

Universidad de Costa Rica
Escuela de Biología

II-2020

Sigla: B-0674

Nombre del Curso: Problemas especiales en Zoología 15: Herramientas prácticas para investigación reproducible

Ciclo: II

Créditos: 4

Horas lectivas: 4

Requisitos: B0304 y B0305 Ecología General (T y L); no tiene correquisitos

Profesores:

Beatriz Willink (beatriz.willink@ucr.ac.cr) Escuela de Biología, oficina #3A

Marcelo Araya (marcelo.araya@ucr.ac.cr) Escuela de Biología, oficina #4A

Horario de atención a estudiantes:

Beatriz: L a V [8.00,17.00] por correo electrónico

Marcelo: L a V [8.00,17.00] por correo electrónico

Descripción

Vivimos en un periodo único en la historia de la producción científica, en el que se generan datos y publicaciones a una tasa nunca antes vista. Esta mega producción surge debido tanto al crecimiento en el número de científicas y científicos a nivel mundial, como a los avances tecnológicos que permiten tomar datos de formas cada vez más eficientes. Sin embargo, una grave desventaja de este crecimiento es que ha generado incentivos profesionales que en ocasiones priorizan el volumen científico sobre su calidad. Uno de los efectos más permisivos es la poca reproducibilidad, producto de la escasa documentación de los métodos y análisis utilizados, así como de la falta de acceso a los datos generados. La biología no ha escapado de esta llamada crisis de reproducibilidad científica, que afecta la credibilidad de nuestra comunidad. Por suerte, en respuesta a esta crisis se ha desarrollado una multitud de herramientas, en la mayoría de los casos libremente disponibles, para aumentar la transparencia y accesibilidad de datos que respaldan las conclusiones de estudios científicos. Este curso pretende profundizar en el uso de estas herramientas así como de buenas prácticas a lo largo de los procesos de investigación en biología que permitan garantizar transparencia, accesibilidad y reproducibilidad de la producción científica. Consideramos que este curso proveerá beneficios tanto para la comunidad científica, al promover altos estándares de calidad, como beneficios personales para las y los estudiantes, al prepararles para un mercado laboral que empieza a privilegiar la capacidad de desarrollar ciencia abierta, colaborativa y reproducible.

Objetivo general

Familiarizar a las y los estudiantes con herramientas y buenas prácticas para garantizar transparencia, accesibilidad de datos y reproducibilidad en la investigación científica biológica.

Objetivos específicos

1. Dar a conocer la problemática actual sobre reproducibilidad científica.

2. Brindar a los estudiantes herramientas computacionales de software libre que faciliten la documentación y accesibilidad de la investigación.
3. Proveer a los estudiantes con experiencia en la aplicación de las herramientas brindadas por medio de prácticas y proyectos individuales.

Contenido y cronograma*

Semana	Tema
1 (13 de agosto)	BW y MA[†]: Introducción al curso BW: Reproducibilidad y otras propiedades deseables en investigación. ¿Por qué importa la reproducibilidad? MA: Introducción a git-github, Repaso de R
2 (20 de agosto)	BW y MA: Introducción a ciencia abierta BW: Publicaciones, datos, recursos educativos, métodos, software, y revisión abiertas. Licencias.
3 (27 de agosto)	BW: Antes de los datos. Causación y modelos gráficos. Predicciones. Consideraciones de diseño. Datos ciegos. Pre-registro de experimentos.
4 (03 de setiembre)	MA: Programación como herramienta indispensable para la reproducibilidad (con énfasis en R). Ventajas de automatización vs “a mano”. for loops, if/else, gsub, grep. Ambientes reproducibles.
5 (10 de setiembre)	MA: Manejo y estructura de bases de datos . Datos compartibles. Formatos de archivo, “Non-data content”, README, tidy data.
6 (17 de setiembre)	BW y MA: Selección de artículo para reproducción. Idea general del proyecto final. BW: Precisión taxonómica . Bases de datos en línea. Taxize. ¿Microbios? vistazo a QIIME 2.
07 (24 de setiembre)	MA y BW: Buenas prácticas en la escritura y documentación de código . Generalidad. Claridad vs eficiencia. Sangrías, paréntesis. Nombres. Funciones. Comentarios. Lintr. Organización de scripts.
8 (01 de octubre)	MA: Control de versiones y colaboración usando git
9 (08 de octubre)	MA: Control de versiones y colaboración usando github
10 (15 de octubre)	MA: Reportes dinámicos (rmarkdown)
11 (22 de octubre)	MA: Gráficos dinámicos MA y BW: Avance de proyecto final
12 (29 de octubre)	BW: Trampas estadísticas y cómo detectarlas. Replicación y pseudoreplicación. Valores de P, Tamaño de efecto. Supuestos. Contrastes apropiados e inapropiados. ¿Por qué considerar estadística Bayesiana?
13 (05 de noviembre)	BW: Redacción para ciencia reproducible. Métodos y Resultados. Citar correctamente. BW y MA: Presentación de artículos replicados I
14	BW: Contribución a repositorios de datos. Principios y

