



PITHIA-NRF First Training School

Rome, 29 May - 1 June 2023

Introduction to the PITHIA e-Science Centre

Tamas Kiss

University of Westminster, London, UK

PITHIA-NRF (Plasmasphere Ionosphere Thermosphere Integrated Research Environment and Access services: a Network of Research Facilities) integrates data, models and physical observing facilities for further advancing European research capacity in this area. A central point of PITHIA-NRF is the PITHIA e-Science Centre (PeSC), a science gateway that provides access to distributed data sources and prediction models to support scientific discovery. This presentation and live demonstration will provide an overview of the current status and capabilities of the PeSC, highlighting the underlying ontology and metadata structure, the registration process for models and datasets, the ontology-based search functionalities and the interaction methods for executing models and processing data.

One of the main objectives of the PeSC is to enable scientists to register their Data Collections, that can be both raw or higher-level datasets and prediction models, using a standard metadata format and a domain ontology. For these purposes, PITHIA builds on the results of the ESPAS FP7 project by adopting and modifying its ontology and metadata specification. The project utilises the ISO 19156 standard on Observations and Measurements (O&M) to describe Data Collections in an XML format that is widely used within the research community. Following the standard, Data Collections are referring to other XML documents, such as Computations that a model used to derive the results, Acquisitions describing how the data was collected, Instruments that were used during the data collection process, or Projects that were responsible for the data/model. Within the XML documents, specific keywords of the Space Physics ontology can be used to describe the various elements. For example, Observed Property can be Field, Particle, Wave, or Mixed, at the top level. When preparing the XML metadata file, only these values are accepted for validation. Once described in XML format, Data Collections can be published in the PeSC and searched using the ontology-based search engine.

Besides large and typically changing/growing Data Collections, PeSC also supports the registration of Catalogues. These are smaller sets of data, originating from a Data Collection and related to specific events, e.g. volcano eruptions. Catalogue Data Subsets can be assigned DOIs to be referenced in publications and provide a permanent set of data for reproducibility.

Additionally to publication and search, the PeSC also provides several mechanisms for interacting with Data Collections, e.g. executing a model or downloading subsets of the data. In the current version two of the four planned interaction methods are implemented: accessing the Data Collection by a direct link and interacting with it via an API and an automatically generated Graphical User Interface (GUI). Data Collections can either be hosted by the local provider or can be deployed on EGI cloud computing resources.



The PITHIA-NRF project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101007599