTA due to differing sky brightness, instrumental throughput, and detector well depths.

Action recommended by reviewer:

Check numbers.

Conclusion of Review Panel:

Covered previously.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON13
Issue (short description):	Astrometric calibration
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 6.2: If you want to determine things to 0.1pixel accuracy (0.03"), where in the system do you account for the effects of atmospheric dispersion and differential refraction?

Action recommended by reviewer:

Check numbers.

Conclusion of Review Panel:

Chip-by-chip astrometric calibration is performed using ~ 60 2MASS sources per frame. Absolute all-sky astrometry is good to ~ 100 mas RMS, but it is possible to do much better than this (~10 mas)

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON14
Issue (short description):	Quality Control
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 7.3: Where are the QC parameters stored?

Action recommended by reviewer:

Clarify

Conclusion of Review Panel:

Stored in FITS headers by pipeline.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	DALTON15
Issue (short description):	QC Zero points
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Gavin Dalton

Reviewer's description of issue:

Section 7.3: QC.ZPTs: Are these in ADU or electrons? (big difference!)

Action recommended by reviewer:

Clarify

Conclusion of Review Panel:

In ADUs. To be corrected in documents.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD2
Issue (short description):	Missing template for off-source observations?

Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Luc Simard

Reviewer's description of issue:

The list of calibration templates does not seem to include a template for off-source observations that will be required in some cases to remove background sky from crowded field and/or extended objects. Are observers expected to construct their own depending on the characteristics of their targets?

Action recommended by reviewer:

Please explain at review

Conclusion of Review Panel:

Two different OBs can be grouped together in P2PP. It remains to be determined where this recipe keyword will be inserted into the headers. It better be inserted!

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD1
Issue (short description):	Version system for acquisition, calibration and observation templates
Document(s):	VIS-SPE-IOA-20000-0002
Reviewer:	Luc Simard

Reviewer's description of issue:

I could not find a description of the version system that will be used to keep track of changes to the acquisition, calibration and observation templates. The names for the various templates do not contain any timestamps. As an example, what would happen if the template VIRCAM_img_cal_twiflat has to be modified after the first year of operations? Where would this change be recorded and how would different versions of a given template be associated with the data taken with it?

Action recommended by reviewer:

Please explain at review. A version system should be designed if not part of current plan.

Conclusion of Review Panel:

ESO has versioning control.

VISTA Data Flow System - Final Review of UK System, October 2006

Issue ref. no:	SIMARD5
Issue (short description):	Asynchronous processing operations
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

It looks like the VISTA/VIRCAM pipeline will run at four different location: Cambridge (CASU), Paranal (VISTA ops), Garching (ESO HQ) and Edinburgh (WFAU). This is dubbed as "asynchronous processing operations in the VDFS documentation. CASU will have the full pipeline. Paranal and Garching will only have a subset of the pipeline to run quality assurance checks. WFAU will need something close to the full pipeline for re-processing / re-calibration of archived data. What steps will be taken to insure that software residing at multiple sites is properly synchronized and that all the pipelines produce *exactly* the same data products where relevant?

Action recommended by reviewer:

An important issue that deserves discussion at the Review

Conclusion of Review Panel:

Emails are sent to WFAU to inform them that an update is available. Tarball is sent via email.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD3
Issue (short description):	Pixel-to-pixel correlated noise in confidence maps?
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

It was not clear from the mathematical description of confidence maps in Section 2.12 that pixel-to-pixel correlated noise would be included from one processing step to the next. I assume that correlated noise will indeed be included, but a clarification on this point would be appreciated.

Action recommended by reviewer:

Please clarify at review

Conclusion of Review Panel:

Yes, it is included.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD4
Issue (short description):	Object detection, deblending and classification
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

The definition of an object in an astronomical survey is crucial, and it is always science-driven. Section 2.13.1 states that a "standard matched filter" technique will be used, but nothing is said about object deblending. I suspect that different VISTA surveys will be optimising object deblending for their own science. Is the data reduction library ready to handle different algorithms if needed?

