PERICLES: a new knowledge management programme applied to SOLAR data on COLUMBUS.

Christian Muller, Didier Moreau, Nadia This. Alice Michel
Belgian User Support and Operations Centre
.Rani Pinchuk, Jelle Pelfrenne, Saliha Klai, Carla Jacobs
Space Applications Services
and the PERICLES consortium
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.

In this paper, the early stages of PERICLES will be described in the context of the space science case study – mainly, the requirements gathering activity, the data survey and the work on domain ontologies. The SOLAR data set and related data and documentation will be presented as they are preserved at B.USOC. In addition, the paper will present an initial roadmap aiming at the reuse of the data in a situation where both current scientists and operators will no longer be available and initial plans for tools to be developed in the project to achieve this objective.

Keywords: Solar, space data, Preservation models, lifecycle, data analytics, semantics.

INTRODUCTION

This paper describes the objectives, approach, use cases and proposed deliverables of the EC FP7 Integrated Project PERICLES: http://www-pericles-project.eu. The PERICLES project was 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.

The project will address these preservation challenges in relation to digital content from two quite different domains: on the one hand, digital artworks, such as interactive software-based installations, and other digital media from Tate’s collections and archives; on the other hand, payload data originating from the European Space Agency and International Space Station. This communication is centred on the space case managed by B.USOC and Space Applications Services

PROBLEM DESCRIPTION

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.

In the specific Space Case, this communication is dedicated to solar observations from the ISS SOLAR package launched in 2008 with COLUMBUS and which evolved from a short-time solar monitor operation to be repeated in the ISS life time to a long-life observatory more and more aimed at the detection of solar variability and its influence on earth climate and atmospheric chemistry. In the case of SOLAR, the final objective of the PERICLES programme will be to produce a data set which will be reusable in seventy years from now when the present instrument scientists and most of their institutions will not exist anymore.

THE SOLAR PACKAGE;

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 modeling, 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, Deutschland) [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] .

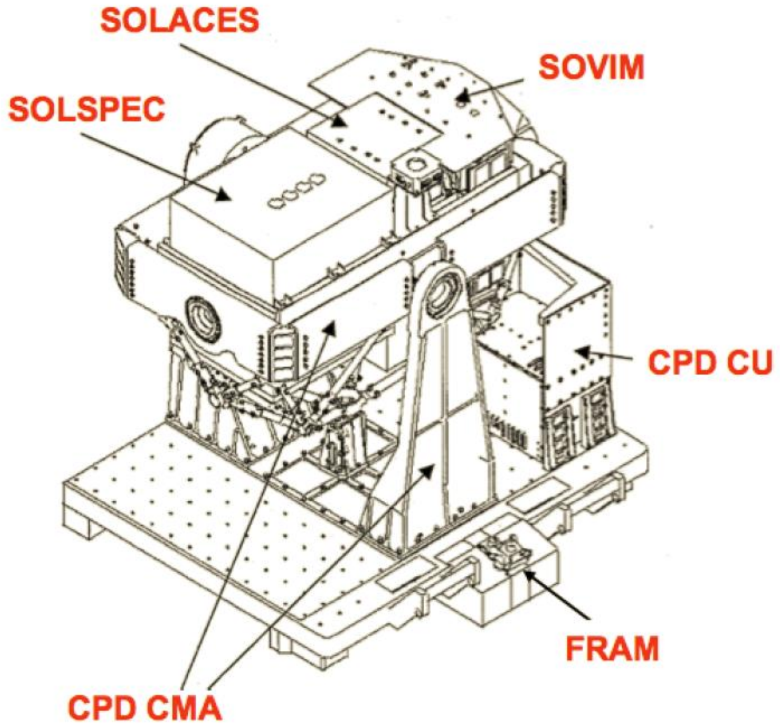

Figure1: the SOLAR package as it flies on the ISS COLUMBUS module, the supporting platform (FRAM and CPD) aims solar pointing during the solar visibility periods.,

