ANNEX I - DESCRIPTION OF WORK

1. TITLE

ASTRONOMICAL WIDE-FIELD IMAGING SYSTEM FOR EUROPE
Short title: ASTRO-WISE

List of participants:

1. Nederlandse Onderzoekschool Voor Astronomie / Kapteyn Institute (NOVA) established in The Netherlands;
   With subcontractor:
   Visible and Infrared Survey Telescope for Astronomy / Queen Mary & Westfield College (VISTA) established in the United Kingdom.
2. European Southern Observatory (ESO) established in Germany;
3. Osservatorio Astronomico di Capodimonte (OAC) established in Italy;
4. Centre National de la Recherche Scientifique – Delegation Paris A / Traitement Élémentaire Réduction et Analyse de PIXels de MEGACAM (TERAPIX) established in France;
5. Ludwig-Maximilians-Universität München (USM) established in Germany.

List of abbreviations used:

ASTROGRID	UK project towards Virtual Observatory
AVO	Astrophysical Virtual Observatory
CCD	Charge Coupled Device
CFHT	Canadian-French Hawaii Telescope
DataGRID	European Grid initiative
EIS	ESO Imaging Survey
OmegaCAM	Wide field imaging camera for the VST
PM	Project Management
PSF	Point Spread Function
SWARP	Software Utility to co-add astronomical images
VLT	Very Large Telescope
VST	VLT Survey Telescope
WP	Work package

2. OBJECTIVES

2.1 Goals and objectives of the Project

The partners of ASTRO-WISE operate software ‘pipelines' and associated software tools and databases at their National datacentres for astronomical wide field imaging. The ASTRO-WISE project co-ordinates the development work to deal with and access wide field imaging data. The project pools the expertise of the participants who all have experience and commitments to optical wide-field survey work. All partners are involved in development of software pipelines for image data reduction. All partner sites will function as (linked) national data centres. In addition, the ASTRO-WISE project involves optional satellite data centres, spread throughout the EU. The code and expertise to run a satellite is a deliverable of the project.

The specific objectives of ASTRO-WISE are:

- To develop, maintain and provide access to a computational environment to process wide-field imaging data; the EU-wide shared environment will house both up-to-date calibration data and software, such as automatic pipelines, to process the raw data.
- To develop and disseminate in the community software tools needed to access the wide-field image data, to perform individual research programmes - tools are essentially search and visualisation tools, scalable to Terabyte regimes.
- The provided infrastructure will be used for the production and dissemination of survey data (multicolour, wide area calibrated images and source catalogues), to be accessed by virtual observatories.

The achievement of these goals involves advances and upgrades of existing infrastructures including:

- The setting of standards and the design and implementation of a federated data model to support the exchange of data, computer code and all data-reduction related administration between the various National data centres engaged in the processing and distribution of the wide-field imaging data.
- Exchange of programmer expertise and collaboration between the experts at the sites.
- A test bed for at-the-edge-of-technology handling of large amounts of data for the European community - in close co-ordination with AVO and ASTROGRID (see paragraph 5).

The huge data volumes require a dynamical approach, in which results can be re-derived with recent versions of calibrations and code, without the operator having to care about version control. Also, the users must have the option of additional fine-tuning, customized to the users specific needs. To this end, the various calibration data and other input files will be distributed over the network, which connects the data centres. The geographical distribution of the key information (both methods and data) will be provided by the system, allowing the various National Data centres to work cooperatively in an efficient manner and optimally profit from the work done at each of the centres.

The infrastructure that will be developed will be capable to both support the generation of typical Public Survey products and also the back-end analysis for individual research projects on dedicated observations. Ultimately, such dedicated observations can be re-processed with the standardised methods in order to obtain additional Public Survey products, in turn accessible by the browsers that will be developed by the complementary project AVO. Qualification of the system will also be done through direct astronomical research. This enhancement, consolidation and dissemination through National data centres, will form the Astronomical Wide-field Imaging System for Europe.

The project will establish a number of data centres where the voluminous wide-field imaging data can be processed to the specifications of the user (who can select on the basis of atmospheric conditions, quality of calibration data, etc., which data to process and how, what type of measurements are to be made on detected sources). The linking of the data centres means that the user will always have access to the most up-to-date data.

