

CodeRefinery - a hub for FAIR Software practices

1: Who we are

2: Our work

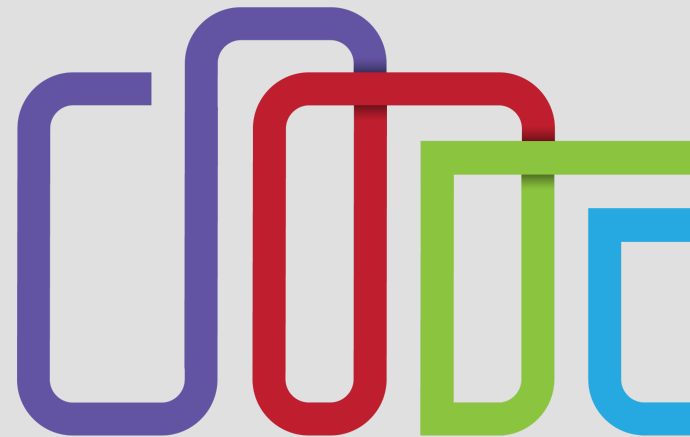
3: What we have learned

4: Unsolved problems

5: How you can help

Radovan Bast [@_radovan](https://twitter.com/_radovan)

UiT The Arctic University of Norway



CODE REFINERY

About me



- Theoretical chemist turned research software engineer.
- I write research software and teach programming to researchers and lead the [CodeRefinery_project](#).
- I lead the high-performance computing group and the research software engineering group at UiT.

NelC: key facts

Nordic e-Infrastructure Collaboration

(slide will be added - waiting for data from a colleague)

FAIR principles and software

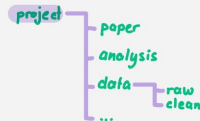
Researchers need to navigate many tools and concepts.



REPRODUCIBLE RESEARCH

6 helpful steps

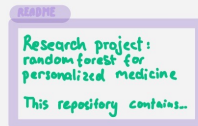
1 Get your files + folders in order



2 Use good names for files, folders, functions, ...

`6-steps-reproducibility.pdf`  `clean.data ← function(...) ? ... ?`

3 Document with care: README, Metadata, code comments, ...



```
README
Research project:
random forest for
personalized medicine
This repository contains...
```



CC-BY 4.0 Heidi Seibold
@HeidiBaya

4 Version control code, text, ...



5 Stabilize computing environment and software



6 Publish your research outputs: Code, data, documents, ...



Team and project: coderefinery.org

- Started in 2016, now we are in phase 3 until 2025
- **Partners:** NeIC (1 FTE), Aalto University, ENCCS, CSC, DeIC, Sigma2/NRIS, SNIC, T1C for interactive HPC (DK), USIT/UiO
- Over 2000 persons trained
- Over 30 instructors/speakers
- Over 100 exercise leads
- **Pioneering teaching methods**
- **Working in public**
- Lesson material
- Video recordings
- Manuals
- Training network
- Brand
- Impact
- Community
- Knowledge



CodeRefinery: bridging the gap



Specialist training

Traditionally run by computing centers

CodeRefinery provides collaboration network

CodeRefinery

Expert training for reusable software

Not broadly taught in all degree programs that need it

The Carpentries

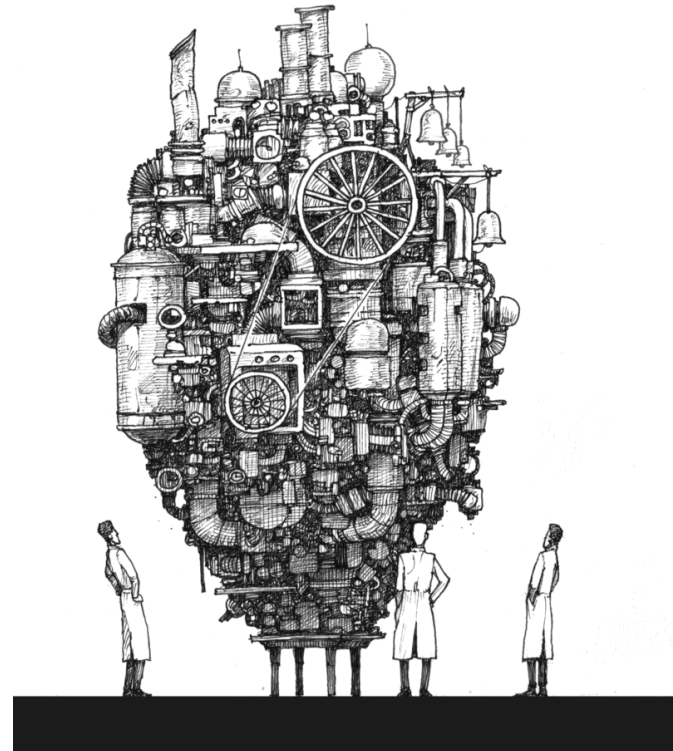
Basics training in programming
and data science for novices

