Low availability of code in ecology: call for urgent action

Dr. Antica Culina
Netherlands Institute of Ecology
Code is essential part of research

*Enables:*

- remembering one’s own analysis after a while
- others to understand and trust your work
- others to re-use the code and extend it
Scientific code:

1) **Analysis code**
   • used to simulate model results, conduct statistical analyses, (also to correct errors in data, create figures).
   • *the majority of code written for ecological studies is analysis code*

2) **Scientific software**
   • more general and is designed to be used in many different projects (e.g., R and Python packages).
There are existing guidelines for code in ecology/general
Code policies Ecology journals, as of June 2015

• More than 75% do not mention scientific code in the author guideline.

• Of the journals that mention scientific code, only 14% require code to be made available.

from Mislan et al, 2016, Trends in Ecology and Evolution
What did we do?

• Looked at code availability in 14 ecological journals with some kind of a code sharing policy / recommendation
• Looked at some aspects of the shared code (where shared, how referred to..)
Search for articles

Randomly select 200
- 4366 articles

Code used?

Code released?
- Yes = 189

Evaluate code
How many papers that used code, released that code?

- Only 14%! (36 out of 189 papers)
- This is worrying, as these papers were published in Journals with a code policy (only 14 % of Ecology journals)
Where was code shared?

Means that:
- Only slightly more than 50% is accessible to all (repository, personal website)
- Long term availability (but not to everyone) for most of the code

Of this:
- 50% Dryad
- 25% GitHub
- 15% Figshare
As expected, Ecology is lagging behind

Perkel, 2016, Nature
Full access to rerun the analysis:

For articles that shared code, is there a full access to rerun the code?

Is CODE shared via repository/website?

Is the DATA shared?

Is the SOFTWARE a free source?
Is code easily discoverable from the article?

• 8 out of 36 articles do not explicitly mention that there is code available

• *discoverability of the code on its own (without knowing about the article)*—not assessed
Is code documented and easy to follow?

• Documentation is usually ok: 90% had some kind of documentation
  ✓ manual
  ✓ brief how to
  ✓ inline comments

• Around 50% has inline comments – easy to follow
Publish your computer code: it is good enough

Freely provided working code — whatever its quality — improves programming and enables others to engage with your research, says Nick Barnes.
What next?

• We want to go back to the analysis, and repeat it for 2018
  ➢ Has anything changed since?

Comments and suggestions welcomed!
Watch this space
a.culina@nioo.knaw.nl
@antica_c
@EvansSre