ASTERICS - H2020 - 653477

Detailed WP3 Project plan

ASTERICS GA DELIVERABLE: D3.1

Document identifier:	ASTERICS-D3.1-final.docx
Date:	**31-08-2015**
Work Package:	**OBELICS – WP3**
Lead Partner:	**LAPP-CNRS**
Document Status:	**FINAL**
Dissemination level:	**PUBLIC**
Document Link:	TBD

Abstract

The document describes the detailed plan for Work Package 3, OBELICS, as far as they are evolved at the end of M4 in the project.

I. COPYRIGHT NOTICE

II. DELIVERY SLIP

	Name	Partner/WP	Date
From	Giovanni Lamanna	CNRS-LAPP/ WP3	20/08/2015
Reviewed by	Giuseppe Cimò	ASTRON	24/08/2015
Approved by	Rob van der Meer	ASTRON	31/08/2015

III. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
v1	04/08/2015	first draft	G. Lamanna/LAPP-CNRS
v3	13/08/2015	second draft	G. Lamanna/LAPP-CNRS
v7	26/08/2015	draft to partners for comments	G. Lamanna/LAPP-CNRS
final	31/08/2015	final submitted	G. Lamanna/LAPP-CNRS

IV. EXECUTIVE SUMMARY

At the time of the project proposal writing, there were many unknowns and uncertainties that made it impossible to create a detailed work plan for Work Package 3, OBELICS. Therefore, the proposal contained the outline of the plan, which had to be elaborated on during the project. This document describes the first iteration. It contains more focussed goals and deliverables, more detailed processes and procedures and a detailed resources list. All of these will help the OBELICS management guide the various groups to coherent results.

Table of contents

Acronym list

List of acronyms used throughout this document

ASTERICS	Astronomy ESFRI & Research Infrastructure Cluster
OBELICS	OBservatory E-environments LInked by common ChallengeS
ESFRI	European Strategy Forum on Research Infrastructures
CTA	Cherenkov Telescope Array
SKA	Square Kilometre Array
KM3NeT	A multi-km3 sized Neutrino Telescope
EUCLID	An ESA space mission to map the Dark Universe
LSST	Large Synoptic Survey Telescope
EGO	European Gravitational Observatory
(E-ELT)	European Extremely Large Telescope
(LOFAR)	Low-Frequency Array for radio astronomy
MAUD	MAnagement, User engagement and data Dissemination
LAPP	Laboratoire d'Annecy-le-Vieux de Physique des Particules
WP3	Work Package 3
RDA	Research Data Alliance
RDA-WG	Research Data Alliance Working Group
RDA-IG	Research Data Alliance Interest Group
EU-T0	International consortium of data and computing centres
EGI	European Grid Initiative

PRACE	Partnership for Advanced Computing in Europe
EUDAT	European Data infrastructure
e-IRG	e-Infrastructure reflection Group
D-GEX	Data Generation and information eXtraction
UCM	Universidad Complutense de Madrid
INAF	Istituto Nazionale di Astrofisica
D-INT	Data systems INTegration
ASTRON	The Netherlands Institute for Radio Astronomy
D-ANA	Data ANAlysis/interpretation
UCAM	University of Cambridge
VIRGO	Gravitational Waves Observatory
e-VLBI	electronic Very Long Baseline Interferometry
HESS	High Energy Stereoscopic System
MAGIC	Major Atmospheric Gamma Imaging Cherenkov
ANTARES	Astronomy with a Neutrino Telescope and Abyss environmental RESearch
ESO	European Southern Observatory
PM	Persons Months
WP	Work Package
TBR	To Be Recruited
ICT	Information and Computing Technology
PP	Project Plan

1. Introduction

The project plan (PP) of ASTERICS WP3, OBELICS, aims at providing a guide to all partners to the execution and the control of the expected work. The PP is expected to evolve over the life of the project.
This WP3-PP document describes:

i) scopes;
ii) roles of partners and stakeholders;
iii) schedule and resources;
iv) work breakdown structure (WBS) and deliverables;
v) management plan;
vi) risk analysis and management.

2. WP3 scope

The WP3-OBELICS work plan targets common *data challenges* of the astroparticle physics ESFRI projects. Major OBELICS scopes are:

- to enable interoperability and software reuse for the data generation, integration and analysis of the ASTERICS ESFRI and pathfinder facilities.
- to create an open innovation environment for establishing open standards and software libraries for multi-wavelength/multi-messenger data.
- to develop common solutions, share prototypes, exchange experience for: streaming data processing, extremely large databases, advanced analysis algorithms, software and middleware frameworks for data management, data processing and data access.

By the end of the ASTERICS project, the main expected impact of OBELICS will be:

- to contribute to the construction and the operation of research infrastructures identified in the ESFRI Roadmap (CTA, SKA, KM3NeT, E-ELT), and other major international projects (namely EUCLID, LSST, EGO, E-VLBI) including precursor experiments such as ANTARES, HESS, LOFAR, MAGIC and VIRGO.
- to initiate a collaborative path among scientific communities and e-infrastructures towards economy of scale, cross-fertilisation and resource optimization.

3. WP3 structure and roles

WP3 represents a major effort in its own within ASTERICS. The size of WP3 and the involvement of large international collaborations from major ESFRI projects in Astronomy demand an autonomous leadership and management to guarantee the best return of the OBELICS actions and to contribute coherently to the ASTERICS scopes.
OBELICS is structured according to four main tasks.

