



# AI4ES

RED DE EXCELENCIA EN  
TECNOLOGÍAS HABILITADORAS  
BASADAS EN EL DATO

## **BASES DEL DATATHON "AI4ES – RED DE EXCELENCIA EN TECNOLOGÍAS HABILITADORAS BASADAS EN EL DATO" 2022**

La Red de Excelencia Cervera AI4ES (red de excelencia en tecnologías habilitadoras basadas en el dato), integrada por los centros tecnológicos CTIC, EURECAT, ITI y TECNALIA, en colaboración con el Departament de Salut de la Generalitat de Catalunya y la empresa BASF, convoca los premios del Datathon AI4ES 2022 con arreglo a las siguientes BASES:

### **PRIMERA. - MOTIVO DEL DATATHON Y LOS RETOS PLANTEADOS**

El Datathon AI4ES tiene como objetivo principal incentivar la utilización de tecnologías de inteligencia artificial (IA) y analítica de datos para resolver retos de importancia empresarial y social, que supongan el procesado de grandes volúmenes de datos y reconocer los mejores proyectos presentados para la resolución de dichos retos. Para ello, se convoca el premio AI4ES 2022 para reconocer a los mejores estudiantes y profesionales en el uso de tecnologías de IA para procesamiento de datos. En su edición de 2022, el Datathon AI4ES plantea dos retos diferentes:

- ***Reto 1: Detección y cuantificación de enfermedades basadas en imágenes de drones para la protección de cultivos***
- ***Reto 2: Incidencia y evolución enfermedades víricas***

La descripción detallada de estos retos se encuentra en la sección QUINTA de estas bases.

Los proyectos propuestos por los participantes del Datathon AI4ES se centrarán en la resolución de alguno de los dos retos definidos anteriormente.

## **SEGUNDA. - PARTICIPANTES**

Podrán participar todas aquellas personas físicas mayores de edad que lo deseen, tanto de forma individual como en equipo. Cada premio se otorgará a cada proyecto con independencia del número de participantes en su elaboración.

## **TERCERA. - PLAZO DE PRESENTACIÓN**

El plazo de presentación de solicitudes se iniciará desde el día de publicación de estas bases en la web de AI4ES y finalizará a las **23.59 horas del 25 de noviembre de 2022**, a través del formulario habilitado en la página web, <https://www.ai4es.com/ai4es-datathon/>. La organización se reserva el derecho a extender la fecha de cierre de las inscripciones para favorecer una mayor participación.

## **CUARTA. - CATEGORÍAS Y PREMIOS CONVOCADOS**

Existen tres categorías de participación.

- a) **Categoría de iniciación (“Futuras promesas”)**, para la presentación de trabajos por estudiantes y graduados noveles (cuando todos los participantes en el proyecto cumplan que: o bien estén cursando un grado actualmente, o bien hayan obtenido su grado con posterioridad al 31/12/2018).
- b) **Categoría de especialización (“Especialistas”)**, para la presentación de trabajos por graduados y profesionales de empresas con más de 3 años de experiencia profesional (esto es, cuando no se cumplan las condiciones de la categoría de iniciación)
- c) **Categoría AI4ES (“Miembros AI4ES”)**, para la presentación de trabajos por investigadores de los centros integrantes de la red AI4ES (cuando el investigador o algún miembro del equipo forme parte de uno de los centros integrantes de la red Ai4ES).

Se otorgará un premio de 1.000 euros al mejor proyecto de cada reto de la categoría de iniciación, y 2.000 euros al mejor proyecto de cada reto de la categoría de especialización.

Cada uno de los premios se repartirá, a partes iguales, entre todas las personas del grupo del proyecto que resulte premiado. Cada uno de los miembros de los equipos de los proyectos premiados deberán aportar la documentación personal y fiscal necesaria para que la red AI4ES pueda entregar los premios correspondientes. Se aplicarán las retenciones de rentas que, en su caso, establezca la legislación vigente.

Todo proyecto, así como la presentación del mismo, con independencia de si ha sido premiado o no, podrá ser comunicado mediante los medios que la red AI4ES estime oportuno con el fin de conseguir la máxima difusión y poder darlo a conocer.

## QUINTA. – RETOS PLANTEADOS

### RETO 1 - DETECCIÓN Y CUANTIFICACIÓN DE ENFERMEDADES BASADAS EN IMÁGENES DE DRONES PARA LA PROTECCIÓN DE CULTIVOS

La división Soluciones Agrícolas de BASF realiza cientos de pruebas de campo cada año para evaluar el rendimiento biológico de los nuevos productos para la protección de cultivos. Actualmente, la mayoría de los datos se generan mediante evaluaciones visuales realizadas por agrónomos expertos en investigación de campo. Por ejemplo, en los ensayos de fungicidas, el agrónomo de campo evalúa en cada parcela el porcentaje de cultivos infectados por cada enfermedad.

Para ser más eficiente y mejorar la calidad de los datos, BASF está digitalizando y modernizando este proceso de evaluación, por ejemplo, utilizando dispositivos portátiles (teléfonos inteligentes) y drones.

