



UNIVERSIDAD  
POLITÉCNICA  
DE MADRID

PROCESO DE  
COORDINACIÓN DE LAS  
ENSEÑANZAS PR/CL/001



Etsi Agronómica, Aliment. y  
Biosistemas

# ANX-PR/CL/001-01

## GUÍA DE APRENDIZAJE

### ASIGNATURA

**203000024 - Retos En Programación Informática**

### PLAN DE ESTUDIOS

20BC - Master Universitario En Biología Computacional

### CURSO ACADÉMICO Y SEMESTRE

2025/26 - Primer semestre

## Índice

---

### Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	2
5. Descripción de la asignatura y temario.....	3
6. Cronograma.....	5
7. Actividades y criterios de evaluación.....	7
8. Recursos didácticos.....	9
9. Otra información.....	10

## 1. Datos descriptivos

---

### 1.1. Datos de la asignatura

<b>Nombre de la asignatura</b>	203000024 - Retos en Programación Informática
<b>No de créditos</b>	3 ECTS
<b>Carácter</b>	Optativa
<b>Curso</b>	Primer curso
<b>Semestre</b>	Primer semestre
<b>Período de impartición</b>	Septiembre-Enero
<b>Idioma de impartición</b>	Castellano
<b>Titulación</b>	20BC - Master Universitario en Biología Computacional
<b>Centro responsable de la titulación</b>	20 - Etsi Agronómica, Aliment. Y Biosistemas
<b>Curso académico</b>	2025-26

## 2. Profesorado

---

### 2.1. Profesorado implicado en la docencia

<b>Nombre</b>	<b>Despacho</b>	<b>Correo electrónico</b>	<b>Horario de tutorías</b> *
Mark Denis Wilkinson (Coordinador/a)	B54 CBGP	mark.wilkinson@upm.es	L - 13:15 - 14:00

\* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

## 3. Conocimientos previos recomendados

---

### 3.1. Asignaturas previas que se recomienda haber cursado

El plan de estudios Master Universitario en Biología Computacional no tiene definidas asignaturas previas recomendadas para esta asignatura.

### 3.2. Otros conocimientos previos recomendados para cursar la asignatura

- High level of experience and competence in any programming language
- Demonstrated ability to solve complex, novel problems by coding (any language)

## 4. Competencias y resultados de aprendizaje

---

### 4.1. Competencias

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.

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

CT02 - Capacidad para aplicar el método científico para la resolución de problemas de forma efectiva y creativa.

CT03 - Tener compromiso bioético y profesional y respeto por la sostenibilidad ambiental.

## 4.2. Resultados del aprendizaje

RA8 - Adquisición de conocimientos de cuales son las mejores prácticas para la publicación de datos científicos

RA7 - Adquirir conocimientos para el manejo de técnicas avanzadas de representación de datos y enfoques de integración masiva de datos

RA9 - Aprendizaje del uso de interfaces públicas de

## 5. Descripción de la asignatura y temario

---

### 5.1. Descripción de la asignatura

Though biology has quickly become a data-centric science, the practitioners of biology have not understood the importance of their data - in particular, the care that must be taken in annotating it, describing and representing it, and publishing it. Therefore, the greatest challenge for bioinformaticians at this time is related to Finding data, Accessing data, Integrating data, and Reusing data (F.A.I.R.). This course will initially focus on problems related to large-scale data integration. We will study the software tools for capturing data, from files and then directly from the Web. We will learn to manipulate data structures, and how to write clear software that represents complex data structures. We will learn about contemporary approaches to data and knowledge modeling in the life sciences. We will learn how to use these contemporary approaches to do global-scale, cross-resource queries, such that we can integrate the massive amounts of data our biologist colleagues require to interpret their results. Finally, we will learn how to publish scholarly data in a FAIR manner, and how to create Web interfaces to make these massive integrated datasets tractable for biologists.

**This course is aimed at students who are already competent programmers**, generally with several years of experience, who want to challenge their skills with a wide range of real-world, complex biological informatics coding and data exploration tasks. Students who do not have significant prior experience in coding **should not consider this course** - for students with insufficient coding skills, this course will be frustratingly time-consuming and interfere with your success in other courses. Assignments are provided approximately every 2 weeks, and have a 2

or 3 week submission deadline (depending on the complexity of the task). These deadlines are firm and will not be extended.

## 5.2. Temario de la asignatura

1. Data Discovery
2. Data Access
3. Data Integration
4. Data Publishing
5. Data Manipulation & Formatting
6. Formal Knowledge Representation

## 6. Cronograma

### 6.1. Cronograma de la asignatura \*

Sem	Actividad tipo 1	Actividad tipo 2	Tele-enseñanza	Actividades de evaluación
1	<b>GitHub and Proper Code Management</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
2	<b>Introduction to the Ruby programming language part 1</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral			
3	<b>Object Oriented Programming</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Assignment 1 - simple objects</b> TI: Técnica del tipo Trabajo Individual Evaluación Progresiva No presencial Duración: 00:00
4	<b>Automated Web Access</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
5	<b>"Representation State Transfer" and RESTful APIs</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Assignment #2 - Intensive Data Integration using REST APIs</b> TI: Técnica del tipo Trabajo Individual Evaluación Progresiva No presencial Duración: 00:00
6	<b>Ruby "gems", Creating documentation, good software publishing and documentation for reusability</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
7	<b>BioRuby Level 1 - sequence manipulation</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Assignment #3 - locate and visualize insertion sites in a genome</b> TI: Técnica del tipo Trabajo Individual Evaluación Progresiva No presencial Duración: 00:00
8	<b>BioRuby Level 2 - Other data types, literature, citations, pathways, ontological annotations, etc.</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
9	<b>BioRuby Level 2 - Other data types, literature, citations, pathways, ontological annotations, etc.</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Assignment #4 Find the orthologues between two genomes</b> TI: Técnica del tipo Trabajo Individual Evaluación Progresiva No presencial Duración: 00:00

