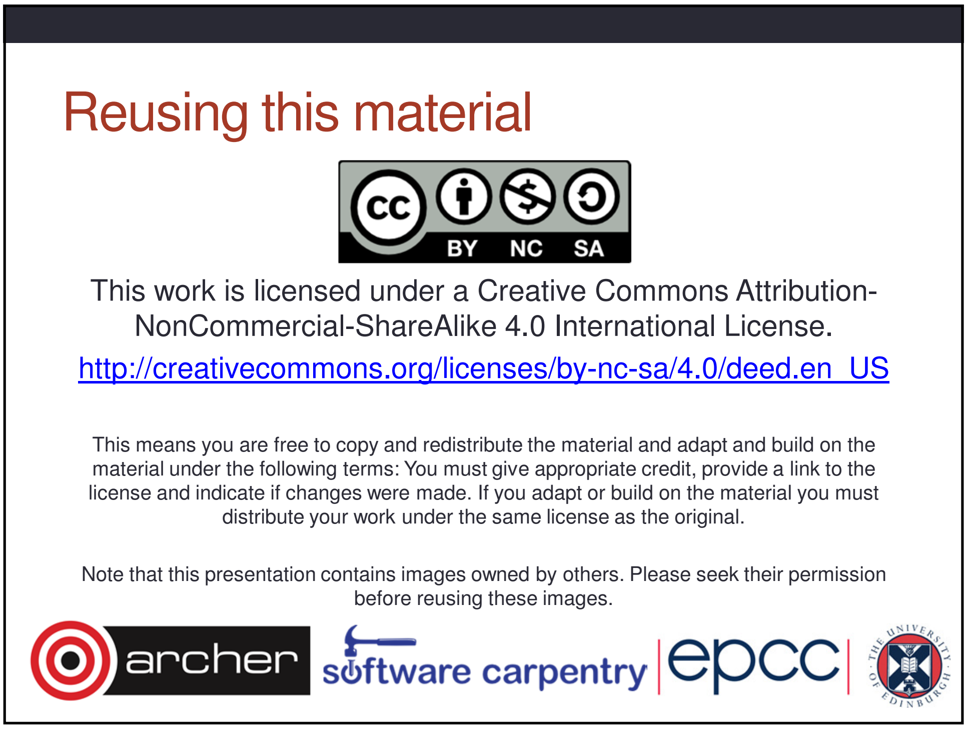
Pulling everything together

Mike Jackson, ARCHER CSE Team
michaelj@epcc.ed.ac.uk



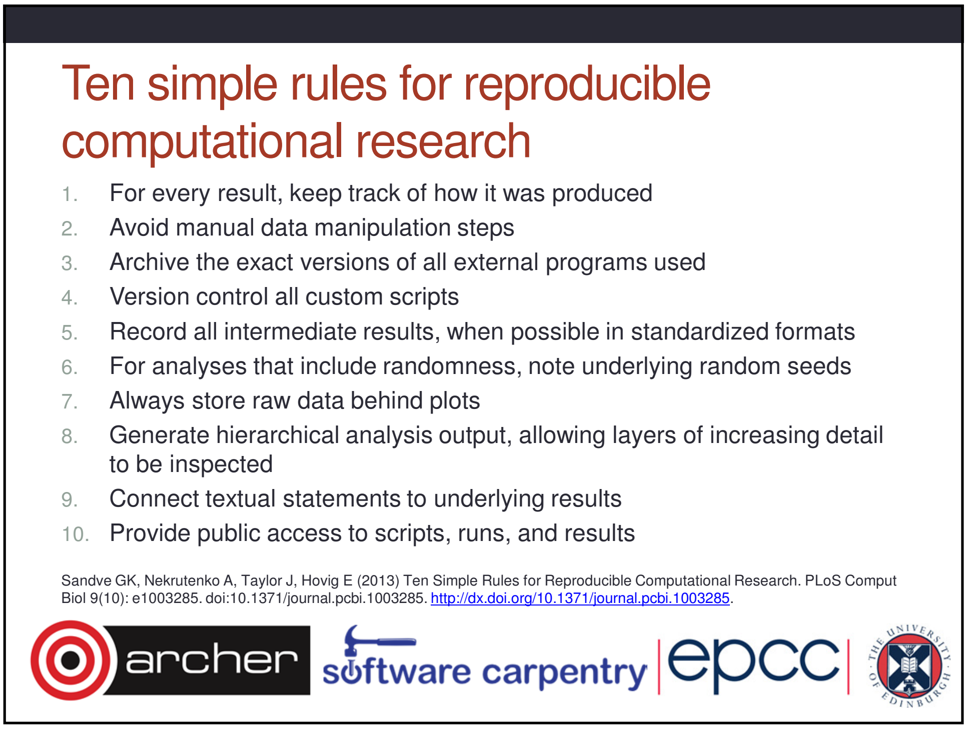
Reusing this material

http://www.archer.ac.uk
support@archer.ac.uk

Best practices for scientific computing

1. Write programs for people, not computers
2. Let the computer do the work
3. Make incremental changes
4. Don’t repeat yourself (or others)