Task 3.1 MAUD, MAnagement, User engagement and data Dissemination.
Leader: **LAPP**.

This task concerns all management activities of the OBELICS work package. All technical tasks are built upon the state-of-the-art in ICT, in cooperation with major European e-infrastructures and are conceived to minimise fragmentation. Therefore MAUD guarantees that communications and links with other communities and e-science service providers are considered in order to contribute to the effectiveness of the proposed objectives.

MAUD actions, supported by the technical task leaders, aim at:

- fostering the federation of Astronomical and Astroparticle scientific communities around data management issues;
- training scientists to new challenges in computing and scientific software;
- preparing the new generation of researchers to acquire new skills for the best usage of multi-messengers new data coming from the next generation ESFRI observatories;
- providing opportunity to the Astronomical and Astroparticle communities to contribute to the common international and inter-domain activities to approach the complexity of discovering, accessing and processing scientific big-data.

In this respect MAUD organizes the three annual WP3 workshops and training events.

The internal and external interfaces are among the most critical actions expected to take over by MAUD:

(i) Internal interfaces:
   a. coherent actions among technical tasks.
   b. involvement of task leaders to encompass the overall needs and the frontier developments required by the data management of the ESFRI projects.
   c. periodical review of the evolution of priorities and the PP itself during the life cycle of OBELICS.

(ii) External interfaces (*here provided as examples*):
   a. commit to potential cooperation and active participation in RDA working groups (WG) and interest groups (IG), such as, for instance:
      i. “publishing data services RDA WG”,
      ii. “big-data analytics RDA IG”,

iii. "preservation e-infrastructures RDA IG", etc.;

b. technical engagement with major computing data centres supporting the ESFRI projects such as EU-T0 for the data identity federation and future exascale data service management;
c. survey available software products and services within major European e-infrastructures, such as EGI, PRACE and EUDAT and interacting with e-IRG providing astronomical ESFRI projects inputs towards the evolution of the *e-infrastructure commons* in Europe;
d. industrial partnership for innovation around new computing architectures test-benches, common trainings, pre-commercial cloud-based computing exploration, etc.

The *Technical Task* activities are planned following two criteria:

1. Three tasks are built around three main stages of the scientific data flow: data generation, data integration and data analysis.
2. In each Task the proposed topics are among the most challenging ones for the ESFRI projects and for which more communalities and potential room for cooperative work among ESFRI communities are identified.

In order to achieve the OBELICS scope the activities are organized in each Task through a matrix approach: transversal topics that associate different partners in working groups aiming at the cross-fertilisation *vs.* teams of ESFRI projects-based key representatives and personnel to support the individual needs or leverage the specific achievements of each Astronomical infrastructure.

Task 3.2 D-GEX, Data GEneration and information eXtraction.
Leaders: **UCM+INAF**.

In this area of the data flow, there are common challenges to create more robust hardware and software solutions for handling the ever-increasing data streams, and to ensure interoperability between varieties of different data-sources. OBELICS will pursue sustained cross-fertilisation via a three-step process:
a) sharing studies and seek synergies,
b) fostering evaluation and adoption of innovative solutions,
c) sharing common prototype frameworks and standards.

One of the practical challenges will be the sharing of individual prototype data models and formats as well as finding ways to share computing hardware for benchmarking purposes.

A few transversal topics or specific sub-tasks are:
Sub-task 3.2.1: Surveying the real-time streaming data architectures applied and envisaged for the ESFRI and pathfinder facilities, to establish best practices and agree on common software frameworks or common software modules, extending e.g. LOFAR, ASKAP, HESS, MAGIC, ANTARES and ALMA frameworks.

Sub-task 3.2.2: Developing new and common data models and high performance formats for data streaming, compatible with interoperability standards beyond the existing FITS, ROOT and HDF5 leading to common standards.
Sub-task 3.2.3: Developing prototype libraries that allow robust and optimised handling of secondary data streams and meta-data (environmental and engineering data, temporary local archive, device control software and observation scheduling), ensuring long-term & shared maintenance of the proposed products.
Sub-task 3.2.4: Benchmarking low-power computer platforms (including Multicore, MIC, Microservers, GPU, FPGA, ARM) and software technologies/methods for data-driven scalable parallel programming. This subtask will also follow a three-step approach, and will last the full ASTERICS project duration, since it will evolve by monitoring the continuous evolution of the technologies and could be also inspired by progress made in other scientific domains. The expected measurable value is the up-take of these new computing and information technologies by the ESFRI facilities and there platforms.

Task 3.3 D-INT, Data systems INTegration.
Leaders: **ASTRON+LAPP**.

The major common challenge addressed in this task is scaling-up existing databases and storage architectures beyond petascale level, while allowing for more complex queries addressing both primary sensor data and secondary data-streams.
Within the D-INT task, the first priority is to profit from the progress accomplished or the developments needed in all the involved ESFRI projects. The data bases issues, the metadata standards and the datasets administration standards, highly depend on the data access requirements of the users, the archive data size and data product catalogue specifications of each ESFRI project. This task will bring together key representatives from the data-archive developments of ESFRI projects. Dedicated workshops and trainings will create a common working environment where D-INT partners can share developments and build together some shared prototype platform. A common repository for individual database prototypes will guarantee the positive outcome of D-INT actions: increasing the capacity of sharing know-how and, where possible, cooperating to avoid repetitions.