2.2 Project methodology

The actual work of the project is defined in 6 work packages (cf. Table 3.1).
The fundamental methodology of the project is based on setting standards, including an exhaustive definition of all data acquisition procedures and classes of data items. This full description of procedures and classes allows the project to define a strict, complete data model.

The technical work of the project will focus on developing tools and code in a co-ordinated fashion, supported by both these standards (WP3) and a strong project management (WP6). The detailed implementation work of pipelines and tools is contained in WP1 and WP2.

The environment takes care of the administration and has knowledge of the full data model. The database will be distributed over the participating sites (WP3), thus allowing any user, programmer, or operator to literally trace every bit of information, which other bits of information have been used to derive it, or would be required to re-derive the result. The same work package also supports the development work. The co-ordination of the development at the different sites, within this system, forms the core of the project methodology.

Tools dedicated to searching, querying and visualising Tbyte databases will be developed. These tools will have an innovative nature in order to cope with the large data volumes. Some of these tools are developed step-by-step from pre-existing ones by eliminating all non-scalable aspects. Others will be implemented using know-how already existing in other fields of research.

An important product of the Wide Field Imaging infrastructure at the observatories together with the ASTRO-WISE project will be a number of large astronomical surveys. Literally tens of millions of astronomical sources will be found and measured per year of operation.

The project methodology includes a strong project oriented management. The expertise and technology developed by participants are bundled into a single well-defined system. The definition of data classes and the associated database form an important management tool to achieve this harmony.

3. ROLE OF PARTICIPANTS

The project pools wide-field imaging expertise of the key research teams:

- NOVA: Scientific qualification, OmegaCAM calibration, pipeline software, database, together with its sub-contractor VISTA: optical and infrared wide-field imaging, pipeline, co-addition, source extraction
- ESO: EIS Public Surveys, optical and infrared wide-field imaging, pipeline, co-addition, source extraction, catalogue construction, survey tools customized to VLT instruments, delivery to public archives.
- OAC: Artificial intelligence/source extraction, catalogue building
- TERAPIX: pipeline, co-addition, source extraction, data volumes, visualisation tools
- USM: pipeline qualification and photometric calibration

All participants are involved in the construction and delivery of new wide field imaging instrumentation. On Hawaii, at the CFHT, the MegaPrime imaging camera is expected

to become operational in 2001 (data reduction will be handled by TERAPIX). In the middle of 2003 the European VLT-Survey telescope (VST), together with its wide-field imaging camera (OmegaCAM), are expected to become operational. Near to the VST, a second infrared and optical sensitive survey telescope with 50 2x4k optical CCD's, and at least 9 2x2k Infrared detectors, will be erected by 2005 (built by the VISTA consortium).
All participants are involved in National data centres handling the data flow from both new and existing wide field imaging instrumentation.
The project will provide mutual exchange of this expertise and infrastructure. The project will provide a vital data processing facility network for the European wide-field imaging telescopes now being constructed.

The provided hardware will fall under the project management. The key distribution of work packages over participants is given in Table 3.1.

Table 3.1 Work packages

WP1 Provide and operate processing facility of raw image data		
objective: – develop, implement and operate a full wide-field imaging pipeline from raw data to astrometrically and photometrically calibrated images – populate the database with the necessary calibration information		
Work description: – implementation of existing pipeline modules – implementation of interface to data base for all calibrations and survey system results – further develop, benchmark astrometric calibration, co-addition methods e.g. SWARP, multi-resolution decomposition of images – further develop techniques for reduction of IR data – operational tasks of populating database and maintaining pipeline		
participant involvement:		
NOVA	- major	focus on pipeline administration
VISTA	- minor,	but major on de-fringing technique and liaison on pipeline
ESO	- major	focus on pipeline administration
OAC	- minor	qualifying, flat fielding and photometric calibration
TERAPIX	- major	co-addition techniques (SWARP), PSF homogenisation Automatic quality assessment, photometric calibration
USM	- minor	pipeline-photometry

WP2 Provide tools for querying, searching and visualization

objective:

- provide tools for the end user and qualify through the involvement of academic research

Work description: design, develop, implement and qualification of:

- extract source list tool, including artificial intelligence methods
- associate source list tool
- search database tools
- visualization tools
- relay to the user *all* existent data in the data base related to a particular astronomical object

participant involvement:

NOVA	- major	focus on source search engines and association tools
VISTA	- minor	liaison on its experience with such tools
ESO	- major	focus on visualisation tools, GUIs
OAC	- major	artificial intelligence/source extraction and data mining tools
TERAPIX	- major	visualisation and source extraction tools, Panorapix, sExtractor
USM	- major	focus on photometry tools, SED fitting, photometric redshifts

WP3 Provide Federated database

objectives:

Design and implement federated databases that distribute documentation, source code, tools, pipeline administration, and calibration data over the participating sites as a support tool for WP1 and WP2.