Semana	Tema
(12 de noviembre)	prácticas FAIR- Convenciones de nomenclatura. Respaldos. Documentación y metadatos. MA: Selección de repositorios y ejemplos. BW y MA: Presentación de artículos replicados II
15 (19 de noviembre)	MA y BW: Uso de repositorios de datos en investigación. BW y MA: Presentación de artículos replicados III
16 (26 de noviembre)	BW: Del capítulo de tesis a la publicación científica. Checklist de reproducibilidad. Preprints. Revisión por pares. Cartas de presentación. Tiempos y expectativas. Retracción.
17 (4 de diciembre)	BW y MA: Entrega de proyectos finales
18 (11 de diciembre)	Ampliación

* El cronograma es una guía para las actividades del curso. Está sujetos a cambios en el orden de los contenidos y es posible que se incluyan contenidos adicionales.

†Indica las iniciales del profe principal a cargo del tema. **BW** = Beatriz Willink, **MA** = Marcelo Araya.

Metodología y actividades para cumplir con los objetivos

Este curso combina presentaciones a cargo de los profesores, discusiones con los participantes y actividades prácticas. Cada semana se cubrirá un tema a cargo de uno o ambos profesores de curso. El o la profesora presentará el tema y su contexto y se llevarán a cabo una o más actividades prácticas para ilustrarlo. Las actividades pueden ser simultáneas a la presentación, pueden alternarse con las presentaciones o pueden hacerse al final, dependiendo del caso. En todas las sesiones, la resolución de las actividades será evaluada como parte del curso. Se promueve el diálogo horizontal y la colaboración entre participantes, pero en la mayoría de los casos, cada estudiante debe entregar su propia solución de las actividades, a más tardar una semana después de que fueran asignadas.

Cada semana también se discutirán artículos científicos y artículos tipo opinión. Una o un estudiante estará a cargo de liderar cada discusión. La responsabilidad de liderar la discusión será asignada de manera voluntaria o aleatoria.

Las y los estudiantes llevarán a cabo dos proyectos independientes. El primero es la reproducción total o parcial de los análisis de un artículo científico de su interés. El o la estudiante debe escoger un artículo con un nivel de complejidad adecuado con ayuda de los profesores. Además debe redactar sus conclusiones sobre su experiencia, indicando las dificultades que tuvo al reproducir el estudio y los aspectos del estudio que facilitaron su reproducción. Las y los estudiantes también llevarán a cabo una breve presentación sobre el artículo seleccionado y experiencia reproduciéndolo.

El segundo proyecto individual consiste en aplicar las herramientas de reproducibilidad aprendidas a una parte de su investigación planeada o en curso, o a un pequeño proyecto de investigación planteado para el curso y con datos pre-existentes o simulados. Para este proyecto individual las y los estudiantes deben elaborar una estrategia o plan de reproducibilidad indicando las hipótesis y predicciones de manera gráfica (si aplica), justificando el diseño experimental o de muestreo y aportando información suficiente para acceder y comprender los datos. Además, el proyecto individual incluirá elaborar código apropiadamente anotado para el manejo y análisis de al menos una parte de los datos y un resumen de métodos (del análisis), redactado al estilo de un artículo científico.

El curso será 100% virtual. Se promueve la participación sincrónica, pero todas las clases serán grabadas y quedarán disponible para los y las estudiantes en la plataforma de Mediación Virtual.

Uso de Mediación Virtual (100% Virtual)

Se utilizará la plataforma virtual de la UCR (sistema de mediación virtual) para colocar todo el material de apoyo para el curso, así como información importante para las y los estudiantes. Las y los estudiantes entregarán algunas tareas y su proyecto final por medio de la plataforma. Es indispensable que las y los estudiantes revisen el entorno 2-3 veces por semana. Las clases serán virtuales y por medio de la plataforma zoom accesada a través de mediación virtual.

Evaluación

Ejercicios/Tareas	30%
Discusión de artículos	10%
Reproducción de artículo	20%
Presentación artículo	10%
Proyecto final	30%

Bibliografía

El curso no cuenta con un libro de texto, sino que se leerán artículos científicos y divulgativos relevantes a cada uno de los temas de manera previa a la discusión.

Nota

Según CIRCULAR-CUSED-025-2009, en acuerdos de la sesión 8-2009, artículo 6:

a. El período de tiempo razonable para guardar los trabajos y exámenes de los estudiantes posterior a la conclusión del ciclo lectivo es de seis meses, concluido este tiempo se pueden eliminar. Esta circunstancia deberá ser comunicada a los estudiantes al inicio de lección por medio del programa del curso o carta al estudiante.