Contribution ID: **133** Type: **Lightning talk**

Growing the data science community by expanding the CODATA/RDA school model

Wednesday, 10 October 2018 16:55 (5 minutes)

Various reports have commented on the shortage of individuals skilled in Research Data Science worldwide, which limits the transformative effect of the data revolution. Given the extent of the shortage, models to rapidly increase the cohort of researchers equipped to do data science and empower them to be ambassadors for their fields in teaching others is required.

The CODATA-RDA School for Research Data Science has established a successful two-week curriculum to provide a foundational level of data science skills to Early Career Researchers from a wide range of disciplinary backgrounds. The course covers the principles and practice of Open Science, research data management, using data platforms and infrastructures, data annotation, analysis, statistics, visualisation and modelling techniques. Students are taught in a computer lab setting with many hands-on exercises using open source tools, allowing them to learn new technologies and return home with access to the software they need.

Since the inaugural school in Trieste, Italy in August 2016, annual events have continued in Italy and other regional hubs have been established in Latin America and Africa. In collaboration with the International Centre for Theoretical Physics (ICTP) and its sister sites, we are primarily bringing data schools to researchers from Lower and Middle Income Countries, with the intention of reducing the digital divide. There is however, a big demand for these schools across Europe, North America and Australasia too, provoking us to consider business models to increase the delivery of schools and grow the community of data scientists worldwide.

The schools have helped many individuals to take their learning further. We run a student helper programme where participants return as classroom assistants to support the tutors facilitate hands-on exercises. This has offered new perspectives and increased the insights gained, enhancing their learning. (See: https://researchdata.springernature.com/users/81866-sara-el-jadid/posts/29719-enriching-my-learning-by-helping-others and https://researchdata.springernature.com/users/81847-marcela-alfaro-cordoba/posts/29656-my-journey-towards-open-science) Many have also gone on to run their own schools locally, increasing the reach of the schools on a train-the-trainer basis. This year, two previous students ran an Urban Data Science school in India, applying the lessons from the CODATA school to their peers. (See: https://shailygandhi.github.io/Urban-Data-Science-Curriculum-Development)

This paper will report on the use of the CODATA school model by others and our plans to expand this. A one-week school was run in Australia this June, taking the course materials as a base. In order to scale up this provision, we are establishing a set of requirements and a process for others to replicate the content (e.g. to retain agreed core elements of the curriculum, to adopt recommended teaching styles, to use open tools so participants retain access etc). A fee structure is also being proposed to endorse / badge these affiliated schools as following the CODATA/RDA model, and to provide a sustainability model for the core LMIC provision.

Type of abstract

Lightning Talk

Summary

This lightning talk will reflect on models to rapidly grow data science skills via a network of formal schools, a train-the-trainer model and franchising schemes to roll out the programme. Details will be shared about the emerging CODATA/RDA business model for sustainability of the schools.

Primary author: JONES, Sarah (UG)

Co-author: QUICK, Rob (IU (OSG))

Presenter: JONES, Sarah (UG)

Session Classification: Lightning Talks

Track Classification: Lightning Talks