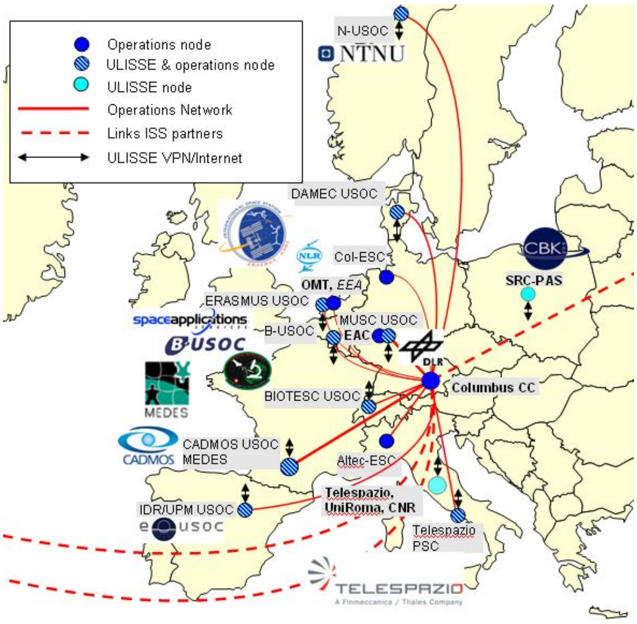
Nationaal Lucht- en Ruimtevaartlaboratorium

National Aerospace Laboratory NLR

Executive summary

The ULISSE environment for collaboration on ISS experiment data and knowledge representation



Report no.
NLR-TP-2010-464

Author(s)
E.A. Kuijpers
L. Carotenuto
C. Cornier
D. Damen
A. Grimbach

Report classification
UNCLASSIFIED

Date
October 2010

Knowledge area(s)
Ruimtevaartinfrastructuur en -payloads

Descriptor(s)
USOC
ULISSE
ISS
experiments

Problem area
Improving experiment data preservation, valorisation and exploitation for a broad range of disciplines is needed for numerous European experiment facilities hosted on the International Space Station (ISS). User Support and Operation Centres (USOCs) operate experiments through European experiment facilities including those hosted on the ISS

Description of work
A basis for improvement is to consider the existing experiment data archives distributed in the USOCs network, and to extract, represent and preserve metadata and knowledge about the experiments as a basis requirement for access by a European community for further scientific uses. The USOCs Knowledge Integration and dissemination for Space Science Experimentation (ULISSE) project aims at defining an approach and developing an infrastructure and tools to provide the required functionalities, services and/or applications.

This report is based on a presentation held at the IAC 2010, Prague, 27 September – 1 October 2010.

Results and conclusions

Solutions to document knowledge about experiments in a uniform way are proposed and being refined. New technologies allow for various improvements in providing access and adding value to existing datasets. A first demonstrator has been developed.

Applicability

The development of a ULISSE demonstrator through a distributed infrastructure and several tools and applications will allow to actually operate and valorise a first set of experiments and scientific data from various disciplines and under several USOCs' responsibilities.

Nationaal Lucht- en Ruimtevaartlaboratorium, National Aerospace Laboratory NLR

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 20 511 31 13, Fax +31 20 511 32 10, Web site: www.nlr.nl

Nationaal Lucht- en Ruimtevaartlaboratorium
National Aerospace Laboratory NLR

NLR-TP-2010-464

The ULISSE environment for collaboration on ISS experiment data and knowledge representation

E.A. Kuijpers, L. Carotenuto[1], C. Cornier[2], D. Damen[3] and A. Grimbach[4]

[1] Telespazio

[2] CNES

[3] Space Applications Services

[4] DLR

.

This report is based on a presentation held at the IAC 2010, Prague, 27 September - 1 October 2010.

Customer	European Commission
Contract number	EU Grant 218815
Owner	NLR + partner(s)
Division NLR	Aerospace Systems and Applications
Distribution	Unlimited
Classification of title	Unclassified
	October 2010

Approved by:

Author	Reviewer	Managing department
EAK/ EK 11-10-2010	ZP/ [signature] 11-10-2010	ASSP [signature] 12/10/10

Summary*

The USOCs Knowledge Integration and dissemination for Space Science Experimentation (ULISSE) project aims at developing an infrastructure and tools for improving preservation, valorisation and exploitation of data produced by European experimentation in space. The USOCs (User Support and Operation Centres) are a network of space operative centres; they have been established in various EU countries with the support of national space agencies and are engaged by the European Space Agency (ESA) to conduct the operations for European scientic experiments on board the International Space Station. ULISSE is a Seventh Framework project funded by the European Commission.

To be able to support and improve utilisation next to the experiment data archives meta data and knowledge about the experiments needs to be extracted, represented and preserved allowing access to a European community for further scientific uses. Tools need to be developed to exploit the use of the information for the broad range of disciplines being represented in the European Columbus experiment facilities on ISS.

As part of the project an inventory has been made of requirements, methodologies, datasets and services which would be considered for ULISSE demonstration and exploitation. Topic map technology was used to develop the knowledge representation. Dedicated workshops have been held for gathering relevant information from scientists and space engineers. All topic maps are merged into one map that will be integrated in the platform to support navigation and data browsing and nding in the ULISSE platform.

The tools being developed include generic web-services for both detailed data and public relations oriented data. A sciencecast collaboration tool is being developed which will be based on topic map technology to disseminate knowledge about experiments. For metadata development dedicated tools will be tailored to application for ISS experiment data. To allow data valorisation of different datasets data integration and organisation tools will be developed. In addition the use of Augmented Reality, planning and validation tools will be explored. The tools will be implemented on a platform consistent with the distributed ground infrastructure architecture implemented for the European USOCs.

The authors are willing to present both the tools and technologies that are going to be developed for ULISSE and the different issues to be addressed for preservation and valorisation of space data, like data and intellectual ownership, data dissemination policy, international cooperation, interfaces with space agencies and operative ground segments.

* Abstract submitted to IAC

Contents

Abbreviations **7**

1. INTRODUCTION
2. COLLABORATION ON EXPERIMENT DATA 8
3. ULISSE PROJECT SUMMARY 9
4. REPRESENTING EXPERIMENT DATA 10
    4.1 Data, metadata and knowledge representation 10
    4.2 Metadata to describe experiment data 10
    4.3 Topic Maps 10
    4.4 Metadata relationships with topic maps 11
5. PLATFORM INFRASTRUCTURE 11
    5.1 Software architecture 11
    5.2 Hardware configuration 12
6. EXPERIMENT DATA VALORISATION 13
    6.1 Improving Information Distribution 13
    6.2 Adding knowledge Representation 13
    6.3 Data Synergies and Simplifying Access 13
    6.4 Improving experiment data generation 14
    6.5 Improving data presentation 15
7. FIRST PHASE DEMONSTRATOR 15
8. CONCLUSIONS 17
9. ACKNOWLEDGEMENTS 17
10. REFERENCES 17

APPENDIX A IAC presentation 18

Abbreviations

API	Application Programming Interface
AR	Augmented Reality
BEST	Beyond EAST
CCSDS	Consultative Committee for Space Data Systems
CD-MCS	Columbus Distributed Monitoring and Control System
CGS	Columbus Ground Software
CMS	Content Management System
Col-CC	Columbus Control Centre
DB	DataBase
DEDSL	Data Entity Dictionary Specification Language
EAST	Enhanced Ada SubseT
EDR	European Drawer Rack
EEA	Erasmus Experiment Archive
ELGRA	European Low Gravity Research Association
ESA	European Space Agency
EU	European Union
EuTEF	European Technology Exposure Facility
FP7	Seventh Framework Programme
FPD	First Public Demonstrator
FSL	Fluid Science Laboratory
FVS	Formal Verification Service
IAC	International Astronautical Congress
IAF	International Astronautical Federation
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
ISS	International Space Station
KB	Knowledge Base
OAIS	Open Archival Information System
P2VS	Planning to Verification Translation Service
PSS	Planning and Scheduling Service
ProcVS	Procedure Validation and Verification Service
PVL	Parameter Value Language
PVS	Planning and Validation Service
PVT	Planning and Validation Tool
REST	Representational State Transfer

SI-Tool	Framework, to give access to data and services
SLURM	Simple Linux Utility for Resource Management
SRC-PAS	Space Research Centre Polish Academy of Sciences
UHB	User Home Base
ULISSE	USOCs knowLedge Integration and dissemination for Space Science Experimentation
USOC	User Support and Operation Centre
ViRLoc	VR Locator tool
VPN	Virtual Private Network
VR	Virtual Reality
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations
YAMCS	Yet Another Monitor and Control System

IAC-10-D5.2.2

THE ULISSE ENVIRONMENT FOR COLLABORATION ON ISS EXPERIMENT DATA AND KNOWLEDGE REPRESENTATION

E.A.Kuijpers
National Aerospace Laboratory NLR, Netherlands, Ed.Kuijpers@nlr.nl

L. Carotenuto
Telespazio s.p.a, Italy, luigi.carotenuto@telespazio.com

Christine Cornier
CNES, France, christine.cornier@cnes.fr

D.Damen
Space Applications Services, Belgium, david.damen@spaceapplications.com

A. Grimbach
MUSC-DLR, Germany, alois.grimbach@dlr.de

Improving experiment data preservation, valorisation and exploitation for a broad range of disciplines is needed for numerous European experiment facilities hosted on the International Space Station (ISS). User Support and Operation Centres (USOCs) operate experiments through European experiment facilities including those hosted on the ISS. A basis for improvement is to consider the existing experiment data archives distributed in the USOCs network, and to extract, represent and preserve metadata and knowledge about the experiments as a basis requirement for access by a European community for further scientific uses. The USOCs Knowledge Integration and dissemination for Space Science Experimentation (ULISSE) project aims at defining an approach and developing an infrastructure and tools to provide the required functionalities, services and/or applications.