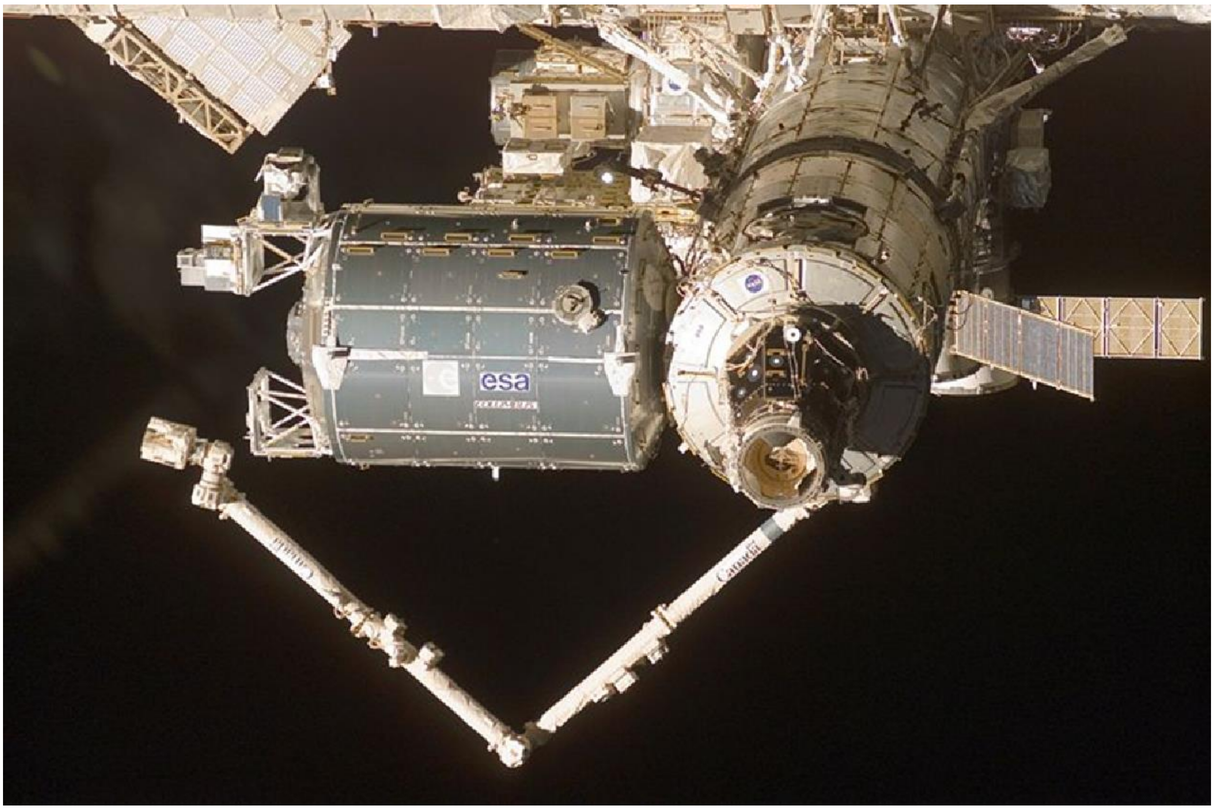

Figure2: the SOLAR package as in operation since February 2008 on the ISS COLUMBUS module, NASA document.

