

Conjuntos de datos inclusivos para el aprendizaje automático

Resistencia antimicrobiana

Seminario web para solicitantes

10 de julio de 2024

Ponentes

MERCY AYUB, ACTS

ADELA BRIANSÓ JUNQUERA, LACUNA FUND

SAMUEL WANJAU, ACTS

DR. OJENGE WINSTON, ACTS

Dra. ELIZABETH MUTUA, ASESORA DE CALIDAD DE DATOS DE ACTS



Objetivos de la sesión

- Proporcionar a los posibles solicitantes información sobre Lacuna Fund y los requisitos de las propuestas.
- Tiempo para preguntas de los solicitantes sobre la solicitud de propuestas (RFP).



Orden del día

- 00:00** Bienvenida, presentación, orden del día
- 00:10** Presentación: Visión general del Fondo Lacuna y requisitos para las solicitudes de propuestas; Elementos de una propuesta sólida
- 00:40** Presentación: Calidad de los datos y consideraciones sobre el hosting
- 00:50** Presentación: Oportunidad de tutoría
- 01:00** Preguntas y respuestas
- 01:25** Próximos pasos
- 01:30** Fin de la sesión

Una 'lacuna' es un hueco o una parte que falta en un manuscrito.

Lacuna Fund apoya la creación, ampliación y mantenimiento de conjuntos de datos de formación y evaluación equitativos que permitan a las herramientas de aprendizaje automático abordar mejor problemas urgentes en contextos de renta baja y media en todo el mundo.

Principios que guían Lacuna Fund:

- Accesibilidad
- Enfoque participativo
- Equidad
- Calidad
- Ética
- Impacto transformador



Financiadores



Financiadores adicionales



Fondo Lacuna

Estructura de gobierno

Comité de Dirección

Orientación estratégica y gobernanza general del Fondo

Financiador Colaborador

Proporciona información sobre las prioridades clave al Comité Directivo y constituye un foro para una mayor colaboración entre financiadores.

Grupos de asesoramiento técnico

Analizar los procesos de financiación y seleccionar las propuestas, aportar información sobre las estrategias y necesidades específicas de cada ámbito.

Secretaría

Gestiona las subvenciones y los informes, proporciona facilitación, comunicaciones y apoyo operativo.



Dominios para conjuntos de datos



El Fondo Lacuna proporciona recursos para la creación, ampliación o mantenimiento de conjuntos de datos.

Convocatorias de propuestas actuales en estos ámbitos:

- **AMR**
- **NLP**

Convocatorias de propuestas anteriores en estos ámbitos:

- Agricultura
- Idioma
- Salud
- Clima

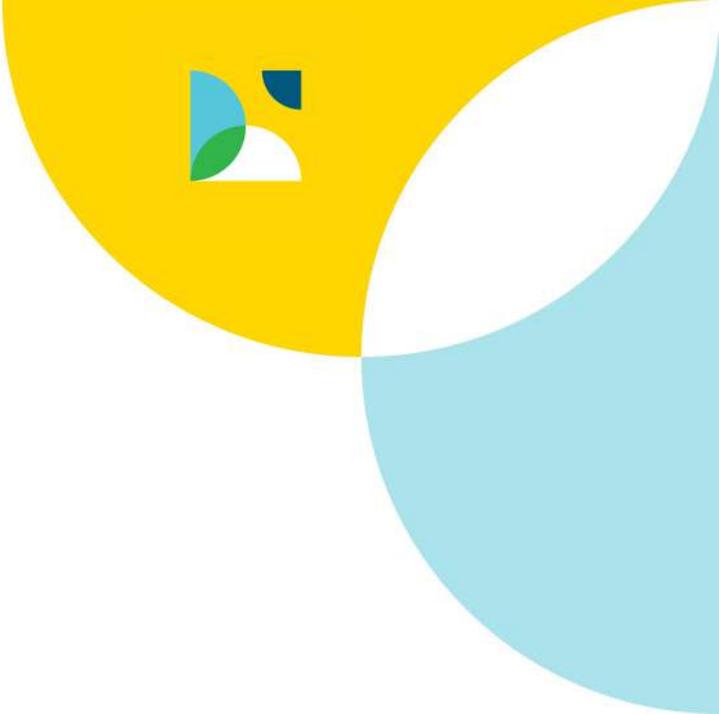
Futuras convocatorias de propuestas:

- Otros ámbitos en los que el Comité de Orientación identifica una necesidad

Financiación de esta llamada a propuestas

La convocatoria de propuestas AMR es posible gracias al generoso apoyo de WELLCOME.





