PERICLES - PROMOTING AND ENHANCING REUSE OF INFORMATION THROUGHOUT THE CONTENT LIFECYCLE TAKING ACCOUNT OF EVOLVING SEMANTICS , APPLICATION TO EARTH AND SPACE SCIENCE DATA.

Christian Muller and the PERICLES consortium.

B.USOC
Belgian Users Support and Operation Centre, Brussels, Belgium
christian.muller@busoc.be

ABSTRACT

PERICLES (Promoting and Enhancing the Reuse of Information throughout the Content Lifecycle exploiting Evolving Semantics) is an FP7 project started on February 2013. It aims at preserving by design large and complex data sets. PERICLES is coordinated by King's College London, UK and its partners are University of Borås (Sweden), CERTH-ITI (Greece), DotSoft (Greece), Georg-August-Universität Göttingen (Germany), University of Liverpool (UK), Space Application Services (Belgium), XEROX France and University of Edinburgh (UK). Two additional partners provide the two case studies: Tate Gallery (UK) brings the digital art and media case study and B.USOC (Belgian Users Support and Operations Centre) brings the space science case study.

PERICLES addresses the life-cycle of large and complex data sets in order to cater for the evolution of context of data sets and user communities, including groups unanticipated when the data was created. Semantics of data sets are thus also expected to evolve and the project includes elements which could address the reuse of data sets at periods where the data providers and even their institutions are not available any more.

B.USOC supports experiments on the International Space Station and is the curator of the collected data and operation history. The B.USOC operation team includes B.USOC and Space Applications Services personnel and is thus ideally configured to participate in this project. As a first test of the concept, B.USOC has chosen to analyse the SOLAR payload operating since 2008 on the ESA COLUMBUS module of the ISS. Observation data are prime candidates for long term data preservation as variabilities of the solar spectral irradiance have an influence on earth climate. The paradigm of these observations has already changed a lot in the last fifty years from a time where scientists were aiming at determining with high accuracy the "solar constant" which was the total solar energy per surface unit received at the top of the earth's atmosphere to the present situation where the same quantity is known as the total solar irradiance and has been shown by thirty years of space observations to vary of about one tenth of a per cent in synchronism with the solar cycle. Right now, larger variations have been detected at UV wavelengths but their effects on climate and atmospheric chemistry are still a matter of scientific discussion.

By creating semantic links between various data bases, the PERICLES process can be applied to various already linked data bases as the current set of earth observation data already managed by ESA in ESRIN to optimise their future use as an element of the observational data base of future earth's system models. PERICLES also presents a fundamental reflexion on the reuse and long term preservation of data which corresponds to the needs of climate research. These arguments on long term data use apply also to the space science data bases hosted by ESA in ESAC. This communication will especially insists on the application of the concept to long duration studies and the possibilities of managing the interactions between related data series.

***Index Terms*—** applications and use cases, data models, lifecycle, management, and databases

1. INTRODUCTION

The PERICLES project is funded through the FP7 ICT Call 9 Digital Preservation. The project involves partners of a range of complementary types, including six academic partners, one multinational corporation, two SMEs and two non-academic public sector organisations.

As digital content and its related metadata are generated and used across different phases of the information lifecycle, and in a continually evolving environment, the concept of a fixed and stable 'final' version that needs to be preserved becomes less appropriate. As well as dealing with technological change and obsolescence, long-term sustainability requires us to address changes in context, such as changes in semantics - for example, the 'semantic drift' that arises from changes in language and meaning - or disciplinary and societal changes that affect the practices, attitudes and interests of the 'stakeholders', whether these be curators, artists, scientists, or indeed a broader public, such as visitors to exhibitions.

Such a changing environment necessitates a corresponding evolution of the strategies and approaches for preservation if stakeholder communities are to be able to continue to use and interpret content appropriately. A key issue is the provision of sufficient contextual information to enable both lifecycle management and preservation on the one hand, and re-use or re-interpretation of content on the other, as well as the facility to model and describe preservation processes, policies and infrastructures as they themselves evolve. Capturing and maintaining this information throughout the lifecycle, together with the complex relationships between the components of the preservation ecosystem as a whole, is key to an approach based on 'preservation by design', through models that capture intents and interpretative contexts associated with digital content, and enable content to remain relevant to new communities of users.
This paper addresses especially the space cases provided by BUSOC and its partner SpaceAps in the frame of ESA and other space agencies programmes.

