



Proyecto Especial Ciencia Abierta 2025

Para la continuación en 2025 de la iniciativa del Proyecto Especial “Ciencia Abierta Inecol” proponemos continuar bajo la premisa de la Ciencia como un Derecho Humano. El planteamiento que hemos venido impulsando busca abrir todo el quehacer científico al público en general. Propicia desarrollar prácticas que involucren la publicación sin fronteras de pago, compartir los datos de investigación y los códigos computacionales empleados en su procesamiento. También se interesa en fomentar la participación ciudadana en proyectos de interés general. En esta tercera fase que inicia en 2025, el proyecto se enfocará en realizar un diagnóstico profundo de las prácticas y el estado de las publicaciones académicas del Inecol, con el ánimo de identificar oportunidades para alinearlas con los valores de la ciencia abierta. Esto implica realizar también un diagnóstico general de las publicaciones académicas a nivel nacional. Por otro lado, se finalizará la configuración de los servicios informáticos en el servidor del Inecol, para ponerlos a disposición de todo el personal del instituto. Continuaremos con los eventos de reflexión en torno a la ciencia abierta y ensayaremos la producción de algunos “documentos impulsados por datos” en el formato de “tableros de control” que permitirán a los usuarios interactuar con datos de interés para propiciar acciones de sustentabilidad, un usuario interesado en este momento es **SEDEMA Veracruz**, por lo que nos proponemos ensayar con ellos este proceso con un enfoque de coconstrucción de este enfoque.

Actividades propuestas

Diagnóstico del proceso editorial de las revistas del Inecol con perspectiva de Ciencia Abierta

Las revistas del Inecol se encuentran todas publicadas en acceso libre total: sin cargos a los autores ni cargos por descarga de los artículos, lo cual es una gran ventaja para el proyecto; además, las revistas utilizan el sistema de publicación Open Journal System, el cual es software de código abierto. Después de un acercamiento y una valoración inicial, proponemos hacer un diagnóstico enfocado en los siguientes objetivos:

1. Conocer el estado de las revistas en cuanto a la reproducibilidad de las publicaciones; es decir, generar estadísticas sobre la práctica actual de publicar los datos en conjunto con



descripciones claras de los procesos y los códigos informáticos utilizados para generar los resultados, además de estadísticas de qué tanto interés genera publicar este tipo de contenido a quienes envían sus trabajos a estas revistas.

2. Identificar puntos de mejora y aeleración del proceso editorial a través de técnicas computacionales: la utilización de herramientas de edición de texto basadas en marcado, tales como Quarto o Pandoc, pueden producir archivos en el formato JATS (Journal Article Tag Suite), el cual es utilizado por las revistas del Inecol para enviar sus artículos a Scielo; el formato JATS de Scielo ha sido ampliado para incluir más tipos de etiquetas, entonces como parte del diagnóstico se revisaría cuáles etiquetas coinciden en el JATS estándar producido por Quarto y el de Scielo y se harían ajustes al código en Quarto para incluir aquellas faltantes.

Diagnóstico de las prácticas de publicación en Gold Open Access

A raíz de la Iniciativa de Acceso Abierto de Budapest, en la cual se manifestaba la necesidad de eliminar los costos por descarga y los tiempos de espera o embargo para que los artículos estuvieran al alcance de todo público, las casas editoriales Wiley, Elsevier, MDPI, Springer, Sage, Taylor & Francias (WEMSST), entre otras, iniciaron una transición hacia esquemas en los que los autores son los que pagan los Cargos de Proceso de Artículo (Article Processing Charges), los cuales a menudo son elevados (más de 4,000 USD) y han llevado a las revistas a la práctica de relajar sus exigencias científicas hacia los autores, con tal de obtener mayores recursos, incurriendo a menudo en prácticas predatorias. El movimiento Open Access ha tenido como consecuencia final que las revistas WEMSST obtengan ganancias millonarias. Actualmente existe un movimiento incipiente de universidades que cancelan sus suscripciones a este tipo de editoriales, sobre todo en **Europa**.

Esta fase del proyecto tiene el interés de generar estadísticas de cuánto dinero se ha destinado a pagar publicaciones en Gold Open Access, paralelamente con un modelación de cómo incrementarían los factores de impacto de las revistas nacionales si los académicos voluntariamente adoptaran el compromiso de preferir publicar en revistas nacionales y dejar de citar o citar un mínimo de referencias de revistas WEMSST; esto como un resultado paralelo de redistribuir los recursos actualmente destinados a pagar publicaciones WEMSST hacia personal que haga trabajo editorial en las revistas nacionales y una consecuente aceleración de los procesos editoriales, considerando que los largos tiempos de espera de las revistas nacionales son una razón principal por la cual los investigadores no envían artículos a estas revistas.