The following preliminary sub-topics have been identified:

Sub-task 3.3.1: Collecting use cases from existing facilities (in particular LOFAR) that place extreme requirements on the databases and the e-infrastructures on which they are deployed, and develop these into benchmarks for future facilities and technologies (in particular the Extremely Large Data Base “XLDB” initiative) aiming for full interoperability. This will allow the ESFRI and related pathfinder facilities to optimally engage with providers of e-infrastructures.
Sub-task 3.3.2: Developing of prototype benchmarks for testing within a larger multidisciplinary context (in particular the XLDB initiative) aiming for interoperability. The

current investigation and prototypes of some partners are around the Qserv solutions or new generation DBs such as Cassandra or MongoDB. These will be used as examples and compared to the use cases of other projects.
Sub-task 3.3.3: Developing a portfolio of open services for data integration, based on existing Data Management System services like FLUME, RUCIO and Hadoop but extending these with optimised software modules to support e.g. VO-integration (in collaboration with WP4), and data interoperability between primary and secondary data-streams. Some services are already well known and applied; therefore a minimal repository can be made available in the first year. The repository will continuously be filled with new results of common work.
Sub-task 3.3.4: Extending software frameworks for data catalogues and query solutions to maximise data integration. This requires the reduction of latency for high data rates and the integration of multi-parameter Instrument Response Functions. Benchmarking and verification will be done through real pathfinder data, as well as Monte Carlo simulations. The complete software will be made available through the service repository.

Task 3.4 D-ANA, Data ANAlysis/interpretation.
Leaders: **INAF+UCAM**.

The data analysis software developments target three main aspects, which are in common among all ESFRI projects, although the specific weight among them is assigned individually by each research infrastructure. These aspects are:

i) scientific software methods for information extraction;
ii) data reduction technique and software for big-data sets management;
iii) orchestration of services for data and metadata sets for effective extraction of archived information.

D-ANA is concerned by the following main topics:
Sub-task 3.4.1: Developing a collection of statistically robust and domain independent open source software libraries for data analysis and data mining on petascale datasets. This will enable a sustained community-based effort towards excellent exploitation of all data generated by the ESFRI and pathfinder facilities. The initial set of libraries developed within this task is in particular:

- Statistically robust approaches (Bayesian and likelihood analyses) to advance cross- matching between catalogues and transients detected via different instruments
- Domain independent image analysis for simultaneous feature classification and extraction in multi-dimensional/multi-resolution data where the data are from multiple instruments.
- Effective likelihood reconstruction methods and new graphical processing approaches (mainly for event-based and signal-based projects but not exclusively) optimised for new computing technologies and maximum efficiency.

Sub-task 3.4.2: Establishing a common set of workflow architectures for the orchestration of compute intensive analysis of petascale datasets on distributed computing infrastructures. This involves providing use cases and technical requirements, designing and testing of workflow engines on distributed compute-intensive systems, and improving existing authorisation, authentication and accounting protocols (e.g. eduGAIN services). In this respect, the cooperation with projects supported by the European Commission through the call EINFRA-7-2014, "Provision of core services across e-infrastructures" aimed to produce a global authorisation and authentication infrastructure is foreseen.

All activities in this task will liaise and be complementary with WP4 (DADI), interfacing the respective activities will allow that the data, managed through the mechanisms identified and built by WP3, are archived, accessed, discovered and interoperated through the mechanisms defined by WP4.

All the ESFRI facilities (CTA, SKA, KM3NeT), and other major international projects or pathfinders, namely EUCLID, LSST, VIRGO/EGO, LOFAR, e-VLBI, HESS, MAGIC and ANTARES take part in all tasks and share responsibilities for task deliverables.

The institutes appointed leaders of the WP3 tasks have formally confirmed their engagement.

Project	Institute	Name	Family Name	role
CTA	LAPP-CNRS	Giovanni	Lamanna	OBELICS coorrdinator and CNRS contact
CTA	LAPP-CNRS	Gilles	Maurin	LAPP contact and HESS contact
CTA	INAF-RM	L. Angelo	Antonelli	INAF contact and CTA contact
CTA	CEA	Karl	Kosack	CEA contact
CTA	IFAE	Javier	Rico	IFAE contact and MAGIC contact
CTA	UCM	Jose Luis	Contreras	UCM contact
LSST	CNRS-IN2P3	Dominique	Boutigny	interim LSST contact
SKA	ASTRON	Marco	de Vos	ASTRON contact and LOFAR contact
SKA	UCAM	Paul	Alexander	SKA contact
SKA	UCAM	Bojan	Nikolic	UCAM contact
SKA	JIVE	Arpad	Szomoru	JIVE contact
EUCLID	INAF-TS	Marco	Molinaro	WP4 interface
EUCLID	IAP-CNRS	Yannick	Mellier	IAP contact and EUCLID contact
KM3NeT	FAU	Kay	Graf	FAU contact
KM3NetT	INFN	Cristiano	Bozza	INFN contact and KM3NeT contact
KM3NeT	CPPM-CNRS	Paschal	Coyle	CPPM contact and ANTARES contact
EGO	APC-CNRS	Eric	Chassande-Mottin	APC contact and EGO contact
E-ELT	ESO	Michael	Sterzik	ESO contact and E-ELT contact

TABLE 1: List of partner contact persons for the ESFRI projects and other roles.

4. WP3 schedule and resources

The OBELICS activities, as currently (M4) known, are scheduled here. The expected allocated resources for the task assignments are also listed.