5. Plan for mistakes
6. Optimize software only after it works correctly
7. Document design and purpose, not mechanics
8. Collaborate

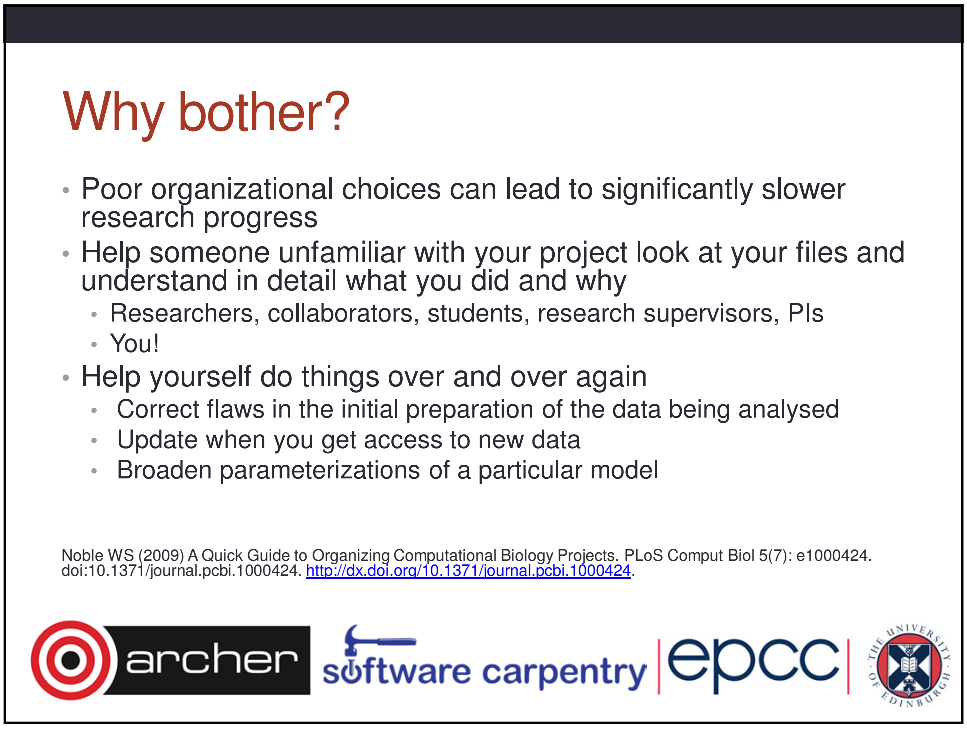
Wilson G, Aruliah DA, Brown CT, Chue Hong NP, Davis M, et al. (2014) Best Practices for Scientific Computing. PLoS Biol 12(1): e1001745. doi:10.1371/journal.pbio.1001745. http://dx.doi.org/10.1371/journal.pbio.1001745.



