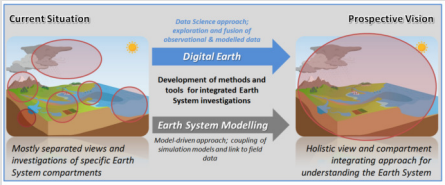
October 2019



Newsletter of the *Digital Earth* Project

Contributions of all Helmholtz Centres to Digital Earth

This newsletter of Digital Earth presents recent results and development of joint activities of all participating Helmholtz Centres within Digital Earth.

Digital Earth – New Themes - New Possibilities

Daniela Henkel[1], Jens Greinert[1], Holger Brix[2], Diana Rechid[2], Stephan Frickenhaus[3], Angela Schäfer[3], Doris Dransch[4], Bruno Merz[4], Wolfgang zu Castell[5], Ralf Kunkel[6], Andreas Petzold[6], Peter Dietrich[7], Thomas Kalbacher[7], Roland Ruhnke[8], Ottmar Moehler[8]

1 GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
2 Helmholtz-Zentrum Geesthacht Zentrum für Material- und Küstenforschung (HZG)
3 Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI)
4 Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum (GFZ)
5 Helmholtz Zentrum München – Deutsches Forschungszentrum für Gesundheit und Umwelt (HMGU)
6 Forschungszentrum Jülich(FZJ)
7 Helmholtz-Zentrum für Umweltforschung (UFZ)
8 Karlsruher Institut für Technologie (KIT)

Since its start in 2018 the Digital Earth project has established excellent knowledge exchange structures and pushed joint developments during the first project year between the eight involved Helmholtz Centres. Digital Earth further strengthened links to MOSES and ESM and at the same time boost specific tasks of Digital Earth with two additional **Themes** that could be defined.

These two Themes are: Theme#1 **"Automated FAIR data processing chains, including metadata standardization and handling interfaces for implementation"** and Theme#2 **"Linking Models and Data"** of work package 2 "Data Exploration".
To directly generate test scenarios and support the Show Case "Flood" of Digital Earth, two tasks have been added to work package 1 "SMART Monitoring" (**sub-task "Cloud in-situ Smart Monitoring"** and **sub -task "Method transfer and application to support campaign planning with digital soil moisture mapping at regional scales"**).

As part of **Theme#1**, a common standard for the documentation of QA/QC procedures will be developed including procedure parameterization, procedure chains and applicability information. Furthermore, a central registry for a unique identification of QA/QC procedures using Persistent Identifiers (PID) will be developed, implemented and coupled to a service, which allows to query and access QA/QC metadata in a standardized way. Common interfaces will be developed and implemented to allow access to the QA/QC procedures either individually or within container-based frameworks; this makes these procedures and its documentation machine-readable and allows application/extension to other data.

In **Theme#2** we added the objective of **sub-task "Earth System Model Diagnosis Visual Front-End"** to establish new diagnostic capacities to directly confront Earth system model simulations with observations on irregular time-space grids. This will help to prioritize future directions for the model development, and will provide data to a wider group of scientists and stakeholders with easy to handle tools. With the second additional **sub-task "Advancing 4D-Visualisation towards a Mission Planning Tool"** the focus is laid on advanced 4D-Visualisation of regular and irregular model data results and measurements with the long-term goal of also using 4D Visualisation as a Mission Planning Tool.

Six of the eight centres involved in Digital Earth are able to support the new tasks with additional personnel. GFZ and HMGU will support within already existing capacities.
With the additional IVF funds, the successes of the future topic funding are to be secured in particular before the start of next round of "Program-Oriented Funding" (phase IV) - PoF-IV. The funds are to be specifically tailored to the PoF-IV and can be used for all measures directly related to the respective future topic to strengthen programme innovation.

Digital Earth at REKLIM Conference 2nd International Conference "Our Climate – Our Future: Regional Perspectives on a Global Challenge"

Daniela Henkel[1]
1 GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel

Digital Earth joined the 2nd REKLIM Conference that took place from 23 - 25 September 2019 in Berlin, Germany and presented first results and tools established after one year of project to more 200 scientific colleagues from all over the world.

Data Exploration Framework Workshop

Christian Scharun[1], Daniel Eggert[2]
1 Karlsruher Institut für Technologie (KIT)
2 Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum (GFZ)

On 30.09.2019 and 01.10.2019 eleven Digital Earth members from six of the involved Helmholtz Centres came together at the GFZ in Potsdam to exchange and report about the progress of their work related to work package 2 "Data Exploration Framework".