- Schedule:
  - Start-up actions
    (Time: from October 2015 to February 2016)
    i. Work plan and Work Breakdown Structure assessment.
    ii. First recruitments at work.
    iii. Task and sub-task assignments.
    iv. First D-GEX meeting and work plan approval.
    v. First D-INT meeting and work plan approval.
    vi. First D-ANA meeting and work plan approval.
    vii. Data dissemination plan (MAUD).
    viii. OBELICS general meeting.

  - First year events
    (Time: from January 2016 to May 2016)
    i. Pursue of co-located event or active participation: EU-T0 & APPEC workshop on 24-25 January 2016;
    ii. Pursue active participation: RDA meeting.
    iii. Plan summer conferences participation.
    iv. Organization of D-INT workshop.
    v. Establishing e-infrastructure engagements for cooperation with OBELICS (MAUD).
    vi. Organize user-engagement meeting towards a *working open project: how to collaborate on coding projects of mutual interest* (MAUD).

  - First year ASTERICS deliverables
    (Time: expected May 2016 + 2 months maximum delay)
    i. D3.2 (MAUD)
    ii. D3.3 (D-GEX)
    iii. D3.4 (D-ANA)
    iv. Second year data dissemination plan (MAUD).

  - Second year activities.
    (Time: from May 2016 to May 2017)
    i. OBELICS general meeting.
    ii. After first 18 months WP3 internal review and project plan upgrading.

iii. Training/Spring-session events dedicated to *Scientific Software Developers*, potential and optional subjects are *Parallel programming and new computing architectures* opening to other scientific domains.
iv. To consider: OBELICS participation to EGI, PRACE, RDA, e-IRG and EUDAT forums.
v. Workshop on *Data and computing infrastructures for open science,* opened to other scientific domains and commercial Cloud providers.
vi. D-GEX meeting: status report and deliverables monitoring.
vii. D-INT meeting: status report and deliverables monitoring.
v. D-ANA meeting: status report and deliverables monitoring.

- Second year ASTERICS deliverables
  (Time: expected May 2017)
  i. D3.5, D3.8 (D-INT)
  ii. D3.7 (D-GEX)
  iii. D3.9 (D-ANA)
  iv. D3.6 (MAUD)

- Innovation actions.
  (Time: from November 2016 to June 2017)
  i. First survey on potential innovation actions to be funded under condition of active participation of industrial partners (D-INT, D-GEX, D-ANA).
  ii. Explore and build-up possible cooperation with other ICT and e-infrastructures innovation projects funded by EC within the H2020 framework.
  iii. Open internal competitive call for the innovation funding allocation.
  iv. Innovation funding allocation (MAUD).

- Third year activities.
  (Time: from May 2017 to May 2018)
  i. OBELICS general meeting.
  ii. Training/Spring-session events dedicated to *Scientific Software Developers*, optional subject*: big-data management and access frameworks*.
  iii. OBELICS participation to co-located conferences and workshops with EU-T0, EGI, PRACE, RDA, e-IRG, EUDAT as well as any new H2020 project for open science funded under H2020 2016-2017 calls.
  iv. Second Workshop on *Data and computing infrastructures for open science*.
  v. D-GEX meeting: status report and deliverables monitoring.
  vi. D-INT meeting: status report and deliverables monitoring.
  vii. D-ANA meeting: status report and deliverables monitoring.

viii. WP3 internal review and project plan upgrading for the last 18 months.
ix. Longer-term cooperative prospective white paper (MAUD).

- Third year ASTERICS deliverables
  (Time: expected May 2018)
  i. D3.11 (D-GEX)
  ii. D3.12 (D-INT)
  iii. D3.13 (D-ANA)
  iv. D3.10 (MAUD)
  v. Innovation funding allocation (if postponed) and/or reports on innovation actions results and feedback (MAUD).

- Fourth year activities.
  (Time: from May 2018 to May 2019)
  i. Main workshops and training event co-organized with ESFRI projects and APPEC / ASTRONET for feedback on ASTERICS-OBELICS cooperation added value, new perspectives and contributions supporting new *data scientist* profile: "Longer-term white paper" presentation (MAUD).
  ii. General OBELICS meeting inviting industrial partners and e-infrastructures representatives, e.g. EU-T0, EGI, PRACE, RDA, e-IRG, EUDAT.
  iii. D-GEX meeting: status report and deliverables monitoring.
  iv. D-INT meeting: status report and deliverables monitoring.
  v. D-ANA meeting: status report and deliverables monitoring.

- Fourth year ASTERICS deliverables
  (Time: expected May 2018)
  i. D3.16 (D-INT)
  ii. D3.15 (D-GEX)
  iii. D3.17 (D-ANA)
  iv. D3.14 (MAUD)

- Resources:
  ESFRI key-representatives and most of staff personnel are implied to some minimum time fraction compatible with the project and with main supervision purpose. Recruitment will be focused on maximizing the project success probability by ensuring that highly skilled and trained personnel is injected in all directly productive positions. In this respect the training actions organized by MAUD are also aimed at providing new skills or consolidate know-how of the recruited personnel.
  - Since a highly qualified personnel has to be recruited for some of the proposed topics within the technical tasks, the recruitment is expected to

take several months. Currently (at Month 4) some 30% of expected positions for the first phase of the project are filled. Most of the remaining resources are expected to be added to the project between Months 7 and 10. Further residual recruitments (about 12%) are planned after Month 10 in response of the specific time-line of the different sub-tasks.