Diagnóstico de uso de inteligencia artificial en el Inecol

La inteligencia artificial (IA) ha dado un salto cualitativo de popularidad en los últimos años, con la llegada de la IA generativa; sin embargo, desde hace cerca de diez años existen modelos de código abierto muy avanzados para tareas que pueden ser de interés para el personal del Inecol, aunque la mayoría de estos modelos son desconocidos en la comunidad, y existe la creencia generalizada de que la IA sólo existe como software de paga o la creencia errónea que aquella de paga es mejor que la de acceso libre. La aparición de la familia de modelos DeepSeek ha puesto de manifiesto que la IA de acceso libre está avanzando más rápidamente que la de paga, además de mostrar la importancia de tener desarrollos nacionales en IA y de su aplicabilidad a la ciencia.

Como parte de esta fase del proyecto, se pretende hacer un diagnóstico del uso de IA en el Inecol, el cual podría extenderse a la Universidad Veracruzana, para saber el grado de conocimiento y deseabilidad de desarrollo de capacidades en tareas de IA tales como: detección de objetos, clasificación de imágenes y videos, segmentación de imágenes, procesamiento de lenguaje natural, clasificación de textos y audios, reconocimiento automático del habla, entre otras.

Configuración de servicios informáticos para impulsar la Ciencia Abierta en el Inecol

Desde la primer fase del proyecto, se identificó la importancia de contar con aplicaciones para facilitar la Ciencia Abierta, en todo el proceso de la creación científica, desde la toma de datos hasta la publicación accesible de datos, códigos informáticos utilizados y visualización comunicativa de la información generada. En la primer fase se hizo la desconexión lógica de un servidor parte del sistema de cómputo de alto rendimiento Huitzilin del Inecol. En esta fase se finalizaría la instalación de estos servicios informáticos, todos de acceso libre, que son: OpenDataKit, para creación de formularios móviles para recolección de datos en campo, Source-forge como repositorio de código, Geoserver para publicación de datos geoespaciales, Dataverse para publicación de datos no geoespaciales y Shiny-server para creación de publicaciones interactivas.



Coconstrucción de documentos impulsados por datos

Se buscará desarrollar un programa que combine la capacitación en el uso de Quarto para la producción de “documentos impulsados por datos” que tiendan puentes entre distintos tipos de actores. Como oportunidad de prototipo nos proponemos desarrollar un piloto con SEDEMA Veracruz. Requeriremos la contratación una persona que nos ayude con la capacitación y conducción del proceso de desarrollo computacional que requiere esta aproximación.



Presupuesto de trabajo

- *Consultoría para diagnóstico y modelado de escenarios de migración a publicaciones de acceso libre en las revistas académicas:* Se requiere una persona, por un periodo de 4 meses, quien se encargará de recopilar la información relevante y de hacer un modelado estadístico de factor de impacto en las revistas del Inecol en distintos niveles de citación de revistas WEMSST vs revistas nacionales. Presupuesto requerido: \$50,000.00 + IVA.
- *Consultoría para diagnóstico sobre uso de Inteligencia Artificial en el Inecol:* Se requiere una consultoría, por un periodo de 4 meses, para diseñar un instrumento de entrevistas donde se recopile el perfil académico, uso actual e intereses sobre diversas tareas de IA del personal del Inecol, así como un análisis estadístico exploratorio de la información generada. Presupuesto requerido: \$50,000.00 + IVA.
- *Consultoría para diagnóstico de los procesos editoriales del Inecol hacia la Ciencia Abierta:* Se requiere una persona, por un periodo de 4 meses, para recabar información de las revistas del Inecol sobre las prácticas de publicación de datos y códigos informáticos que permitan la reproducibilidad, así como una identificación de cuáles actividades serían necesarias para fomentar estas prácticas, así como de prácticas de escritura usando herramientas que faciliten el etiquetado JATS para los editores de las revistas. Presupuesto requerido: \$50,000.00 + IVA.
- *Coconstrucción de documentos impulsados por datos:* Se prevé la contratación de una persona que nos ayude con la capacitación y conducción del proceso de desarrollo computacional que requiere esta aproximación. Presupuesto requerido: \$50,000.00 + IVA.
- *Consultoría para instalación de servicios computacionales para Ciencia Abierta:* Se requiere una persona, por un periodo de 4 meses, que instale las aplicaciones descritas anteriormente en el servidor dedicado del proyecto. Presupuesto requerido: \$50,000.00 + IVA.
- *Consultoría administrativa:* Para llevar a cabo en tiempo y forma las contrataciones, llenado de formatos, facturación y demás partes del proceso administrativo, se requiere una persona. Presupuesto requerido: \$40,000.00 + IVA.



Reumen presupuestal (montos + IVA, pesos mn)

Actividad	Monto
Diagnóstico sobre migración a revistas nacionales de acceso libre	\$58,000.00
Diagnóstico sobre uso de IA en Inecol	\$58,000.00
Diagnóstico procesos editoriales	\$58,000.00
Coconstrucción de documentos impulsados por datos	\$58,000.00
Consultoría para instalación de servicios computacionales para Ciencia Abierta	\$58,000.00
Consultoría administrativa	\$46,400.00
Total	\$336,400.00