2. THE ESA USOC NETWORK

The ISS European payload operations are conducted through the distributed USOC network.

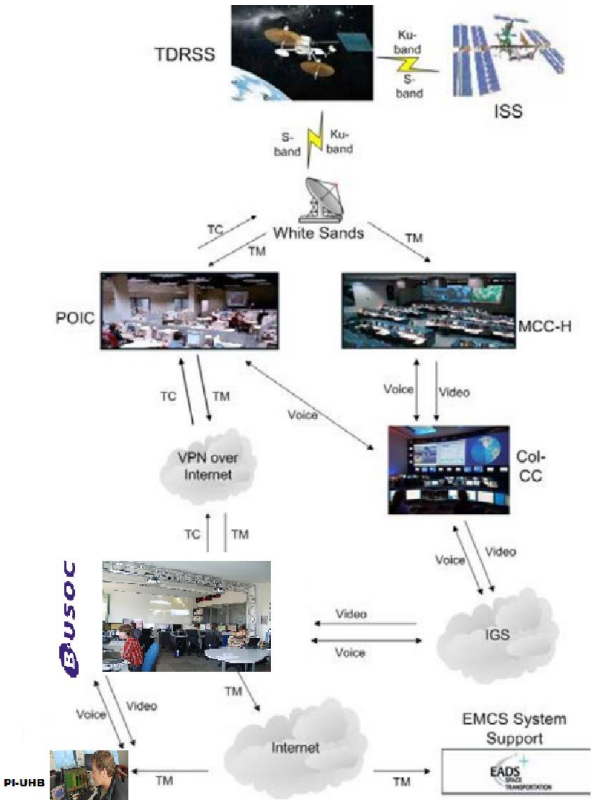

Fig.1: Schematics of European USOC operations, payloads are not only located in ESA COLUMBUS module but could also be in the U.S. segment requesting data and commands to transit through NASA centres while the ESA data flows through the ESA IGS (Interconnection Ground Subnetwork).

3. THE SOLAR PAYLOAD ON THE ISS

The SOLAR payload is built from three complementary space science instruments that measure the solar spectral irradiance with an unprecedented accuracy across almost the whole spectrum: 17-3000 nm. This range carries 99% of the Sun's energy emission. Apart from the contributions to solar and stellar physics, knowledge of the solar energy flux (and its variations) entering the Earth's atmosphere is of great importance for atmospheric modelling, atmospheric chemistry and climatology. The three instruments are: SOLSPEC (Solar Spectra Irradiance Measurements, developed by CNRS,France and IASB/BIRA, Belgium) [1], SOL-ACES (Auto-Calibrating Extreme Ultraviolet and Ultraviolet Spectrophotometers, developed by the Fraunhofer Institute, Germany) [2], SOVIM (Solar Variable and Irradiance Monitor, jointly developed by the Observatory of DAVOS, Switzerland and the Royal Meteorological Institute, Belgium). [3]. The three original PI's agreed before flight to a synergistic treatment of the data [4].

SOLAR has in fact a much longer history than its current flight on COLUMBUS. The precise measurement of the solar irradiance as input to the earth system began one hundred years ago when this parameter was known as the "solar constant", space borne instruments in the last thirty years have shown variations of the total solar irradiance while spectral irradiance especially in the UV has confirmed early balloon and rocket observations of high variations. The SOLAR instruments SOLSPEC and SOVIM were first designed for the SPACELAB 1 payload which flew on the US space shuttle in 1983, the decision to fly and the first design studies dating from 1975. After SPACELAB-1, ESA transferred the SPACELAB equipment to NASA and NASA flew these payloads several times in order to cover the solar cycle until the last COLUMBIA mission in 2003. Ideally, at least this set of missions should be regrouped with the SOLAR ISS data set in order to build a coherent series.

During all these years and even during the ISS SOLAR mission the paradigm of the observations has changed. In 1975, the objective was still to determine accurately the solar constant together with a precise spectrum ranging from the UV to the near infrared. In the next flights, it was to perform the same determinations at specific periods of the 11 year solar cycle as the minimum or maximum. In 2012, 2013 and 2014, new operation modes aim at detecting variations during a full solar rotation which observed from the earth is a time scale of about 27 days, the purpose of this

exercise which requests a 7° attitude change of the ISS is to aim at the detection of even shorter variations related to sunspot activity.