- OBELICS partners have confirmed engagement of staff personnel according to Table 2 at Month 4. Further evolution is expected and will be closely monitored.
- Assignment of resources by task is shown in Tables 3. An evolution of this assignment per sub-task as a function of reassessment of the project plan and ESFRI project schedules is expected. The reassessment will take into account the priority evolution dictated by the ESFRI projects and coordinated by the ESFRI project contact persons. The OBELICS leader will monitor that the changes in the task assignment and work plan do not negatively affect the achievement of committed deliverables as discussed in section 6.

Project	Institute	Name	Family Name	% FTE	EC funded	Task	start Month	end Month	effort planned (PM)
CTA	LAPP-CNRS	Giovanni	Lamanna	0,2	0	3.1	1	48	9,4
CTA	LAPP-CNRS	Gilles	Maurin	0,1	0	3.2	6	48	4,2
CTA	LAPP-CNRS	Pierre	Aubert	1	1.0	3.2, 3.4	6	42	36
CTA	LAPP-CNRS	Jean	Jacquemier	0,5	0	3.3	6	48	21
CTA	LAPP-CNRS	X	Y	1	TBR	3.3, 3.4	10	46	36
CTA	LAPP-CNRS	X	Y	1	TBR	3.1. 3.3	6	48	42
LSST	LAPP-CNRS	X	Y	1	TBR	3.3	14	38	24
CTA	INAF-RM	L. Angelo	Antonelli	0,2	0	3.2	1	48	9,4
CTA	INAF-RM	Denis	Bastieri	0,1	0	3.2	4	48	4,4
CTA	INAF-RM	Matteo	Perri	0,3	0	3.2	4	48	13,2
CTA	INAF-RM	Stefano	Gallozzi	0,1	0	3.2	4	48	4,4
CTA	INAF-RM	Saverio	Lombardi	0,1	0	3.2	4	48	4,4
CTA	INAF-RM	X	Y	1	TBR	3.3, 3.2	6	42	36
CTA	INAF-RM	X	Y	1	TBR	3.3	6	42	36
CTA	CEA	Karl	Kosack	0,1	0	3.4	1	48	4,7
CTA	CEA	Thierry	Stolarczyk	0,1	0	3.4	1	48	4,7
CTA	CEA	Fabio	Acero	0,1	0	3.4	1	48	4,7
CTA	CEA	X	Y	1	TBR	3.4	9	45	36
CTA	IFAE	Javier	Rico	0,1	0	3.2, 3.3	1	48	4,7
CTA	IFAE	Tharek	Hassan	0,8	0,8	3.3	6	36	24
CTA	UCM	Jose Luis	Contreras	0,1	0	3.2, 3.3	1	48	4,7
CTA	UCM	Marcos	Lopez	0,1	0	3.2, 3.3	1	48	4,7
CTA	UCM	X	Y	1	TBR	3.2	8	32	24
SKA	ASTRON	Marco	de Vos	0,1	0		1	48	4,7
SKA	ASTRON	X	Y	1	TBR	3.2, 3.3	6	36	30
SKA	ASTRON	X	Y	1	TBR	3.4, 3.3	6	36	30
SKA	ASTRON	X	Y	1	TBR	3.3	6	42	36

Project	Institute	Name	Family Name	% FTE	EC funded	Task	start Month	end Month	effort planned (PM)
SKA	UCAM	Paul	Alexander	0,1	0	3.4	1	48	4,7
SKA	UCAM	Bojan	Nikolic	0,5	0	3.4, 3.3	1	48	23,5
SKA	UCAM	Gerry	Gilmore	0,1	0	3.4	6	48	4,2
SKA	UCAM	Richard	McMahon	0,1	0	3.4	6	48	4,2
SKA	UCAM	X	Y	1	TBR	3.2, 3.3	10	42	32
SKA	UCAM	X	Y	1	TBR	3.3	6	42	36
SKA	UCAM	X	Y	1	TBR	3.4	8	40	32
LSST	UCAM	X	Y	1	TBR	3.4	8	40	32
SKA	JIVE	Arpad	Szomoru	0,1	0	3.4	4	48	4,4
SKA	JIVE	Des	Small	0,5	0,5	3.4	4	24	10
SKA	JIVE	X	Y	0,5	TBR	3.4	16	44	14
EUCLID	INAF-TS	Fabio	Pasian	0,3	0		1	48	14,1
EUCLID	INAF-TS	Giuliano	Taffoni	0,3	0	3.4.2	4	48	13,2
EUCLID	INAF-TS	X	Y	1	TBR	3.4.2	6	42	36
EUCLID	INAF-TS	Marco	Molinaro	0,2	0	3.3.3 (VO)	4	48	8,8
SKA	INAF-TS	Cristina	Knapic	0,2	0	3.4.2 (A&A)	4	48	8,8
EUCLID	IAP-CNRS	Yannick	Mellier	0,2	0	3.4	1	48	9,4
EUCLID	IAP-CNRS	Henry Joy	McCracken	0,1	0	3.4	4	48	4,4
EUCLID	IAP-CNRS	Karim	Benabed	0,1	0	3.4	4	48	4,4
EUCLID	IAP-CNRS	Patrick	Hudelot	0,2	0	3.4	4	48	8,8
EUCLID	IAP-CNRS	X	Y	1	TBR	3.4	6	42	36
KM3NeT	FAU	Tamas	Gal	0,3	0	3.3	4	48	13,2
KM3NeT	FAU	Kay	Graf	0,1	0	3.3	4	48	4,4
KM3NeT	FAU	Thomas	Heid	0,1	0	3.3	4	48	4,4
KM3NeT	FAU	Clancy	James	0,2	0	3.3	4	48	8,8
KM3NeT	FAU	X	Y	1	TBR	3.3	6	42	36
KM3NeT	INFN	Cristiano	Bozza	0,2	0	3.2, 3.4	4	48	8,8
KM3NeT	INFN	Agnese	Martini	0,2	0	3.2, 3.4	4	48	8,8
KM3NeT	INFN	Rosa	Coniglione	0,2	0	3.2, 3.4	4	48	8,8
KM3NeT	INFN	X	Y	1	TBR	3.4	6	30	24
KM3NeT	INFN	X	Y	1	TBR	3.2, 3.4	6	18	12
KM3NeT	INFN	X	Y	1	TBR	3.2, 3.3	6	18	12
KM3NeT	INFN	X	Y	1	TBR	3.3, 3.2.4	12	24	12
KM3NeT	CPPM-CNRS	Paschal	Coyle	0,1	0	3.4	4	48	4,4
KM3NeT	CPPM-CNRS	Jurgen	Brunner	0,1	0	3.4	4	24	2
KM3NeT	CPPM-CNRS	Liam	Quinn	0,665	1	3.2, 3.4	6	42	23,94
EGO	APC-CNRS	Eric	Chassande-Mottin	0,1	0	3.4	4	48	4,4
EGO	APC-CNRS	X	Y	1	TBR	3.4	6	30	24
E-ELT	ESO	Michael	Sterzik	0,1	0	3.2, 3.3, 3.4	0	0	0