This WP provides the project management's backbone of the survey system.

Work description:

- define classes/standards
- select data base engines
- define interfaces to pipeline and tools environment /implement
- define query language/method /implement
- define scripting /implement
- implement classes interfaces to data bases

participant involvement:

NOVA	- major	define, implement and qualify
ESO	- major	development, operate database, focus on pipeline administration
OAC	- minor	operate database locally
TERAPIX	- major	operate database locally and exchange expertise
USM	- minor	operate database locally

WP4 provide parallel processing power to ASTRO-WISE		
objectives: provide the necessary computing power to the prime data centers; install and operate Linux farm Gigabit PC parallel clusters		
Work description: – hardware procurement and installation – upgrading and export to other main sites of existing 4 PC clusters – maintenance		
participant involvement:		
NOVA	- major	qualify and provide expertise
VISTA	- minor	liaison on operational aspects of such systems
ESO	- major	qualify and provide expertise
OAC	- minor	only maintain local system
TERAPIX	- minor	only maintain local system
USM	- minor	only maintain local system

WP5 Provide data storage to ASTRO-WISE		
objectives: provide the necessary direct access data storage to the prime data centers; typically 1Tbyte/ site direct access, upgradable to 10 Tbyte/site within two years.		
Work description: – hardware procurement and installation – hardware maintenance		
participant involvement:		
NOVA	- major	qualify and provide expertise
VISTA	- minor	liaison on operational aspects of such systems
ESO	- major	qualify and provide expertise
OAC	- minor	only maintain local system
TERAPIX	- minor	only maintain local system
USM	- minor	only maintain local system

WP6 Coordination
objective: Coordinate WP1-WP5 into a coherent survey system; exportable to satellite sites, accessible to the community, and coordinated with future Virtual Observatory enterprises.
work description: – Provide strong Project management structure – Supervise the coordination of the components of WP1-WP5 into a coherent survey system – meetings, travel and report on meetings – coordinate with future surveys such as VISTA – coordinate with future customers, such as AVO, ASTROGRID – coordinate with future Infrastructure providers such as DataGRID and AVO.
participant involvement: All participants and the subcontractor participate in the Project management team NOVA - RTD contact person and co-ordinator VISTA - liaison person/ local principle investigator ESO - local principle investigator OAC - local principle investigator TERAPIX - local principle investigator USM - local principle investigator

4. DELIVERABLES AND WORK PLANNING/SCHEDULE

4.1 Deliverables and milestones

The main deliverable will be qualified software for the operations of a set of national data centres in France, Netherlands, Germany, Italy, UK and ESO, each facilitating the processing, calibrating, and archiving wide-field imaging data from ESO telescopes, CFHT and VISTA. The centres will be outfitted with computer systems with mass storage and software pipelines, and linked in federated databases to exchange administration, code, data and calibration results.
The infrastructure that will be delivered will be capable to both support the generation of typical Public Survey products and also the back-end analysis for individual research projects on dedicated wide-field imaging observations, including the tools to re-process and data mine (e.g. search, extract, visualise) the images. The network facility will be available to individual researchers wishing to derive survey products, EU-wide. Prime access will be to the sites operated by the National data centres, but also satellite hosts with less hardware infrastructure, will be supported. EU institutions can call for satellite station status and deliveries, thus providing potential EU wide access to the system.

The project is split into three phases, each ending with a deliverable which forms the prime objective of that phase. Phase I is characterised by design reviews and creating class definitions as soon as possible to receive wide-field imaging data, in order to avoid possible later backward compatibility problems. During Phase II the system will be developed to prepare for mass production, while during Phase III the system will be fully qualified, populated, and delivered.