As an initial phase an inventory has been made of requirements, datasets and services which would be considered for demonstration and exploitation. A metadata standard has been defined to capture the information on experiments and their datasets. Topic Maps technology is used to represent knowledge about experiments accessible via a ScienceCast collaboration tool. Other tools relate to planning and validation, presentation, data organisation, data access, data integration and distribution. The development of a ULISSE demonstrator through a distributed infrastructure and several tools and applications will allow to actually operate and valorise a first set of experiments and scientific data from various disciplines and under several USOCs' responsibilities.

1. INTRODUCTION

The USOCs (User Support and Operation Centres) are a network of space operative centres. They have been established in various EU countries with the support of national space agencies and are engaged by the European Space Agency (ESA) to conduct the operations for European scientific experiments on board the International Space Station. Many experiments have been performed successfully. The need to improve the infrastructure for using the data becomes more urgent with the growing number of ISS experiments.

To fully exploit the data generated in space experiments, a dedicated infrastructure is needed, firstly to make data accessible and secondly to preserve and disseminate information and related knowledge. The USOCs knowLedge Integration and dissemination for Space Science Experimentation (ULISSE) project [1] within the European's Seventh Framework programme aims at developing an infrastructure and tools for improving preservation, valorisation and exploitation of data produced by European experimentation in space. ULISSE activities and infrastructure are totally complementary to the on-going USOC activities of payload operations on the ISS.

This paper first introduces the collaboration issues on experiment data and the ULISSE project. An approach to metadata and knowledge representation is discussed in section 4. The infrastructure issues on the platform in section 5 and the tools in section 6. A demonstrator is being developed and discussed in section 7.

2. COLLABORATION ON EXPERIMENT DATA

Many space experiments have been executed to study phenomena in low gravity conditions, but the results are distributed in many ways and are difficult to access. The experiment data are dispersed in many ways and require detailed knowledge for understanding. Typically, the experiment data can only be understood by a few specialised people that participate at the space mission.

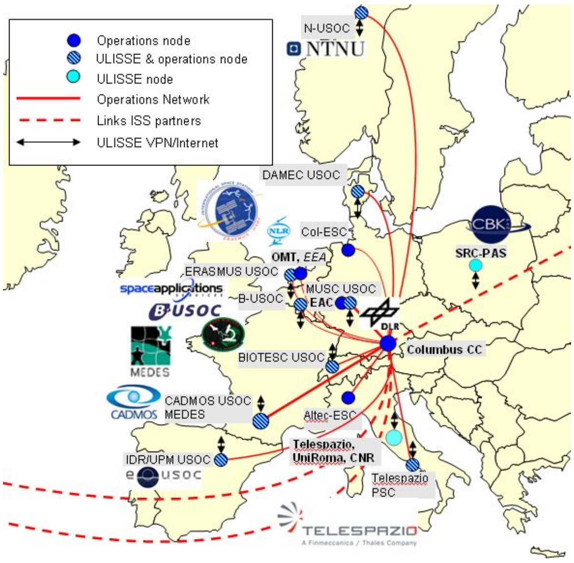


Fig. 1: European distributed infrastructure ISS data.

While the experiment preparation and execution takes long time and relevant effort, generally more limited resources and time are available for post-analysis. The consequence of the present organisation for data archiving and knowledge dissemination is a strong limitation of:

- Re-utilisation of data for further deepening of scientific analysis;
- Cross-fertilisation between different scientific areas;
- Generation of spin-off for terrestrial and /or space applications;
- Information utilisation from the perspective of non-scientific stake-holders (as decision-makers and educational bodies).

These limitations can be overcome by an integrated system for the dissemination and access of experimental data. This is a complex task. In the Erasmus Experiment Archive (EEA) [2] a total of 2282 experiments are listed. Around 369 experiments are mentioned for space station platforms, but the number is increasing in view of the continuous use of the ISS and Columbus facilities. The experiment results are not included and are dispersed over many entities and publications. Fig. 1 depicts various entities involved in Europe for the ISS operations and the overlap with the partners in the ULISSE project..

The state of the art of space data exploitation needs to be improved in a number of areas:

Data: The interpretation of scientific data requires additional information (like facility status data, local environment data as microgravity levels, etc.) and further information (as description of experimental set-up and procedure), often distributed among different units of space agencies.

Archiving: the whole set of all data and information is generally collected by the investigator only, who is the sole responsible for data analysis and scientific evaluation.

Knowledge dissemination: scientific results are published by the investigator in one or more papers on scientific journals, thus oriented toward a selected public specialized in the same scientific discipline. Only summaries of ESA experiments are published on ESA journals and on the EEA.

Moreover, the data distribution policies are experiment and platform dependent. ESA has developed a data policy[3] that identifies a number of relevant aspects to be considered in relation with the possible distribution and utilization of science data. Main aspects are related to the relation of the requiring country with the Space Agency, the role of the Space Agency and of the investigator for the data utilization and possible commercial exploitation, possible ethical issues, etc.

To be able to support and improve utilisation of the experiment data next to the experiment data archives, meta data and knowledge about the experiments needs to be extracted, represented and preserved, thus allowing access to a European community for further scientific use. Tools need to be developed to exploit the use of the information for the broad range of disciplines being represented in the European Columbus experiment facilities on ISS. Different issues related to preservation and valorisation of space data need to be addressed, like data and intellectual ownership, data dissemination policy, international co operation, interfaces with space agencies and operative ground segments. The concepts for an Open Archival Information System (or OAIS) are applicable [4]

Many aspects can be identified for science data archiving and access in general [5]. For space there are a number of constraints generating additional complexities such as: long lead times for instrument/experiment development, the combination of national and international funding for experiments, the constraints imposed by human space flight, constraints imposed by space industry and space agencies and broad range of disciplines involved.

The previous limitations can be overcome only with an integrated system for the dissemination and access to experimental data. This is the goal for the ULISSE project.

3. ULISSE PROJECT SUMMARY

The purpose of the project is to collect and provide reliable access to scientific data and related information in the long term. Hence, ULISSE will provide scientific and technical data concerning the wide ensemble of scientific disciplines investigated on the International Space Station, such as: Life Sciences, including Biology, Space Medicine, Exobiology and Biotechnology, Material Science, Fluid Science, Solar Physics and Fundamental Physics.

The ULISSE consortium is co-ordinated by Telespazio S.p.A. and involves participation of Telespazio PSC (formerly MARS Center S.r.l.) B-USOC, CNR-ISTC, MEDES-IMPS, Space Applications Services, CNES, UniRoma, DAMEC, National Aerospace Laboratory NLR, MUSC-DLR, ETH Zurich, N-USOC, SRC PAS, UPM, WERUM and ELGRA.

As a first step, ULISSE will make available data from previous space experiments on ISS as well as data from other space platforms, like sounding rockets, Foton capsule, Space Shuttle, etc. These datasets, coming from a large number of experiments, represent a relevant resource that, within ULISSE, will be exploited for testing and demonstrative purposes.

ULISSE services are oriented towards different typologies of users, mainly in the framework of scientific community, but taking into account also space industries, space agencies, decision-makers, educational bodies and general public. Accordingly, a large number of services have been defined together with the basic functionalities that the platform has to implement for the services provision.

A data survey has been conducted and a first version of the map of ULISSE datasets has been drawn. Successively, data coming from future experiments will be inserted in ULISSE as soon as they are available for release. ULISSE will ensure access to the data, in any case compliant with the applicable data policy (typically the data policy of the space agency that funds the experiment).

In parallel, a special effort has been devoted to the representation of the huge amount of information related to space science experimentation. To this aim, ULISSE has produced a standard metadata inventory. The metadata standard will be used for a full description of the ULISSE datasets.

As the datasets coming from space experimentation are typically multi-disciplinary and including both scientific and technical knowledge, ULISSE has followed the topic map approach, to represent the contained knowledge. The approach was chosen because topic maps provide support in representing complex data structures.

Accordingly, workshops dedicated to each scientific discipline have been held to introduce the principles of topic maps to the data providers and for gathering relevant information from scientists and space engineers. The output comprised dedicated topic maps for each of the numerous scientific fields of space research as well as one topic map for general space experimentation. Successively, all topics maps will be merged and serve as a basis for navigating and browsing and retrieving datasets within ULISSE.