Will the detection algorithm use images taken in different bandpasses in the same region of the sky to produce the best possible detection image?

The "morphological classification" is essentially a measure of sharpness. Could the reduction library be extended to other types of source classifications such as the ones implemented in SDSS (may not be part of the VISTA science requirements?).

Action recommended by reviewer:

Please explain at review

Conclusion of Review Panel:

Covered previously.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD6
Issue (short description):	Header curation mechanism?
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 5.1 states that “derived parameters from the processing will be stored as FITS keyword/value pairs in the appropriate FITS header units”. What mechanism will be used to curate the FITS headers in the event that an error is found with the headers and retrieving/re-ingesting all affected FITS images is found to be prohibitive? For example, astrometric calibrations for hundred of thousand of images taken over many years may need to be revised along the way.

Action recommended by reviewer:

A database-driven header update mechanism may need to be devised in some cases. We had to go through such an exercise for the CFHT Megaprime/CFH12k archives.

What has been the WFCAM experience been so far with this?

Conclusion of Review Panel:

Full flat pixel fiels do not have to be read/written. Mechanism is already in place in the database to keep track of old and new values.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD7
Issue (short description):	Dummy Data Products
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 5.1, p. 45: Dummy data products will be generated in the event of failures such as malfunctioning detectors in the VIRCAM mosaic. Will these dummy data products need to be populated with actual dummy pixel values to pass through downstream processing steps or will header flags be sufficient?

Action recommended by reviewer:

To be clarified at review.

Conclusion of Review Panel:

Dummy pixel values are set to zero.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD8
Issue (short description):	Iterative improvement of confidence maps in VIRCAM image stacking?
Document(s):	VIS-SPE-IOA-20000-0010
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 6.11 – vircam_imstack: This is not an issue but a suggestion. Have you considered iteratively improving the confidence maps during stacking? We used this technique to produce stacked images from HST WFPC2/ACS associations with great success. More details can be found at http://cadcwww.dao.nrc.ca/wfpc2/WFPC2_pipe.html.

Action recommended by reviewer:

Something to consider

Conclusion of Review Panel:

Algorithm already performs at least one iteration.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD13
Issue (short description):	Intermediate Data Products
Document(s):	VDF-WSA-VSA-002
Reviewer:	Luc Simard

Reviewer's description of issue:

In Section 4, p. 9, requirement T7 and p. 19, AD01 5.13: “Once released, each data release shall remain indefinitely available, and the image and catalogue data shall not be modified”.

This is a *crucial* issue because it is directly connected to *refereed publications*. Situations where data products used for a paper are no longer available must be absolutely avoided. The “pragmatic” policy described here of discarding the oldest versions once the system runs out of storage space is not an acceptable mitigation strategy.

Action recommended by reviewer:

Consider a system that carefully archives all processing history (software, configuration parameters, recipes, etc. etc.) so that oldest versions can be generated “on-the-fly” instead of being stored on-line.

Conclusion of Review Panel:

Requirement should be re-formulated/clarified. If it remains in effect, then resources required to meet it should be provided.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD14
Issue (short description):	Observing Conditions
Document(s):	VDF-WSA-VSA-002
Reviewer:	Luc Simard

Reviewer's description of issue:

In Section 4.2, p. 10, C1: Will observing conditions (temp, humidity, seeing FWHM, transparency, etc. etc.) be part of the VSA science archive content or will they be obtained directly from the ESO Paranal weather station?

Action recommended by reviewer:

Knowing observing conditions is very important for science users to properly use the data. They should be provided to users by VDFS.

Conclusion of Review Panel:

Already stored in FITS headers. Archive metadata model tracks everything.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD15
Issue (short description):	Astrometric and Photometric Re-calibrations of archived data
Document(s):	VDF-WSA-VSA-002, VDF-WFA-VSA-009
Reviewer:	Luc Simard

Reviewer's description of issue:

In Section 4.2, p. 11, C3-C4: “Science archive must be designed from the start to enable astrometric and photometric calibrations”.