Ten simple rules for reproducible computational research

1. For every result, keep track of how it was produced
2. Avoid manual data manipulation steps
3. Archive the exact versions of all external programs used
4. Version control all custom scripts
5. Record all intermediate results, when possible in standardized formats
6. For analyses that include randomness, note underlying random seeds
7. Always store raw data behind plots
8. Generate hierarchical analysis output, allowing layers of increasing detail to be inspected
9. Connect textual statements to underlying results
10. Provide public access to scripts, runs, and results

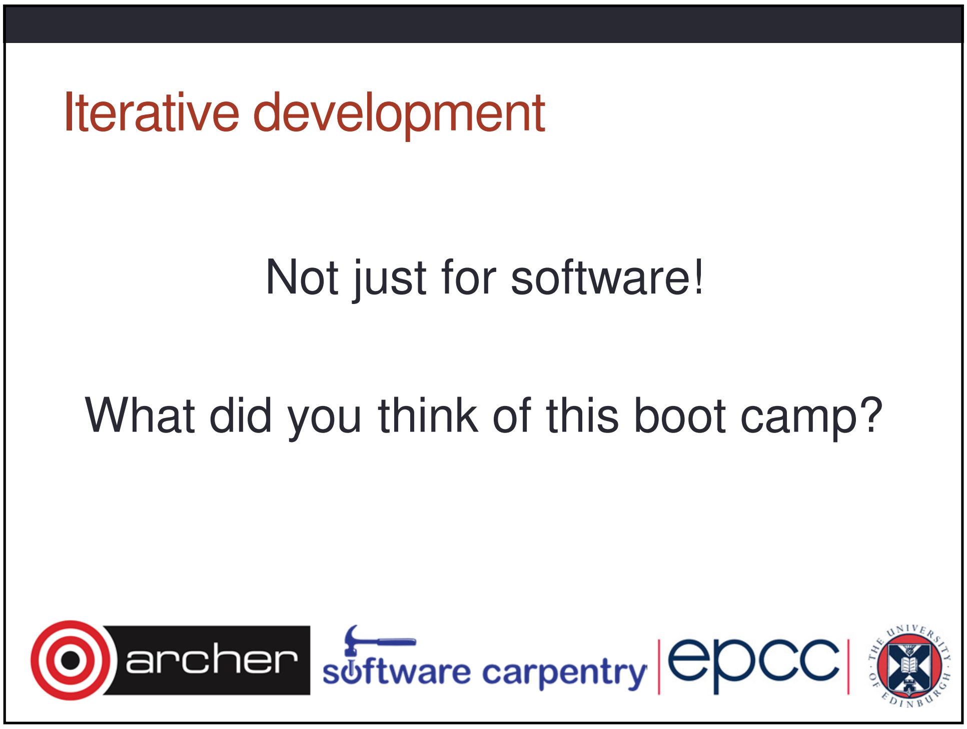
Sandve GK, Nekrutenko A, Taylor J, Hovig E (2013) Ten Simple Rules for Reproducible Computational Research. PLoS Comput Biol 9(10): e1003285. doi:10.1371/journal.pcbi.1003285. http://dx.doi.org/10.1371/journal.pcbi.1003285.



Why bother?

- Poor organizational choices can lead to significantly slower research progress
- Help someone unfamiliar with your project look at your files and understand in detail what you did and why
  - Researchers, collaborators, students, research supervisors, PIs
  - You!
- Help yourself do things over and over again
  - Correct flaws in the initial preparation of the data being analysed
  - Update when you get access to new data
  - Broaden parameterizations of a particular model

