

To: NeIC Board

From: Radovan Bast

Executive officer: Michaela Barth

Regarding: CodeRefinery Project Midway Report

Date: 20.11.2020

Introduction and summary

The CodeRefinery project is in its second phase (October 2018 to October 2021) operating with a budget of 10.3 FTE distributed over 3 years and 5 partners.

CodeRefinery is working with students, researchers, and research software engineers from all disciplines to advance FAIRness of software management and development practices so that research groups can collaboratively develop, review, discuss, test, share, and reuse their codes.

This project offers **training opportunities to researchers from Nordic research groups and projects** to learn basic-to-advanced research computing skills and become confident in using state-of-the-art tools and practices used within modern **collaborative software engineering**.

In addition to workshops in software engineering tools and workflows we offer **instructor training** and support [The Carpentries community](#) to increase the number of Software and Data Carpentry workshops and Carpentries instructor training opportunities in the Nordics.

The CodeRefinery project has been also the catalyst for the [Nordic HPC](#) initiative as well as the nucleus for [Nordic research software engineers](#) community and its activities. As part of the project a [GitLab code repository hosting](#) for Nordic research software is offered.

*Summary: At the midway point, the project has quickly **adapted to online teaching**, the project **grew during the COVID-19 period of 2020** and is on track. The identified risks are not threatening the overall outlook of the project.*

Progress and next steps

Training events. In the second phase of the project **we have grown significantly** (Tab. 1). The number of participants by countries for full workshops are summarized in Tab. 2.

Table 1: *Number of workshops and persons trained over time.*

	Phase 1 (Sep 2016 - Aug 2018)	Phase 2 (Sep 2018 - Nov 2020)
Full workshops	13	17
Instructor training events		3
Other workshops	5	14
Number of persons trained	477	1073

Table 2: Number of participants by countries for full workshops in phase 2.

Country	Total	SE	FI	NO	NL	DK	FR	EE	DE	UK	US	Other
Nr. participants	663	201	187	112	43	33	32	25	7	5	4	14

The very different numbers among Nordic countries should not be interpreted as being proportional to the training need or training interest. These numbers mostly correlate with the location of committed workshop organizers and the efficiency of announcement channels. For more details see the list of [past workshops and events](#) and [workshop data](#).

Transition to online teaching: The CodeRefinery project managed to **grow and expand scale and reach during the pandemic**. The project team has reacted very quickly to the restrictions related to COVID-19 and we have transitioned to online teaching within weeks and delivered our first online training event on April 7 and 8. We have used this as an opportunity to **reach a wider audience** and by the end of May we were able to deliver an **online training event to over 100 participants**, with live streaming the event and recording the sessions. Transitioning to online teaching has enabled us to on-board many more new and future instructors and volunteer instructors and helpers compared to earlier years, in particular from **Aalto Scientific Computing group**. We have found it very beneficial to allow teams to register and participate in a training event (“bring your own breakout room”). Our lessons learned are documented in two blog posts (see references).

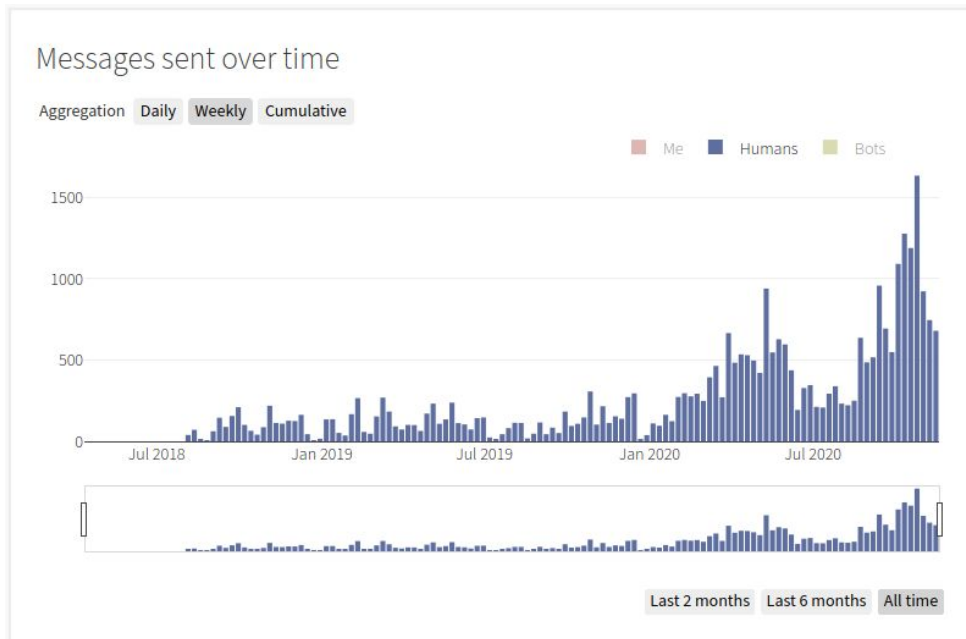
Carpentries membership: workshops and instructor training: NelC has a 3-year Carpentries Platinum membership until Nov 1, 2021. Major membership benefits are 15 seats for Carpentries instructor training per year and 6 centrally organized Carpentries workshops per year without coordination fee. Since November 2019, Naoe Tatara has been acting as Carpentries Regional Coordinator in the Nordic countries and has significantly increased the visibility of Carpentries activities within the Nordics and visibility of CodeRefinery and NelC within the Carpentries community. The uptake of the instructor training is good. However, centrally organized workshops are generally underused, and opportunities for online Carpentries workshops should be more widely disseminated. We have presented our lessons learned from online workshops of increasing scale online workshops to the Carpentries in a blog post ([“Lessons Learned from Running Code Refinery’s First Online Workshop”](#)) and a conference presentation ([Online workshops - How to scale up your online workshop to reach 100 learners?](#)) at [CarpentryCon2020@Home](#).