Our course portfolio via coderefinery.org

- Version control
- Collaboration using Git
- Testing
- Documentation
- Notebooks
- Modular code development
- Reproducible research
- Software licensing
- How to share and publish code
- ...

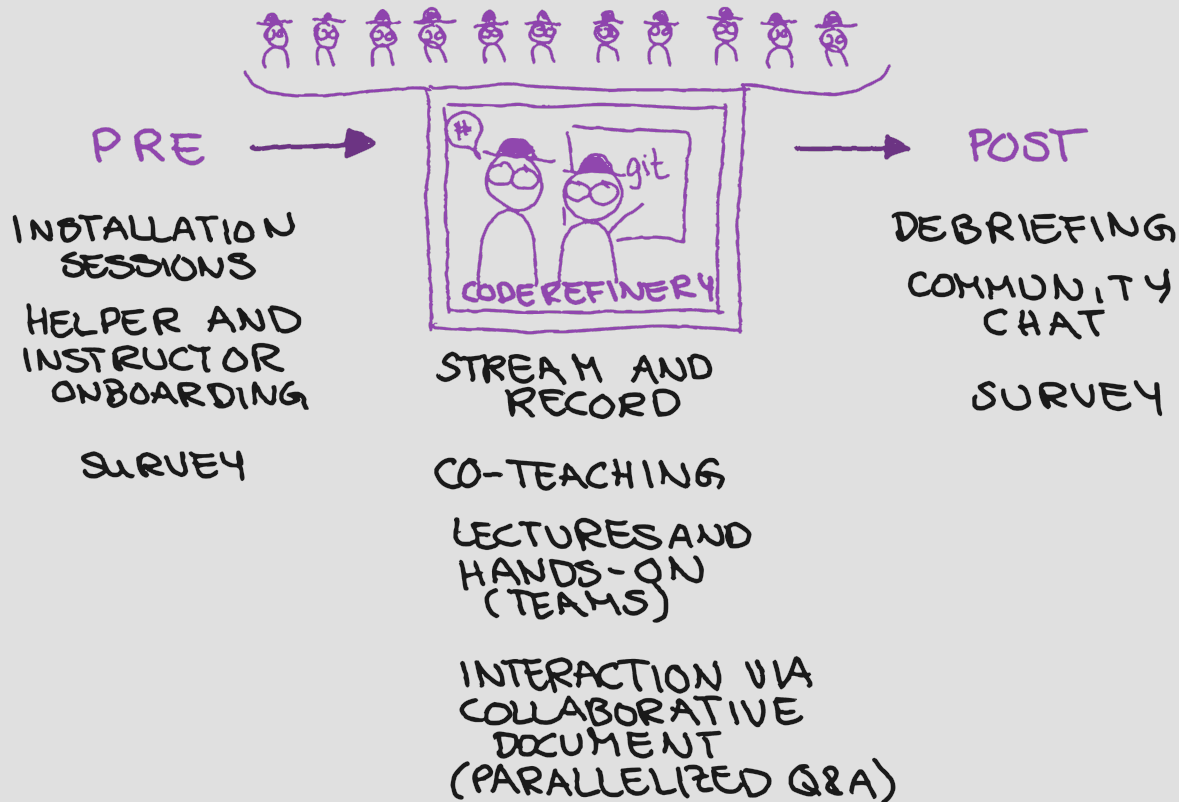
Lessons and recordings:

