



## **Towards Supporting Climate Scientists and Impact Assessment Analysts with the Big Data Europe Platform**

Iraklis Klampanos (1), Diamando Vlachogiannis (1), Spyros Andronopoulos (1), Antonio Cofiño (2), Angelos Charalambidis (1), Rob Lokers (3), Stasinios Konstantopoulos (1), and Vangelis Karkaletsis (1)

(1) National Center for Scientific Research "Demokritos", Greece, (2) Universidad de Cantabria, Spain, (3) Alterra, Wageningen UR, The Netherlands

The EU, Horizon 2020, project Big Data Europe (BDE)<sup>1</sup> aims to support European companies and institutions in effectively managing and making use of big data in activities critical to their progress and success. BDE focuses on seven areas of societal impact: Health, Food and Agriculture, Energy, Transport, Climate, Social Sciences and Security. By reaching out to partners and stakeholders, BDE aims to elicit data-intensive requirements for, and deliver an ICT platform to cover aspects of publishing and consuming semantically interoperable, large-scale, multi-lingual data assets and knowledge.

In this presentation we will describe the first BDE pilot for Climate, focusing on SemaGrow<sup>2</sup>, its core component, which provides data querying and management based on data semantics.

Over the last few decades, extended scientific effort in understanding climate change has resulted in a huge volume of model and observational data. Large international global and regional model inter-comparison projects have focused on creating a framework in support of climate model diagnosis, validation, documentation and data access. The application of climate model ensembles, a system consisting of different possible realisations of a climate model, has further significantly increased the amount of climate and weather data generated. The provision of such models satisfies the crucial objective of assessing potential impacts of climate change on well-being for adaptation, prevention and mitigation.

One of the methodologies applied by the climate research and impact assessment communities is that of dynamical downscaling. This calculates values of atmospheric variables in smaller spatial and temporal scales, given a global model. On the company or institution level, this process can be greatly improved in terms of querying, data ingestion from various sources and formats, automatic data mapping, etc.

The first Climate BDE pilot will facilitate the process of dynamical downscaling by

1. providing a semantics-based interface to climate open data, to ESGF<sup>3</sup> services,
2. searching, downloading and indexing climate model and observational data, according to user requirements, such as coverage and experimental scenarios,
3. executing dynamical downscaling models on institutional computing resources, and
4. establishing a framework for metadata mappings and data lineage.

The objectives of this pilot will be met building on the SemaGrow system and tools, which have been developed as part of the SemaGrow project<sup>4</sup> in order to scale data intensive techniques up to extremely large data volumes and improve real time performance for agricultural experiments and analyses. SemaGrow is a query resolution and ingestion system for data and semantics. It is able to extract semantic features from data, index them and expose APIs to other BDE platform components. Moreover, SemaGrow provides tools for transforming and managing data in various formats (e.g. NetCDF), and their metadata. It can also interface between users and distributed, external data sources via SPARQL endpoints. This has been demonstrated as part of the SemaGrow project, on diverse and large-scale scientific use-cases. SemaGrow is an active data service in agINFRA<sup>5</sup>, a data infrastructure for agriculture.

<sup>1</sup>Big Data Europe (<http://www.big-data-europe.eu>) – grant agreement no.644564.

<sup>2</sup><https://github.com/semagrow/semagrow>

<sup>3</sup>Earth System Grid Federation: <http://esgf.llnl.gov>

<sup>4</sup><http://www.semagrow.eu>

<sup>5</sup><http://aginfra.eu>