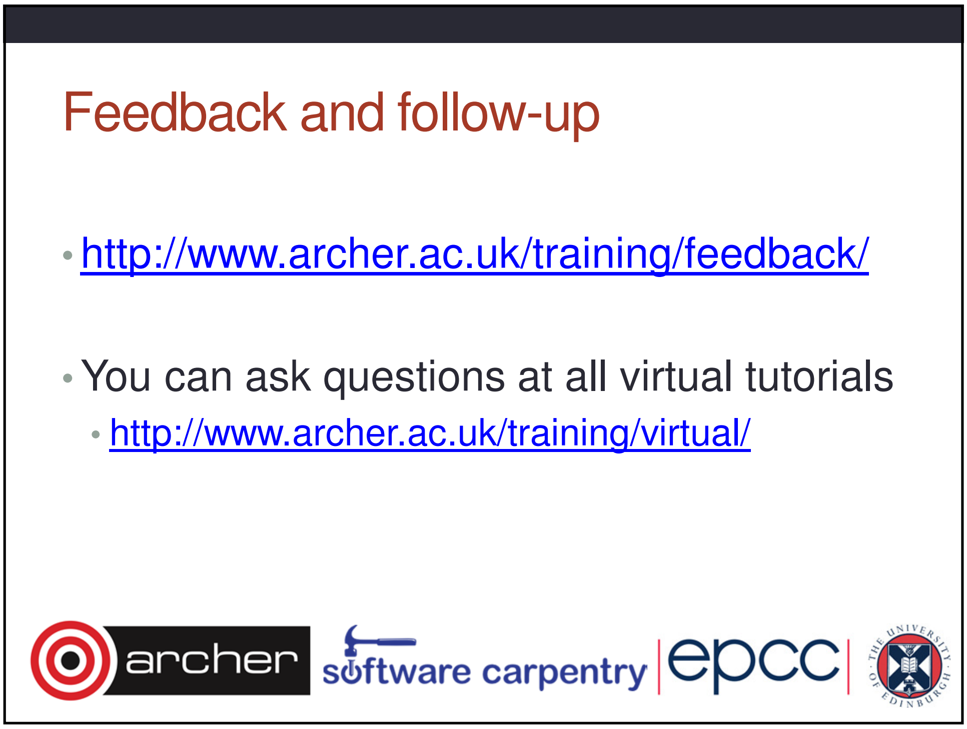
Noble WS (2009) A Quick Guide to Organizing Computational Biology Projects. PLoS Comput Biol 5(7): e1000424. doi:10.1371/journal.pcbi.1000424. http://dx.doi.org/10.1371/journal.pcbi.1000424.



Iterative development

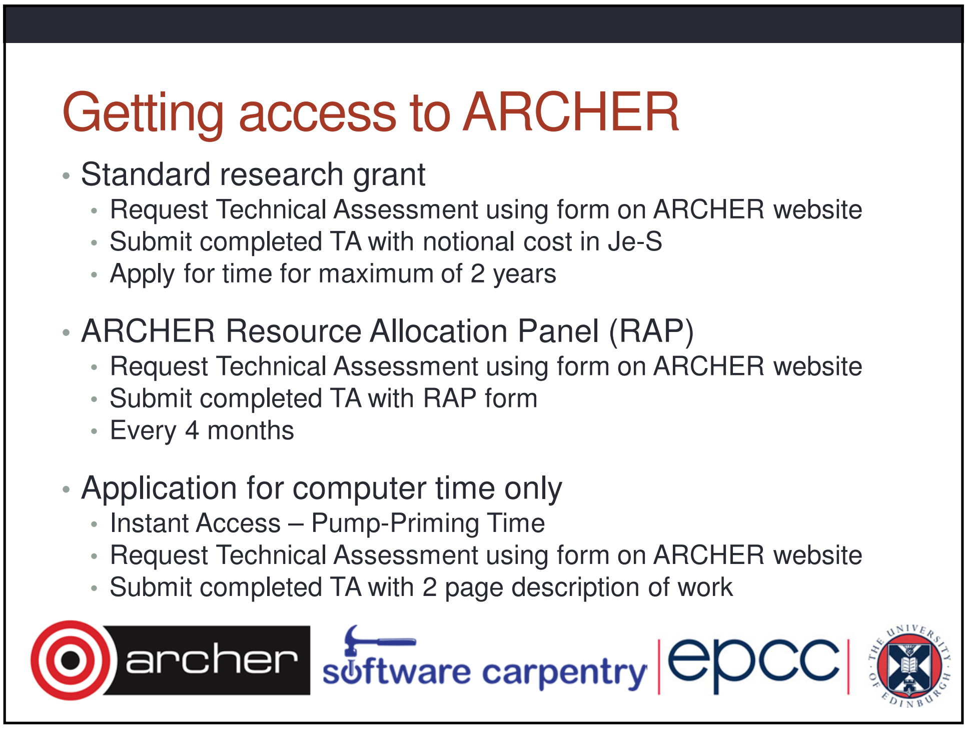
Not just for software!