ULISSE will make use of a distributed infrastructure: a middleware will integrate distributed resources (as data and tools) for the provision of services to the users. The middleware will also ensure a continuous enlargement of the ULISSE network by providing the means to integrate any further node that would become available in the future, providing additional data and/or tools. On this basis, a network architecture for connecting the different ULISSE services has been defined, being compliant with the security requirements of each node and avoiding any overlap with the operative infrastructure of the USOCs.

The project includes specific dissemination activities: scientific as well as publications for a broader audience, public events and educational activities on space research. In this way the project intends to increase the involvement of specialized communities and the awareness of general public. In this context, a number of dissemination actions, as congress presentations, publications, presentations at press/special events, have been realized in 2009 and 2010.

Once available, ULISSE will allow a systematic accumulation and preservation of data coming from space experiments; the growth of the data sets and the availability of ULISSE services for accessing data are expected to produce a strong impact not only on space research, but also on ground-based research and education activities; ULISSE will also contribute to harmonise the legal context for collecting and distributing data generated by international programmes like the space missions. Finally ULISSE will promote the generation of spin-off and applications of space research results and to the awareness of the public about space research.

4. REPRESENTING EXPERIMENT DATA

4.1 Data, metadata and knowledge representation

As part of the project an inventory has been made of requirements, datasets and services which would be considered for ULISSE demonstration and exploitation. Next to the raw experiment data a lot of information needs to be documented which can be considered data about data, i.e. metadata. Moreover knowledge needs to be captured about the relationships between various concepts and which evolves during the elaboration and use of data.

4.2 Metadata to describe experiment data

The initial version of the ULISSE metadata standard was based on the ISO19115 standard implemented as ISO19139, defined for Geographical data (from Earth Observation) [6]. This standard has been fitted and formalized for the ULISSE context consistent with the constraints and needs discussed in section 2 keeping the strongly structured principle :

- Metadata information
- Experiment
- Maintenance
- Constraints
- Acquisition Information
- Data sets Info
- Data distribution info

Some sections from initial ISO19115 appear less relevant for the ULISSE context or are covered via external references to further information.

This metadata standard supports the reference description of the experiment and datasets from the custodian point of view. The metadata provided should be relatively stable. As part of a test phase the insertion of metadata by various USOCs has been evaluated using this standard and schema.

The GeoNetwork tool was chosen for use in ULISSE to implement the metadata standard and offer :

- the management of the associated metadata base and associated access rights,
- an "editing" interface for template and metadata,
- a search interface to "discover" information within the specific database of concerned experiments.
- an access interface to metadata through API.

This tool is initially based on the ISO19139 standard. But one of its main interests is the configurability and evolutionary capacities. A ULISSE template has been implemented with in this tool consistent with the data model fitted and actual needs for ULISSE, and used by USOCs describing experiment data.

Issues which have been addressed in the design and updates following testing:

- Various structure and organisation for the datasets
- Strict isolation of the datasets from the metadata base in GeoNetwork
- Maintenance
- Sensitivity of some metadata and information
- Many data formats being used
- Centralized vs local editing
- Backup requirements and options
- Differences in thesauri per discipline
- Trade-offs in levels of details

4.3 Topic Maps

Topic Maps technology was used to develop the knowledge representation. Topic Maps are a standardized technology (ISO/IEC 13250) for representing knowledge. It lets us represent the meaning of the data that is stored by including semantics together with the knowledge itself. Having both the technical and scientific information of the experiments in a semantically rich structure like topic maps allows for the creation of smarter applications. This standard of description will support the description of complementary knowledge associated to the experiment and datasets as to the scientific topic from the scientific point of view. The generation of the concerned information could be more dynamic.

In Topic Maps, every subject of a knowledge domain is represented as a topic. Every topic may have a type (which is a topic itself), any number of names and any number of occurrences. An occurrence can be seen as a property of a topic, for instance, a person topic might have a phone number occurrence. Topics can be linked together with associations and each topic that plays a part in an association will be assigned a role. Furthermore, there is the concept of scope. Scopes can be assigned to names, occurrences and associations, and they indicate when the item should or should not be taken in to account when interpreting the topic map. Finally, a topic is identified through a set of identifiers (not through its names!) which allow automatic merging of different topic maps.

Workshops dedicated to each scientific discipline have been held for gathering relevant information from scientists and space engineers. During these workshops, Topic Maps ontologies were created for the scientific disciplines that were analyzed and for space science experimentation in general.

All topics maps are available on their own and merged together from a centralized knowledge base that is integrated in the platform to support navigation, data browsing and searching in the ULISSE platform.

4.4 Metadata relationships with topic maps

Metadata entered in GeoNetwork is of a more technical nature (people involved, files produced, etc.) and misses scientific information and explicit relationships between the concepts it describes. Topic Maps are used to bridge this gap. Thus two standards are planned to be used in ULISSE as presented previously and two sets of metadata generated and managed. Both verify the XML syntax. But it appears necessary to constraint both definitions to keep consistency and information sustainability of the descriptions for a given experiment.

The implementation in ULISSE is based on the identification of the metadata base as the reference base for description by custodian, only source to update if necessary. This base will then be ingested to the Topic Maps Knowledge base, available through ScienceCast, but these concerned "topics" (or information) will there be locked against modification. Besides this initial set complimentary information will be inserted representing additional knowledge.

Some others relationships could be managed in ULISSE with others description bases to populate this "reference" metadata base. For example, a basic description of the experiments is available in the Erasmus Experiment Archive (EEA) [2] will be referenced in the ULISSE metadata but could also be ingested as part of information. Moreover, some partners already manage some knowledge bases as their references and will populate the ULISSE bases (metadata and/or topic maps) from these. Population of some specific metadata bases can also be extracted from this reference ULISSE base.

The metadata from GeoNetwork is automatically converted to topic maps using a set of XSLT stylesheets, see section 3. These topic maps are used as a basis for scientists to load up in ScienceCast and to add scientific information tool.

The technical metadata is a key reference taking a considerable documentation effort. Therefore, the original data is marked as originating from the metadata files and locked while importing in to the Topic Map knowledge database.

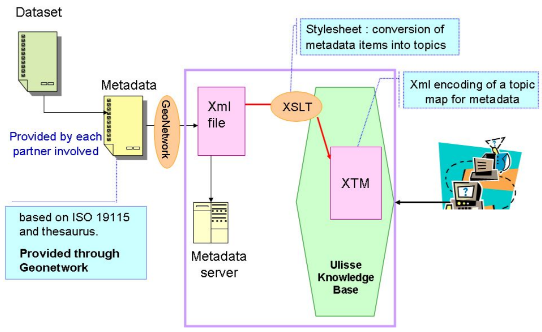


Fig. 2: Conversion of Metadata to Topic Maps

5. PLATFORM INFRASTRUCTURE

5.1 Software architecture

The platform had been defined in order to implement an integrated services architecture. In particular a middleware layer to integrate distributed heterogeneous resources has been designed on the basis of the identified user needs:

- Identifying and analysing the existing resources and existing data, defining associated services and available interfaces.
- Identifying new classes of services and additional resources which will be useful to implement a cooperative approach.
- Identifying the service flows for fruition of the domain knowledge represented in the KB as Topic Maps formalism

The ULISSE middleware platform is based on a 3-tiers architecture, composed by:

- Acquisition Services Layer
- Integration Services Layer
- Management Services Layer.

This layer guarantees the interface standardisation with heterogeneous information sources establishing a common communication protocol with respect to the integration layer. This approach proves very effective in case that an additional source should be integrated, so a real modularity is guaranteed to the system.

The integration layer, interfaced to the acquisition layer, receives and organizes the information associated with the sources. The information organization is accomplished through an explicit representation of the services which need that information.

The management layer guarantees the access to all the services of the system. The services are defined at this level and they are activated at this level on the basis of the resource availability and users requests.

A dedicated Application Programming Interface (API) has been defined to manage the interface development. The Representational State Transfer (REST) Framework is the basis for the interaction between clients and servers.

Some services require considerable computing resources. This motivated the use of an open-source resource manager (SLURM) which provides a framework for starting, executing, and monitoring the planning and validation which can be time consuming.

Several web-services will be accessed independently. All user requests addressed to one of the web servers are routed through a proxy server.

5.2 Hardware configuration

USOCs will act as data provider allowing remote access via ULISSE. The servers at the USOCs will be strictly isolated from the USOC operational infrastructure.

For the development and testing of concepts a service node is foreseen. The service node hardware consists of a number of servers:

- ULISSE DB Server dedicated to Knowledge Base
- ULISSE Web Server dedicated to web oriented application (as Geonetwork, ScienceCast, etc.) and the ULISSE WEB Portal
- ULISSE Service Framework dedicated to: 3-tiers ULISSE platform; tools for computing resources management (as SLURM); tools for software package that maintains list of issues and allows the management of customer support request, defect reports, access monitoring, etc. (ticketing system)
- ULISSE Networking monitoring system

Moreover the hardware configuration provides some workstations dedicated to ULISSE Tools (as P&V tool) and Telespazio sources.