Nordic research software engineers: A number of CodeRefinery members are involved in Nordic research software engineers (RSE) community building. We are preparing for the [first Nordic research software engineers conference](#), Stockholm, May 27-28, 2021. This conference was postponed twice due to COVID-19. To keep the momentum of the growing Nordic RSE community, we have organized the [Nordic-RSE get-together online event, Nov 30 - Dec 2, 2020](#). Topics include both the experience of being a RSE and tech tools useful for research.

Research Software Hour: Richard Darst, Anne Fouilloux, and Radovan Bast are streaming a weekly [Twitch](#) show about scientific computing and research software: <https://researchsoftwarehour.github.io/>.

Twitter, newsletter, and chat: The project has 552 followers on Twitter, 213 subscribers to the CodeRefinery newsletter and 171 users on the team chat. In 2020, with the shift to more online work, we have observed a significant increase in messages exchanged (Fig. 1).

Figure 1: Number of weekly messages over time.



GitLab code repository hosting: 580 users with projects, 108 groups, 986 projects. Notable projects: ARC/NT1, Earth Observation Laboratory, Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, Symbioses, Nordic Language Processing Laboratory, EISCAT, UiT Master of Applied Computer Science, UiT Narvik R&D - Umbrella.

Outlook

We anticipate that the **need for training in research software engineering will not decrease** after the end of the project period. We find it extremely important that our training activities and community development work can continue after October 2021 and **we urge the NelC board to support efforts to keep this project active** beyond the funding period. The future of this project is also coupled to the **future of the Carpentries membership** and thus will affect the number of Carpentries workshop and instructor training opportunities in the Nordics.

We see the following options and are preparing a position paper and discussions with relevant partners:

- a. **National e-infrastructure providers and/or local institutions** support a continuation of the project by allowing those project members who are invested and motivated in this project to use a portion of their person months for training, lesson development, and project coordination.
- b. We implement a **payment and sponsorship model** where institutions pay for workshops and thus finance the work time of e-infrastructure provider staff for training, lesson development, and project coordination.
- c. The project manages to secure a follow-up funding either at the institutional, national, Nordic, or European level.

If we pursue option b) or wish to simplify participation in European projects, it may be important to set up a formal organization with its own accounting which can participate in funding proposals as its own entity.

To secure the continuation of the [GitLab code repository hosting](#) for Nordic research software we are negotiating with DeiC, ETAIS, and Sigma2 to anchor this service in a long-term model financed by either NelC or participating national e-infrastructure providers.

Risks

The risks listed in the [project plan](#) are still valid, but the main risks for the project related to finishing on time, budget, and with quality are:

- Project partners not being able to contribute staff.
- Staff being moved to other projects by their local organizations which can affect continuity, their identification with the project, and puts on-boarding strain on the project.
- Staff changing jobs and project partners not being able to fill their role and commitment.
- New staff not being able to take up larger-percentage contributions because the funding period ends in less than one year.
- With COVID-19 it may be impossible to organize an in-person conference within the funding period. This may limit the benefit realization for the Nordic RSE community.

References

Documents and project reports

- [Project directive \(phase 1\)](#)
- [Collaboration agreement \(phase 1\)](#)
- [Project plan \(phase 1\)](#)
- [Report \(phase 1\)](#)
- [Project directive \(phase 2\)](#)
- [Collaboration agreement \(phase 2\)](#)
- [Project plan \(phase 2\)](#)

Presentations

- ["Online workshops - How to scale up your online workshop to reach 100 learners?" at CarpentryCon2020@Home](#)
- ["Opening Eyes and Turning Lights On", "RSE Stories" podcast interview with Radovan Bast](#)
- ["Carpentries next level - CodeRefinery model", Lightning talk at CarpentryConnect 2020 \(Sabry Razick\)](#)
- [Lightning talk at CarpentryConnect 2019, Manchester \(Radovan Bast\)](#)
- ["Coding like it's 1969", Toulouse, April 2019 \(Radovan Bast\)](#)
- [Albanova seminar, Stockholm, April 2019 \(Thor Wikfeldt\)](#)
- [Short talk at the CarpentryCon 2018 conference, Dublin \(Thor Wikfeldt\)](#)

Articles about CodeRefinery

- "Better Science Through Better Software", <https://neic.no/news/2020/06/29/better-science-through-better-software/>
- "A FAIRer future", [Nature Physics 15, 728–730 \(2019\)](#)
- "Bättre datahantering gör forskning mer reproducerbar", [article in the Curie newsletter](#) from the Swedish Research Council
- "CodeRefinery 2.0", [PDC December 2018 newsletter](#) and [Titan.uio.no blog post](#)
- "Teaching researchers to write better code", <https://www.inthefieldstories.net/teaching-researchers-to-write-better-code/>
- "Better software leads to better science", [Nordforsk January 2018 news article](#)
- [EUNIS-19 NeIC's CodeRefinery project](#)

Blog posts

- [Community discussion in Nordic and Baltic countries](#)
- [git-pr: painless small pull requests](#)
- [Report from the Mega-Coderefinery workshop](#)
- [CodeRefinery tools in action: NordicHPC](#)
- [Research Software Hour](#)
- [Rebase vs merge](#)
- [CodeRefinery workshops moving online](#)
- [Announcing the first Nordic-RSE conference](#)
- [Lessons learned from running our first online workshop](#)