What did you think of this boot camp?



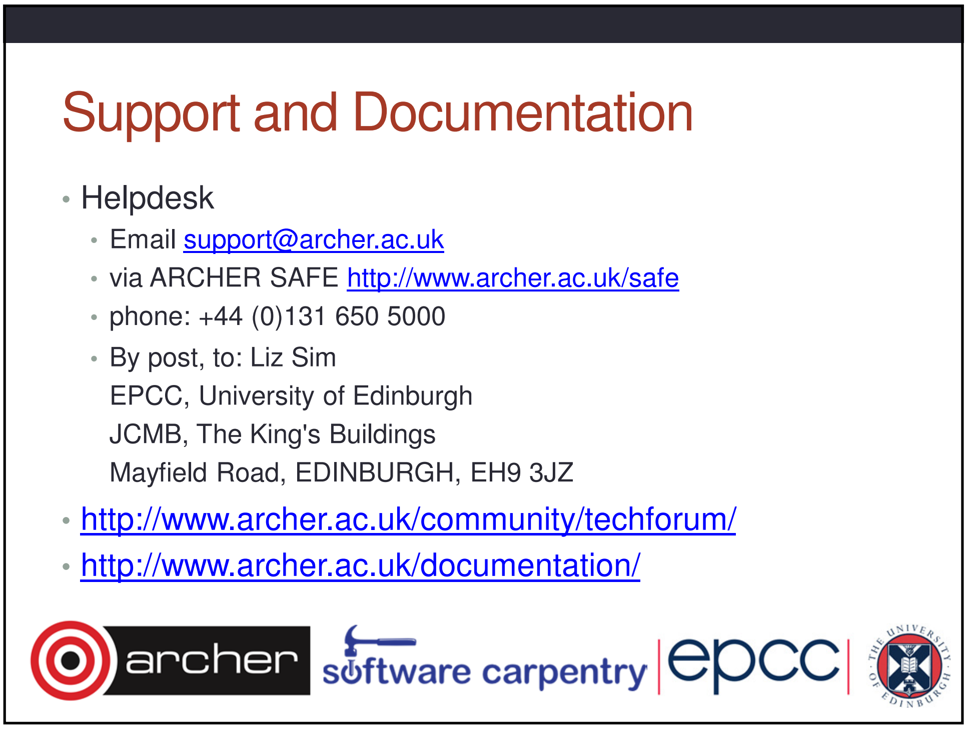
Feedback and follow-up

- http://www.archer.ac.uk/training/feedback/
- You can ask questions at all virtual tutorials
  - http://www.archer.ac.uk/training/virtual/