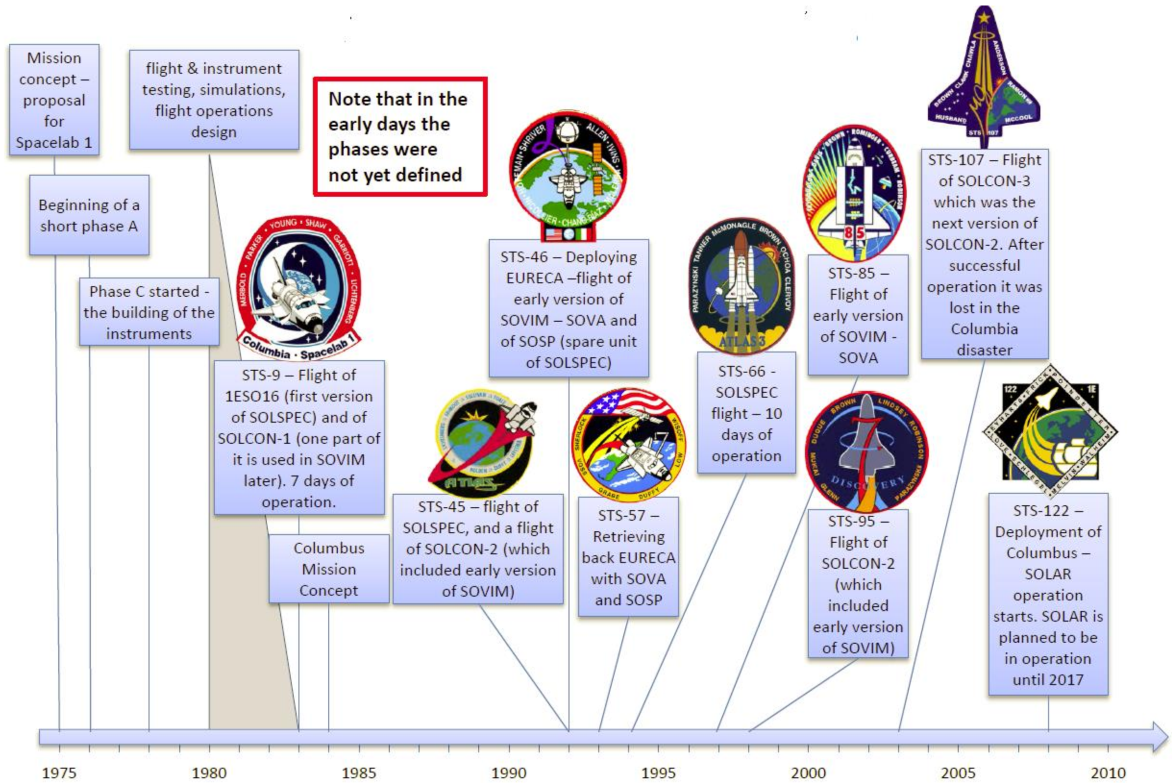

Figure3: flight history of SOLAR instruments and their precursors.

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 reflew 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. Recently in 2012 and 2013, 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.

Thus the objective of SOLAR moves from delivering a reference solar data set to the monitoring of solar weather and climate. This will help to understand the influence of the natural solar forcing on the earth system. The PERICLES exercise should allow the future user communities to revisit the data set for new objectives which were not perceived at design time.

CONTENT OF THE DATA SET

The user communities have different perceptions of the data collection, these communities are currently: the engineer's user group, operations user group, sciences users and agencies and data owners. PERICLES has also to address a fifth community: the scientists of the future which is a very prospective task as those may not be consulted.

Engineering documents

From an engineering point of view, a space project follows now the phasing recommended by the "European Cooperation for Space Standardization. The different phases are Phase 0 – Mission analysis/needs identification: This is mainly an activity conducted by the project initiator, the top level customer and representatives of the end users. The subsequent phases are: Phase A – feasibility: This is mainly an activity conducted by the top level customer and one or several first level suppliers with the outcome being reported to the project initiator, and representatives of the end users for consideration, Phase B – preliminary definition, phase C –detailed definition, Phase D – qualification and production, Phase E – operations/utilisation, phase F – disposal. After phase F, a still to be defined phase of Long Term Data Preservation will follow, PERICLES addresses not only this last phase but also all the documents and data generated before.

Phase E is sometime divided between the commissioning phase and the operational phase, this division is not only administrative, the transition is marked by the end of the mandate of the industrial

contractors still present during commissioning and the full transmission of the payload management to the operation team. This point it involves a handover of documentation between two teams.
The first document of this engineering data base should be the ESR : Experiment Science Requirements, drafted by ESA project Scientists detailing the requirements from the Principal Investigators, actually for SOLAR this ESR was only drafted when the payload was already operational for several years. Generally this is document is part of phase A. The a posteriori generation of this document is related to SOLAR history. In the early SPACELAB days, payloads were not perceived as parts of the space segment, the original SPACELAB concept was that any payload developed in a sciencet laboratory and following specifications similar to airborne payloads could fly in a rack in the SPACELAB module or on the exposed pallets. After SPACELAB-1, NASA accepted the payloads on the basis of the previous flight analysis and thus the phasing and hierarchy of documents existing for SOLAR ISS is not applicable to the previous missions.
For SOLAR-ISS, the Engineering DB counts 923 documents from the phases 0 to D, these include: specifications and design documents, acceptance data packages, test plans, acceptance, functional and other test reports ,instrument/PL ICDs, Other ICDs (Interface Control Document), assessment reports, safety data packages, design Reports, CIDL (Configuration Item Design List)/ User/Operations manuals, certifications, thermal analysis, verification control document.

Operation documents prepared in advance of the operations. .

In 1998, the ESA Manned Space Program board decided to adopt a decentralised infrastructure for the support of European payloads on-board the International Space Station (ISS). This concept was based on operating multiple User Support and Operations Centres (USOC's), each assigned to supporting a majority of tasks related to the preparation and in-flight operations of European payloads. The USOC's are based in national centres distributed throughout Europe. Depending on the tasks assigned to a USOC, they have the responsibility of a Facility Responsible Centre (FRC) or Facility Support Centre (FSC). While FRC is delegated the overall responsibility for a multi-user rack facility or class-1 payload an FSC takes up the responsibility for a sub-rack facility, class-2 payload (e.g. experiment container, drawer payload etc.) and/or self-standing experiments utilising experiment specific equipment. For example the BioLab is operated from the Microgravity User Support Centre (MUSC) in Cologne. The SOLAR payload is operated by the Belgian USOC located in Brussels.