<https://coderefinery.org/lessons/>

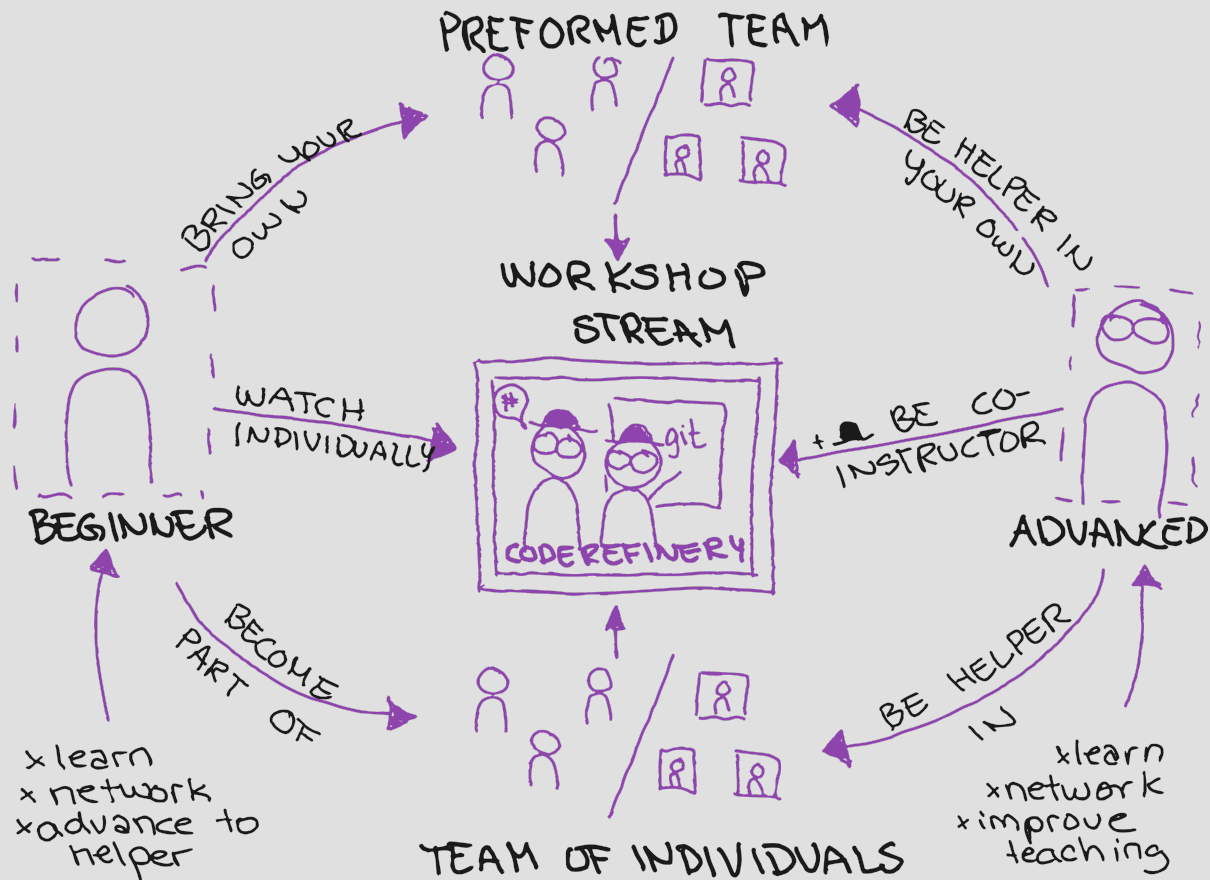


[citation needed]