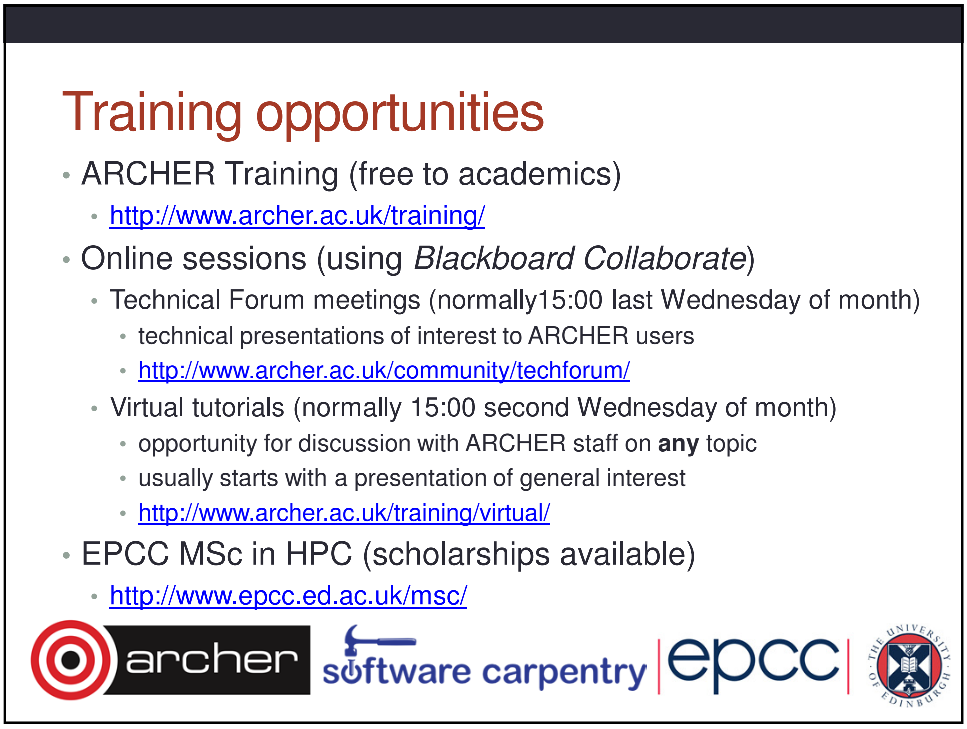
Getting access to ARCHER

- Standard research grant
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with notional cost in Je-S
  - Apply for time for maximum of 2 years
- ARCHER Resource Allocation Panel (RAP)
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with RAP form
  - Every 4 months
- Application for computer time only
  - Instant Access – Pump-Priming Time
  - Request Technical Assessment using form on ARCHER website
  - Submit completed TA with 2 page description of work



