

Integrating data infrastructure to facilitate holistic approach to the investigations of anthropogenic hazards evoked by exploration and exploitation of geo-resources

Beata Orlecka-Sikora (1), Grzegorz Kwiatek (2), Dorota Olszewska (1), Stanisław Lasocki (1), Paolo Gasparini (3), Elena Kozlovskaya (4), Jouni Nevalainen (4), Jean Schmittbuhl (5), Jean Robert Grasso (6), Marc Schaming (5), Pascal Biggare (7), Gilberto Saccarotti (8), Alexander Garcia (3), Nigel Cassidy (9), Sam Toon (9), Grzegorz Mutke (10), Mariusz Sterzel (11), and Tomasz Szeppeniec (11)

(1) Institute of Geophysics Polish Academy of Sciences, Warsaw, Poland (orlecka@igf.edu.pl), (2) Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum GFZ, Potsdam, Germany, (3) Center for the Analysis and Monitoring of Environmental Risk (AMRA) Università di Napoli "Federico II", Napoli, Italy, (4) Sodankylä Geophysical Observatory/Oulu Unit, University of Oulu, Finland, (5) Institut de Physique du Globe de Strasbourg, Ecole et Observatoire des Sciences de la Terre, Strasbourg, France, (6) Physicien du Globe, Isterre, Grenoble Observatory, Grenoble, France, (7) INERIS Ecole des Mines de Nancy Campus ARTEM, France, (8) Istituto Nazionale di Geofisica e Vulcanologia Direttore Sezione di Pisa, Italy, (9) School of Physical and Geographical Sciences, Keele University, UK, (10) Central Mining Institute, Katowice, Poland, (11) ACK Cyfronet, AGH, Poland

The EPOS integration plan assumes a significant contribution to the research on anthropogenic hazards (AH) associated with the exploration and exploitation of geo-resources. These problems will be dealt in Thematic Core Service “Anthropogenic Hazards” (TCS AH). TCS AH is based on the prototype built in the framework of the IS-EPOS platform project (<https://tcs.ah-epos.eu/>), financed from Polish structural funds (POIG.02.03.00-14-090/13-00), with will be further developed within EPOS IP project (H2020-INFRADEV-1-2015-1, INFRADEV-3-2015). TCS AH aims to have a measurable impact on innovative research and development as well as on society by providing comprehensive, wide-scale and high quality AH research infrastructure. One of the main deliverables are numerous comprehensive induced seismicity datasets called “episodes”. The episode is a comprehensive data description of a geophysical process, induced or triggered by technological activity, which under certain circumstances can become hazardous for people, infrastructure and the environment. In addition to the six episodes already implemented during the mentioned IS-EPOS project, at least 20 new episodes related to conventional hydrocarbon extraction, reservoir treatment, underground mining and geothermal energy production are being integrated into the e-environment of the TCS AH. The heterogeneous multi-disciplinary data are transformed to unified structures developed within IS-EPOS project, to form integrated and validated datasets. Dedicated visualization tools for multidisciplinary data comprising episodes are also implemented. These tools are capable to aggregate and combine different data types and facilitating specific visualization possibilities (e.g. combining seismic and technological information).

The implementation process, tailored for each episode, consists of four steps: (i) Data revision, determination of its accuracy and limitations; (ii) Data preparation and homogenization to follow the TCS AH standards; (iii) Data collection - uploading the data to local data centres and (iv) Metadata preparation. Some web services for efficient data integration and reduction of the possible mistakes, have been already developed. The data sets will be also passing through quality control according to the established quality control scheme. The data quality control workflow includes five steps; the first three are done in the episode provider's local data centre, whereas the next two are accomplished on the TCS AH side: (1) Episode data are transferred to the local data centre; the control group roles are distributed and the workflow observer is appointed; (2) The data are standardized and formats are validated; the completeness and quality of the data is checked; (3) Metadata are prepared according to TCS AH metadata scheme; the published data and metadata are checked; (4) Contextual quality of the data is analysed and episode appears in TCS AH maintenance area; (5) The new episode data is available in TCS AH.

TCS AH will also serve as an integration platform for the episodes gathered in the framework of “Shale gas exploration and exploitation induced risks” project (Horizon 2020, call LCE 16-2014).