ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

203000022 - Fair Data Management

DEGREE PROGRAMME

20BC - Master Universitario En Biologia Computacional

ACADEMIC YEAR & SEMESTER

2025/26 - Semester 1

Index

Learning guide

1. Description....1
2. Faculty....1
3. Prior knowledge recommended to take the subject....2
4. Skills and learning outcomes....2
5. Brief description of the subject and syllabus....3
6. Schedule....5
7. Activities and assessment criteria....7
8. Teaching resources....9

1. Description

1.1. Subject details

Name of the subject	203000022 - Fair Data Management
No of credits	3 ECTS
Type	Optional/elective
Academic year ot the programme	First year
Semester of tuition	Semester 1
Tuition period	September-January
Tuition languages	English
Degree programme	20BC - Master Universitario en Biologia Computacional
Centre	20 - Etsi Agronómica, Aliment. Y Biosistemas
Academic year	2025-26

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Mark Denis Wilkinson (Subject coordinator)		mark.wilkinson@upm.es	Sin horario. Any time via prior arrangement by email

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- Linux command line

4. Skills and learning outcomes *

4.1. Skills to be learned

CE01 - Comprender las bases moleculares y las técnicas experimentales estándares más comunes en las investigaciones ómicas (genómica, transcriptómica, proteómica, metabolómica, interactómica, etc.).

CE04 - Utilizar diferentes bases de datos (incluidos los bigdata), conocer sus estructuras y ontologías, aplicar la estadística a su análisis, siendo capaz de utilizar herramientas de representación y visualización.

CE08 - Capacidad de integrar tecnologías y sistemas propios de la Inteligencia Artificial, con carácter generalista, y en contextos más amplios y multidisciplinares.

CE09 - Capacidad de interpretar los modelos de clasificación supervisada y no supervisada obtenidos al aplicar las técnicas de Aprendizaje Automático para un conjunto de datos.

CG02 - Familiarizarse con el trabajo y los métodos de la Biología Computacional en condiciones reales, adquiriendo la capacidad de diseñar aplicaciones/experimentos de forma independiente y describir, cuantificar, analizar y evaluar críticamente los resultados obtenidos.

CG03 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con el area de la Biología Computacional.

CG05 - Que los estudiantes sean capaces de integrar conocimientos en el area de la Biología Computacional, de formular conclusiones, hipótesis o líneas de trabajo a partir de la información disponible, y de formarse una opinion fundamentada sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos.

CT01 - Capacidad para aplicar de forma profesional a su trabajo los conocimientos adquiridos considerando sus impactos en un contexto global y social.

CT07 - Ser capaz de manejar las tecnologías de la información y comunicación en un contexto profesional.

CT08 - Tener capacidad de análisis y síntesis para interpretar datos relevantes y abordar los problemas desde diferentes perspectivas.

4.2. Learning outcomes

RA5 - Ser capaz de desplegar aplicaciones en estos sistemas, así como saber identificar el tipo de tecnología más adecuado para el problema a resolver.

RA4 - Conocer distintas tecnologías para el almacenamiento y procesamiento de grandes cantidades de información en sistemas distribuidos y/o en la nube

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

This course presents the concepts of FAIR - Findability, Accessibility, Interoperability, and Reusability - as applied to scholarly data publishing. The FAIR Principles are discussed in detail, and then over a series of weeks a variety of technologies are taught and explored in a hands-on manner that provide technical solutions to each of the FAIR Principles.

5.2. Syllabus

1. Introduction to the course
    1.1. Introduction to the reusability problem
2. The FAIR Principles
3. Data and Metadata Modelling
    3.1. Introduction to Linked Data
    3.2. Ontologies and Ontology search engines
    3.3. Modelling Principles for Machines
    3.4. Modelling and transformation tools
4. Interacting with the Web
    4.1. REST
    4.2. The Linked Data Platform
5. Publishing Data and Metadata "FAIRly"
    5.1. DCAT, DCAT2, SKOS, and FDP
    5.2. Triplestores
6. Querying FAIR Data
7. Testing FAIRness

6. Schedule

6.1. Subject schedule*

Week	Type 1 activities	Type 2 activities	Distant / On-line	Assessment activities
1	**Introducción** Duration: 02:00 Lecture			
2	**Globally Unique Identifiers & Docker Tutorial** Duration: 02:00 Lecture			
3	**Identifiers and Linked Data** Duration: 02:00 Lecture			
4	**Ontology Search Engines and Data Modelling** Duration: 02:00 Lecture			
5	**Data Transformation RML YARRRML** Duration: 02:00 Lecture			**Assignment on Modeling and Synthesis** Problem-solving test Progressive assessment Not Presential Duration: 00:00
6	**REST and Linked Data Platform** Duration: 02:00 Lecture			
7	**Metadata Modelling** Duration: 02:00 Lecture			**Assignment on Web Interactions** Problem-solving test Progressive assessment Not Presential Duration: 00:00
8	**Simple Knowledge Organization System (SKOS)** Duration: 02:00 Lecture			
9	**Publishing FAIR data: the FAIR Data Point** Duration: 02:00 Lecture			
10	**Querying FAIR Data - SPARQL** Duration: 02:00 Lecture			**Assignment on publishing FAIR data** Problem-solving test Progressive assessment Not Presential Duration: 00:00
11	**Querying FAIR Data** Duration: 02:00 Lecture			

12	**Exploration of Metadata harvesting, FAIR Testing, and FAIR Signposting** Duration: 02:00 Lecture			
13				**Doing it all** Problem-solving test Progressive assessment Not Presential Duration: 00:00
14				
15				
16				
17				**Doing it all** Problem-solving test Global examination Not Presential Duration: 00:00

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
5	Assignment on Modeling and Synthesis	Problem-solving test	No Presential	00:00	20%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03
7	Assignment on Web Interactions	Problem-solving test	No Presential	00:00	20%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03
10	Assignment on publishing FAIR data	Problem-solving test	No Presential	00:00	20%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03
13	Doing it all	Problem-solving test	No Presential	00:00	40%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03

7.1.2. Global examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
17	Doing it all	Problem-solving test	No Presential	00:00	100%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Doing it all	Problem-solving test	Face-to-face	30:00	100%	5 / 10	CT08 CE04 CT01 CT07 CE08 CG05 CG02 CG03

7.2. Assessment criteria

Ongoing evaluation: 60%

Final Assignment: 40%

Global Assessment: Same as final assignment, but worth 100%

Extraordinary Evaluation: Same as final assignment, but worth 100%

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
The FAIR Guiding Principles for scientific data management and stewardship	Bibliography	https://doi.org/10.1038/sdata.2016.18
Linked Data	Web resource	https://www.w3.org/wiki/LinkedData
Introduction to Ontologies	Bibliography	https://oa.upm.es/10381/1/An_Introduction.pdf
Introduction to REST Design	Web resource	https://www.infoq.com/articles/rest-introduction/
REST in detail	Web resource	https://www.ics.uci.edu/~fielding/pubs/dissertation/rest_arch_style.htm (a chapter from Roy Fielding's thesis - highly detailed, for geeks only!)
DCAT-2 Specification	Web resource	https://www.w3.org/TR/vocab-dcat-2/
SKOS Primer	Web resource	https://www.w3.org/TR/skos-primer/
Introduction to SPARQL	Web resource	https://www.dataversity.net/introduction-to-sparql/

Evaluating FAIRness	Bibliography	https://www.nature.com/articles/s41597-019-0184-5