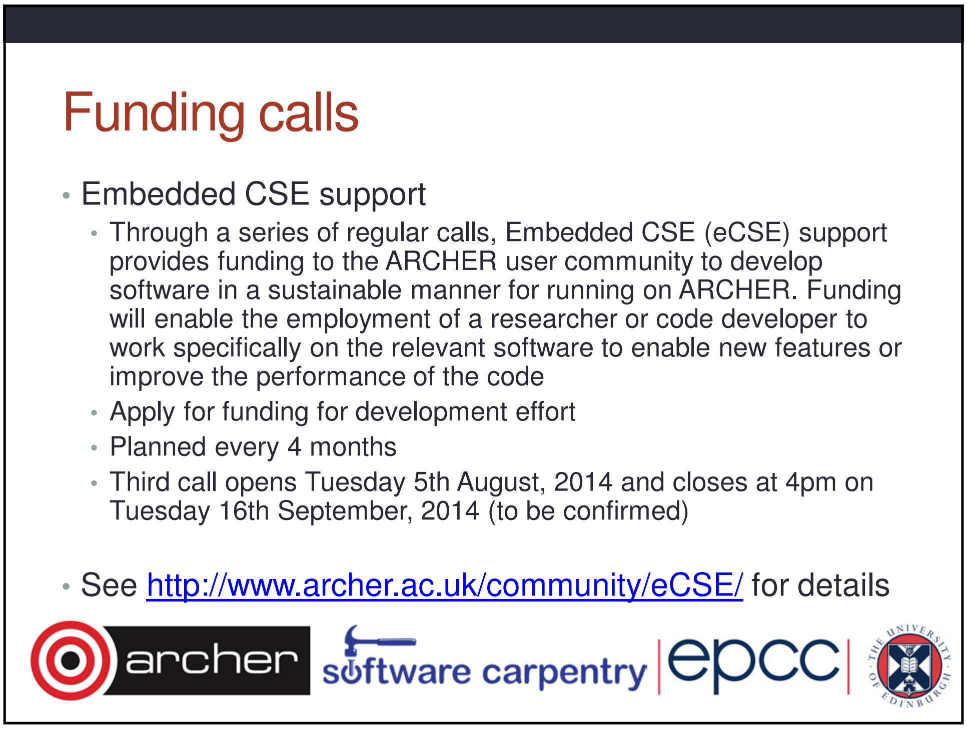
Support and Documentation

- Helpdesk
  - Email support@archer.ac.uk
  - via ARCHER SAFE http://www.archer.ac.uk/safe
  - phone: +44 (0)131 650 5000
  - By post, to: Liz Sim
    EPCC, University of Edinburgh
    JCMB, The King's Buildings
    Mayfield Road, EDINBURGH, EH9 3JZ
- http://www.archer.ac.uk/community/techforum/
- http://www.archer.ac.uk/documentation/



Training opportunities

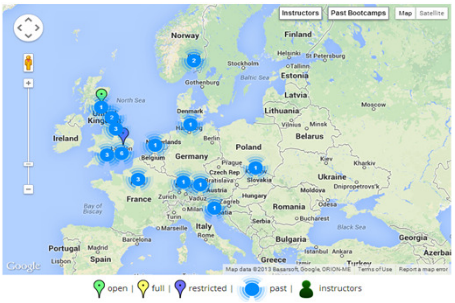
- ARCHER Training (free to academics)
  - http://www.archer.ac.uk/training/
- Online sessions (using *Blackboard Collaborate*)
  - Technical Forum meetings (normally15:00 last Wednesday of month)
    - technical presentations of interest to ARCHER users
    - http://www.archer.ac.uk/community/techforum/
  - Virtual tutorials (normally 15:00 second Wednesday of month)
    - opportunity for discussion with ARCHER staff on **any** topic
    - usually starts with a presentation of general interest
    - http://www.archer.ac.uk/training/virtual/
- EPCC MSc in HPC (scholarships available)
  - http://www.epcc.ed.ac.uk/msc/



Funding calls

- Embedded CSE support
  - Through a series of regular calls, Embedded CSE (eCSE) support provides funding to the ARCHER user community to develop software in a sustainable manner for running on ARCHER. Funding will enable the employment of a researcher or code developer to work specifically on the relevant software to enable new features or improve the performance of the code
  - Apply for funding for development effort
  - Planned every 4 months
  - Third call opens Tuesday 5th August, 2014 and closes at 4pm on Tuesday 16th September, 2014 (to be confirmed)
- See http://www.archer.ac.uk/community/eCSE/ for details

Getting involved in Software Carpentry



admin-uk@software-carpentry.org

http://software-carpentry.org/bootcamps

http://software-carpentry.org/lessons.html

archer | software carpentry | epcc | THE UNIVERSITY OF EDINBURGH

Software Sustainability Institute

www.software.ac.uk

We want the research community to

- Recognise software as a fundamental research output
- Recognise the value of research software engineers
- Receive better software training
- Recognise the role of software in reproducibility

What we do

- **Fellowship**: we find and nurture researchers with combined talents in research and software
- **Software Carpentry**: teaching basic software skills to researchers from all disciplines
- **Consultancy**: providing software expertise to advance research software
- **Campaigning**: raising the profile of software in research and highlighting the issues it faces

BETTER
SOFTWARE
BETTER
RESEARCH
www.software.ac.uk

We are supported by EPSRC under grant EP/H043160/1

Thank you!