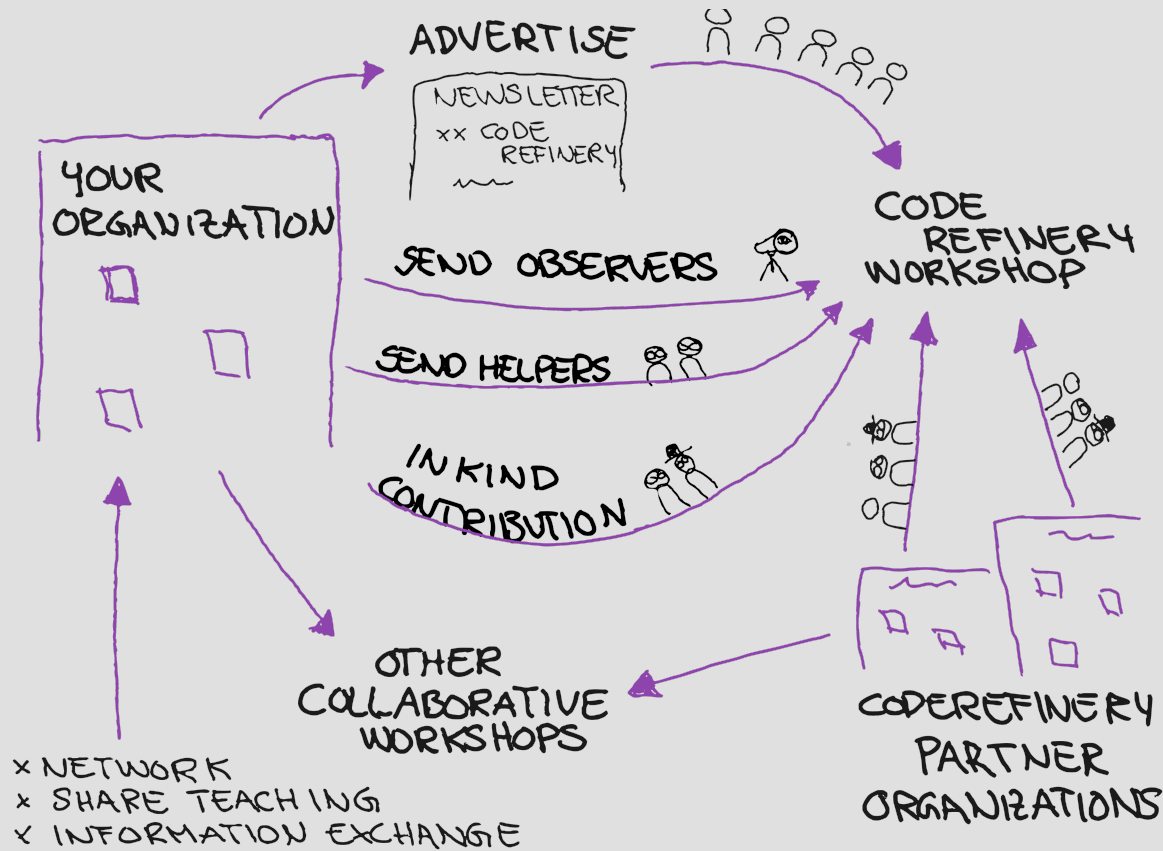
A typical CodeRefinery workshop



Participating as a learner



Participating as an organization



How to give a concert

- Assemble band
- On-board new band members
- Organize venue
- ~~Tour bus and hotel~~
- Announce
- Print posters and advertise
- Set playlist
- Deal with tickets and cancellations
- Provide rehearsal space
- Rehearse and update the set
- Provide security and manage audience
- Sound check
- Tune the instruments
- Play the show
- Broadcast, light, and record
- Clean up
- Give interviews
- Edit and publish recording
- Publish memoirs
- Get money for the next tour

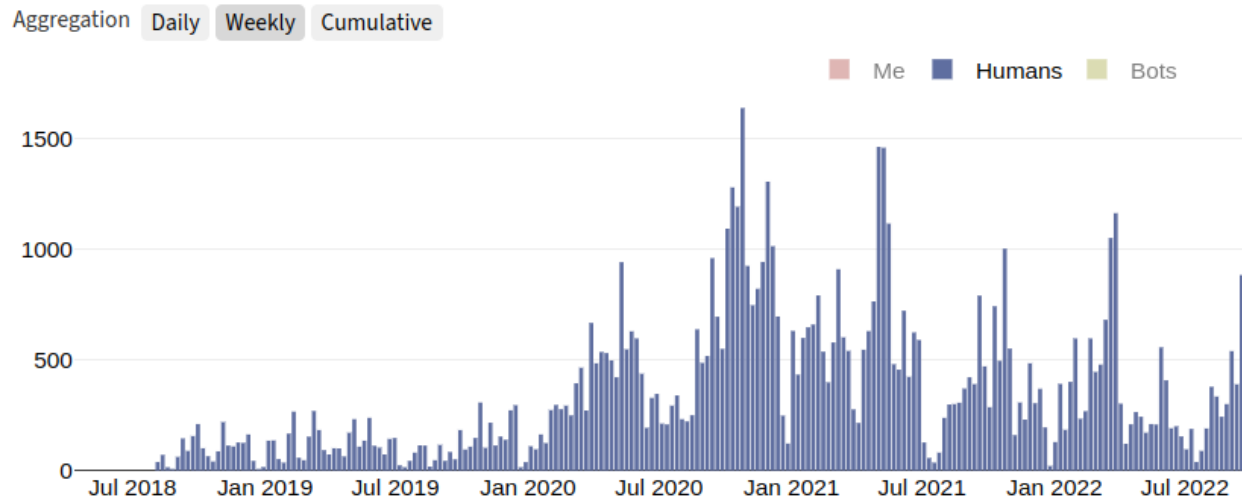


[Copyright 1980 Atlantic Records, fair use]