Table 4.1 Milestones and deliverables

Mile-stone	Months after kick-off	Description of milestones, actions and deliverables
Phase I: be ready to receive first data without later re-definitions or re-shuffling		
1	0	milestone: kick-off WP6
2	0	milestone: kick-off- prepare design review for all WP1, 2, 3, 4 and 5
3	3	milestone: design review WP1, 2 and 3; procurement WP4 and 5
4	6	action: import pipeline
5	6	action: procure WP4, WP5 start sub-component installation
6	9	deliverable: WP1 fully operational with first version of WP4 and 5 i.e. operational pipelines at participant centres/WP1
Phase II: evaluate and prepare for mass production		
7	15	milestone: WP2 Evaluation - given experience with operational WP1
8	21	milestone:WP1/WP3 1 year operations and population of database review deliverable: set of tools to users and populated calibration database deliverable: export Beta version to satellite sites
9	27	action: upgrade of mass storage WP4- ready for bulk operations
10	30	milestone: system full mass production proof (i.e. 100 Tbyte data volume) deliverable: operational pipeline, tools, data, results in Tbyte regime
Phase III: qualify, operate and populate the mass production system		
11	33	milestone: review and set goals and schedule for final system
12	48	milestone: end of implementation/development work of the project deliverable: key objective of the project - a qualified survey system to be used for the derivation of survey and research results; including pipeline, software tools and databases.

4.2 Management structure, techniques used for co-ordination

WP6 and WP3 provide the backbone for the project internal co-ordination.
WP6 settles a project management (PM) team which is in charge of the supervision and coordination of the project. Each participant assigns a local principle investigator and his/her deputy. The deputy functions as the local contact person. NOVA provides a project co-ordinator who heads the PM team and a contact person.
All persons and hardware that are assigned to this project will be co-ordinated by the PM team.
The local co-ordination at the participants sites is done by the local principle investigator or his/her deputy.
The PM team will, at all times, overview the coherence of the developed system, and the team has the final responsibility for achieving coherence. Particularly, given the spread of human resources over sites and the importance of a coherent approach, the

role of the PM team is crucial for the success of the project. The PM team has final responsibility over all staff and hardware dedicated to this project.
The PM team meets at least two times a year. The project management organisation includes the following ingredients to support the communication between participants spread over different sites:

- Weekly teleconferencing on fixed days and times. The local contact persons will always participate in these teleconferences
- Two-monthly meetings reviewing work packages progress and setting priorities for the next quarter
- 6-monthly project review with PM team.

There will be an extensive infrastructure channelling, procedurising and defining fixed formats (classes) of all digital information exchanged between participants. Most of this is handled by WP3, which also provides an important management tool. To sum up, the exchange of digital information is supported by the following infrastructure:

- Project internet web-site, with pages for all official documentation, manuals, arrangements, names of associated persons. Password protected informal pages for intra consortium communications
- Federated database for text files, such as source code, manuals, etc.
- Federated database for all relevant wide-field imaging data, like calibration files.

4.3 Participants arrangements

Astronomers have data rights over their data for a limited period (typically one year). ASTRO-WISE shall not modify those rights, which will apply to the derived data products also. Thus, each data centre will be confronted with local data proprietary rights, which, however, will typically not last longer than 1 year.
The distribution of the EC funding over the various ASTRO-WISE participants is approximately proportional to the own contribution of each participant.
All source code generated in the project has an open status in the project and is distributed freely over the sites.
ESO and VISTA are expected to focus on Public Surveys, i.e. delivery of wide-field imaging products to the ESO community, with no proprietary period for either original data or survey products, including the tools to generate them.
ESO will also involve the EIS Visitors programme into the ASTRO-WISE, which is not accounted on the ASTRO-WISE budget.
Participants will bring in their own hardware (parallel processors farm and mass storage) and will operate and maintain this hardware as a contribution to the project. The provided hardware will fall under the project management and is indicated in section 4.5 Durable equipment.

4.4 Human resources by participant and work package

Table 4.2 presents a detailed breakdown of the manpower to be assigned to the project. For the participants and the sub-contractor the table specifies their contribution to human resources for the various work packages. Table 4.2 also specifies the

distribution of EC funded resources over the different sites and over the different work packages. The equivalent of 12 person-months for work to be done on final implementations of WP3 is not yet allocated to a particular site and is tentatively placed on the budget of the co-ordinator.
For each participant the own human resource contribution to the project are summed. The total human resource effort of the project is 671 person-months.

