Software and Computing Training
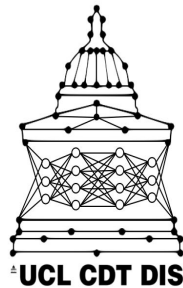
Nikos Konstantinidis (UCL)
Tim Scanlon (UCL)
Neil Chue Hong (Edinburgh)

Overview

❖ Introduction

❖ S&C training landscape in the UK

➢ Centres for Doctoral Training in Data Intensive Science

❖ International training offerings

❖ Future of training in UK

❖ Summary

❖ Statement

❖ Discussion Points

Introduction

❖ Introduction
  ➢ S&C is critical to the success of our entire programme
  ➢ Essential we have:
    ■ Well trained researchers in software/computing
      ● Produce robust, efficient and easily maintainable code/frameworks
    ■ Training and expertise on latest developments
      ● Exploit latest techniques and architectures

❖ A software training programme that provides these skills is vital
  ➢ Should provide a more coherent and structured training offering
    ■ Tailored for the S&C skills necessary to fulfil our physics goals
    ■ Also prepare researchers to work in other fields
  ➢ Should be a core part of the training we offer to all researchers
    ■ PhD, PDRA, core staff and academics

Impact and Career Progression

- ❖ Effective skills training programme can also bolster our impact
  - ➢ Especially if with a significant industry input

- ❖ Produce highly skilled individuals with unique skillset
  - ➢ Mix of software engineering, data wrangling/cleaning, statistics, ML…
    - ■ 'ML engineers'
  - ➢ Many go to industry
    - ■ Improved outcomes for key parts of economy

- ❖ Provides strong case to successfully bid for UKRI innovation/impact funds
  - ➢ Complements and strengthens current activates
    - ■ Allow us to do more and better fundamental research
  - ➢ CDTs are a good example of this
    - ■ Extends the Doctoral Training Programme (DTP)
    - ■ CDT 'Halo-effect' enhances DTP student's training
    - ■ Gives the potential for better/more research outcomes

- ❖ Many of the skills we develop vital to long-term economic growth
  - ➢ Areas the government are looking to bolster in post-CV19 growth
  - ➢ Can engage with this with a wider training programme

Current UK S&C Training Landscape

❖ Landscape across the UK somewhat ad-hoc and disjointed
  ➢ Undergraduate training
    ■ Programming usually taught (but no always)
  ➢ Scientific computing MSc starting to become popular
  ➢ Training provided by universities and experiments to PhD students
    ■ Varies depending on the UK
  ➢ Software carpentries
    ■ Usually only focuses on a few basic skills
  ➢ Swift-HEP
    ■ Only a relatively some fund provided for training
  ➢ STFC Data Science focussed summer school
    ■ Curriculum varies every year
  ➢ Industry
    ■ Occasional courses provided by Partners (e.g. FPGAs by Intel)
  ➢ International options
    ■ HEP Software Foundation workshops, CERN and other Summer Schools

❖ A more coherent and formal training programme seems very desirable
  ➢ STFC's Centres for Doctoral Training (CDT) in Data Intensive Science (DIS) an example in this direction

CDT Programmes

❖ Designed to enhance standard training with more focus on DIS skills in academia and industry

❖ In addition to standard HEP/Astro courses:
- ➢ Introductory software carpentry
- ➢ 120 hours of taught courses
  - ■ Choices from: software engineering, HPC, computational/simulation models, machine learning, numerical optimisation, statistics
- ➢ Industry Group projects
- ➢ 6-month industry placements
- ➢ Training workshops/seminars
  - ■ Data privacy; data security; FPGAs for DIS
- ➢ National CDT Training Event
- ➢ STFC DIS Summer School

❖ Almost all opportunities open to non-CDT students
- ➢ Seen significant 'Halo-effect'

	Activities
Year 1	• Taught courses • Group project • Exams • PhD project assignment • Software (SW) Carpentry • CDT Summer School **Transferable Skills** *Communication skills, Scientific writing, Media training*
Year 2	• **MPhil to PhD transfer** • Placement assignment • SW Carpentry (tutor) **Transferable Skills** *Entrepreneurship, Intellectual property, Science in the economy*
Year 3	• Placement • International training school • CDT Summer School (tutor) **Transferable Skills** *Research planning, Proposal writing*
Year 4	• International conference • **PhD Award** **Transferable Skills** *Interview skills, Careers workshop*

Industry Placements

❖ As part of CDTs, PhD students or PDRAs undertake 3/6-month placements
  - ➢ DIS related projects with CDT Partners
    - ■ Greatly enhances training
      - ● Tools, techniques, work-flows and language used in industry
    - ■ Promotes knowledge transfer
    - ■ Build strong links with industry
      - ● Creates further opportunities
    - ■ Impact skills/expertise developed in fundamental physics

❖ We've now been running placements for several years
  - ➢ Industry typically pays for the privilege

❖ Would be good to build similar scheme UK-wide
  - ➢ Allow PDRAs and academics in this programme to take part
  - ➢ Strong benefits to field: knowledge transfer, research outcomes, impact