Firewalls will protect the central server and the USOCs which will act as data providers. VPN technology will be used on interface the various entities.

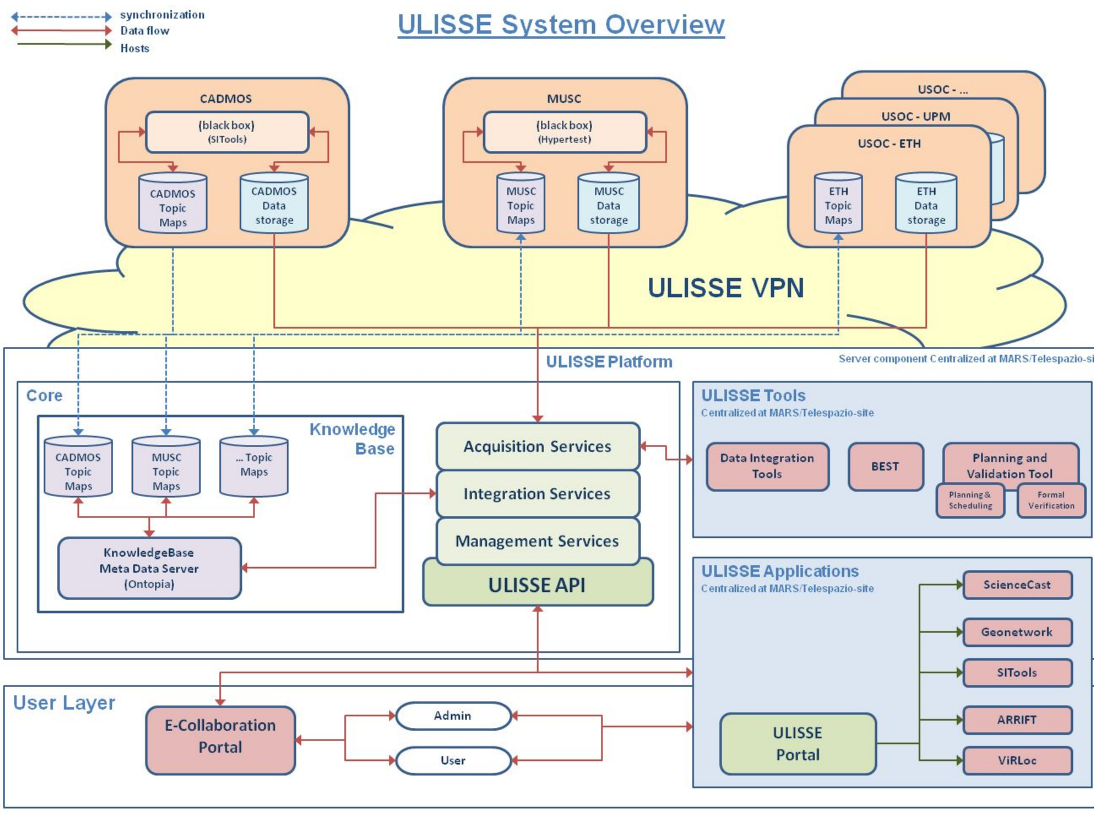


Fig. 3: Platform concept

6. EXPERIMENT DATA VALORISATION

6.1 Improving Information Distribution

An E-collaboration portal is being developed with Drupal based on recognized open source CMS (Content Management System) technology. This tool provides facilities for creating and editing articles without a need for specific background. The content editor supports individuals and communities of users to easily publish, manage and organize a wide range of contents on a website. The portal aims at disseminating and promoting space research related to ULISSE fields towards the general public.

SITools is a framework developed by CNES to give access to data and services and are being used to provide access to data-files for measurement sets provided and has been demonstrated for a B-USOC dataset being accessed remotely.

6.2 Adding Knowledge Representation

Both a static (metadata) and dynamic (topic map) formalism is supported for knowledge representation as discussed in section 4. The GeoNetwork tool is supported for adding metadata. The ScienceCast is a browser-based tool that supports:

- Editing topic maps and topic maps ontologies.
- Browsing and navigating through knowledge stored as topic maps.
- Natural language querying of information stored in topic maps.

ScienceCast is meant to support knowledge management processes and procedures while still providing a natural and flexible way to represent disparate sources of information. Fig. 4 depicts the current user interface. Related experience has been gained for parabolic flight data [7].

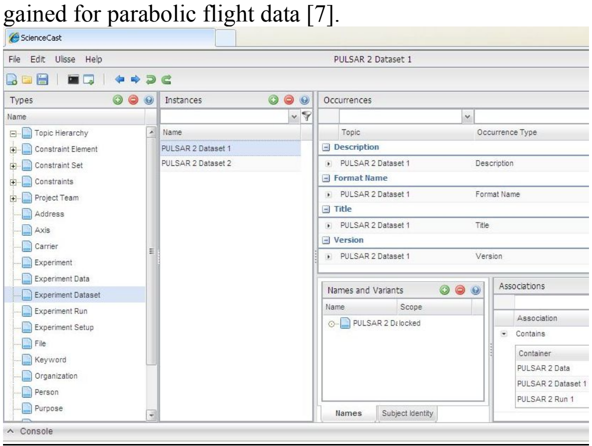


Fig. 4: User interface topic maps based ScienceCast.

6.3 Data Synergies And Simplifying Access

A key to simplifying access is improving data organisation and improving data organisation.

The data integration tool is a set of practical algorithms to compare and to enhance datasets allowing to explore synergies. Of special interest is to explore the combination of different datasets. Typically it is of interest to compare instruments accommodated within a single facility, to correlate experiment data with system data and to compare data from different facilities.

Considering that most experiment facilities have different data format, simplifying access to the data is a big challenge. Various dataformats are being used:

- Format Hypertext database parabolic flight data.
- Excel sheets for science data
- CCSDS-format descriptions
- Generic tools
- Telescience Resource Kit (TReK) data for European experiments using Nasa facilities.
- Processed Parameters for system data

This is a major bottleneck for simplifying access to the data and exploring data synergies.

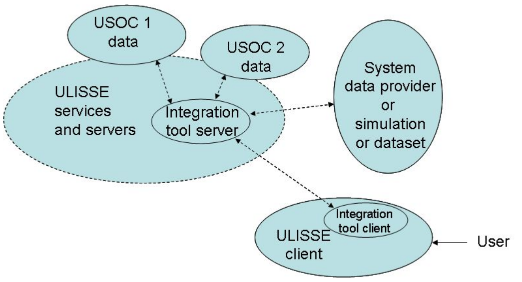


Fig. 5: Concept of distributed data integration.

Note that describing the internal structure requires a dedicated effort. Two approaches are considered. The first approach is to use a generic toolset called BEST and the second approach is to reuse existing libraries and database.

BEST (Beyond EAST) is a data framework. It relies on two recommendations, EAST and DEDSL:

- EAST (Enhanced Ada SubseT) primarily designed by CNES in the framework of CCSDS Panel II (CCSDS 644.0-B-1 and ISO 15889:2000). EAST is designed to create non-ambiguous descriptions of data formats including syntactic (logical and physical) information.

- DEDSL (Data Entity Dictionary Specification Language) designed in the framework of CCSDS Panel II. DEDSL allows you to add semantic information to data by the means of semantic attributes. Two implementations are available: one in PVL (Parameter Value Language), the other one in XML (eXtensible Markup Language).

The second approach is to reuse existing databases and libraries. In view of the long duration prospect to using ISS data and CCSDS based data formats are of major interest. The YaMCS library developed by Space Application Services is being used to have a generic format description compatible with Columbus ISS data.

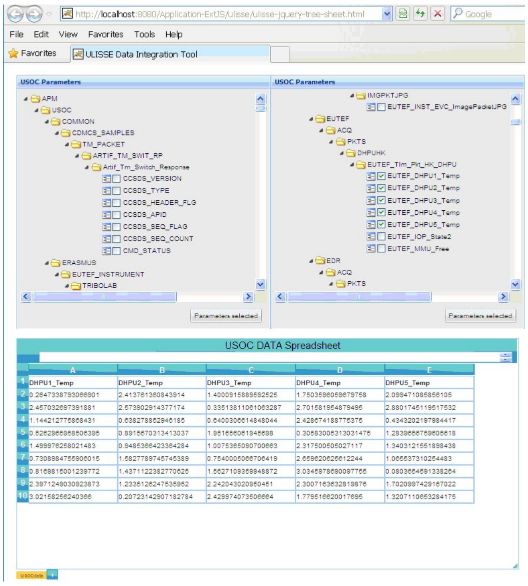


Fig. 6: Data integration using spreadsheet.

The tool is being demonstrated and tested with EDR and EuTEF commissioning data [8]. To integrate with Columbus/ISS system data will be considered for position data and accelerometer data using processed parameters. Via a drag and drop the users can select parameters and apply basic functions in a spreadsheet (Fig. 6).

