# Nextflow Tower exemplar

| **Initial motivation** | To be able to perform bioinformatics analysis on WGS data from bacterial pathogens in order to be able to determine genetic relatedness and the presence of AMR determinants |
| --- | --- |
| **Major topics** | Usability, portability, scalability |
| **Constraints** | Unreliable internet transfer speeds |
| **Previous implementation** | Bioinformaticians ran command line tools in a series of bash scripts. This was performed on a local workstation accessible to   * staff from one department * those who had received command line training. |
| **Current implementation (workflow figure)** | The previous implementation migrated through several alternative configurations with iterative improvements   1. Firstly the multiple steps making up a bioinformatics workflow were wrapped into a series of declarative steps in a pipeline recognised by the workflow manager Nextflow. The software dependencies were provided through a single Docker/Singularity container. 2. Nextflow Tower was used initially for monitoring purposes 3. Nextflow was used to launch pipelines from the web interface and run the computing on cloud computing (AWS) |
| **Advantages of current implementation** | * A workflow manager + dependencies pre-built in containers enabled high throughput processing of multiple samples through several pipelines each producing a unique set of outputs pertaining to pathogen characterisation * The use of Nextflow tower for monitoring allowed for auditable pipeline running whereby the command used, including parameters, data inputs and data destinations, could be traced. * Launching of Nextflow tower using the interface created from the nf-core compatible schema allowed   + Non command line users to launch a pipeline   + Use of distributed cloud computing removed dependency on localised on-premise computing   + Long term full cost likely to be lower   + Full auditing of pipelines and costs |
| **Disadvantages** | Data generated on local sequencers needs to be transferred to the cloud. Local internet connectivity may mean that this is slow/interrupted. |