Following the concepts of the show cases and the exploration framework, small groups of domain and data science experts from different Helmholtz Centres develop individual, but yet linked, **Toolboxes** supporting new or existing workflows. One example is the **Anomaly**

Detection Toolbox of Viktoria Wichert (HZG) and Daniela Rabe (GFZ) where seawater measurements acquired along ferrylines are used to model the origin of water masses and evaluate parameters like temperature and salinity. A Toolbox supporting the exploration of multidimensional data is developed by Everardo Gonzalez (GEOMAR). This **4D Data Viewer** allows the exploration of big model data such as netCDF files via slice and dice methods provided by Thredds/Opendap. The prototypes of the developed Toolboxes can be tested and evaluated online.

The focus of the work done in work package 2 is to give domain experts a better understanding of the system earth. This shall be reached via a holistic view on various scales, e.g. a holistic view for an entire show case / domain or a holistic view need for a particular question within a show case. Next to the discussion on how to provide a holistic view, the workshop members tried to identify overlapping methods used by multiple toolboxes. Once common methods are identified, they can be implemented in a more re-usable way, saving additional effort. Finally a roadmap with milestones for the upcoming months was compiled.

Evaluation of Success: Results from the first survey and inventory of categorised requirements

Laurens Bouwer[1], Diana Rechid[1], Bernadette Fritzsch[2], Thomas Kalbacher[3], Werner Köckeritz[4], Andreas Petzold[5], Roland Ruhnke[6]

1 Climate Service Center Germany, Helmholtz-Zentrum Geesthacht (HZG)
2 Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI)
3 Helmholtz-Zentrum für Umweltforschung (UFZ)
4 Helmholtz-Zentrum Potsdam - Deutsches GeoForschungsZentrum (GFZ)
5 Forschungszentrum Jülich (FZJ)
6 Karlsruher Institut für Technologie (KIT)

The Digital Earth project aims at implementing data science methods in the field of Earth & Environment. During the project, the success of applying these methods, and related requirements of researchers are being measured and assessed. Thereby the project can provide overall recommendations on how to improve the application of data science. This report provides the criteria and evaluation results from a questionnaire at the start of the project. It includes the First set of criteria and indicators, as well as the initial set-up of the survey/evaluation scheme.

The evaluation questions that are addressed are: what are the requirements for Data Science? What is the Scientific Progress? What is the Usability of the Workflows? And: what is the Success of Digital Earth? Several criteria are presented for the assessment of the requirements for the application of Data Science methods. Through a questionnaire, we have assessed the data science methods that are currently being applied, and that will be applied in the project, in particular data exploration and data visualization methods, including Artificial Intelligence (AI) and Machine Learning. The requirements for data science, and collaboration between partners are described. The envisaged goals and criteria for measuring success of the project are presented. Finally, usability of results and the developed workflows are assessed using the FAIR principles, and the transfer of data and information to other users (scientists and other parties) is described.

The most important results and conclusions from the first questionnaire and evaluation, and categorized requirements include the following:

Several **collaborations** have already been established in Digital Earth, during the proposal writing and since the kick-off. This is documented through the responses from the researchers in Digital Earth on their collaborations and exchanges with other Helmholtz Centres within the project. There is even more potential and wishes for collaboration which is also reported in the responses. Collaboration with a few specific centres are highlighted as these are desired for their competences.

The project has also progressed on **interdisciplinary approaches**, showing collaborations between different fields of Earth and Environmental Science, but also tying in Data Science expertise, in particular related to visual data exploration as well as Artificial Intelligence and Machine Learning. Many centres at the beginning of the project however still show limited experience with many of these methods. There is also more need for more support on Data Science methods, especially related to Machine Learning.

With regard to **requirements for data science**, it is reported that good observational and data collection and data infrastructure are indispensable for doing Data Science. The Digital Earth project therefore would not have been possible without this basis. It builds on and profits from several complementary efforts focused on field observations and data infrastructure at the individual Helmholtz Centres, as well as targeted collaborative projects, such as TERENO and MOSES.

With regard to science practices, several aspects of **FAIR Science** were reported. Many centres make their data and code openly available, but licenses and policies for the publication and use are still lacking in many cases. Scientific and observational data, tools and information are made available beyond academia, to a surprisingly large extent. This availability is complemented by the publication of guidelines for use, tailoring for specific applications, and quality assessment.

With regard to **project success**, the researchers report several important indicators, that can be used to later evaluate the process and scientific success of Digital Earth. First of all, researchers hope for better integration and collaboration between Earth and Environment disciplines. Digital Earth should also help to increase the usability of data, information and workflows. Joint scientific publications, new research proposals, and new data science activities are regarded as most important signs of success for Digital Earth, and should be regarded as key performance indicators for the project later on.

The implications for Data Science application in the field of Earth & Environment, including the **programming in PoF-IV** include the following: the application for Data Science methods depend on collaboration between centres with different expertises. Also, Data Science is dependant on a good observational and data infrastructure, which needs to complement any development in data science applications. In addition, a close tie between research projects and activities in Earth & Environment and initiatives that promote the use of AI, such as HAICU, is needed.