
Overview

This standard identifies the competencies you need to develop algorithms and pipelines for the analysis of life science data.

You will be required to demonstrate that you can recognise and understand the format, scope and limitations of different types of life science-related data. This must be combined with the ability to identify the right computational (programming and infrastructure) needs so that you can specify or develop the correct analytical approach. If necessary you must be able to contribute to (sometimes leading) the development of new methods.

This activity is likely to be undertaken by individuals working in Life Science, Pharmaceutical, Chemical Biology, Agritech & Biotech industries. This could include job titles described as bioinformatics, computational biology, computational toxicology, Cheminformatics, Health informatics, Medical informatics, Agri-informatics for example.

Performance criteria

- You must be able to:*
- P1 determine the most suitable method for bioinformatics analysis of different biological and chemical life science data types.
 - P2 select appropriate statistical tests for the data.
 - P3 consider the research question and limitations of the experimental design when choosing an analytical approach.
 - P4 identify appropriate computing infrastructure requirements for the analysis of such data.
 - P5 apply current techniques, skills and tools necessary for computational biology practice.
 - P6 use a range of programming languages as required by different analytical approaches and data types.
 - P7 contribute to research to develop new analytical approaches.
 - P8 where appropriate, lead the development of new analytical approaches.
 - P9 build and implement analytical pipelines, or write new algorithms as necessary for the analysis of life science data.
 - P10 appropriately test analytical pipelines or new algorithms developed for the analysis of life science data.
 - P11 document novel algorithms and pipelines so that others can reproduce and make use of them.

Knowledge and understanding

You need to know and understand:

- K1 bioinformatics analysis methodologies and common bioinformatics software packages, tools and algorithms.
- K2 common workflow management tools.
- K3 programming and scripting languages.
- K4 the use of suitable version control tools.
- K5 emerging software sustainability practices.
- K6 open source software repositories.
- K7 local policy on the sharing or open release of code.
- K8 different types of licensing restrictions for established bioinformatics or statistical code that could form part of your method.
- K9 system administration and networks.
- K10 the computing infrastructure requirements of different analytical approaches.
- K11 relevant high performance computing platforms including Linux, Unix, local and remote HPC, and cloud computing.
- K12 documentation requirements in your organisation for the description of new methods.

Developed by	Cogent
Version Number	1
Date Approved	September 2018
Indicative Review Date	September 2022
Validity	Current
Status	Original
Originating Organisation	Cogent
Original URN	COGBIO-04
Relevant Occupations	Information and Communication Technology Professionals; Science; Science and mathematics Science; Science Professionals
Suite	Bioinformatics
Keywords	computing, platforms, Linux/Unix, local, remote, HPC, cloud