<u>6.4 Improving experiment data generation</u>

As a part of the ULISSE project, timeline-based planning and scheduling [9, 10] are considered as a means to improve the experiment data generation. A fully automated planning and scheduling tool will assist in synthesizing flexible and robust experiment procedures starting from a description of past experiments and researchers goals. In addition, an automatic validation and verification tool is to verify that the generated procedures meet the experiment safety requirements. The following planning and validation services have been designed based on previous experiences:

1. PSS (Planning and Scheduling Service) which synthesizes flexible experiment plans and schedules;
2. FVS (Formal Verification Service) which verifies correctness of plans;
3. P2VS (Planning to Verification Translation Service), which translates from a planning input specification to model checker input language, see reference [10] for details of the approach;
4. PVS (Planning and Validation Service) which synthesizes flexible experiment plans and schedules and verifies their correctness;
5. ProcVS (Procedure Validation and Verification Service) which validates and verifies on-board operational procedures.

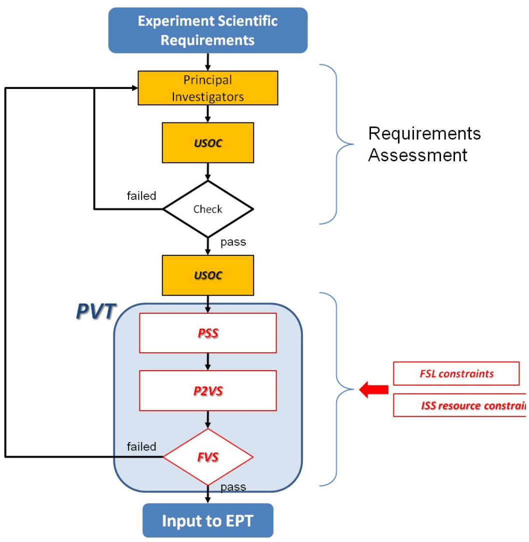


Fig. 7: Planning and validation.

Currently, a first prototype for the ULISSE APIs regarding the services required has been developed implementing the logic flow (Fig. 7). The first case study scenario was developed based on past experiments (in particular, GEOFLOW) performed on board on the ISS. The PSS prototype is able to produce a schedule of a set of activities that are to be executed on the Fluid Science Laboratory (FSL) in Columbus. The prototype receives a set of tasks that are to be executed on the FSL and a set of constraints from both FSL constraints (experiment constraints) and Space Agency (resource constraints). The prototype produces (if possible) a sequence of tasks ith an associated temporal allocation, satisfying all the given constraints.

In order to integrate in a deeper way the planning and validation tool (PVT) with in the ULISSE framework, different abstracted description levels have been identify to exploit Topic Maps describing planning and validation concepts. A first Topic Maps (high-abstracted level) ontology has been designed and implemented to integrate planning and validation concepts inside the ULISSE knowledge base (KB). In particular, this ontology describes the PVT components, inputs and outputs, relating them with the ULISSE's KB. Within the end of the project, more concrete ontology will be defined and exploited to couple in a tighter way the PVT and the ULISSE framework.

6.5 Improving data presentation

Augmented Reality (AR) and Virtual Reality (VR) as future user interface technologies offer a means to improve data presentation. Towards visualisation ULISSE metadata, two different approaches are being studied.

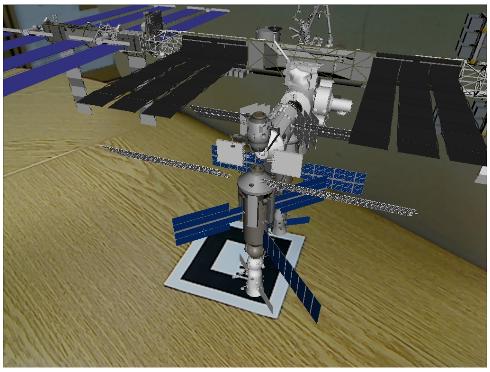


Fig. 8: Adding 3-D models in Augmented Reality.

Augmented Reality. AR allows the combination of physical real-world data and computer-generated data, where computer graphic object are blended into reality in real-time. In the context of ULISSE an AR viewer enables the visualisation of 3D models of an experiment platform or facility in a physical environment. A printed marker pattern in the view of a camera guides the rendering of 3D data into the camera view. That allows an intuitive exploration of 3D models in combination with real physical models (Fig. 8). The AR Realtime Interactive Filter Tool (ARRIFT) uses a camera setup and markers to integrate data and can be used to explore potential applications.

Virtual Reality. The VR Locator tool (ViRLoc) is a location-based approach for presentation of ULISSE (meta-)data in 3D virtual reality models. ViRLoc has different viewer modes. One mode using a GlobeViewer allows for example to navigate through the ULISSE data providers. Choosing a data provider or a certain research topic via 2D menus, information about the institution itself or about the associated experiments are places onto the globe with the aid of labels. The user can select a label and zoom in. Depending on the zoom level, detailed information are displayed. The Globe Viewer allows, for example an quick location-based overview of different USOCs, who are involved in a certain research topic. Another viewer mode of ViRLoc can be used to replay datasets in combination with a 3-D visualization of the Space platform position.

AR is more difficult to implement as dedicated provisions are needed (e.g. camera and viewer setups).

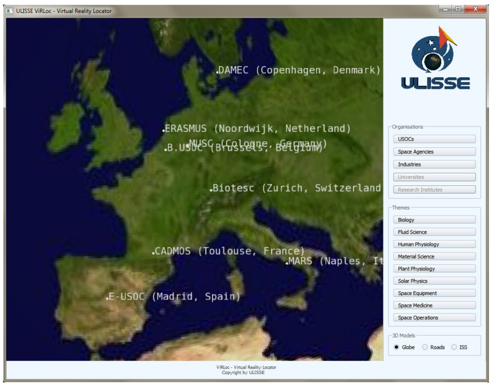


Fig. 9: Browsing using VR models.

7. FIRST PHASE DEMONSTRATOR

The first major milestone in releasing the platform will be the First Public Demonstrator (FPD) session which will terminate the first development phase. It is scheduled for November 2010 during the Space Weather Week in Bruges, Belgium. A second demonstrator session will be held in 2011 shortly before the release of the final platform, comprising its full functionality and including demonstrator versions for all envisaged tools and applications.

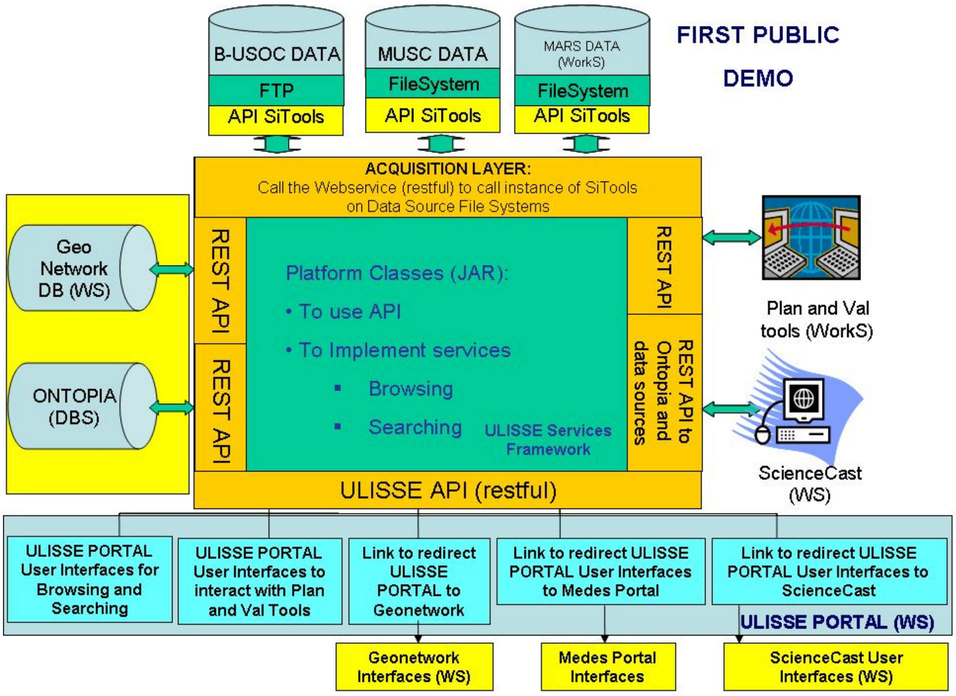


Fig.10: First Phase Demonstrator Components

An overview of the contained components is depicted in Fig. 10. Components have been assessed with respect to availability, stability and scalability. The assessments served as input for the project plan, primarily towards the FPD but also defining long-term activities with respect to the final implementation. Accordingly, alpha versions of selected tools are currently being integrated into the test platform. Parts of the API can be accessed through a prototyped web interface.

The metadata for the datasets is stored via GeoNetwork. The metadata is also transformed into topic maps to provide advanced services for the knowledge base. Several core features will be ready for the FPD session. They will enable the consortium to show and make clear ULISSE's capabilities and potential which can be considered the core objective of the FPD session.