There are many rules and guidelines to follow: **JOIP**: Joint Operations and Interface Procedures. It outlines the operational interfaces and guidelines for all mission phases of the ESA Human Spaceflight programme. It covers all processes and interactions of real-time operations in Columbus such as anomaly handling, ground segment requests, voice interaction, commanding, real-time change requests , **OIP**: Operations Interface Procedure. It outlines operational interfaces and guidelines for all mission phases of the USOS programme, Flight Rules: a set or pre-planned decisions concerning off nominal situations; also a set of definitions, delegations of authority and responsibilities of all organizations involved in the conduct of mission operations. The purpose of flight rules is to minimize the amount of real time rationalization and discussion when a contingency situations or decision points occur, **Payload regulations**: definition is the same as flight Rules, with the difference that payload regulations address only nominal and off-nominal payload operations that do not have safety and vehicle integrity implications, **PODF**: Payload Operations Data File is a XML file, providing a pre-defined sequence of commands and checks to operate the payload. For payloads interfacing Columbus every ground operator and crew member needs to follow a PODF to execute a payload activity. PODFs are retrieved from IPV. IPV is a web application for managing and displaying the database of on-board and ground procedures. , **Operations Manual:** B.USOC specific guidelines and procedures for the on console operations. It consists of references to documents, information for on console, Point of contacts and ground procedures. Ground procedures are procedures describing typical processes actions such as File transfer preparation or start up of the workstation**.**

Living operation documents generated by B.USOC

BUSOC wiki : the BUSOC wikipages provide an overview of the roles and responsibilities, guidelines on Increment preparation, use of the Predictor tool etc., **eRoom:** the eRoom is a Documentum installation that serves as a repository of interface control documents, user manuals, specifications, technical notes, procedures, protocols, minutes, presentation, reports, operational products, databases and emails, **Minutes of Meetings**: on a weekly basis there are meetings focused on the preparation and on-going operations. These can be internal only or with external parties from the Flight operations or the scientists (European Weekly operations Conference (EWOC), Mission Science Office telecon, BUSOC Ops days, ...).

Living operation documents generated by the agencies managing the ISS

Other data sources originating from the ISS international management are also used during SOLAR operations, these are the **OSTPV** (On board Short Term Planner Viewer) – OSTPV is a web application that displays timeline information about ground and on board procedures, schedules and activities. Data can be received from a remote site in a textual, structured format. However, the primary OSTP user interface is horizontally scrollable web page with a timeline over several days and various events scattered on it. BUSOC Operator views OSTPV in read-only mode, a dedicated user account from NASA is required to access the OSTP. **MDB** – The Mission Data Base contains a machine-readable description of the telemetry, including the size of various parameters sent in telemetry packets and their interpretation from binary to engineering values, soft and hard limits as defined in the Design phase.
B.USOC manages the telemetry data with YAMCS is a mission control system. It allows acquiring telemetry and dispatching telecommands. It can parse the telemetry stream and store it. YAMCS is open source software developed by SpaceApps and BUSOC. Information about it is available in http://www.yamcs.org. [5]

Telemetry data originating from the payload and ISS.

The actual data coming from SOLAR and the ISS represent the telemetry. It is data organized into packets that is sent by the payload to the control centre. Three main different kind of telemetry can be considered**: Housekeeping data**: Typical engineering data of the payload directly sent from the ISS to the control centre: it contains temperature measurements, voltage and current readings, various operational states and reports, such as rotation axes of moving parts). In order to understand the housekeeping data, the operators use the SOLAR User Manual, **Health and status**: Limited set of engineering data from sensors ‘outside’ the actual payload which are sent from the ISS to Col CC and forwarded from there.