El objetivo de este reto es entrenar un modelo robusto de regresión de imágenes, que procese imágenes de parcelas de campo individuales y prediga su valor de evaluación.

Para este reto, se cuenta con un conjunto de datos con imágenes multiespectrales captadas por drones y su *ground-truth* (evaluaciones visuales realizadas por el agrónomo de campo) que cubren diferentes años de ensayos en campos de trigo de varias localidades, con diferentes enfermedades y niveles de enfermedad.

Los aspectos a tener en cuenta dentro de este reto son:

- 1) Según el ensayo y su localización, se utilizan diferentes sensores con una resolución de imagen variable. Algunas pruebas se capturan con sensores de 10 bandas (incluidos los canales infrarrojos y la banda de borde rojo), otras se capturan con sensores de 5 bandas y, para algunas pruebas, solo están disponibles las imágenes RGB estándar.
- 2) Las etiquetas, que representan el "ground-truth", son de hecho estimaciones humanas ruidosas del estado real de la enfermedad. Por lo tanto, puede haber valores atípicos o errores en diversos grados en los valores de evaluación proporcionados.
- 3) Los datos proporcionados no están equilibrados, las calificaciones de evaluación más comunes están cerca de 0 y los valores de calificación más altos de 50 a 100 son menos frecuentes en el conjunto de datos.

Además, algunas parcelas se ven afectadas por una enfermedad, y otras parcelas se ven afectadas por una combinación de enfermedades.

Para este reto, se proporcionará:

- Los datos de entrenamiento, que incluyen imágenes RGB, de 5 bandas y de 10 bandas para parcelas de campo y sus evaluaciones objetivo (etiquetas) en formato CSV.

- Un proceso de entrenamiento de ejemplo, basado en Python y XGBoost, que sirve como punto de partida para la experimentación y el desarrollo, aunque cada participante es libre de utilizarlo o comenzar desde cero su solución.

Los participantes que se inscriban en este reto recibirán un enlace desde el cual descargar y ejecutar el ejemplo de proceso de entrenamiento y el conjunto de datos. Los participantes solo podrán utilizar los datos y el código proporcionados para participar en el Datathon y eliminarán los datos cuando este finalice. BASF podrá proporcionar un permiso de uso de los datos más amplio bajo un acuerdo específico.

### RETO 2 - INCIDENCIA Y EVOLUCIÓN ENFERMEDADES VÍRICAS

Durante la pandemia COVID-19 hemos aprendido como disponer de datos de calidad y a tiempo es clave para tomar las mejores decisiones. Recientemente, el Departament de Salut de la Generalitat de Catalunya ha puesto en marcha el SIVIC (Sistema de Información para la Vigilancia de Infecciones en Catalunya – <http://https://sivic.salut.gencat.cat/>) que actualmente ya proporciona información precisa, completa y en abierto para el seguimiento de las infecciones respiratorias en Atención Primaria y Hospitales, desde datos diagnósticos de pacientes hasta secuenciación de los virus detectados. Actualmente SIVIC cubre las principales infecciones respiratorias, incluyendo por supuesto COVID-19, pero está previsto evolucionar a todas las patologías infecciosas y potencialmente a todas las problemáticas de Salud Pública para las que el acceso de calidad a los datos impacte positivamente en su gestión.

Con los datos abiertos de la evolución de la pandemia COVID-19 en Catalunya, obtenidos a través de la web del SIVIC, el presente reto tiene como objetivo avanzar en la mejora de la creación de modelos predictivos capaces de ajustar/predecir el futuro de la progresión del COVID-19.

Para ello, se propone usar como datos de entrenamiento el conjunto disponible hasta la semana 22 de 2022, y como datos de test el resto del presente año. Este periodo de test (entre las semanas 22 y 38 aproximadamente) es de especial interés, ya que es cuando se produjo el último repunte de casos COVID-19 (la sexta ola), a partir de dicha semana 22.

El objetivo de este reto es por tanto crear modelos capaces de proporcionar 4 tipos de pronósticos para los datos Catalunya:

- Incidencia,
- Ingresos hospitalarios,
- Ingresos en UCI,
- Defunciones.

Las métricas a reportar en el set de test deben incluir al menos MAE y MAPE. Estas serán computadas a 1, 3, y 5 semanas vista.

Ya que todos los datos (incluidos los de test) están públicamente disponibles, se valorarán 5 aspectos además de la efectividad de los modelos en dicho periodo de test, como son: la solidez de la metodología utilizada, la novedad y calidad de la solución, la interpretabilidad a partir de los modelos (justificación del porqué de los resultados reportados por el modelo), la visualización (variedad y calidad) de los resultados, y por último la claridad y objetividad de la presentación.

A este reto no se podrán presentar personas y equipo pertenecientes a SIVIC y su Red de Inteligencia Epidemiológica que ya han tenido exposición previa al uso y aplicación de estos datos.