Requisitos

Resistencia a los antimicrobianos (AMR)

Resistencia a los antimicrobianos

Objetivo de la llamada

El objetivo de esta convocatoria de propuestas es apoyar los esfuerzos para desarrollar conjuntos de datos abiertos y accesibles para aplicaciones de aprendizaje automático relacionadas con la comprensión de la resistencia a los antimicrobianos (RAM) en los países de renta baja y media (PRMB) de África, América Latina y Asia meridional y sudoriental.



Necesidad

Vemos la necesidad de conjuntos de datos de formación y evaluación que ayuden a comprender mejor las repercusiones de la RAM y a desarrollar intervenciones para mitigarlas. Buscamos conjuntos de datos identificados por expertos locales diseñados para abordar necesidades identificadas localmente, por lo que los siguientes son solo ejemplos ilustrativos.

- Conjuntos de datos para mejorar la prevención y el control de las infecciones (CIP)
- Conjuntos de datos para el descubrimiento de fármacos
- Conjuntos de datos para la vigilancia mundial de la RAM y enfoques anticipatorios
- Conjuntos de datos de apoyo al diagnóstico
- Conjuntos de datos que contienen información sobre el acceso a los antimicrobianos y la falta del mismo
- Conjuntos de datos que mejoran la comprensión biológica básica de los patógenos y la RAM
- En todos los conjuntos de datos, la perspectiva de género y la inclusión de los principales grupos vulnerables, especialmente los más afectados por la RAM.



Qué buscamos

- Recopilación y anotación de nuevos datos
- Anotación y publicación de datos existentes
- Ampliar o mantener los conjuntos de datos existentes y aumentar su utilidad
- Vincular y armonizar los conjuntos de datos existentes



Componentes de la propuesta

- Información sobre el solicitante
- Descripción de la propuesta
- Calendario del proyecto
- Presupuesto y descripción del presupuesto
- Formulario de autoinforme de diligencia debida





Requisitos

Los requisitos específicos del financiador son los siguientes

- Ser una entidad sin ánimo de lucro, una institución de investigación, una empresa social con ánimo de lucro o un equipo de estas organizaciones. Las personas físicas deben presentar su solicitud a través de un patrocinador institucional. Se recomienda encarecidamente la asociación para reforzar la colaboración y maximizar los beneficios derivados del uso de los conjuntos de datos, pero sólo el solicitante principal recibirá fondos.
- La entidad debe tener una misión de apoyo al bien social, definido en sentido amplio.



Requisitos

- Tener su sede en el país o región donde se recopilarán los datos. El solicitante principal debe tener su sede en un país de renta baja o media, y debe proponer trabajar en el mismo país, o en un país de renta baja o media de África, América Latina o Asia Meridional o Sudoriental. La lista de países elegibles puede consultarse en: <https://wellcome.org/grant-funding/guidance/low-and-middle-income-countries>
- Las instituciones con sede en otros países o regiones pueden presentar su candidatura como socios de la institución principal. Como ya se ha indicado, solo el solicitante principal recibirá fondos.

Requisitos

- Contar con todas las autorizaciones nacionales o de otro tipo necesarias para llevar a cabo la investigación propuesta. En caso necesario, el proceso de aprobación podrá realizarse paralelamente a la solicitud de subvención. Los costes de aprobación, en su caso, correrán a cargo del solicitante.
- Tener la capacidad técnica -o la capacidad de crear esta capacidad a través de una asociación descrita en la propuesta- para llevar a cabo el etiquetado, la creación, la agregación, la expansión y/o el mantenimiento del conjunto de datos, incluida la capacidad de aplicar las mejores prácticas y las normas establecidas en materia de resistencia a los antimicrobianos para permitir que múltiples entidades realicen análisis de IA/ML de alta calidad.

