

# Abstract: Sharing Data Collections and Models for Ionosphere, Thermosphere and Plasmasphere Research

Tamas Kiss, Gabriele Pierantoni, Dimitris Kagialis,  
Alexander Bolotov, David Chan You Fee  
University of Westminster, UK  
{t.kiss, g.pierantoni, d.kagialis1, a.bolotov,  
d.chanyoufee2}@westminster.ac.uk

Anna Belehaki  
National Observatory of Athens, Greece  
IAASARS – Ionospheric Group  
belehaki@noa.gr

Ivan Galkin  
Borealis Global Designs, Bulgaria  
Ivan.Galkin@borealis-designs.org

## ABSTRACT

PITHIA-NRF (Plasmasphere Ionosphere Thermosphere Integrated Research Environment and Access services: a Network of Research Facilities) is a project funded by the European Commission's H2020 programme to build a distributed network of observing facilities, data processing tools and prediction models dedicated to ionosphere, thermosphere and plasmasphere research. One of the core components of PITHIA-NRF is the PITHIA e-Science Centre that supports access to distributed data resources and facilitates the execution of various models on local infrastructures and remote cloud computing resources.

There are two major types of resources to be registered with the e-Science Centre: Data Collections and Models. Data Collections are either generated as direct outcome of an observation facility (e.g. radars, radio telescopes, meteor cameras, etc.) or can also be generated by various scientific Models. Models are scientific applications that take either raw or cleaned data from observation facilities and produce higher level datasets with predicted characteristics to facilitate further scientific research.

Both Data Collections and Models are registered with the PITHIA e-Science Centre using a rich set of metadata that is based on the ISO 19156 standard on Observations and Measurements (O&M), and specifically augmented and tailored for the requirements of space physics. The metadata structure and the related ontology were originally developed in the FP7 ESPAS project [1] and are currently being modified for the specific requirements of PITHIA. PITHIA-NRF decided to describe and register data collections only, instead of the central registration of every individual data granule, as in previous projects such as ESPAS. Such simplification enables easier management of the e-Science Centre and can lead to longer term sustainability with feasible amount of

maintenance effort required. On the other hand, local searchability of individual data pieces still remains, not restricting the scientists to access the required details at the necessary granularity.

When it comes to the execution of models, the PITHIA e-Science Centre supports three types of model execution and access scenarios, all provided from a single entry-point. Models can be executed on local resources of the various PITHIA nodes (institutions sharing Data Collections and Models). Additionally, some Models can be deployed and executed on cloud computing resources on-demand. Finally, nodes can also offer Models to be downloaded and executed on the users' own resources. Model providers can select the most suitable execution mechanism, based on the specific characteristics of the models and the resources (both human and computational resources) they have.

The implementation of the PITHIA e-Science Centre is work in progress. This presentation will report on the current state of this development work. The ESPAS metadata structure and ontology, tailored for the specific requirements of the project, have already been demonstrated to the research community on the example of some Data Collections and Models. Based on this metadata structure, work is currently ongoing to enable the registration and the ontology-based search facility of both Models and Data Collections. Proof of concept implementations [2] of the various Model access and execution mechanisms have also been implemented and demonstrated to the research community.

## Acknowledgement

This work was funded by the PITHIA-NRF - Plasmasphere Ionosphere Thermosphere Integrated Research Environment and Access services: a Network of Research Facilities (No. 101007599) EU H2020 project.

**Keywords – *e-Science Centre, ontology, metadata, Data Collection, Model execution.***

REFERENCES

- [1] Anna Belehaki, Sarah James, Mike Hapgood, Spiros Ventouras, Ivan Galkin, Antonis Lembesis, Ioanna Tsagouri, Anna Charisi, Luca Spogli, Jens Berdermann, Ingemar Häggström, The ESPAS e-infrastructure: Access to data from near-Earth space, *Advances in Space Research*,

Volume 58, Issue 7, 2016, Pages 1177-1200, ISSN 0273-1177, <https://doi.org/10.1016/j.asr.2016.06.014>.

- [2] Gabriele Pierantoni, Tamas Kiss, Alexander Bolotov, Dimitrios Kagialis, James DesLauriers, Amjad Ullah, Huankai Chen, David Chan You Fee, Hai-Van Dang, Jozsef Kovacs, Anna Belehaki, Themistocles Herekakis, Ioanna Tsagouri, Sandra Gesing: Towards a Reference Architecture based Science Gateway Framework with Embedded E-Learning Support, *Concurrency and Computation, Practice and Experience*, Wiley, 2022, <https://doi.org/10.1002/cpe.6872>