Science data

Science data belongs to the scientific community, represented by the PI. The B.USOC operators have no insight in the science data; they provide time-tagged telemetry of the science channels to the User Home Bases (UHB) of the PI’s providing them exactly the data flux they required before the mission, this pre-treatment uses the functionalities of the YAMCS.
The raw science data must be calibrated. For example, the instruments use reference lamps as calibrated sources, this has been performed extremely carefully before flight. Internal lamps and radiation sources allow continuing this process during flight. The SOLSPEC detector for example is calibrated once in 24 hours during sun visibility window [1]. Another factor influencing the signal is the South Atlantic Anomaly. This anomaly causes an increased flux of energetic particles in the South Atlantic region and exposes orbiting satellites to higher than usual levels of radiation. These disturbances have an effect on detector performances and must be taken into account when calibrating the results. As for all satellites, magnetic storms related to solar flares may lead to put all non-essential systems in standby.
The scientists generate a science product from the data they receive and publish and archive it, one of the PERICLES objectives is to collect also these data together with the metadata (calibration parameters, spectral alignments, correction procedures) actually used by the science teams to allow future reuse of the data. This point constitutes the most difficult point as the scientists consider their own verifications together with external peer review at publication stage to be a definitive data quality criterion.

Data generated by B.USOC during operations.

Other data are generated by B.USOC in the process of operating the payload: **Telecommands** are structured data sent to payloads during the operations. They may contain control structures for shutting up or starting various modules, as well as uploads of data and scripts. A complete history of telecommands over the operating live of the payload is saved and is made available to the operational environment and to the scientific partners. Activities are usually performed following PODF’s (Payload Operations Data Files), where a manual stack of commands is used. These manual stacks are just lists of telecommands, but under configuration control and following a PODF
Auxiliary data : most of auxiliary data comes from public sources. For instance, current B.USOC operations related to the SOLAR payload heavily depend on TLE (two-line elements) to predict the position of the ISS and on the ISS attitude timeline (ATL) to predict the orientation of ISS towards the Sun. The two external data sources are combined in order to create a full prediction of the upcoming month allowing to create a clear science planning and optimal operations support plan. In more details, the data are automatically downloaded to the server. The Predictor Tool (see later) extracts the relevant information to a database. These data are used by the Predictor to calculate the ISS altitude and beta angle. Together with the yaw, pitch, roll data from the attitude database the sun passes can be predicted. **CEFN** Tool– Columbus Electronic Flight Notes (CEFNs) Tool provides an automated workflow for written communication, review, processing and tracking. It is a tracking web application accessible for all Columbus Flight control team members. It is coded in PHP and does not provide any functionality to export contents of the notes in machine-readable format. A CEFN may refer to draft, Inter-console or Flight Note.

Console Logs : Console logs are short time stamped messages that the operators update every time they make an operation or retrieve information on console which is useful for the operations. Console logs are primarily used for passing over information in between shifts and can also be used for forensic analysis. **SOLAR Mission Tool Light** – the SOLAR Mission Tool light was the precursor of the Predictor Tool. At the time of the development it covered planning of the activities, automatically identifying constraints or conflicts with respect to external factors such as visiting vehicles or SAA passes. It consists of an excel sheet with different tabs for different features. Currently it is still used for short and long term handover information and to gather inputs for the BUSOC optimization board. **Checklist :** the checklist remains on console and provides the operator a list of routine tasks to be performed during his/her shift (sample check List v2.5). **DOR** : Daily Operations Report provides an overview of the 24 hours of SOLAR science activities. It includes a summary of the operations, planning of the coming days within a Sun Window as well as reports on discrepancies and anomalies. The DOR is automatically generated by the Predictor Tool. The DOR represents one of the main sources to follow the history of the payload.

Conclusions.

PERICLES aims at the preservation of all the described data for future re-use and new treatment of the mission. The main objective of the mission is scientific and this reuse will also be centred on scientific objectives. The first will be to provide a continuous series with a level of coherence such that cyclic variations could be determined. This was already done since the beginning of the space age around 1975 for the 11 years solar cycles where data from different solar monitors were regrouped and show for the first time the 11 year cycle in measurements of total solar irradiance. [6]. New products as solar indexes are already deduced from the data as the Mg II solar index [7].

Acknowledgments

The authors wish to thank the European Space Agency (ESA) Directorate of Human Spaceflight) and the Belgian Science Policy Office (BELSPO) (ESA Prodex and other programs) for SOLAR funding. 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.

1 1

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

[6] – Lean, J., Cycles and trends in solar irradiance and climate, Climate Change, 1, 111–122, 2010.

[7] - Thuillier G., Deland M., Shapiro A., Schmutz W., Bolsée D., Melo S.M.L., . The solar spectral irradiance as a function of the Mg ii index for atmosphere and climate modelling, , Solar Physics 277, 2 (2012) 245-266, 2012.