

Chapter 9

Lessons Learned in the Digital Earth Project



Jens Greinert, Daniela Henkel, Doris Dransch, Laurens M. Bouwer, Holger Brix, Peter Dietrich, Stephan Frickenhaus, Andreas Petzold, Diana Rechid, Roland Ruhnke, and Wolfgang zu Castell

Abstract The Digital Earth project aimed for the integration of data science and Earth science. Here, we reflect on the main lessons learned that include the need for interdisciplinary collaboration, thinking out of the box, the concept of ‘thinking in workflows’ and models for the sustainable implementation of scientific software, data infrastructure and policies.

Keywords Digitalization · Earth System Science · Interdisciplinary · Collaboration · Workflows · Sustainability · Software · Data infrastructure · Policies

J. Greinert (✉) · D. Henkel
GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany
e-mail: jgreinert@geomar.de

D. Dransch · W. zu Castell
Helmholtz Centre Potsdam-GFZ German Research Centre for Geosciences, Potsdam, Germany

L. M. Bouwer · D. Rechid
Climate Service Center Germany (GERICS), Helmholtz-Zentrum Hereon, Hamburg, Germany

H. Brix
Helmholtz-Zentrum Hereon, Geesthacht, Germany

P. Dietrich
Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany

S. Frickenhaus
Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

A. Petzold
Forschungszentrum Jülich GmbH, Jülich, Germany

R. Ruhnke
Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

W. zu Castell
Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

9.1 Introduction

A central question of the Digital Earth project is: How can data science contribute to improving scientific results within the Earth sciences? This fundamental question was posed by the scientists involved in the project, which has led to the research set-up and aims as described in Chap. 2. Within the project, several methods and tools have been developed and applied (see Chaps. 3, 4, 5, and 6). In addition, the collaboration and success of the project were assessed and evaluated (see Chaps. 7 and 8). In this chapter, we present as a conclusion the four lessons learned that we regard to be an essential basis for a fruitful interrelationship of data science and Earth System Science.

9.2 Lesson 1: Interdisciplinary Collaboration

Moving from multidisciplinary to interdisciplinary collaboration is essential for the adoption of data science methods and for making progress with digitalization in Earth System Science. One obvious success of Digital Earth is the established interdisciplinary collaboration. The results achieved in the project have been created by many scientists that before Digital Earth did not work together, did not know each other and might not even have seen the need and advantage of extending their own expertise before the project. Today, all project members agree that a sustainable collaboration across many disciplines, with different ways of working and despite the high number of obstacles and difficulties particularly in communicating with each other and finding a common ground, is a precondition for novel solutions and it is for sure worth the effort. A kick-start action, like the Digital Earth project, to such an endeavour is essential. Digital Earth provides the needed time to develop the collaboration and to establish a nucleus of knowledge and trust. Only this enables joint problem solving and develops new research ideas and opportunities.

9.3 Lesson 2: Thinking Out of the Box

Investing in working ‘outside the box’, beyond your own comfort zone, research centre and discipline is crucial. Communicating ‘my’ science to others and learning from other disciplines and approaches are important and the only way to expand our knowledge in Earth System Sciences. Cultural, language and importantly knowledge shortcomings hinder an effective communication and collaboration between data scientists and Earth scientists. This needs to be overcome by suitable means and strategies to help both sides acquire a good understanding of the other disciplines (see Chaps. 5 and 7). These shortcomings need to be done even if the

process takes time and does not immediately lead to the envisioned success. Scientists with the explicit aim of bridging between disciplines, Earth compartments and institutions have been identified as good nuclei and multipliers for developing and adapting novel data science approaches to improve Earth System Science. The Digital Earth project provided a frame to enable the distribution of knowledge within and across scientific disciplines and created an environment where people advanced beyond their typical realm.

9.4 Lesson 3: Thinking in Workflows

To set a common ground for interdisciplinary collaboration and ‘thinking out of the box’, the concept of scientific workflows was used in Digital Earth as a base for communication. After initial hesitation primarily by the Earth scientists, the concept helped to structure the processes of knowledge generation and to break it down into exchangeable and reusable steps. These workflows made it much easier to create a common ground between Earth and data scientists, to identify bottlenecks in specific steps in the workflow and to find alternatives for methods and tools. ‘Thinking in workflows’ (see Chap. 5) became the guiding principle in the project, where natural scientists define their needs, identify the available input data and present their wishes for output to the data and computer scientists. The computer scientists add their expertise with regard to methods and approaches in artificial intelligence, visualization, exploration of distributed data and software engineering. Thinking in workflows and formalizing the way Earth scientists generate knowledge allows an effective way of sharing and implementing scientific approaches and data science methods. It supports the reuse of scientific software and enables a component-based and collaborative framework for data-driven science. We identified the approach of ‘Thinking in workflows’ as a suitable and modular way of communication and scientific collaboration. Based on this approach, the next collaboration in smaller and larger projects will be much easier.

9.5 Lesson 4: Sustainable Implementation of Scientific Software, Data Infrastructure and Policies

The need for joined and professional software development and its maintenance is obvious when data science should become a cornerstone in Earth System Science. So far, such software is developed by small groups or individuals who train themselves. There is a need for more professional and standardized scientific software development. Software needs to be reusable and maintained to prevent scientists from inventing the wheel again and again. Research centres need to acknowledge that

software development and maintenance is an ongoing and important effort similar to data management, running analytical facilities and the science itself.

Clear guidelines, policies and licensing rules for joint software development and provision, and the use of data are still ‘under construction’. This creates problems when it is envisioned that developed software tools should be shared with others. More effort is required here.

There is no progress in data-driven science to be expected if infrastructure hurdles exist. Examples are data access (authentication) difficulties and the transfer of large data sets. This was experienced in the project and made collaboration and interdisciplinary research unnecessarily complicated.

Finally, the work and effort that are related to the sustainable implementation and development of scientific software, data infrastructure and policies have to be appreciated and counted as valuable scientific contributions.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