Fechas clave

27 de junio: Solicitud de propuestas abierta - [Presente su candidatura - Fondo Lacuna](#)

12 de julio: plazo de preguntas

Envíe sus preguntas a secretariat@lacunafund.org

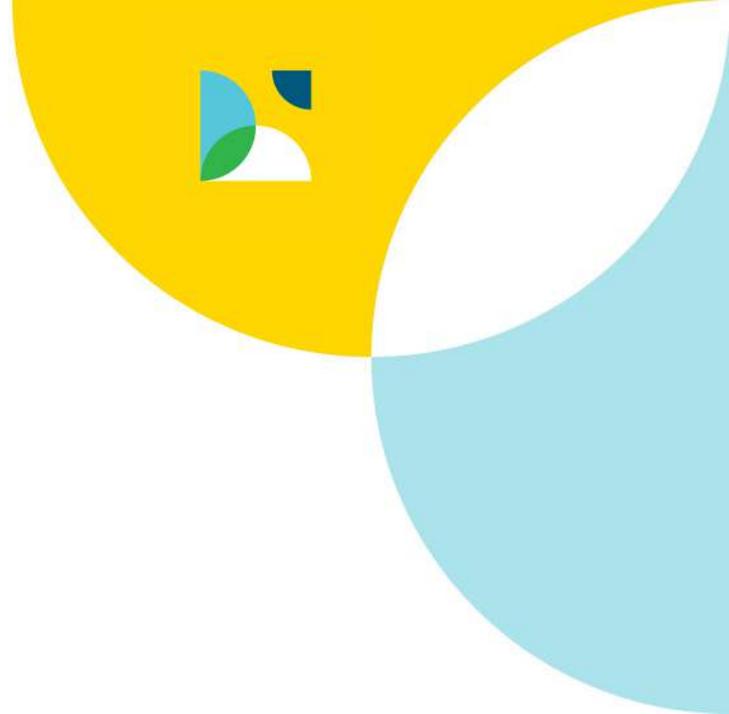
29 de julio: Respuestas publicadas

26 de julio: Última oportunidad para solicitar una tutoría

23 de agosto: Fecha limite para enviar propuestas

- **Noviembre:** La Junta de Asesoramiento Técnico decide sobre la concesión de subvenciones
- **Diciembre:** Notificación a los beneficiarios
- **Mayo:** Anuncios de premios

Elementos de una propuesta sólida



Elementos de una propuesta sólida

- **Equipo multidisciplinar** - experiencia en ciencia de datos y aprendizaje automático, RAM
- **Casos prácticos y participación ciudadana**
- **Planteamiento** claro **del** problema y cómo el conjunto de datos o la agregación propuestos ayudarán a resolverlo.
- Especificidad sobre el tamaño del conjunto de datos: debe tener un **tamaño y una calidad suficientes** para ser útil en futuras aplicaciones.
- **El presupuesto** es adecuado para el tamaño del conjunto de datos producido



Elementos de una propuesta sólida



- Consideraciones de **equidad** (género, estatus socioeconómico, etnia, etc.)
- Consideración y plan de posibles problemas de **privacidad y ética**
- Plan de **gestión de datos y licencias**
- Trabajar en todas las áreas en las que sea posible y pertinente: tiempo, regiones, patógenos, resultados de los pacientes...
- **El presupuesto** es adecuado para el tamaño del conjunto de datos producido

Participación comunitaria

- Describa las consultas previas y/o la colaboración propuesta con los beneficiarios/as previstos.
 - Cuándo y dónde se ha reunido/se reunirá con los socios/as
 - Cómo participarán los socios/as en:
 - Determinación de las necesidades de datos
 - Recogida de datos, etiquetado
 - Gobernanza de datos
 - Uso del conjunto de datos
 - Cómo se beneficiarán los socios/as del conjunto de datos nuevo/ampliado

Privacidad y ética

Explique cómo abordará su equipo:

- a) cuestiones de privacidad,
- b) potencial de uso indebido posterior,
- c) posibles vectores de discriminación (por ejemplo, el género), y
- d) condiciones de trabajo justas y equitativas, si en el proyecto participan etiquetadores remunerados.

Describa el proceso que utilizará para detectar posibles problemas éticos (por ejemplo, una junta de revisión institucional, etc.).

Normas e intercambio de datos

- Localizable: fácil de encontrar en una plataforma pública y ampliamente utilizada.
- Accesible - formato abierto (CCBY 4.0)
- Interoperabilidad - formato de los datos
- Reutilizables - metadatos
- Sostenible: plan de mantenimiento
- Compartidos: comunidad de usuarios comprometida y plan para compartir el conjunto de datos una vez completado.