TABLE 2: List of WP3 participants including pending positions (to be recruited: TBR) per partner and per project.

project	CTA	CTA	CTA	CTA	CTA	SKA	SKA	SKA
partner / Task	LAPP-CNRS	INAF-RM	CEA	IFAE	UCM	ASTRON	UCAM	JIVE
Management	44							
D-GEX				24	24	24	24	
D-INT	48	36				48	48	
D-ANA	36	36	36			24	36	24
total per partner	128	72	36	24	24	96	108	24
total per project					**240**			**228**

project	KM3NeT	KM3NeT	KM3NeT	EUCLID	EUCLID	LSST	LSST	EGO
partner / Task	FAU	INFN	CPPM-CNRS	INAF-TS	IAP-CNRS	UCAM	LAPP-CNRS	APC-CNRS
Management								
D-GEX		12	24	12				
D-INT	36	6		24			24	
D-ANA		18			36	24		24
total per partner	36	36		36	36	24	24	24
total per project			**96**		**72**		**48**	**24**

TABLES 3: Matrix of the allocated resources in Person Months (PM) for each ESFRI or other major research project per partner and per OBELICS task. The table shows the distribution in the GA, defined at the proposal stage of the project. The distribution can be rearranged is necessary for the project.

5. Work Breakdown Structure

The Work Breakdown Structure (WBS) of OBELICS is presented here. Its higher-level part (the main Working Units) corresponds to the main tasks and subs-tasks in the Grant Agreement (GA). The WBS will be further detailed/updated in due course. The WBS approval and the work assignment are planned during a dedicated meeting of OBELICS partner representatives (as listed in Table 1) M8 and are considered among the "Start-up actions" listed in Section 4.

For each technical task (D-GEX, D-INT, D-ANA), the corresponding WBS includes also management work (under the responsibilities of the task coordinators), such as

- *task management* (task meeting organization, planning, internal interface coordination, providing inputs to MAUD for global management and data dissemination plan);
- *schedule control* (which focuses mainly on the progress achievements and task deliverables production);
- *documentation* (mainly producing and archiving the documents, the software libraries or any other relevant product under MAUD request as part of the global OBELICS data dissemination platform);
- *ESFRI projects inputs and plan evolution* (through the ESFRI projects contact persons to monitor the coherent evolution of the work in line with the progress of the ESFRI project implementation).

The MAUD WBS includes the central system and risk administration. In addition the MAUD WBS foresees the actions for the benefit of the ESFRI projects, the results and data dissemination, and the central planning and organization of all major workshops and training events, the execution of which will be conducted together with the coordinators of the three technical tasks.

WBS (first and second levels):

- 3.1 MAUD: MAnagement, User engagement and data Dissemination (LAPP)
  - 3.1.1 Planning, reports, meetings, workshops, training.
  - 3.1.2 Dissemination and communication, software libraries gateway.
  - 3.1.3 E-infrastructure interface plans and actions (some of which are already mentioned in the WP3-schedule presented in Section 2).
  - 3.1.4 Industrial contracts management.
  - *3.1.5 System and risk management.*
  - *3.1.6 ESFRI scientific communities engagement (including interfaces with APPEC, ASTRONET, ESO and other scientific stakeholders).*

- 3.2 D-GEX: Data GEneration and information eXtraction (UCM +INAF)
  - 3.2.1 Surveying real-time or close-to-detector data streaming frameworks.