The response of addressed and potential users will be an important asset to improve platform usability as well as quality of the offered services. To achieve that task, a web based tool is being conceived and tailored questionnaires will optimally support feedback gathering from the community and its evaluation. A further objective of the FPD session comprises the test and the assessment of the deployed technologies and tools. In that sense, the performance of hard- and software components will be monitored to assess load factors and platform scalability. The results will be used to improve the platform architecture and to fix bugs respectively.

Various tools covering ULISSE high-level features will be available for the FPD session. The SI tool suite been designed to ease data access in distributed and heterogeneous setups. The GeoNetwork database will be available as a web application or through its API to access the various metadata sets generated. The ScienceCast/Ontopia offers a topic maps representation of the corresponding experiment and domain descriptions. The Planning and Validation Tools and Promotional portal to the general public will be included as a prototype.

The entire set of experiment descriptions will cover several disciplines from different data providers. The chosen scenarios are based on real space data and on the evaluated service requirements. According to the chosen event, experiment descriptions from the domains of Space Weather and Solar Physics will be visible in the ULISSE platform.

It will be possible to download real data sets as part of an online service, thus demonstrating a possible online access to data coming from a real data archive. In view of ESA data distribution policies compliance with data access restrictions is needed. A final solution, fully complying with the ESA data policies of Human Spaceflight, is envisaged for the second public demonstrator session, together with more advanced and diversified tools and applications as mentioned in the previous chapter.

8. CONCLUSIONS

The ULISSE project tries to address many issues needed to improve utilisation of space experiment data. Solutions to document knowledge about experiments in a uniform way are proposed and being refined. New technologies allow for various improvements in providing access and adding value to existing datasets.

9. ACKNOWLEDGEMENTS

Contributions of the ULISSE team are gratefully acknowledged. The YaMCS library was developed by Space Applications Services and kindly made available for the project. The work has received funding from the European Commission through the Seventh Framework Programme (FP7/ 2007-2013) under the Space Theme, under grant agreement n°218815 within the ULISSE Project.

10. REFERENCES

[1] ULISSE project page: http://www.ULISSE-space.eu/.

[2] Erasmus Experiment Archive http://eea.spaceflight.esa.int/ .

[3] Human Spaceflight, Research And Applications Programme Board, Human Spaceflight Data Policy, ESA/PB-HSR(2004)29, rev.2, corr.1.

[4] OAIS reference model CCSDS 650.B-1.

[5] Kleppner, D., P.A. Sharp et al., Ensuring the integrity, asscessibility, and stewardship of research data in the digital age, National Academies Press, Washington, 2009.

[6] Metadata ISO 19115:2003(E), Technical Corrigendum (2006), ISO 19115:2003.

[7] Grimbach, A. , P. Wever, S.Schneider, R.Willnecker , Knowledge Preservation: A Semantic Approach To Visualizing And Reusing Microgravity Material Science Data, IAF conference paper, IAC-10-D5.2.6.

[8] Pronk, Z., L. Steinicke, J.M. Wislez, P. Dujardin, J. C. Degavre, Columbus payload operations from remote decentralized user support and operations centres, IAF conference paper, IAC-08-A2.5.4.

[9] Cesta, S. Fratini, and A. Oddi. Planning with Concurrency, Time and Resources: A CSP-Based Approach. In I. Vlahavas and D. Vrakas, editors, Intelligent Techniques for Planning. Idea Group Publishing, 2004.

[10] Cesta, A. Finzi, S. Fratini, A. Orlandini, E. Tronci. Analyzing Flexible Timeline Plan. In proc. of the 19th European Conference on Artificial Intelligence (ECAI 2010), LNAI, Vol. 215, IOS Press, ISBN 978-1-60750-605-8. 2010.

This page is intentionally left blank.

APPENDIX A: IAC presentation

THE ULISSE ENVIRONMENT FOR COLLABORATION ON ISS EXPERIMENT DATA AND KNOWLEDGE REPRESENTATION

Ed Kuijpers, National Aerospace Laboratory NLR, NL
Luigi Carotenuto, Telespazio s.p.a, Italy,
Christine Cornier, CNES, France,
David Damen, Space Applications Services, Belgium
Alois Grimbach, MUSC-DLR, Germany

IAC Prague,
IAC-10-D5.2.2

USOCs KnowLedge Integration and dissemination for Space Science and Exploration

30-Sept. 2010
V. 0.5

ULISSE

SEVENTH FRAMEWORK PROGRAMME

Overview

ULISSE

- Collaboration on experiment data
- Introduction ULISSE project
- Metadata and knowledge representation
- Platform and data valorisation tools
- Demonstrator and outlook

30 Sept. 2010 IAC-2010-D5.2.2 2

Collaboration on experiment data(1)

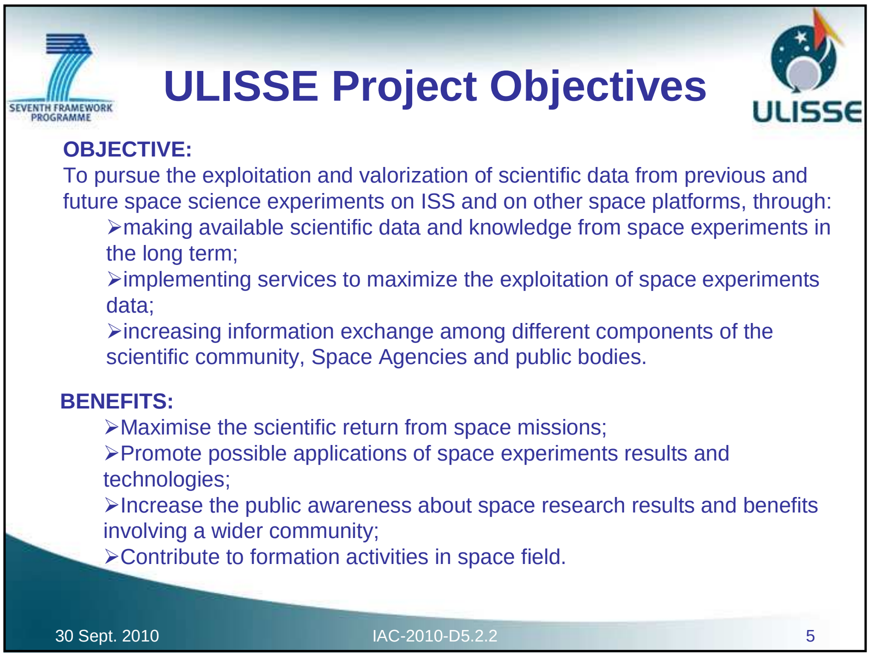
- ISS experiment data distribution complex
- Operations and raw data archiving via dedicated User Support and Operations Centre (USOCs)
- Experiment evaluation distributed over many science teams and disciplines
- Current limitations exploitation
  - Re-utilisation
  - Cross-fertilization between different scientific domains
  - Spin-off generation
  - European utilisation information

