Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 58 (2016) 1177–1200

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

The ESPAS e-infrastructure: Access to data from near-Earth space

Anna Belehaki [a,b,*], Sarah James [c], Mike Hapgood [c], Spiros Ventouras [c], Ivan Galkin [d], Antonis Lembesis [e], Ioanna Tsagouri [a,b], Anna Charisi [a,b], Luca Spogli [f], Jens Berdermann [g], Ingemar Häggström [h], the ESPAS Consortium [1]

[a] National Observatory of Athens, Greece
[b] National and Kapodistrian University of Athens, Greece
[c] RAL Space, STFC Rutherford Appleton Laboratory, UK
[d] University of Massachusetts Lowell, USA
[e] ATHENA Research Center, Greece
[f] Istituto Nazionale di Geofisica e Vulcanologia, Italy
[g] Deutsches Zentrum für Luft- und Raumfahrt e.V., Germany
[h] EISCAT Scientific Association, Sweden

Received 10 March 2016; received in revised form 7 June 2016; accepted 8 June 2016
Available online 17 June 2016

Abstract

ESPAS, the "near-Earth space data infrastructure for e-science" is a data e-infrastructure facilitating discovery and access to observations, ground-based and space borne, and to model predictions of the near-Earth space environment, a region extending from the Earth's atmosphere up to the outer radiation belts. ESPAS provides access to metadata and/or data from an extended network of data providers distributed globally. The interoperability of the heterogeneous data collections is achieved with the adoption and adaption of the ESPAS data model which is built entirely on ISO 19100 series geographic information standards. The ESPAS data portal manages a vocabulary of space physics keywords that can be used to narrow down data searches to observations of specific physical content. Such content-targeted search is an ESPAS innovation provided in addition to the commonly practiced data selection by time, location, and instrument. The article presents an overview of the architectural design of the ESPAS system, of its data model and ontology, and of interoperable services that allow the discovery, access and download of registered data. Emphasis is given to the standardization, and expandability concepts which represent also the main elements that support the building of long-term sustainability activities of the ESPAS e-infrastructure.

Keywords: Near-Earth space data; Data infrastructure; Data model; Ontology

1. Introduction

Near-Earth space is the region that extends from the middle atmosphere up to the outer radiation belts. This region is of significant interest because of its potentially undesired effects on human life and on technological systems, whose understanding, modeling and prediction requires continuous scientific exploration and advances. Consequently a number of observing systems have been set up to acquire observations from the near-Earth space, producing a wealth of diverse types of data which still need to be homogenized and organized in order to become widely accessible.

The exploitation of multi-instrument data from a large number of distributed observing sites is the requirement

* Corresponding author at: National Observatory of Athens, Metaxa and Vas. Pavlou, Palaia, Penteli 15236, Greece.
E-mail address: belehaki@noa.gr (A. Belehaki).
[1] http://www.espas-fp7.eu/trac/wiki/PublicPages/ESPASConsortium.

http://dx.doi.org/10.1016/j.asr.2016.06.014
0273-1177/

for accurate predictions of the near-Earth space environment. As the near-Earth space is part of the complex Sun-Earth system, data from the Sun, the interplanetary medium but also from the upper and lower layers of the atmosphere, are needed to drive near-Earth prediction models. In space physics, predictions are made via physical, semi-empirical or empirical models. The models are fed with observed values (e.g., measured solar wind speed and density) or with typical values for specific environmental parameters (e.g., average speed and density of the slow solar wind during solar minimum), and the model output provides values which can be compared to other observations or parameters derived from observations (e.g., local or global geomagnetic activity index). A comprehensive comparison between model results and observed data enables the community to distinguish between models with good and with poor performance under certain geophysical conditions. Space physics models with good predictive capabilities may be used to forecast accurately the state of the space environment and to enable the end user communities to mitigate the effects of major disturbances on humans and technological systems. Results obtained from model runs depend to a large extent on the boundary conditions. Sometimes the problem can be solved by specifying boundary conditions over the entire globe and running the model on a global scale. However, specification of global boundary conditions requires data from many observational sites. Ionospheric total electron content (TEC) maps are a typical example for the dependency of maps on global data coverage in order to be realistic.

This specific need has led the space science community to work intensively for the development of systems that can facilitate data discovery and processing.

The Inter-university Upper atmosphere Global Observation NETwork (IUGONET) has been implemented by Japanese universities and institutes and aims at providing new research platforms, metadata database and analysis software tools, to facilitate the use and distribution of the long-term observation data for upper atmospheric physics (Hayashi et al., 2013). In addition to the open search service based on the metadata database, IUGONET provides: automatic data download; data analysis without regard to the file format of the data; parallel display of different types of data; utilization of various analysis functions (e.g., frequency analysis, filtering); output into an ASCII file or image files. The generated metadata are archived as XML files for interoperability with other metadata databases and future expandability. As the base of the metadata format, IUGONET selected the Space Physics Archive Search and Extract (SPASE) data model/metadata format (Merka et al., 2008), that has been modified to best match the upper atmosphere data, to create the IUGONET common metadata format.

The Automated Multi Dataset Analysis (AMDA) is provided by the Centre de Données de la Physique des Plasmas (CDPP) supported by CNRS, CNES, Observatoire de Paris and Université Paul Sabatier, Toulouse (Jacquey et al., 2010). AMDA is a web-based facility for on line analysis of space physics time series data coming from either its local database or distant ones. AMDA offers functionalities to access and analyze multi-point and multi-instrument data in a transparent way by the user. More precisely, AMDA provides functionalities for: performing search of events; performing automated and semi-automated characterization of events; extracting sub-database from an input time table; performing basic data treatment in order to provide to the user data ready to use with her/his favorite software. AMDA provides direct access to data from distant databases in a transparent way and includes a connection layer compliant with the SPASE standards.

The HELiophysics Integrated Observatory (HELIO) has been developed in the framework of an EU-FP7 research infrastructure project (Bentley et al., 2010). HELIO adopts the concept of distributed network of services that addresses the needs of a broad community of researchers in heliophysics. It coordinates access to the resources needed by the community, and provides services to mine and analyse the data. HELIO has been developed as a set of independent services. Several ways are provided to access them. The services can be used individually, within a workflow or scripting language, or through the HELIO Front-End web user interface. HELIO provides the scientist with an operational scenario for heliophysical data handling (Pérez-Suárez et al., 2012). It relieves the user from the burden of data source identification and data integration, as its web interface makes it possible to place complex searches on multiple data repositories relevant to heliospheric data in a unified, user-transparent way. This facilitates research, and creates a favorable operational environment for knowledge discovery.

The Integrated Space Weather Analysis System (iSWA) is a U.S. Government Computer Server that provides access to space weather data products and tools for both real-time as well as historical analysis (http://iswa.gsfc.nasa.gov/iswa/iSWA.html). Users are provided with the capability to specify and/or forecast the large scale and local space environment. Certain data products may be in experimental or evaluation phases of development. iSWA is customer configurable and adaptable for use as a powerful decision making tool, providing mission managers and decision-makers with personalized "quick look" space weather information, detailed insight into space weather forecasts, and tools for historical impact analysis. iSWA data management is based on a comprehensive data model that drives the system and it is supported by the Cygnet software.

Driven from these developments and in order to reach further advances in the discovery and uniform access to data, the European Commission has funded the near-Earth space data infrastructure for e-science (ESPAS) project from Framework Programme 7. ESPAS provides a one-stop shop for researchers and users of research results who wish to exploit multi-instrument multi-point science data for analysis.