

5th AGILE PhD School Proceedings Foreword

Alexander Kmoch 
University of Tartu, Estonia
alexander.kmoch@ut.ee

Daniel Nüst 
University of Münster, Germany
daniel.nuest@uni-muenster.de

Evelyn Uemaa 
University of Tartu, Estonia
evelyn.uemaa@ut.ee

Abstract

The 5th Agile PhD School was held at the Chair of Geoinformatics, Department of Geography, the University of Tartu, Estonia. 16 PhD students attended from universities across Europe worked on the main themes of geo-visualisation, reproducibility and science communication. The students were challenged to question preconceived ideas, own perceptions and habits in regards to the process of conducting and communicating scientific research. In this editorial we provide a few key takeaways for the main themes and sketch the process of reproducible proceedings.

1 Introduction

The 5th Agile PhD School took place at the Chair of Geoinformatics, Department of Geography, the University of Tartu, Estonia. Overall 16 highly motivated PhD students attended from universities from Austria, the Czech Republic, Cyprus, Estonia, Germany, Greece, Poland, Spain and the UK.

The main themes of the PhD school were geo-visualisation, reproducibility and science communication. For each of these themes we had invited an expert from the respective domain. The students were trained

Copyright © 2020 by the paper's authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In: A. Kmoch, E. Uemaa, D. Nüst (eds.): Proceedings of the 5th AGILE PhD School 2019, Tartu, Estonia, November 2019, published at <https://doi.org/10.5281/zenodo.3835767>.

via lectures, exercises and lab sessions and individual work.

Over the course of 3 full working days the students were challenged to question preconceived ideas, own perceptions and habits in particular in regards to the process of conducting research, processing data and generating results as well as communicating the scientific methods planned or applied and their results.

2 Visualisation

On Tuesday we had Topi Tjukanov from GISPO OY, Finland teaching cartographic, styling and animation-based capabilities to the students. He introduced general visual cartographic and data visualisation principles and then conducted three exercises:

- a static map with a custom globe projection and dynamic coloring based on attribute tables
- spatio-temporal visualisation of mobility data, based on the example of movement trains in Finland
- spatio-temporal visualisation in terms of changing shape of the same feature over time on the example of variations in the ice extent of the Arctic.

My key takeaways as member of the organising committee and host to the students:

1. Dynamic visualisation in scientific papers is yet not possible. However, graphical abstracts and supplemental materials provide already means of adding dynamic visualisation capabilities to a scientific paper. Also, accompanying popular science articles, often published and presented as blog

posts provide all the flexibility to include dynamic visualisations.

2. Dynamic visualisations falls in to major categories: interactive and non-interactive. Interactive visualisations can be maps or plots where a user can zoom in or out, click on features and figures to receive more information via pop-overs or tooltips or similar means. Non-interactive visualisations are typically animations. Short movies or dynamic GIFs are te most widely used types.
3. Not all fields of research, even within Geoinformatics/GIS will benefit of spatio-temporal visualisation in the same way. The fields of research of the PhD students ranged from agent-based modelling and socio-cultural studies, over mobility and remote sensing, to digital cartography and human-computer-interactions (e.g. eye-tracking on digital maps)

From a practicing researcher's point of view, my key takeaway is that animations and maps are exciting and the ease with which to create these in current QGIS versions.

For my latest research output, an extended soil dataset for Estonia, I am currently creating an interactive webtool for soil scientists and experts to test my computational framework that derives values from original text codes. Furthermore, I intend to also develop a small web mapping application for interested experts but also the general public to explore the new soil datasets in Estonia.

3 Reproducibility

On Wednesday Carlos Granell Canut and Frank Ostermann in their role as representatives of the AGILE Reproducibility Initiative presented vision, mission and initiatives of AGILE and how reproducibility is intended to become more required in submissions to the AGILE conference.

After their conclusion Alexander Kmoch guided the students through 3 practical exercises of The Carpentries reproducibility curricula.

- Version control with Git¹
- RStudio² project management and Git integration
- RStudio and R Markdown³ to combine reproducibility concepts from data processing, coding and document writing as a single workflow, so called literate programming⁴.

¹<https://en.wikipedia.org/wiki/Git>

²<https://rstudio.com/products/rstudio/>

³<https://rmarkdown.rstudio.com/>

⁴https://en.wikipedia.org/wiki/Literate_programming

The key takeaways for me as the trainer and practitioner myself:

1. Git and the a concept of version control, online code sharing and collaboration was generally greatly appreciated and its value acknowledged.
2. But the mental overhead to get started make it hard to include regular practice of using Git (status, add, commit) into the research process.
3. Preparing data and source code deposits and decorating such repositories with useful metadata, README and LICENSE information creates an 'ah-ha' moment, but the general incentive is low. Being able to improve and control public appearance and engage with peers can be one possible incentives.
4. I am very convinced that this practice helps me in having my paper accepted more straightforwardly, and I can point critics more directly to the point of discussion. This improves the reviewers' feedback, thus, the quality of the publication and produced new dataset and code, and subsequently, the value for end-users, such as other scientists and experts in the Ministry of Environment and the Landboard of Estonia.

4 Science Communication

On Thursday, the students got a condensed summary of points to consider when writing with the focus on scientific communication. The students did a series of freewriting exercises on exploring their audience, how to structure their scientific article and how to work along this structure to improve the writing and thinking process.

My key takeaway from this session is:

The more we ask various questions and respective counter-questions, the more we guide our thinking process and in turn can produce useful and interesting text. We can anticipate questions from the supervisor, the reviewers and eventually the intended target audience. This helps us to shape the text in order tell the story and persuade the readers of the value of our writing, and thus, our research.

5 Proceedings

The PhD school introduced participants to R Markdown as a powerful tool for reproducible research. In the preparation for the CEUR Workshop Proceedings, the editors took advantage of this format and created a semi-automatic workflow to create high-quality PDF documents. The workflow is based on an R Markdown document itself and includes R scripting, the

Figure 1: DOI (Zenodo archive)
<https://doi.org/10.5281/zenodo.3784142>

conversion of a publisher template to an R Markdown template, and adding structured metadata in the header documents. All manuscripts and the proceedings repository are published on GitHub in the organisation `agilephdschool2019`⁵. The proceedings main repository⁶ provides a boilerplate for future PhD schools. The required changes to the documents show participants how to apply a publisher template to their manuscript document, which enables them to effectively publish computational research in various styles and formats in their future career.

6 Conclusion

It was a fantastic experience for me as one of the hosts and trainers of the PhD school. I still will take a lot of the lessons from this PhD school as useful input for my own research habits as well as training materials for future generations of students and emerging young researchers. I hope the collected reflections and notes will be of value to everyone interested.

Best regards,
Alexander Kmoch
Tartu, 28th of November 2019

⁵<https://github.com/agilephdschool2019/>

⁶https://github.com/agilephdschool2019/agile_2019_phd_school_proceedings/