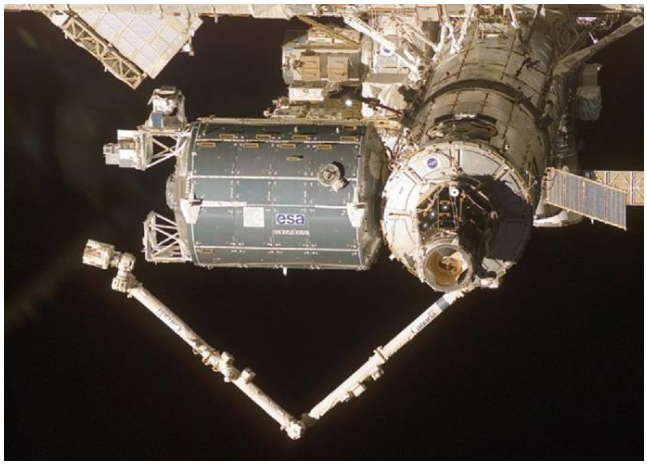

Figure2: the SOLAR package as in operation since February 2008 on the ISS COLUMBUS module, NASA document.

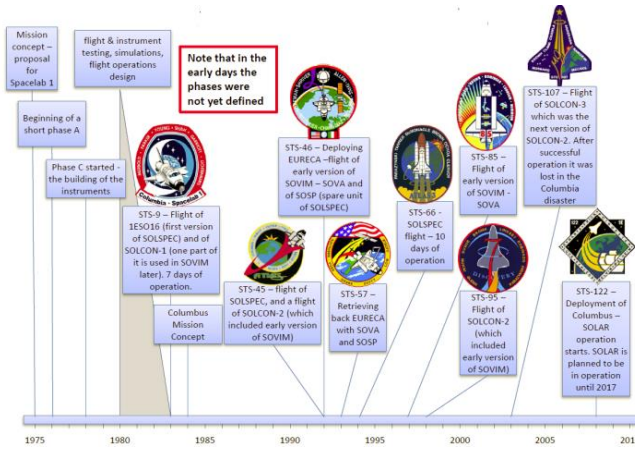

Figure3: flight history of SOLAR instruments and their precursors.

4.ROLE OF B.USOC IN THE ARCHIVING OF DATA

B.USOC as an operation centre of the ISS ESA distributed operation network acts the Facility Responsible Centre for SOLAR, it has thus an ESA mandate to keep a repository of all documents and data generated by SOLAR before and during operations, it provides also the requested data flows to the Principal Investigators User Home Bases (UHB's), the scientists generate at their UHB's the higher level products and the science publications and communications.

An inventory of these documents in the PERICLES perspective was already described [5] and presented at the PV2013 conference in 2013.

The security of the operations is an absolute requirement of ESA and the USOC network and thus, no development or tests of PERICLES new products can be performed in the operational environments before these products are mature and obtain the approval of all concerned parties, thus B.USOC has decided to develop new products on a data server used for the distribution of data from a ground based network of solar monitors [ulisse.busoc.be]; this server acts as a mirror synchronised once a day of the original ground based data which thus remains protected from possible user intervention. In a following step, a similar mirror will be established for slices of the ISS mission data and will be used to design the transition of the B.USOC data and document repository to a real reusable archive ready for long term data preservation. Again, this last step will be performed in agreement with all parties and the space agencies.

B.USOC currently uses a management and command software (YAMCS) [6] designed in a LINUX environment to parse the data so that the scientists receive their own data flow and that the operators receive the necessary ancillary information on their monitors, it is planned using the PERICLES developments to evolve this software into a full data archiving tool so that the entire mission could be replayed from the repository and produce a final archive. The capability of new uses of the YAMCS has recently been shown by the implementation of TYNA, a YAMCS based operation software diminishing the pressure on operators during non-essential phases where no science data are acquired. This final archive should also include the scientific products already published and available at the scientists UHB's. This science part is important as the interpretation process is based on the instrument knowledge and a good interpretation of the calibration processes, changes of procedures have sometime in the past not been documented and even worse, the prefight calibration data is often not archived with the data. It could thus be lost diminishing the possibility of analysis of the final results in order of understanding differences between similar observations. This process requires the full cooperation of the scientific teams.

This final archive will then be reusable for the generation of the specific products requested by the data centres of the future for the study of long term variations.

For new projects, B.USOC intends to use the lessons learnt from PERICLES in its data management plans so that the archiving takes already place at acquisition time.