How is this exactly supposed to work? Will the WFAU archive have a copy of the full CASU pipeline or will it trigger processing at CASU when needed?

This requirement is identical to the one we reviewed in 2003 for WFCAM. We are now in 2006, and they are still listed as "in development" or "to be developed" (document VDF-WFA-VSA-009, p. 13, table 2, CU8 and C11). These tasks are incredibly hard to implement as routine archive curation tasks. I have watched a large team working on the precise photometric calibration of multi-year survey data from CFHT MegaPrime for many years, and they are still not reaching their goal. Personally, I abandoned the idea of routine photometric curation of archive data a long time ago. Astrometric curation turned out to be a massive undertaking all by itself.

Action recommended by reviewer:

VDFS project should be prepared to put this on hold given schedule and funding constraints.

Conclusion of Review Panel:

Empirical approach will be taken, and users will be informed as soon as any operational issues are identified.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD16
Issue (short description):	Mapping Archive Content
Document(s):	VDF-WSA-VSA-002
Reviewer:	Luc Simard

Reviewer's description of issue:

In Section 4.4.2, p. 16, D17: “Science Archive must allow trial-and-error searches (e.g., return the number of source hits rather the output results) for any valid query”. This is an important feature. However, great care must be taken in its implementation to avoid generating unnecessary user frustration. Is anything beyond simple SQL query subsampling planned?

Action recommended by reviewer:

Maps of archive content should be constructed (e.g., data density in some multivariate query space) and presented to users to give them a starting point for their trial-and-error searches.

Conclusion of Review Panel:

Aggregate SQL functions to perform content mapping will be described in archive cookbook.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD17
Issue (short description):	Additional Requirements from VISTA Public Surveys
Document(s):	VDF-WSA-VSA-002
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 8.1, p. 26: This whole section is extremely scary from a scope point of view. For example, a PSF-matched image subtraction pipeline to *reliably* find variables is a HUGE undertaking. It takes a large team to get something like this going. The CFHT Supernova Legacy Survey is the perfect example. They now have a powerful pipeline dedicated to this task, but they put in an incredible amount of work over the past few years.

Action recommended by reviewer:

Desired requirements in this section should not be pursued.

Conclusion of Review Panel:

These "requirements" should be considered as a wish list. Level of effort should be estimated for all of them before deciding to implement.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS4
Issue (short description):	Incorporation of VST data
Document(s):	
Reviewer:	Matt Jarvis

Reviewer's description of issue:

This is possibly beyond the scope of this review, however what are the plans for incorporating data from the various VST surveys into the science archive. Obviously a broad range of science aims would benefit from being to efficiently cross-match VST and VISTA public survey data.

Action recommended by reviewer:

Explore possibilities of incorporating VST data into the science archive and enable cross-matching with VISTA survey data

Conclusion of Review Panel:

External collections are already included as an interim solution. Collection-by-collection case will depend on availability of VO protocols.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS6
Issue (short description):	Data Releases
Document(s):	VDF-WFA-VSA-002
Reviewer:	Matt Jarvis

Reviewer's description of issue:

How will data releases happen? This is an ESO run telescope with specific data products to be released by the various public survey leaders at a time agreed with ESO. How does this overlap with the VISTA PI giving the final go-ahead to a data release? One would want stepped data releases to be worthwhile to the community otherwise there isn't much point in having a specific data release. The timing of which will undoubtedly vary from survey to survey. This is not like UKIDSS where the various survey heads are working together to some extent.

Action recommended by reviewer:

What is the role of ESO, the survey heads and VISTA PI in giving the go-ahead for various data releases? Discuss at review.

