Why/How/When to release (research) software publicly

Manodeep Sinha

Swinburne University

What will you learn?

- **Why** you should release (reproducibility/robustness/versatility)

What will you learn?

- **Why** you should release (reproducibility/robustness/versatility)

- **When** you should release (as early as possible)

What will you learn?

- **Why** you should release (reproducibility/robustness/versatility)

- **When** you should release (as early as possible)

- **How** you should release (license/permanent DOI)

Why you should publicly release research software

- Others can replicate* your software

Why you should publicly release research software

- Others can replicate* your software

- Adapt to their needs -> increases productivity

Why you should publicly release research software

- Others can replicate* your software

- Adapt to their needs -> increases productivity

- Bug-finding becomes distributed -> more robust

Why you should publicly release research software

- Others can replicate* your software

- Adapt to their needs -> increases productivity

- Bug-finding becomes distributed -> more robust

- Others can contribute code -> increases versatility

Why you should publicly release research software

- Others can replicate* your software

- Adapt to their needs -> increases productivity

- Bug-finding becomes distributed -> more robust

- Others can contribute code -> increases versatility

- Aspirational -> All publicly funded research should be open

Some Myths about Open-Source

- You do not have to support user queries/requests

Some Myths about Open-Source

- You do not have to support user queries/requests

- You are free to move on from that project (but others can take over if open-source)

Some Myths about Open-Source

- You do not have to support user queries/requests

- You are free to move on from that project (but others can take over if open-source)

- Open-source does not mean open-development

When should you release code?

- Timing depends on the nature of your research, go for earliest feasible

When should you release code?

- Timing depends on the nature of your research, go for earliest feasible

- Recommended workflow -> create GitHub repo (private if need be), clone repo, commit and push frequently

When should you release code?

- Timing depends on the nature of your research, go for earliest feasible

- Recommended workflow -> create GitHub repo (private if need be), clone repo, commit and push frequently

  - open development

When should you release code?

- Timing depends on the nature of your research, go for earliest feasible

- Recommended workflow -> create GitHub repo (private if need be), clone repo, commit and push frequently

  - open development

- Beware: STRONG inertia in moving from private->public

When should you release code?

- Timing depends on the nature of your research, go for earliest feasible

- Recommended workflow -> create GitHub repo (private if need be), clone repo, commit and push frequently

  - open development

- Beware: STRONG inertia in moving from private->public

- Public code forces me to write better code

How to release code publicly?

- Releasing needs a license

How to release code publicly?

- Releasing needs a license

- License grants permission to inspect, modify and (re)-distribute code

How to release code publicly?

- Releasing needs a license

- License grants permission to inspect, modify and (re)-distribute code

- Absolves you of liabilities from bugs in the code

How to release code publicly?

- Releasing code does not necessarily mean releasing the source, possible (but unusual) to release binaries

How to release code publicly?

- Releasing code does not necessarily mean releasing the source, possible (but unusual) to release binaries

- Take 3 scenarios to demonstrate the possible range

How to release code publicly? "Minimal" mode

- Put the binary/source(s) file on your personal website

How to release code publicly? "Minimal" mode

- Put the binary/source(s) file on your personal website

- *Always* include a license

How to release code publicly? "Minimal" mode

- Put the binary/source(s) file on your personal website

- *Always* include a license

- GADGET-2 was released this way

How to release code publicly?
"Minimal" mode

- Put the binary/source(s) file on your personal website

- *Always* include a license

- GADGET-2 was released this way

TITLE	CITED BY	YEAR
The cosmological simulation code GADGET-2 V Springel Monthly notices of the royal astronomical society 364 (4), 1105-1134	5666	2005

How to release code publicly?
"Minimal" mode

- Put the binary/source(s) file on your personal website

- *Always* include a license

- GADGET-2 was released this way

TITLE	CITED BY	YEAR
The cosmological simulation code GADGET-2 V Springel Monthly notices of the royal astronomical society 364 (4), 1105-1134	5666	2005

How to release code publicly?
"Minimal" mode

- Very little effort required

How to release code publicly?
"Minimal" mode

- Very little effort required

- Enables a way for others to run/build from your work

How to release code publicly?
"Minimal" mode

- Very little effort required

- Enables a way for others to run/build from your work

- Unclear how to cite

How to release code publicly? "Minimal" mode

- Very little effort required

- Enables a way for others to run/build from your work

- Unclear how to cite

- Lack of transparency -> no explicit Changelog, progress tracker, milestones

How to release code publicly? "Minimal" mode

- Very little effort required

- Enables a way for others to run/build from your work

- Unclear how to cite

- Lack of transparency -> no explicit Changelog, progress tracker, milestones

- Can be interpreted as less trust-worthy and limit uptake (there are exceptions)