10	<b>Linked Open Data</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
11	<b>SPARQL Query over Global Databases</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Assignment #5 Answer biologists questions as SPARQL queries</b> TI: Técnica del tipo Trabajo Individual Evaluación Progresiva No presencial Duración: 00:00
12	<b>Web Programming and basic Server administration</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
13	<b>Contemporary Linked Data solutions</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			<b>Build a Web interface over a linked data database representing Polyadenylation site information</b> EP: Técnica del tipo Examen de Prácticas Evaluación Global No presencial Duración: 20:00
14	<b>Contemporary Knowledge Representation and FAIR Data Publishing</b> Duración: 02:00 AC: Actividad del tipo Acciones Cooperativas			
15				
16				
17				

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

## 7. Actividades y criterios de evaluación

### 7.1. Actividades de evaluación de la asignatura

#### 7.1.1. Evaluación (progresiva)

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
3	Assignment 1 - simple objects	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	20%	5 / 10	CG02 CG05
5	Assignment #2 - Intensive Data Integration using REST APIs	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	20%	5 / 10	CE04 CE08 CT02 CT03
7	Assignment #3 - locate and visualize insertion sites in a genome	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	20%	5 / 10	CG02 CE01 CE09
9	Assignment #4 Find the orthologues between two genomes	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	20%	5 / 10	CE01 CE04 CE08 CT02
11	Assignment #5 Answer biologists questions as SPARQL queries	TI: Técnica del tipo Trabajo Individual	No Presencial	00:00	20%	5 / 10	CE04

#### 7.1.2. Prueba evaluación global

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
13	Build a Web interface over a linked data database representing Polyadenylation site information	EP: Técnica del tipo Examen de Prácticas	No Presencial	20:00	100%	5 / 10	CG02 CG05 CE01 CE04 CE08 CE09 CT02 CT03

### 7.1.3. Evaluación convocatoria extraordinaria

Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
Build a Web interface over a linked data database representing Polyadenylation site information	EP: Técnica del tipo Examen de Prácticas	Presencial	20:00	100%	5 / 10	CG02 CG05 CE01 CE04 CE08 CE09 CT02 CT03

## 7.2. Criterios de evaluación

For all assignments the criteria for evaluation are:

- 1) Is the answer correct and complete? Biologists will not accept incorrect answers. Regardless of the "elegance" of your software, if the answer is incorrect, the software is not useful.
- 2) Is the output easy to understand (from the perspective of the biologist)?
- 3) Does the software demonstrate that the learning objectives have been achieved? Every assignment has associated competencies - you should demonstrate these competencies by highlighting (in your code documentation) which sections of your code address each of the learning objectives.
- 4) Is the software well-documented and easy to use? Will I be able to understand how it works 2 years later?
- 5) Is the software well-architected?
  - Does it use sensible objects?
  - Does it use sensible and efficient algorithms?
  - Does it consider memory-management (especially for big-data)?

- Does it re-use, rather than re-invent?

- Is the solution "elegant"?

6) Was it submitted on-time.

## 8. Recursos didácticos

### 8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Principled Web Architecture	Bibliografía	<a href="http://portal.acm.org/citation.cfm?doid=514183.514185">http://portal.acm.org/citation.cfm?doid=514183.514185</a> Roy Fielding's description of his REST design patterns for Web interfaces
Data Integration in the era of 'omics	Bibliografía	doi: 10.1186/1752-0509-8-S2-I1 General discussion of biological data integration challenges
Identification of Orthologues - algorithms	Bibliografía	Methods Mol Biol. 2015;1231:203-32. doi: 10.1007/978-1-4939-1720-4_14.
Semantic Web Core Technologies	Recursos web	<a href="https://www.w3.org/RDF/">https://www.w3.org/RDF/</a> <a href="https://www.w3.org/OWL/">https://www.w3.org/OWL/</a>
Use of Semantic Web in biomedicine	Bibliografía	Samadian S; McManus B; Wilkinson MD. 2012. "Extending and encoding existing biological terminologies and datasets for use in the reasoned semantic web.". J. Biomed. Semantics 3:6.

SPARQL to query biomedical databases	Bibliografía	PLoS Comput Biol. 2016 Jun 23;12(6):e1004989. doi: 10.1371/journal.pcbi.1004989. eCollection 2016.
The FAIR Principles	Bibliografía	<a href="https://www.nature.com/articles/sdata201618">https://www.nature.com/articles/sdata201618</a>
FAIR Principles revisited	Bibliografía	<a href="http://content.iospress.com/articles/information-services-and-use/isu824">http://content.iospress.com/articles/information-services-and-use/isu824</a>
The FAIR Principles Implemented	Bibliografía	<a href="https://peerj.com/articles/cs-110/">https://peerj.com/articles/cs-110/</a>

## 9. Otra información

---

### 9.1. Otra información sobre la asignatura

La asignatura se relaciona con los ODS3, ODS4, y ODS9