<https://www.go-fair.org/fair-principles/>

Presupuesto y costes autorizados

- Debe presentarse en la plantilla Excel.
- En caso de duda, ¡explíquese!
- Los costes permitidos incluyen:
 - Formación/creación de capacidad;
 - Recogida de datos; etiquetado; QA/QC o verificación; y post-procesamiento;
 - Modelo de referencia;
 - Concesión de licencias;
 - Alojamiento del conjunto de datos;
 - Publicación de los resultados en acceso abierto;
 - Metadatos o declaración de datos;
 - Esfuerzos de crowdsourcing, como los maratones de etiquetado.



Calidad de los datos

- Profiling de datos: consiste en analizar los datos para obtener información, incluidos los tipos de datos, las distribuciones, los valores que faltan, los valores atípicos y las posibles incoherencias.
- Proceso de depuración de datos: consiste en tratar los valores que faltan, corregir errores, eliminar valores atípicos, resolver incoherencias y normalizar los formatos de los datos.
- Exactitud de los datos: verificación de la exactitud de los datos comparándolos con fuentes de confianza o conocimientos de expertos.
- Completitud de los datos - Garantizar que el conjunto de datos contiene todos los elementos necesarios para la tarea de aprendizaje automático.

Consideraciones sobre la calidad de los datos

1. Alojamiento y accesibilidad

- Cómo y dónde alojamos los datos es realmente importante.
- La aceptación de un conjunto de datos puede depender de cómo y dónde se aloje.

¿Dónde está alojado actualmente su conjunto de datos? ¿En qué se basó esa decisión?

Alojamiento y accesibilidad

¿Qué hay que tener en cuenta a la hora de elegir una plataforma de alojamiento?

- Dominio
- Caso práctico
- Coste

Política del Fondo Lacuna:

Flexible para permitir la propiedad local de los datos y responder a los casos de uso previstos para los datos. Guiados por los principios del Fondo Lacuna

Alojamiento y accesibilidad - Principios Lacuna



| Accessibility

| Equity

| Ethics

| Participatory Approach

| Quality

| Transformational Impact

2. Documentación del conjunto de datos

"En la industria electrónica, es norma acompañar cada componente de una hoja de datos que proporcione características de funcionamiento estándar, resultados de pruebas, uso recomendado y otra información.

Del mismo modo, recomendamos que cada conjunto de datos vaya acompañado de una ficha técnica que documente su creación, composición, usos previstos, mantenimiento y otras propiedades."

Ventajas de la documentación de conjuntos de datos

- Reproducibilidad
- Transparencia
- Usabilidad

La documentación de los conjuntos de datos mostrará el contexto, la estructura y el contenido



Documentación de conjuntos de datos - Directrices Lacuna

- Ficha de datos
- Fichero de metadatos
- Identificador de objeto digital (DOI)

3. Licencias - Directrices Lacuna

- Política de propiedad intelectual de Lacuna. Los conjuntos de datos y la propiedad intelectual relacionada desarrollados con fondos de la subvención:
- Ser propiedad de la entidad subdonataria
- Ser objeto de licencia abierta por parte de la entidad subcesionaria para maximizar el potencial de liberación responsable y uso posterior de dicha propiedad intelectual.



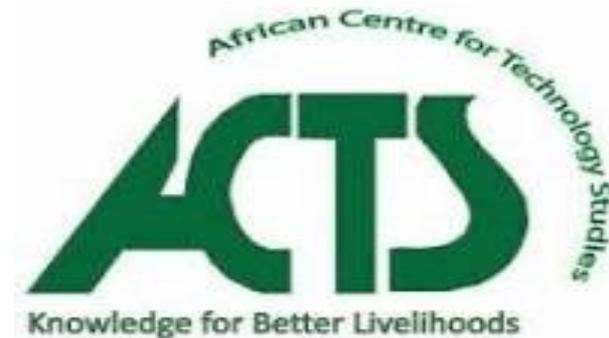


Licencias... ¿qué cubre?

- Metadatos, fichas técnicas u otra información.
- Con el conjunto de datos se incluyen orientaciones para guiar su uso.
- Modelos de muestra u otras herramientas informativas publicadas con los datos.