3.2.2 Standards on data model and data format.
3.2.3 Prototype libraries handling secondary data streams.
3.2.4 Benchmarking low-power computer platforms.
3.2.5 Task management, schedule control and documentation
3.2.6 ESFRI projects inputs and plan evolution

- 3.3 D-INT: Data systems INTegration (ASTRON + LAPP)
  - 3.3.1 Scaling-up existing databases and storage architectures beyond the petascale level for complex queries.
  - 3.3.2 Archive solutions.
  - *3.3.3 Task management, schedule control and documentation*
  - *3.3.4 ESFRI projects inputs and plan evolution*

- 3.4 D-ANA: Data ANAlysis/interpretation (INAF + UCAM)
  - 3.4.1 Open source software for data analysis.
  - 3.4.2 Workflow architectures for petascale datasets on distributed computing infrastructures.
  - *3.4.3 Task management, schedule control and documentation*
  - *3.4.4 ESFRI projects inputs and plan evolution*

- Deliverables:
  - The deliverables listed in Table 4 are a set of documents or actions that the project is committed to in the GA of the ASTERICS project.

    They were defined at the time of the proposal.
  - For monitoring processes, progress, evolution of the plan and for collecting information for these deliverables a set of running documents and services is defined that will be updated and reviewed continuously, or intermittently during the project. Summaries or extracts of these documents and services will be published in the deliverable reports or on the website. The documents are listed in Table 5.

Deliverables

Nr	Description	Task	Month
D3.1	Detailed WP3 Project plan	3.1	4
D3.2, 3.6, 3.10	Annual user engagement forum, workshops and training events	3.1	12, 24, 36
D3.3	Analysis Report on Standards and Libraries	3.2	12
D3.4, 3.17	Release of Software Libraries	3.4	12, 48
D3.5	Analysis Report on Resource Requirements	3.3	18
D3.7, 3.15	Processing Platform Technology Benchmark Report	3.2	24, 48
D3.8, 3.16	Database Technology Benchmark Report	3.3	24, 48
D3.9	Statistical Solvers Technology Benchmark Report	3.4	24
D3.11	Analysis Report on Frameworks and Architectures	3.2	36
D3.12	Repository of Services	3.3	36
D3.13	Repository of WMS Services	3.4	36
D3.14	Final Integral WP3 Report	3.1	48

TABLE 4: OBELICS deliverables, corresponding task working-group responsible for and expected time.

Working documents and services	Description	planned delivery or update
Tools and services	Flexible project management web application	M5 + continuously
Hiring plan	Reporting strategy and plans for funded personnel recruitment	M6-M12-M24
ESFRI projects data challenges white paper	Collecting descriptions of challenges in Data Management and Computing Models used for the ESFRI projects. The paper will be used for communication purpose and projects evolution monitoring.	M7-M12-M36
E-Infrastructure cooperation agreements	Any document reporting agreements or memoranda with ICT providers and e-infrastructures for dissemination and common developments.	M8-M18
Innovation report	A document summarizing the innovation plan, the conducted actions, achieved results and lessons learned in cooperation with industry.	M12-M24
Training strategy	Training programme and reports	M12-M48 continuously
Long-term prospective report	A report describing main achievements to identify and explore new paths for future collaborations (after the project life) in the context of multi-messengers astronomy and sustainable approach for e-infrastructures for research.	M30

TABLE 5: Further internal OBELICS deliverables for management and monitoring purpose.

6. Management plan

The OBELICS management plan provides partners with planned procedures and organization to achieve the goals of the project. Four main elements composing the plan for the management of the OBELICS project are here proposed.

1. **How to do**: actions to achieve the goals and objectives.
2. **Organization**: roles and commitments of participants involved in the project, as well as how these roles might change throughout the project.
3. **Management timeline**: a timeline for various stages of the project
4. **Execution**: a process to handle possible project modifications and steer project's broader impacts.

1. Actions.

   - ESFRI projects contact persons provide OBELICS partners with the results of past and current ESFRI developments for open shared knowledge. A list of competences and products/services is produced. OBELICS Task groups provide to ESFRI projects the know-how from other experiments and defines common paths for common development to maximize the allocated resources and the consequent impact of the achieved results.
   - Running precursor experiments are aimed to be platforms for testing purposes or source of datasets and methods for further developments.
   - OBELICS task leaders will act for the benefit of all ESFRI partner project. They will pursue the dissemination and discussions of technical task objectives within larger frameworks and other multi-domains forums, if applicable.
   - Cooperation with major European data and computing centres/resource providers, e-infrastructures and Integrating Activity projects are considered of prior relevance. They will guarantee that up-to-date technical developments in terms of services, architectures and workflow management tools, scaled at the size of the new ESFRI project in Astronomy, are shared. The aim is to provide through OBELICS a critical mass of *data scientists* able to work together with infrastructure providers to explore and tests new solutions for the ESFRI projects. The success of this action will be monitored through documents establishing goals, roles and resources of these collaborations and disseminating the results. Cooperation agreement would be needed also in view of gathering some hardware components and services provisions from third party computing centres.
   - Technical tasks will promote actions within new multi-domain contexts for open science purpose (e.g. XLDB, Open software foundations, etc.). Participation to dedicated electronic forums and some e-infrastructures

transversal working groups will be pursued. The open character of these contexts guarantees easy monitor of the success of the OBELICS participation.

- Cooperation with industry for innovation. At the starting phase, all potential or under investigation/consolidation cooperative relationships between industry and ESFRI projects will be considered. Duplicate entries will be linked as much as possible. Further new actions are expected to be proposed by task coordinators after the first year of the project. The planned internal call for funding of innovative actions has a limited allocated budget. This requires minimizing the risk when the investment for innovation will be decided.
- Open Science: training, workshops and dissemination (also toward other scientific communities). The conferences, press communications and active participations in already consolidated forums will be supported. When the training and workshops aims are not focused on astronomy and astroparticle physics or collect potentially the interest from other domains, other domain-based clusters and/or known projects (i.e. in particle physics, geo-science, biology) will be involved by the OBELICS management.