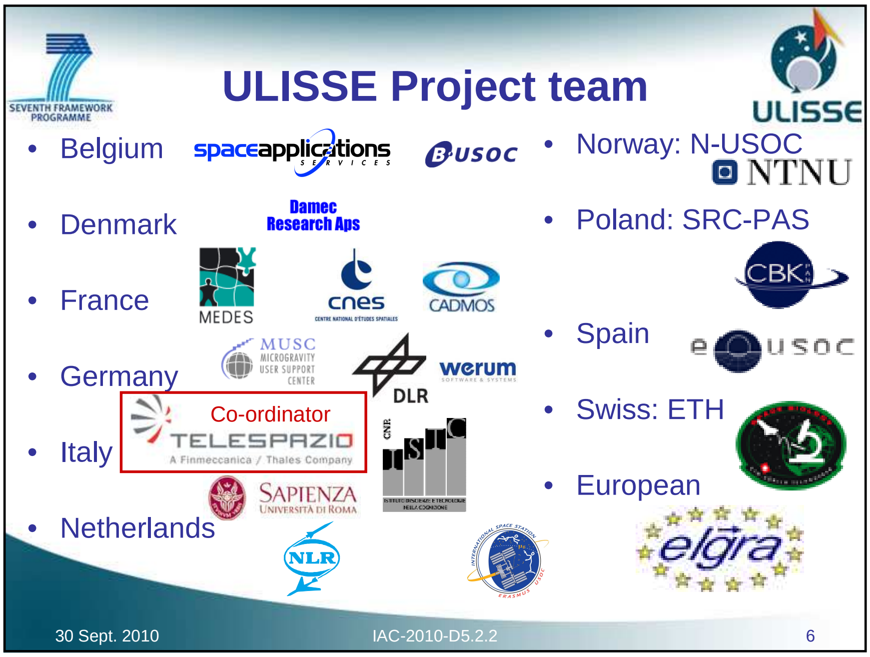
30 Sept. 2010 IAC-10-D5.2.2 3

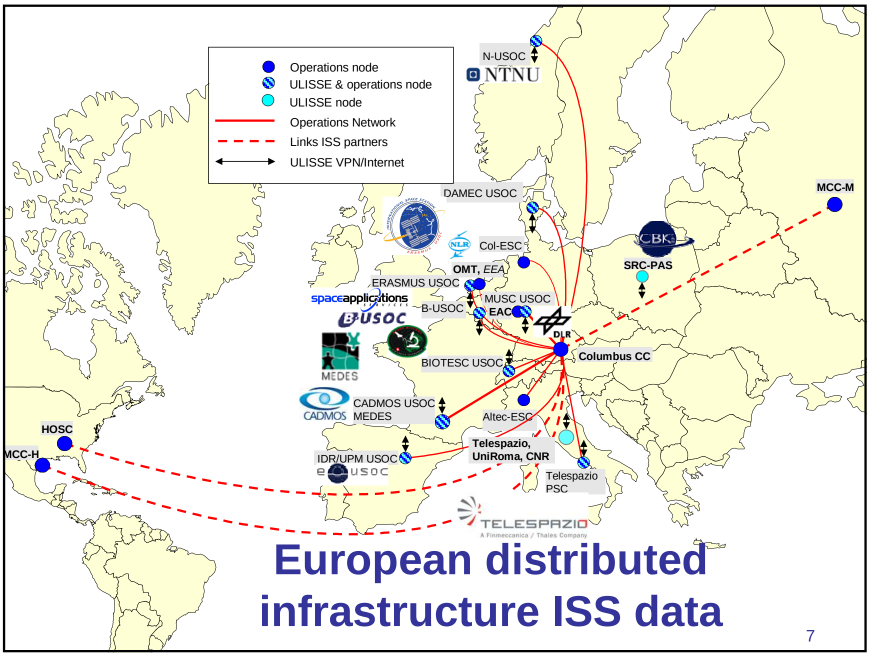
Collaboration on experiment data(2)

- Infrastructure for exploitation needs improvement:
  - Data
  - Archiving
  - Knowledge dissemination
- USOCs Knowledge Integration and dissemination for Space Science Experimentation (ULISSE) project within the European's Seventh Framework programme aims at developing an infrastructure and tools for improving preservation, valorisation and exploitation of data produced by European experimentation in space
  - Co-ordinated by Telespazio Napels (formerly MARS Center S.r.l.)
  - Team based on European USOCs and industry

30 Sept. 2010 IAC-2010-D5.2.2 4


SEVENTH FRAMEWORK PROGRAMME
ULISSE Project Objectives
ULISSE
OBJECTIVE:
To pursue the exploitation and valorization of scientific data from previous and future space science experiments on ISS and on other space platforms, through:
➢making available scientific data and knowledge from space experiments in the long term;
➢implementing services to maximize the exploitation of space experiments data;
➢increasing information exchange among different components of the scientific community, Space Agencies and public bodies.
BENEFITS:
➢Maximise the scientific return from space missions;
➢Promote possible applications of space experiments results and technologies;
➢Increase the public awareness about space research results and benefits involving a wider community;
➢Contribute to formation activities in space field.
30 Sept. 2010
IAC-2010-D5.2.2
5


SEVENTH FRAMEWORK PROGRAMME
ULISSE Project team
ULISSE
• Belgium
spaceapplications SERVICES
B.USOC
• Denmark
Damec Research Aps
• France
MEDES
cnes
CADMOS
• Germany
MUSC MICROGRAVITY USER SUPPORT CENTER
DLR
werum
• Italy
Co-ordinator
TELESPAZIO
A Finmeccanica / Thales Company
CNR
SAPIENZA UNIVERSITÀ DI ROMA
• Netherlands
NLR
• Norway: N-USOC
NTNU
• Poland: SRC-PAS
CBK
• Spain
e USOC
• Swiss: ETH
• European
elgra
30 Sept. 2010
IAC-2010-D5.2.2
6



METADATA Definition

SEVENTH FRAMEWORK PROGRAMME

ULISSE

Metadata descriptions needed for collaboration

Metadata is loosely defined as data about data.

Electronically archived data concept to describe:

- *definition*
- *structure*
- *administration*

of experiments and datasets in context to ease search and use of the captured and archived data for further use.

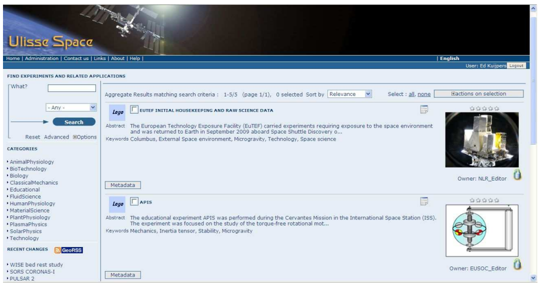
30 Sept. 2010 IAC-2010-D5.2.2 8

SEVENTH FRAMEWORK PROGRAMME

Metadata template

ULISSE

- ISO 19115
- Contents
  - Metadata information
  - Experiment
  - Maintenance
  - Constraints
  - Acquisition Information
  - Data sets Info
  - Data distribution info



- Metadata testing
  - USOC expertise on experiments

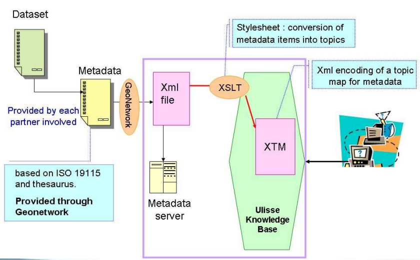
30 Sept. 2010 IAC-2010-D5.2.2 9

SEVENTH FRAMEWORK PROGRAMME

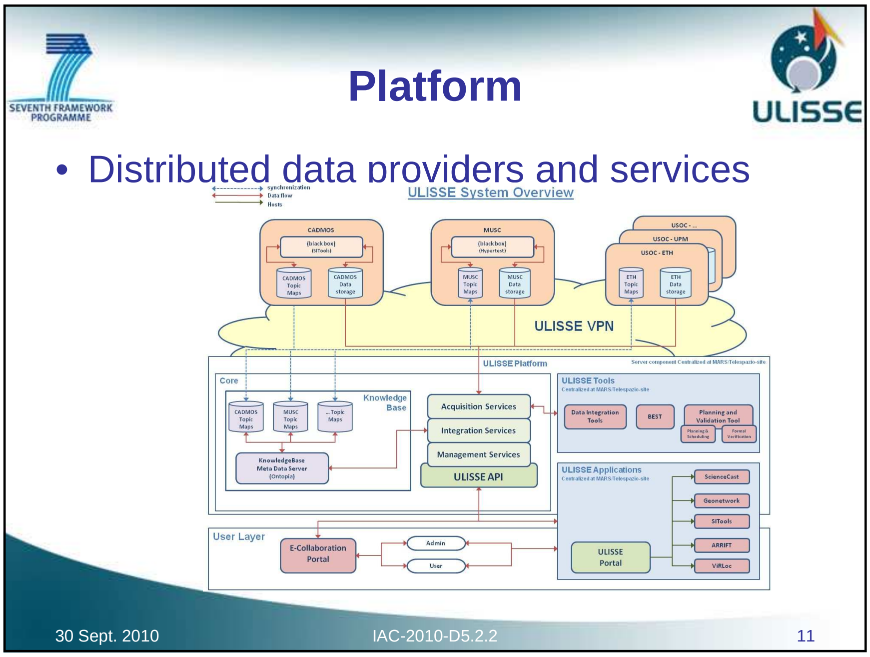
Knowledge representation

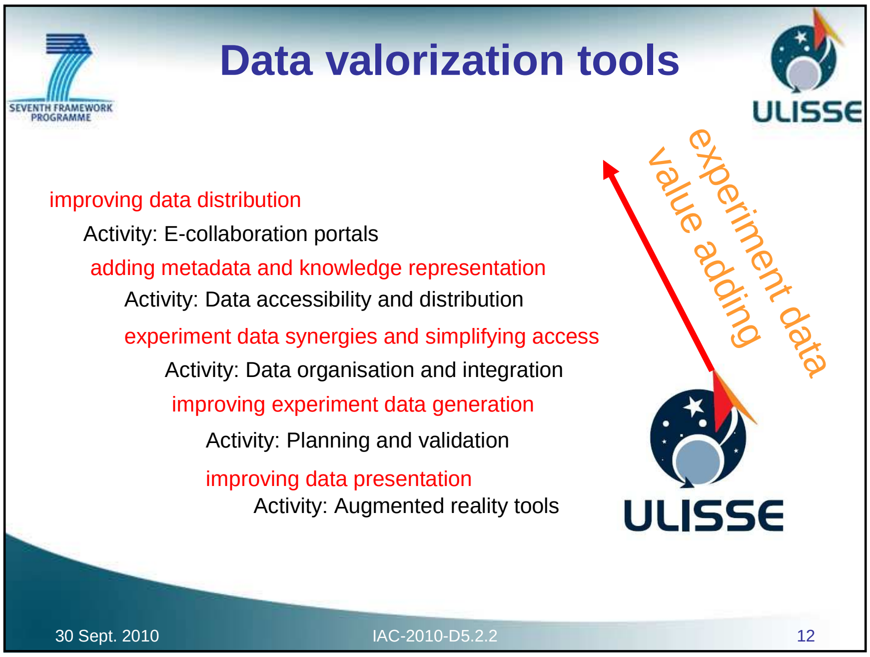
ULISSE

- Topic Map ontology descriptions:
  - Topic Type, Scope, Occurrence Type, Association Type (with role types), Unary Association Type, Supertype-Subtype Association, Type-Instance Association
- Dynamic part and Static part
  - Metadata import
  - Locking concept