5. SPECIFICITIES OF THE PERICLES PROJECT

The PERICLES project goes much further than its application to the single SOLAR case, it intends to develop itself into a new scheme in acquiring the data from new missions attributed by the space agencies to B.USOC.

One of its main elements will be to constitute already the basis of data preservation at acquisition time instead of having to replay the mission, as planned for SOLAR, at the end of the operations of its space segment.

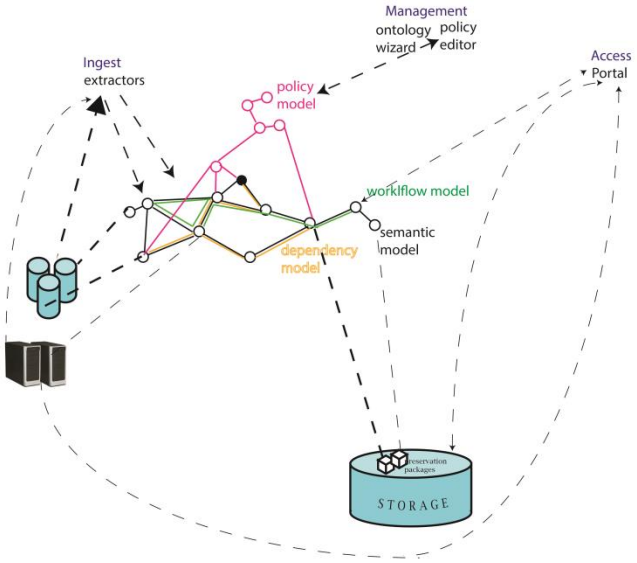

Fig. 4: the "big picture" describing the PERICLES project, the storage box is a living element where new data constantly arrive from the space segments.

The different elements of the PERICLES project are described in Fig. 4, they are now tested on the ground based IASB-BIRA UV network data before requesting the ESA and PI's authorisation to use them as a new B.USOC tool in the management of ISS data. In particular, the "Process Extractor Tool" and the "Anomaly Detector" have been recently evaluated in the B.USOC environment using the ground based UV network.

6. FURTHER DEVELOPMENTS

The PERICLES project lasts until 2018 while SOLAR will be financed until 2017 and must not be removed before 2020, leaving thus place for mission extension. At the time of SOLAR wrap-up, the PERICLES products will have reached maturity and will participate in the constitution of the SOLAR global archive. Other space agencies will continue solar monitoring with different or improved instrument, the PERICLES process will then assist in ensuring continuity of the series using the current SOLAR observations.

REFERENCES

[1] - Thuillier G., Foujols T., Bolsée D., Gillotay D., Hersé M., Peetermans W., Decuyper W., Mandel H., Sperfeld P., Pape S. et al, SOLAR/SOLSPEC: Scientific Objectives, Instrument Performance and Its Absolute Calibration Using a Blackbody as Primary Standard Source, Solar Physics 157, 185-213, 2010

[2] - Schmidtke, G.; Brunner, R.; Eberhard, D.; Halford, B.; Klocke, U.; Knothe, M.; Konz, W.; Riedel, W.-J.; Wolf, H. SOL–ACES: Auto-calibrating EUV/UV spectrometers for measurements onboard the International Space Station., Advances in Space Research ,37, 273-282, 2006.

[3] - Mekaoui , S., Dewitte, S. ,Conscience, C., Chevalier, A., Total solar irradiance absolute level from DIARAD/SOVIM on the International Space Station, Advances in Space Research , 45 1393–1406, 2010.

[4] - ISS-SOLAR: Total (TSI) and spectral (SSI) irradiance measurements. Schmidtke, G.; Fröhlich, C.; Thuillier, G. Advances in Space Research , 37, . 255-264, 2006.

[5] Muller, C. et al, PERICLES: a new knowledge management programme applied to SOLAR data on COLUMBUS, PV2013 proceedings, ESA, 2013..

[6] - Sela, A. , Mihalache, M and Moreau, D, YAMCS - A Mission Control System, SpaceOps2012, American Institute of Astronautics, 2012.

.

ACKNOWLEDGMENTS

The authors wish to thank the European Space Agency (ESA,Directorate of Human Spaceflight and Operations) and the Belgian Science Policy Office (BELSPO) (ESA Prodex and other programmes) for SOLAR funding and support of the ground based solar UV network. PERICLES is funded under FP7 ICT Call 9, and addresses work programme objective ICT-2011.4.3 Digital Preservation and we are grateful to the European Commission for its support.