DIS CDTs - Status

❖ The CDTs have been a very successful scheme so far
➢ UCL will welcome its 5th cohort this year
■ The 1st cohort are now graduating
➢ Will have enrolled > 50 students in the CDT

❖ Benefits
➢ The formal S&C training significantly enhances the student's skillset
■ Excellent feedback from industry partners on aptitude of students
➢ Mixing of students from different areas very beneficial
➢ Enabled strong links with industry to be built
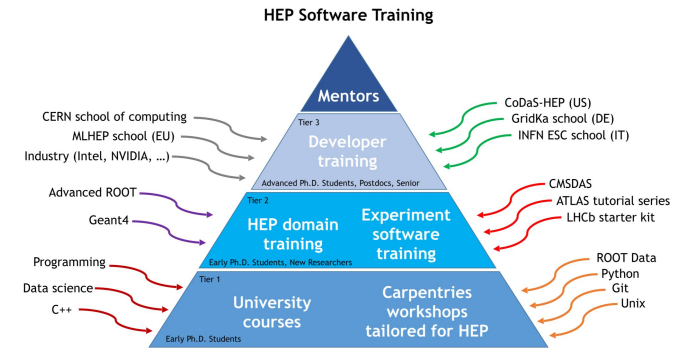■ Bring in industry funding to create more studentships

❖ However….
➢ STFC was not able to provide any studentships this year
■ Need better and increased long-term funding
➢ Scheme only impacts upon a small fraction of STFC studentships
■ Need to ensure the benefits are available to all students

S&C Training Internationally

❖ **Several international endeavours in this area**

➢ HSF: https://hepsoftwarefoundation.org/workinggroups/training.html

➢ IRIS-HEP: https://iris-hep.org/ssc.html

➢ First-hep: https://first-hep.org/about/

➢ Software carpentry: https://software-carpentry.org

HEP Software Training

❖ **All have a similar aim: to help the research community to provide training in the computing skills needed for researchers to produce high quality and sustainable software**

➢ Introduce a common HEP software curriculum and collaborative framework

■ Basic programming skills (Python, C++)

■ Basic software engineering skills (Unix, git/GitHub/GitLab, CI)

■ Skills in the core data tools (Root)

■ More advanced training when needed (parallel programming, machine learning, efficient software implementations)

❖ **Large degree of collaboration between all these different programmes**

Possible Future of Training in UK

- **Current S&C training offerings in UK need to implement a similar endeavour**
  - ➤ Need a platform for wider upskilling of field
    - ■ PhDs, PDRAs, core staff and academics
  - ➤ Build upon and expand existing training where possible (CDT etc.)
    - ■ Implement a UK-wide coherent/integrated training programme

- **Such a plan for the UK could include:**
  - ➤ Set of S&C standards for UG courses across UK (in conjunction with the IoP)
  - ➤ Recommended S&C curriculum for 1st year experimental HEP PhD students
  - ➤ Menu of regular UK workshops/courses (carpentry style?)
    - ■ Collaborate with HSF, IRIS-HEP, HEP-FIRST and others to develop courses
    - ■ Catering to all levels of seniority and expertise
    - ■ More advanced courses where specific skill-sets are needed:
      - ● Software engineering, firmware engineering, recent language developments (C++/Python), programming for heterogeneous architectures, parallel/optimised scientific software, machine learning, tools for software development/testing/profiling, ….
    - ■ Industry led workshops
      - ● Exploiting the latest hardware developments: FPGAs, GPUs, CPUs, intelligent networks working with Intel, NVidia, ARM and Mellanox

Summary

❖ Software/Computing expertise vital for future of HEP programmes
- ➤ Effective training in this area essential to future of field
    - ■ Enable and improve our research programmes
    - ■ Can also enhance impact/industry engagement

❖ Essential we build such a programme
- ➤ Create field-wide programme focussing on critical areas of expertise
    - ■ Available to PhDs, PDRAs, core staff and academics
    - ■ Significant engagement with industry is important
- ➤ Build current initiatives into a national scheme available to everyone
    - ■ Menu of options with regular software carpentry style workshops
- ➤ Build upon and expand investment in CDTs
- ➤ More closely link current research programmes

❖ Software/computing skills training programme should be integral to this endeavour
- ➤ Vital for the long-term goals of the field

R&D Training Statement

❖ Statement:

➢ Effective and relevant skills training is essential to this roadmap. However, HEP S&C training varies significantly, is constructed from a patchwork of training offerings, lacks coherence and is lagging offerings in other countries. The UK needs a coherent, integrated and planned S&C training offering, which caters for people from early PhDs through to senior academics. This training offering should engage with similar international endeavours, but should offer a forward looking training vision tailored for the UK's specific landscape and needs, including UK-based in-person training to build stronger cohorts across the UK.

❖ Discussion points:

➢ Could UG common curricula be influenced by working with the IoP?
➢ Could STFC recommend a common baseline S&C training for studentships?
➢ Is the provision of a more coherent UK-wide programme a Q of funding or planning?
➢ How should a more rigorous training offering be run and funded?
➢ How do we ensure that current PhD training endeavours are scaled up so they are available to all relevant PhD students?
➢ Should we just adopt a similar approach to IRIS-HEP for international engagement?
➢ Any key training areas that have been missed or required in the future?

Backup