30 Sept. 2010 IAC-2010-D5.2.2 10

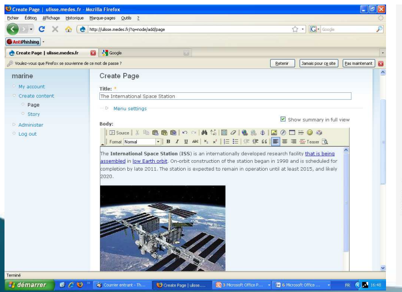

SEVENTH FRAMEWORK PROGRAMME
Platform
ULISSE
• Distributed data providers and services
synchronization
Data flow
Hosts
ULISSE System Overview
CADMOS
(black box)
(SITools)
CADMOS Topic Maps
CADMOS Data storage
MUSC
(black box)
(Hypertest)
MUSC Topic Maps
MUSC Data storage
USOC - ...
USOC - UPM
USOC - ETH
ETH Topic Maps
ETH Data storage
ULISSE VPN
ULISSE Platform
Server component Centralized at MARS/Telespazio-site
Core
Knowledge Base
CADMOS Topic Maps
MUSC Topic Maps
... Topic Maps
KnowledgeBase Meta Data Server (Ontopia)
Acquisition Services
Integration Services
Management Services
ULISSE API
ULISSE Tools
Centralized at MARS/Telespazio-site
Data Integration Tools
BEST
Planning and Validation Tool
Planning & Scheduling
Formal Verification
ULISSE Applications
Centralized at MARS/Telespazio-site
ScienceCast
Geonetwork
SITools
ARRIFT
ViRLoc
ULISSE Portal
User Layer
E-Collaboration Portal
Admin
User
30 Sept. 2010
IAC-2010-D5.2.2
11

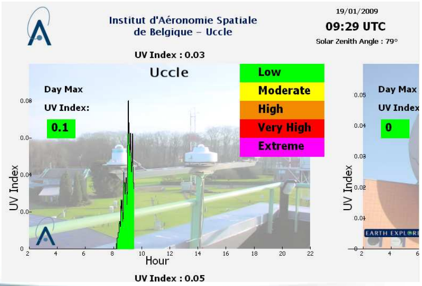

SEVENTH FRAMEWORK PROGRAMME
Data valorization tools
ULISSE
improving data distribution
Activity: E-collaboration portals
adding metadata and knowledge representation
Activity: Data accessibility and distribution
experiment data synergies and simplifying access
Activity: Data organisation and integration
improving experiment data generation
Activity: Planning and validation
improving data presentation
Activity: Augmented reality tools
experiment data
value adding
ULISSE
30 Sept. 2010
IAC-2010-D5.2.2
12

improving data distribution

•E-collaboration portal developed with Drupal (Content Management System) allows individuals and communities of users to easily publish, manage and organize a wide range of content on a website.

• Remote access of data demonstrated with **UV measuring network of BIRA-IASB** dataset

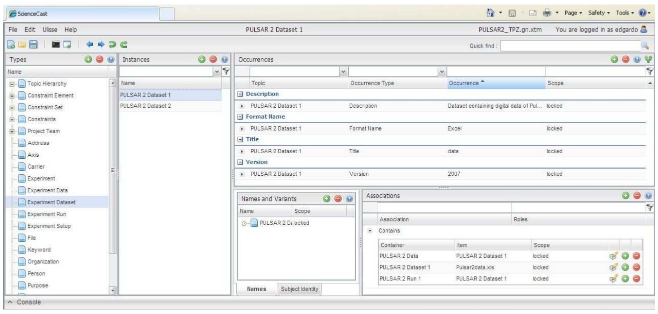




30 Sept. 2010 IAC-2010-D5.2.2 13

adding metadata and knowledge representation

• ScienceCast



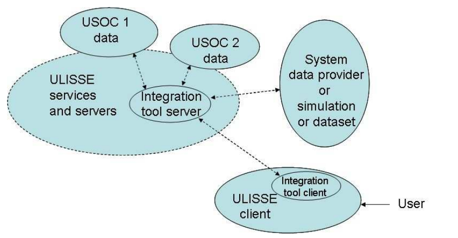
• GeoNetwork metadata

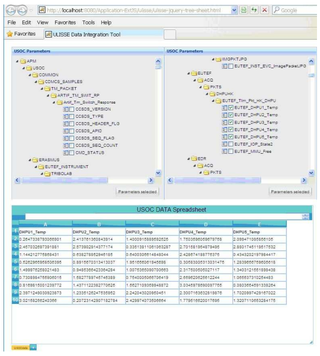
30 Sept. 2010 IAC-2010-D5.2.2 14

experiment data synergies and simplifying access

SEVENTH FRAMEWORK PROGRAMME — ULISSE

- Integrating data from various sources
  - dedicated libraries/databases





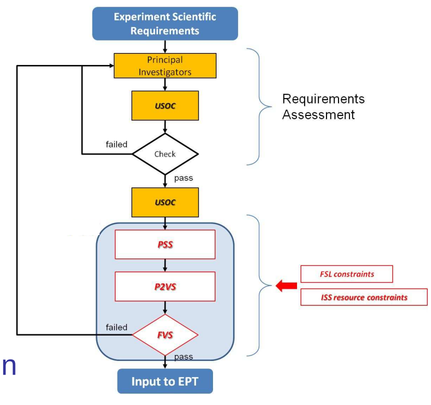
- Structure descriptions data
  - EAST, DESDL (BEST = Beyond EAST)

30 Sept. 2010 IAC-2010-D5.2.2 15

improving experiment data generation

SEVENTH FRAMEWORK PROGRAMME — ULISSE

- PSS: Planning and Scheduling Service
- FVS: Formal Verification Service
- P2VS: Planning to Verification Translation Service
- PVS: Planning and Validation Service
- ProcVS: Procedure Validation and Verification Service



Columbus FSL Study case

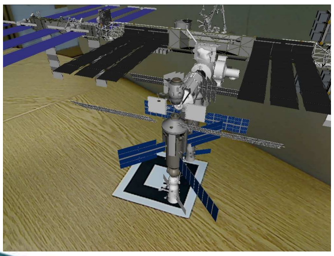
30 Sept. 2010 IAC-2010-D5.2.2 16

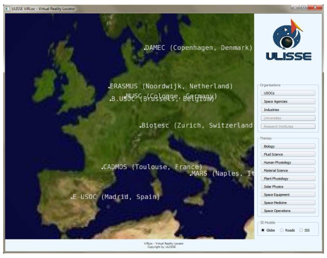
SEVENTH FRAMEWORK PROGRAMME

improving data presentation

ULISSE

- Augmented reality
  - Marker recognized in camera guides 3-D overlay
- Virtual reality
  - 3-D models for user interfacing





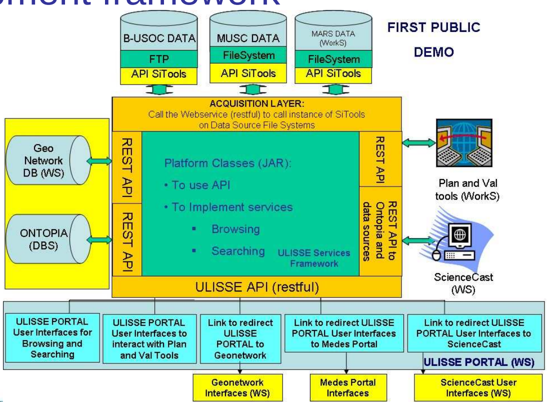
30 Sept. 2010 IAC-2010-D5.2.2 17

SEVENTH FRAMEWORK PROGRAMME

First demonstrator

ULISSE

- Development framework



30 Sept. 2010 IAC-2010-D5.2.2 18

Outlook

- First iteration on metadata and knowledge representation of ISS experiments performed
- Platform and tool solutions for data representation, distributed archiving and knowledge dissemination explored
- Planning
  - First demonstrator Nov. 2010 (Spaceweather week)
  - Second demonstrator 2011 (ELGRA Symposium)
  - Evaluation for exploitation scenario

30 Sept. 2010 IAC-2010-D5.2.2 19

- Acknowledgements:
  - EU FP7
  - ULISSE team
- Thank you for your attention
- Questions?

30 Sept. 2010 IAC-2010-D5.2.2 20