How to release code publicly? "Normal" mode

- Public GitHub repo + license -> create a release

How to release code publicly? "Normal" mode

https://guides.github.com/activities/citable-code/

- Public GitHub repo + license -> create a release

How to release code publicly? "Normal" mode

- Public GitHub repo + license -> create a release

- Get a Zenodo DOI -> permanent archive, and link

How to release code publicly? "Normal" mode

https://guides.github.com/activities/citable-code/

- Public GitHub repo + license -> create a release

- Get a Zenodo DOI -> permanent archive, and link

- DOI can be cited (check on datacite.org)

How to release code publicly?
"Normal" mode

- Public GitHub repo + license -> create a release

- Get a Zenodo DOI -> permanent archive, and link

- DOI can be cited (check on datacite.org)

- Reference DOI in your associated paper

How to release code publicly? "Normal" mode

https://guides.github.com/activities/citable-code/

- Public GitHub repo + license -> create a release

- Get a Zenodo DOI -> permanent archive, and link

- DOI can be cited (check on datacite.org)

- Reference DOI in your associated paper

- Repeat for every new version

How to release code publicly? "Normal" mode

- More effort required

How to release code publicly? "Normal" mode

- More effort required

- Higher chances of getting credit for your work

How to release code publicly? "Normal" mode

- More effort required

- Higher chances of getting credit for your work

- Shows more software competency -> implies software might be more trustworthy

How to release code publicly? "Maximal" mode

- Normal mode +

How to release code publicly?
"Maximal" mode

- Normal mode +

- Good coding practices

How to release code publicly? "Maximal" mode

- Normal mode +

- Good coding practices

- Dedicated software paper (JOSS, etc)

How to release code publicly? "Maximal" mode

- Normal mode +

- Good coding practices

- Dedicated software paper (JOSS, etc)

- Documentation (user and developer)

How to release code publicly? "Maximal" mode

- Normal mode +

- Good coding practices

- Dedicated software paper (JOSS, etc)

- Documentation (user and developer)

- Testing

How to release code publicly?
"Maximal" mode

- Normal mode +

- Good coding practices

- Dedicated software paper (JOSS, etc)

- Documentation (user and developer)

- Testing

- User-support

How to release code publicly? "Maximal" mode

- Normal mode +

- Good coding practices

- Dedicated software paper (JOSS, etc)

- Documentation (user and developer)

- Testing

- User-support

- Well-optimised for use-cases, runtime etc

How to release code publicly? "Maximal" mode

- Truly improves research - reliable results, template to learn good coding practices

How to release code publicly? "Maximal" mode

- Truly improves research - reliable results, template to learn good coding practices

- **Do not attempt** unless already very proficient

How to release code publicly? "Maximal" mode

- Truly improves research - reliable results, template to learn good coding practices

- Do not attempt unless already very proficient

- Creating high-quality software takes time -> time **not** spent on research

How to release code publicly? "Maximal" mode

- Truly improves research - reliable results, template to learn good coding practices

- Do not attempt unless already very proficient

- Creating high-quality software takes time -> time not spent on research

- If you proceed, still need to balance with research expectations

Releasing code publicly: CAS Edition

- Barriers to releasing code

Releasing code publicly: CAS Edition

- Barriers to releasing code

  - code not good enough -> bring to code-review

Releasing code publicly: CAS Edition

- Barriers to releasing code

  - code not good enough -> bring to code-review

  - too much effort to spruce up -> start early

Releasing code publicly: CAS Edition

- Barriers to releasing code

  - code not good enough -> bring to code-review

  - too much effort to spruce up -> start early

  - is this code-base worth releasing -> why not public by default

Releasing code publicly: CAS Edition

- Barriers to releasing code

  - code not good enough -> bring to code-review

  - too much effort to spruce up -> start early

  - is this code-base worth releasing -> why not public by default

- We have a strong open-source culture at CAS

Releasing code publicly: CAS Edition

- Barriers to releasing code

  - code not good enough -> bring to code-review

  - too much effort to spruce up -> start early

  - is this code-base worth releasing -> why not public by default

- We have a strong open-source culture at CAS

- **IDEA** - Start peer-based code-mentoring at CAS? Share credit, add collaborators to review source code