Conclusion of Review Panel:

This should be ironed out when survey PIs have been informed by ESO. Time frame is a concern.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS7
Issue (short description):	Requirements over and above UKIDSS
Document(s):	VDF-WFA-VSA-002
Reviewer:	Matt Jarvis

Reviewer's description of issue:

There are many requirements over and above those required for WFCAM. What is the current status of these developments. Have they already been incorporated into the WFCAM archive as a test?

Action recommended by reviewer:

What is the current status of the additional requirements for VISTA? Discuss at review.

Conclusion of Review Panel:

Under control in existing system.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS8
Issue (short description):	Requirements over and above standard VISTA archive
Document(s):	VDF-WFA-VSA-002
Reviewer:	Matt Jarvis

Reviewer's description of issue:

There are other requirements that specific surveys have asked for, particularly VVV. Which of these are possible within the time and manpower of the VSA? Obviously the aim is to maximise scientific output from all surveys so any that can be accommodated trivially should be.

Action recommended by reviewer:

What additional requirements could be met with the VSA? Discuss at review.

Conclusion of Review Panel:

See SIMARD17.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD19
Issue (short description):	Funding for VDFS development
Document(s):	VDF-WSA-VSA-003
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 4.1, p. 8: Funding for VDFS development runs out in September 2007. The team is expected to disperse after that. It leaves no time for shaking down in Phase 5. This is not acceptable, and I do not think that the extra 6 months of funding (if granted) will be nearly enough. One year should be an absolute minimum.

Action recommended by reviewer:

Ask for funding for more than 6 months on the basis of a solid, detailed plan. Plan could be shown at review.

Conclusion of Review Panel:

Development funds will be part of operations budget. At least level funding from development into operations should be provided.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS5
Issue (short description):	Funding beyond current cut-off date
Document(s):	VDF-WFA-VSA-003
Reviewer:	Matt Jarvis

Reviewer's description of issue:

The schedule for the final shake down with experience of VISTA data seems very short. Is this because of the WFCAM experience and is it really realistic? 6 months of shake down just seems pretty short to me for this type of work although I have no personal experience of it.

Action recommended by reviewer:

Make a detailed plan of what really needs to be done with the arrival of the VISTA data. If 6 months isn't really enough then this would be catastrophic for the VISTA surveys.

Conclusion of Review Panel:

See SIMARD19.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD18
Issue (short description):	System administration staff
Document(s):	VDF-WSA-VSA-003
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 4.1, p. 7: I was surprised by the level of effort planned for system administration of the VSA (30% of Holliman's time). In the 2003 WFCAM science archive review documentation, 70% of a person had been budgeted for system administration. This is closer to my expectation for VDFS. Is the revised 30% figure based on past WFCAM experience?

Action recommended by reviewer:

To be clarified at review.

Conclusion of Review Panel:

ATC IT support has been extremely good. Archive ops include some sysadmin such as RAID maintenance and software installation.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD20
Issue (short description):	Risk register for science archive
Document(s):	VDF-WSA-VSA-003
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 5.2, p. 10: Where is the risk register mentioned here? 2007 is nearly upon us, and the register should already be in place.

Action recommended by reviewer:

Present risk register at review.

Conclusion of Review Panel:

Risk register did not print properly for reviewer's personal copy. Risk register was there.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD21
Issue (short description):	Data Compression Factors
Document(s):	VDF-WSA-VSA-004
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 3.3, p. 5: RICE compression factors of 4x are claimed here. The best compression factors I have ever obtained with RICE was 2x for science frames (darks did compress by 3.4x). Is the factor of 4x based on past WFCAM experience?

Action recommended by reviewer:

Present some actual compression statistics for various types of data at review.

Conclusion of Review Panel:

Compression factors have been verified. OK.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	SIMARD22
Issue (short description):	Network transfer failures
Document(s):	VDF-WSA-VSA-004
Reviewer:	Luc Simard

Reviewer's description of issue:

Section 7.1, p. 10: How are network transfer failures handled? Automatically or manually? What fraction of these failures requires interventions by a human operator? Is this fraction acceptable when scaled to the larger VISTA data volumes or does it quickly become unmanageable?

