2019 YCCSA SUMMER SCHOLARSHIP PROJECT SUBMISSION

This form is for prospective project supervisors to submit their projects to be included in the YCCSA Summer Scholarships Programme for 2019.

It is the purpose of the YCCSA Summer School that any projects submitted are novel and interdisciplinary in nature.

Date	***8/1/19***
Supervisors' Names and Departments / Affiliation and Contact Email	*Dr Jon Hill, Environment and Geography, jon.hill@york.ac.uk* *Dr Katie Davis, Biology, katie.davis@york.ac.uk*
Project Title	***Automating the construction of the Tree of Life***
Project Description	The Tree of Life (ToL) is the record of evolutionary history on Earth, detailing how all species are related. This underpins a vast amount of biological and environmental sciences. Making this tree is clearly a huge undertaking, but is vital as knowledge of shared evolutionary history is essential for understanding the potential impacts of human-induced climate change on today's biodiversity. There are a number of methods of creating subsets of the ToL (phylogenies), using molecular data, morphological data, even data based on behaviour or other ecological characteristics. These subsets can then be combined into larger trees as a faster way to obtain large, inclusive trees. There are complete, or near complete, phylogenies now available for groups such as mammals [1], birds [2] and fish [3], each of which took enormous effort to build, but there are no inclusive phylogenies for the majority of living organisms. This project will explore novel, interdisciplinary, methods of automatically building sections of the ToL. Using publicly available smaller phylogenies [4], taxonomic databases [5] and open-source software [6], we will explore ways in which large trees can be built quickly and efficiently in an automated manner. We can compare our results against those larger, published phylogenies and explore how best to use available data to create phylogenies for groups of organisms for which available data are limited. The project will involve writing new algorithms and scripts to pull data from online sources and collating them in an appropriate format. It will also involve exploring methods of most robustly generating a consensus tree of available data, e.g. how to use published phylogenies alongside new genetic data that is available. The student will receive thorough training in bioinformatics and modern phylogenetic methods. They will also gain experience of applying mathematical and/or computational methods to real world biological questions. This is a innovative and novel idea which could lead to new software and underpin new research in the University and beyond.

Required Skills	Python, R, software development process. The project will require programming skills and ans suit a student with an interest in biology.
Supervision and Collaboration Arrangements	The student will be primarily supervised by Hill and will be treated as part of Hill's research group (https://envmodellinggroup.github.io/) and weekly progress meetings will be held, as well as an open door policy. Meetings of the whole team, including Davis will be held at least 5 times during the project. Hill will supervise the programming and computational side of the project, whilst Davis will supervise the phylogenetics and evolution aspects of the project. All work will be made available on GitHub and be publicly available.
Project Dates	*The summer school runs for 9 weeks, starting on Monday, 08 July 2019 and finishing on Friday, 06 September 2019.*
Other Information	
References	

When complete, please email the form to sarah.christmas@york.ac.uk