What we have learned

- Coordination effort starts to outweigh lesson preparation and teaching
- The importance of communicating value to organizations
- Outreach and marketing takes time and work
- Learners and volunteers typically join for a short time
- Building long lasting relationships takes effort

Messages sent over time



Big problem: visibility and outreach

[] check this box if you would like to be informed about events in future organized by us. This is how we will store your contact information: ...

Ideas:

- roll-ups
- info-screens
- events
- stands
- office hours
- newsletter

GitLab: "adopt-a-service"

- Code repository hosting for Nordic research software:
<https://coderefinery.org/repository/>
- Over 300 active users, 175 groups, with over 400 active projects
- Operated by DeiC
- ~50k DKK/ year: **ca. 2 USD per active user and month**
- Funding beyond 2024/2025 unclear

Lessons learned/ observations

- From local computing centers towards national (e.g. [Sigma2/NRIS](#)) and international compute resources (e.g. [EuroHPC/LUMI](#))
- Distributed support staff: Organizations will have to collaborate on training and stay close to the users
- Good enough practices that are used better than perfect practices not applied
- Instead of "good for others": "good for your future you" and as side effect good for others"

Where to start: Code review

```
205 - def groups(account):
206     l = shell_command(f"id -Gn {account}").split()
207     # removing these two because we treat them separately
208     outside
209     l.remove(account)
210     @@ -221,13 +238,12 @@ def dont_colorize(text, color):
211     return text
212
213
214 - user = getpass.getuser()
215
216
217     @click.command()
218 - @click.option(
219     "-u", "--user", default=user, help=f"The username to check
220     (default: {user})."
221 )
222
223     @click.option(
224     "--csv",
225     is_flag=True,
226
227     @@ -238,7 +254,7 @@ def dont_colorize(text, color):
228     is_flag=True,
229     help="Disable colors.",
230 )
231
232 - def main(user, csv, no_colors):
233     cf.update_palette({"blue": "#2e54ff"})
234     cf.update_palette({"green": "#08a91e"})
235     cf.update_palette({"orange": "#ff5733"})
236
237     @@ -247,12 +263,24 @@ def main(user, csv, no_colors):
238     # redefine the colorize function to do nothing
239     colorize.__code__ = dont_colorize.__code__
240
241
242
243
244
245 + def collect_groups(account):
246     l = shell_command(f"id -Gn {account}").split()
247     # removing these two because we treat them separately
248     outside
249     l.remove(account)
250
251     return text
252
253
254 + default_user = getpass.getuser()
255
256
257     @click.command()
258 + @click.option("-u", "--user", help=f"The username to check
259     (default: {default_user}).")
260 + @click.option("-p", "--project", help=f"The allocation
261     project.")
262
263     @click.option(
264     "--csv",
265     is_flag=True,
266
267     is_flag=True,
268     help="Disable colors.",
269 )
270
271 + def main(user, project, csv, no_colors):
272     cf.update_palette({"blue": "#2e54ff"})
273     cf.update_palette({"green": "#08a91e"})
274     cf.update_palette({"orange": "#ff5733"})
275
276     @@ -247,12 +263,24 @@ def main(user, csv, no_colors):
277     # redefine the colorize function to do nothing
278     colorize.__code__ = dont_colorize.__code__
279
280
281 + if user and project:
282 +     sys.exit(
283 +         colorize("ERROR: ", "red")
284 +         + "please specify user (-u) or project (-p) but not
285         both"
286     )
287
288 + if user is None:
289 +     user = default_user
```

"but the code is not ready"

Future

- Put the project on the map
- Non-profit organization
- Better communicate expectations, needs, and goals
- Make this something that people put on their CVs
- Make this something that organizations put on their homepages

How you can help

1: Join as **observer** and give feedback and suggestions

2: Co-organize, co-create, **co-teach** with us

3: Outreach, advertizing, **marketing**, social media

4: **GitLab** future

5: Know-how about organizational **legal forms** in the Nordics