**Computational science competency**

**Version: 0.001**

**Updated: VJ 3rd Apr 2025**

**There are 5 groups:**

1. High performance computing
2. Computing
3. Law and compliance
4. Research
5. Soft skills

-----------------------------------------------------------------------------------------------------------------------------------------------------

1. **High performance computing**
   1. HPC knowledge
   2. Estimate benefit of HPC on specific use case
   3. Running specific software/tool/workflow on HPC environments
   4. Parallel programming
   5. Performance optimization on HPC
   6. Debugging on HPC
2. **Computing**
   1. Working remotely in Linux environment via command line interface
   2. Managing Python and its packages via package manager
   3. Managing scientific software and preparing a related environment
   4. Using specific tools, software, or IDEs as each individual or together
   5. Estimating computing resource requirement
   6. Distribute software, data, or trained model
3. **Law and compliance**
   1. Compliance with license, policy, and ethics
4. **Research**
   1. Methodology, process, and tools
   2. Literature Review
   3. Data management
   4. Scientific data analysis
5. **Soft skills** 
   1. Teamwork
   2. Communication
   3. Human-driven service provision and support

**Reference:**

<https://competency.ebi.ac.uk/framework/bioexcel/3.0/competencies> and <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005772>

Red – AI generated content [using local LLM] (will be refined later)

Blue – fast thinking and subjected to be edited later

Black – may has minor change layer

**A.1. HPC knowledge**

**Remember**

* Remember common HPC glossary (e.g. node, partition, cluster, core, thread)
* Remember HPC component

**Understand**

* Able to describe the meaning of HPC glossary

**Apply**

* Match HPC glossary with HPC component

**Analyze**

* Describe the relation between glossary and component

**Evaluate**

* Able to estimate the impact of change in Hardware or component to the usage and performance

**Create**

* Able to sketch a concept of an HPC system for a given demand

**Skill**

- HPC knowledge [1-4]

- Application knowledge [5,6]

**Course**

- Introduction to HPC

- Profiling

**A.2. Estimate benefit of HPC on specific use case**

**A.3. Running specific software/tool/workflow on HPC environments**

**Remember**

* Able to run a job by using template script to run a job with only change in input filename

**Understand**

* Able to describe the meaning of each line of the job submission script
* Understand how job schedule works

**Apply**

* Able to modify the script according to the guideline

**Analyze**

* Understand how each line of script related to others and also with the HPC hardware

**Evaluate**

* Evaluate the effect of change in the script

**Create**

* Design a script that easy to use

**Skill**

- Slurm

- batch and offline job

**Course**

- Introduction to Slurm [prerequisite: A.1 understand]

**A.4. Parallel programming**

**A.5. Performance optimization on HPC**

**A.6. Debugging on HPC**

**B.1. Working remotely in Linux environment via command line interface**

**Remember**

* Remember common Linux command line

**Understand**

* Understand how each command works

**Apply**

* Using command to perform tasks

**Analyze**



**Evaluate**

**Create**

**Skill**

- Linux command line interface

- Linux file system

- Remote connection to a Linux system

**Course**

- Introduction to Linux for HPC

**B.2.** **Managing Python and its packages via package manager**

**Remember**

* Remember tools and command

**Understand**

* Understand how each command works

**Apply**

* Using command to perform tasks

**Analyze**



**Evaluate**

**Create**

**Skill**

- how to use pip

- how to use conda

**Course**

- Managing Python packages on HPC [prerequisite: B.1 (70%)]

**B.3.** **Managing scientific software and preparing a related environment**

**B.4.** **Using specific tools, software, or IDEs as each individual or together**

**B.5.** **Estimating computing resource requirement**

**B.6.** **Distribute software, data, or trained model**

**C.1. Compliance with license, policy, and ethics**

**D.1. Methodology, process, and tools**

**D.2. Literature Review**

**D.3. Data management**

**D.4. Scientific data analysis**

**E.1. Teamwork**

**E.2. Communication**

**E.3. Human-driven service provision and support**