2. Organizations.

- OBELICS coordinator and the appointed OBELICS project manager are together the main responsible for the MAUD task.
- The three technical task leaders will support MAUD in the project execution.
- OBELICS coordinator, project manager and task leaders represent the OBELICS Management Team.
- All partner representatives and ESFRI contact persons listed in Table 1 represent the OBELICS Steering Board.
- Task leaders assign the responsibility of the main working units of their WBS to the OBELICS participants.
- Task leaders have the first responsibilities of the corresponding deliverables, validation and monitoring of their WBS.
- The OBELICS management team have periodical meeting every two weeks.
- The project manager constantly monitoring and managing the project will keep constant contacts with task leaders.
- Task leaders call dedicated meeting every month.
- Sub-tasks and topical sub-groups are aimed to exchange among them more frequently, on weekly base.
- ESFRI contact persons are invited to report the OBELICS progress to their projects in occasion of all ESFRI projects' periodical general meetings.
- One person is appointed as interface responsible with WP4.
- The WP3 project manager will be responsible to cooperate with the ASTERICS project manager and establish periodical meetings.

- The allocation of recruited resources assigned to partners is based on the original proposal and in respect of the GA. Their monitoring or reassignment per task or within a task will be matter dealt directly by task leaders and partner contact persons.
- A periodical OBELICS Steering Board meeting will be called every three months. The main objective will be the evaluation of the project evolution.

3. Timeline.

- The Task coordinators assign Timeline for any detailed work foreseen in each technical task as a function of the available resources and to fulfil the global OBELICS schedule (see section 2).
- Main schedule is already discussed in section 2. Here below a table summarizing the resulting milestones is proposed.

Action	Time line
OBELICS general meeting	M8; M14; M20; M26; M36; M40; M48.
Task meetings	From M8 and then every month.
Co-located international events and workshops	M12; M24; M30; M36; M40; M48
Training events	M14; M24; M36
Industrial contracts: call issue	M12-M16
Industrial contracts: execution	M16-M40
Sub-task meetings or topical cooperative exchanges	From M6 and then every one-two weeks.

4. Execution.

- Cooperative work is expected from task leaders to support the global management execution (including the objectives of Task 3.1 - MAUD). When needed the WP3 coordinator for specific management or dissemination action, in agreement with the management team, can appoint an OBELICS participant.
- The project plan will be updated prior to every OBELICS general meeting by the OBELICS Management Team and approved by the OBELICS Steering Board.
- Proposed modifications with high impact, for instance changes to the first level of the WBS (i.e. redefinition of one of the main task and corresponding purpose) or re-scope of some main actions, e.g. training, innovation, users engagements, will be submitted to an approval process that could also pass through a dedicated review under request and responsibility of the OBELICS coordinator. The ASTERICS project manager will be involved in this process to guarantee the respect of the grant and consortium agreements.
- Changes of sub-tasks responsibilities within each Task WBS is proposed by Task coordinators to the OBELICS coordinator and discussed if needed with the OBELICS Management Team.
- The ESFRI project contact persons in OBELICS will assure the coordination with the evolution of the ESFRI projects and their users engagements.
- The partners provide their first hiring plans and their evolution to the OBELICS project manager.
- It is the responsibility of partners (namely through the partner contact person) to validate and monitor the participation of their own institute members, while the task leaders will verify and include their declared participation in the scheduled work plans.
- All project information will be stored in a unique web-based common repository, which will act, according to specific assigned access rights, also as web-platform for data and results dissemination.
- The management of the main tasks includes the definition and (communication to the OBELICS coordinator) of all major risks and uncertainties encountered or estimated.
- The time-line of the execution of the project must guarantee the achievements of the project deliverables described in the GA. However, any resource allocation issue (i.e. hiring plan execution delay) will be managed by focusing the resources on some actions and differentiating the timeline of the task execution independently to one another.

7. Risk analysis and management

The following table lists some of the main risks that can be identified at the starting stage of the OBELICS project and some actions foreseen for their mitigation.

Risk description	Impact	Mitigation action
1) Work not in line with ESFRI needs.	OBELICS project, although producing important results does not contribute to the evolution of the ESFRI partner projects.	The WBS includes engagement with the ESFRI projects, task monitoring and periodical dissemination of results. There will be agile reassessments of the work plan as a function of evolving input received from the ESFRI projects.
2) Limited cross-fertilisation among projects.	Scope of OBELICS partially not achieved	The task division has been built to maximize communalities and cross-work. Sub-tasks and related actions will be built starting from the major challenges of one ESFRI project and federating at least one more to promote cross-fertilization. Partnerships and leadership assignments follow this approach.
3) OBELICS dissemination actions are not effective.	Limited dissemination	Task leaders and sub-task coordinators are aimed to work for the dissemination of OBELICS results and actions as well. To minimize the risk impact the OBELICS management team looks for co-located events with other projects, e-infrastructures or scientific clusters. Increasing resources can be required.
4) Innovation budget for industrial cooperation is not effective	Not enough resources to tender a full product.	Find extra sponsor to invest in a full tender. Change the goal to seed money instead of a full product.