Action recommended by reviewer:

Describe current WFCAM network transfer protocol and statistics at review.

Conclusion of Review Panel:

Everything is done automatically. Operates like a sophisticated rsync. Everything has already been put in place for WFCAM. VDFS will use same software.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL8
Issue (short description):	Running the pixel server
Document(s):	VDF-WFA-VSA-006
Reviewer:	Bob Nichol

Reviewer's description of issue:

How will the pixel server hardware be purchased and how will it be monitored and maintained? Buying these machines is only half the battle, where will they be stored and who pays for the airconditioning and power? What happens if one of their RAID controllers fails? Also, what is the procedure for failed disks? They will have at least one failed disk at any time?

Action recommended by reviewer:

Conclusion of Review Panel:

Taken care of already.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL1
Issue (short description):	Database architecture: RDBMS alternatives
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

The authors say they are not wedded to SQL Server abd could change to other RDBMS if required in the future. My question is: how will that decision be made and what criteria are important to move to another database like Oracle as mentioned.

Also, I fear it's not as easy as they portray? Moving the SDSS to Oracle has been tried and it's difficult because of the data schema and data variable types.

Therefore, I suspect they are stuck with SQL Server; is that an issue if the case?

Action recommended by reviewer:

Define a set of criteria that must be met to conider moving to another RDBMS.

Conclusion of Review Panel:

Migration is not anticipated. Document was being pessimistic. MS SQL Server's specific functions have been confined into small modules that can be easily isolated in the event of a server migration.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL2
Issue (short description):	Table Indexing
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

As the authors say, this is not an exact science. But I would like more detail on how they plan to optimize their database based on their evolving experience? Will they monitor the types of queries used? Will they develop a database of queries and execution times (and tables accessed)? Mining that data of usage would be highly informative. What are their plans?

Action recommended by reviewer:

Action plan for optimizing database

Conclusion of Review Panel:

Types of queries and execution times will be monitored.

So far, "on-the-fly" index optimization has been performed in response to queries that have been challenging. The Panel notes that it may not be possible to provide this as a service

when VISTA is at full throttle. This is in the operational plan for the archive.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL3
Issue (short description):	Table Indexing
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

As the authors say, this is not an exact science. But I would like more detail on how they plan to optimize their database based on their evolving experience? Will they monitor the types of queries used? Will they develop a database of queries and execution times (and tables accessed)? Mining that data of usage would be highly informative. What are their plans?

Will this database of queries and results be available? Any plans for caching the results of common queries?

Action recommended by reviewer:

Action plan for optimizing database

Conclusion of Review Panel:

Same as NICHOL2.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL4
Issue (short description):	Database tutorials and examples
Document(s):	VDF-WFA-VSA-007 VDF-WFA-VSA-008
Reviewer:	Bob Nichol

Reviewer's description of issue:

What are plans to present the VSA users with documentation? This is always the last priority and notoriously incomplete! The SDSS has extensive SQL tutorials and schema browers and information. There is also a growing database of example queries.

Will similar efforts be put into an extensive “user-friendly” database of examples and online help?

Action recommended by reviewer:

Details of the online help for VSA users and how this evolve over the years.

Conclusion of Review Panel:

An on-line cookbook is already available for WFCAM. It can be re-used for VISTA. Real life user queries could be included after obtaining permissions from query authors.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL5
Issue (short description):	Casjobs?
Document(s):	VDF-WFA-VSA-007 VDF-WFA-VSA-008
Reviewer:	Bob Nichol

Reviewer's description of issue:

SDSS extensively uses casjobs, which allows users to schedule big jobs until resources are available. Most queries end up on casjobs. Is there something similar here (did I miss it?)

This maybe off loaded to the VO/AG.

Action recommended by reviewer:

Details of the online help for VSA users and how this evolve over the years.