Oportunidad de tutoría: Programa de tutoría ACTS

Centro Africano de Estudios Tecnológicos
(ACTS)



Programa de tutoría ACTS

- **Misión:** La misión de ACTS es reforzar la capacidad y las políticas de los países e instituciones africanos para aprovechar la ciencia, la tecnología y la innovación en favor del desarrollo sostenible. El programa de tutoría ayudará a los solicitantes a desarrollar solicitudes de calidad y les guiará para que comprendan y sigan las mejores prácticas en la ejecución de sus proyectos.
- Con carácter **excepcional** para esta convocatoria del Fondo Lacuna, ACTS ofrecerá a los solicitantes de todo el mundo una oportunidad especial de unirse al programa de mentores, en el que los solicitantes y los equipos serán emparejados con un mentor.

Programa de tutoría ACTS

- Presente su candidatura en **el sitio web de ACTS**
[Programa de tutoría de ACTS](#)
- Seleccione "**Solicitar tutoría**" o "**Solicitar ser tutor**".
- Ambos formularios tienen campos similares en los que se le pedirá que facilite su información personal, zona horaria, su idioma, área de especialización, área temática de la propuesta
- Le animamos a que prepare un esquema detallado o un borrador de propuesta antes de solicitar la tutoría, lo que le permitirá conocer su área temática
- Presente su solicitud al **menos 4 semanas antes de la fecha límite, antes del 26 de julio de 2024**
- Los mentores se asignarán por orden de **llegada**.
- Todos los solicitantes deben **leer y respetar el código ético y de conducta del programa de tutoría**.

Código de conducta ACTS

- No causar daño
- Consentimiento libre, previo e informado
- Confidencialidad
- Seguridad
- Sin coacción
- Participación de niños (menores de 18 años)
- Género
- Inclusividad

Programa de tutoría

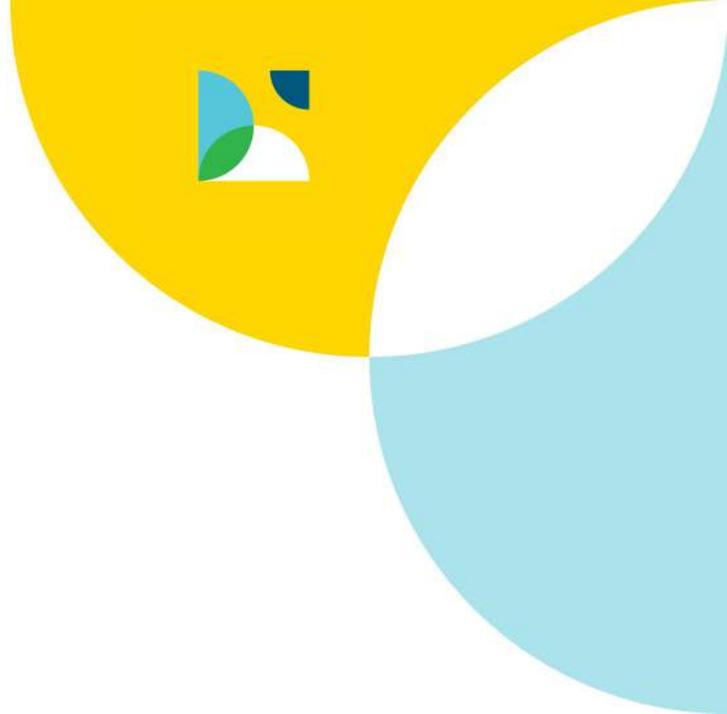
ACTS



Funciones del tutor

- Proporcionar tutoría y orientar a los candidatos para que comprendan el objetivo de la convocatoria
- Establecer relaciones con los alumnos y comprender sus puntos débiles y fuertes
- Ayudar a los alumnos a revisar las propuestas en función de sus necesidades.
- Orientar a los alumnos sobre las funciones del grupo y los conflictos que puedan surgir.
- Asistir a las sesiones o debates del grupo de la propuesta o reunirse por separado con los investigadores principales.
- Discutir trabajos relevantes en el área de la propuesta para garantizar que la propuesta esté contextualizada.

¿Preguntas?



Próximos pasos

- Envíe preguntas adicionales a secretariat@lacunafund.org antes del 12 de julio
- **Respuestas** publicadas en la página de solicitud de Lacuna Fund el **29 de julio**
- **Solicite la tutoría lo antes posible y antes del 26 de julio**
- **Propuestas hasta el 23 de agosto**