Table 4.2 Human resources by participant and work package

Human resource breakdown (person-month)									
	WP1	WP2	WP3	WP4	WP5	WP6	Total	Funded	Not funded
	Proc	tools	db	Beo	sto	coor		by project	by project
NOVA	41	74	52	16	9	28	220	60	160
NOVA/co-or			12			12	24	24	0
VISTA						22	22	6	16
ESO	30	72	18	6	6	8	140	36	104
OAC	33	34	18	7	14	12	118	30	88
TERAPIX	10	28	27			14	79	48	31
USM	20	24	3	7	6	8	68	18	50
Total	134	232	130	36	35	104	671	222	449

4.5 Durable equipment

Participants will bring in their own hardware (parallel processor farm and mass storage) and will operate and maintain this hardware as a contribution to the project.

Participant	Durable equipment	Charged to project	Work packages
NOVA	Processor farm/mass storage	provided to project (0%)	WP4+5
ESO	Processor farm/storage-new	80%	WP4+5
ESO	Processor farm/storage-old	100%	WP4+5
OAC	Processor farm/mass storage	provided to project (0%)	WP4+5
TERAPIX	Processor farm/mass storage	provided to project (0%)	WP4+5
USM	Processor farm/mass storage	provided to project (0%)	WP4+5

4.6 Subcontracts

A liaison task between ASTRO-WISE and the UK community has been specified to ensure co-ordination and exchange of expertise. To facilitate a direct and flexible link to the project VISTA will be subcontracted by NOVA. VISTA's initial prime tasks will be to provide advise on operating survey systems. The PM team will assign tasks from the WPs to VISTAs human resources funded by the EC.

4.7 Other specific costs

Specific database oriented software licenses are required to facilitate the computing network of the project. The co-ordinator will centralize the purchasing of licenses for participants who do not yet have the relevant licences. Approval is hereby given for the above costs to be incurred to the contract.

4.8 Exploitation of results

The network facility will be used to derive Public Survey results and will be available to individual researchers wishing to derive survey products, EU-wide. Prime access will be given to the sites operated by the National data centres, but also satellite hosts with less hardware infrastructure, will be supported. EU institutions can call for satellite station status and deliveries.
Scientific publications using the project infrastructure are expected, and will be supported when possible. The qualification of the system is expected to result into scientific publications.
The ASTRO-WISE will cooperate closely with the complementary programme AVO, which aims to establish an European Astrophysical Virtual Observatory.
Meetings with connected projects, such as OPTICON, DATAGRID, AstroGRID and AVO will be attended or organized.
ASTRO-WISE seeks advise from external scientists, amongst others by organizing meetings.
The project maintains a public outreach programme with particular focus on using the new media for disseminating wide field astronomical image data to the public.

5. COMPLEMENTARY PROJECTS

This project will be executed in a complementary way with the following projects:

5.1 Funded by the EC

AVO
Proposal number: HPRI-2001-50058.
AVO is responsible for medium and long-term development of processing and accessing all kinds of astronomical data.
ASTRO-WISE seeks close coordination with AVO.

OPTICON
Contract number: HPRI-CT-1999-40002.
ASTRO-WISE seeks close coordination with OPTICON and joint workshops will be planned. OPTICON will be informed of progress made.

5.2 Not funded by the EC

DataGRID, ASTROGRID
DataGRID and ASTROGRID are long term projects to enable processing of and access to massive sets of various kinds of experimental and astronomical data.
ASTRO-WISE seeks regular communications with DataGRID and ASTROGRID

Main DataGRID Participants:

The European Organization for Nuclear Research – CERN
Le Comité National de la Recherche Scientifique – CNRS
The European Space Agency's Centre in Frascati – ESRIN
Istituto Nazionale di Fisica Nucleare – INFN
The Dutch National Institute for Nuclear Physics and High Energy Physics – NIKHEF
Particle Physics and Astronomy Research Council - PPARC

ASTROGRID Participants:

Institute for Astronomy, University of Edinburgh
Institute of Astronomy, University of Cambridge
Dept. of Physics & Astronomy, University of Leicester
Space Data Division, Rutherford Appleton Laboratory
School of Computer Science, Queens University Belfast
Mullard Space Science Laboratory, UCL
Jodrell Bank Observatory, University of Manchester