Conclusion of Review Panel:

Batch query server and queue system are in place. However, queries are not automatically off-loaded to a “SDSS CasJobs” equivalent.

This remains work in progress for WFCAM, but it will be fully implemented for VDFS.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL6
Issue (short description):	Value added catalogues?
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

Most of the interesting science from the SDSS has come from data derived from data extracted from the SDSS archives. These are known as “value added catalogs” and there are many available (MPA, NYU, CMU-PITT). Casjobs also has a facility to upload your own set of derived quantities and match them to the database and publish them to other SDSS users. Any plans for this type of facility. It would be very useful and increase the productivity of researchers and the usefulness of the VSA.

Action recommended by reviewer:

Conclusion of Review Panel:

This is a comment and not a goal. Should it be done through WFAU’s rolling grant and not VDFS?

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL7
Issue (short description):	Mask and LSS information?
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

The latest SDSS database included mask information which is vital for all statistical analyses of SDSS sources. Basically it tells you the angular mask of the survey as a function of several variables. How do users get this information from the VSA? Do they have to do it themselves? If so, can they then publish that information via the VSA for other users (see my comment about value-added catalogues)

Action recommended by reviewer:

Conclusion of Review Panel:

Again, this is a comment. Examples include areal footprint of the survey. It was not a requirement. The Panel notes that data products produced by the user community could eventually be ingested into the archive.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL10
Issue (short	Database mirrors

description):	
Document(s):	VDF-WFA-VSA-007
Reviewer:	Bob Nichol

Reviewer's description of issue:

This is off topic, but will mirrors of the VSA be available for others? The SDSS kindly provide software and the SQL DBs so people can mirror their CAS (see http://www.sdss.org.uk/dr5 for an example). This is good for everyone and I can imagine that providing such a facility would be of great benefit. For example, eventually there should be a US mirror of the VSA?

Action recommended by reviewer:

Conclusion of Review Panel:

There are no requirements to do this, but it should be kept under review.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS9
Issue (short description):	Image cut-outs
Document(s):	VDF-WFA-VSA-008
Reviewer:	Matt Jarvis

Reviewer's description of issue:

I'm wondering whether the small image extraction from an individual extension is really what is wanted. Will it be possible to obtain cut-outs of stacked data? Or will the user always be asked which files to combine. I imagine the typical use will not know enough about the data to do this efficiently.

Action recommended by reviewer:

Can there be an option of obtain the `best' stacked cut-out? Discuss at review.

Conclusion of Review Panel:

Possible.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	JARVIS10

Issue (short description):	Catalogue and image data
Document(s):	VDF-WFA-VSA-008
Reviewer:	Matt Jarvis

Reviewer's description of issue:

This is probably classed as a general comment but it would be really nice if you could query both the catalogue archive and image archive at the same time. Many people will want to do this, at least initially to gain trust of the catalogue data.

Action recommended by reviewer:

Can the catalogue query and the image query be linked in some way? Discuss at review.

Conclusion of Review Panel:

This is already implemented.

VISTA Data Flow System - Final Review of UK System, October 2006	
Issue ref. no:	NICHOL9
Issue (short description):	Virtual Observatory
Document(s):	VDF-WFA-VSA-010
Reviewer:	Bob Nichol

Reviewer's description of issue:

I was confused as to the relationship between the VSA and VO/AG. Surely, the VSA is only required to deliver VISTA data in VO compatible formats and interfaces? Today this means VOtables and Skynodes. Any additional functionality of the VO can then develop regardless of the VSA, as long as it keep up to date with the interfaces. More details on the boundaries between VSA and VO responsibilities would be useful.

Why not publish public VISTA catalogues via OpenSkyNodes? I appreciate that security is an issue for proprietary datasets, but one the data is public is should be published via the VO standard?

Action recommended by reviewer:

Conclusion of Review Panel:

If skynode is accepted by the IVOA and implemented in Astrogrid, then VISTA surveys should be made available through open skynodes.