#### Datasets

Links a los datos abiertos del SIVIC, a la web del SIVIC que muestra el histórico y la evolución de los datos, y más información sobre datos abiertos en Cataluña en:

[https://sivic.salut.gencat.cat/dades\\_obertes](https://sivic.salut.gencat.cat/dades_obertes)

<https://sivic.salut.gencat.cat/>

<https://analisi.transparenciacatalunya.cat/>

#### **SEXTA. - PRESENTACIÓN DE LOS PROYECTOS**

Los participantes deberán enviar por correo electrónico a la dirección [datathon@ai4es.es](mailto:datathon@ai4es.es) un informe resumen del proyecto realizado en formato PDF, con un máximo de 5 páginas, y con anterioridad a las **23:59 horas del día 17 de enero de 2023**. Dicho informe resumen deberá explicar: el enfoque adoptado para la resolución del reto, las características de su desarrollo, y los principales resultados y conclusiones.

Asimismo, cada participante deberá realizar una presentación oral de la solución desarrollada con una duración máxima de 10 minutos, durante una sesión de evaluación en formato online que será convocada entre los días 18 y 20 de enero de 2023, en un slot a determinar para cada participante entre las 9:00 y las 15:00 horas.

#### **SÉPTIMA. - JURADO, FALLO Y PUBLICACIÓN DE LOS PREMIOS**

El jurado estará compuesto por uno o varios equipos de evaluación, compuestos por miembros de la red AI4ES y por miembros del Comité Asesor de la red AI4ES. Cada equipo evaluador estará integrado por cinco personas: una por cada centro tecnológico integrante de la red AI4ES (CTIC, EURECAT, ITI y TECNALIA), y un representante del Comité Asesor de AI4ES. Para las categorías del reto 1 "Detección y cuantificación automatizadas de enfermedades basadas en imágenes de drones para la investigación de protección de cultivos", se contará adicionalmente con un representante de la empresa BASF.

El fallo de los premiados se realizará el **24 de enero de 2023**, tras finalizar las sesiones de evaluación y reuniones de consenso entre los miembros del jurado AI4ES. Los proyectos premiados se harán públicos a través de los canales de comunicación de la red AI4ES. El fallo será también comunicado a los premiados por correo electrónico.

El jurado podrá decidir, en caso de que los proyectos presentados no tengan la suficiente calidad o haya una baja participación, que uno o varios de los premios queden desiertos.

### **OCTAVA- CRITERIOS DE VALORACIÓN**

Los proyectos ganadores se elegirán teniendo en cuenta los siguientes criterios:

1. El impacto e importancia del reto solucionado.
2. La creatividad y la innovación.
3. El uso de los datos.
4. El grado de acabado conseguido.
5. La claridad de la presentación oral.

### **NOVENA. - CONDICIONES DE PARTICIPACIÓN**

La participación en estos premios queda supeditada a las siguientes condiciones:

- Es responsabilidad de las personas participantes asegurar que la propuesta sea original y que solucione uno de los retos propuestos en la sección QUINTA.
- La red AI4ES podrá publicar en todos aquellos medios que estime oportuno los resultados de los proyectos desarrollados y presentados al datathon AI4ES 2022. Además, podrán difundir cualquier información relacionada con los resultados de cada proyecto y las personas que la propusieron para darla a conocer por cualquier medio que estimen conveniente, así como imágenes y vídeos de las y los participantes con fines de divulgación y promoción del evento.

### **DÉCIMA. - PROTECCIÓN DE DATOS**

De conformidad con lo que dispone la legislación vigente en materia de protección de datos, la red AI4ES tratará los datos aportados en la solicitud, así como aquellos contenidos en la documentación que la acompañe, cuya finalidad es la gestión de su inscripción, participación, y resolución de la convocatoria de los Premios del Datathon AI4ES 2022.

La legitimación de este tratamiento se basa en el interés público de la convocatoria y en el consentimiento de la persona interesada en su solicitud de participación en ella. Los datos relativos a los ganadores de los premios se podrán publicar en la web de la red AI4ES y en redes sociales.

Los datos de los premiados en las modalidades que incluyen prestación económica serán comunicados a las entidades bancarias con objeto de realizar su abono, a la

Agencia Estatal de Administración Tributaria, a la Intervención General del Estado y al Tribunal de Cuentas para el control contable.

Los datos serán conservados durante la tramitación del procedimiento de concesión del premio, siendo de aplicación la normativa de archivos y patrimonio documental español. Los datos de los premiados en las modalidades que incluyen prestación económica se conservarán conforme a lo dispuesto en la Ley 58/2003, de 17 de diciembre, General Tributaria.

Le informamos que puede ejercer sus derechos en relación con el tratamiento de sus datos personales de acceso, rectificación y supresión, entre otros, mediante solicitud dirigida a la dirección de correo electrónico [datathon@ai4es.es](mailto:datathon@ai4es.es). Puede consultar la información adicional y detallada sobre Protección de datos, sus derechos y la Política de Privacidad de la red AI4ES en el siguiente enlace: <https://www.ai4es.com>

#### **UNDÉCIMA. - ACEPTACIÓN DE LAS BASES**

La participación en esta convocatoria comporta la aceptación de estas bases, que podrán ser interpretadas por el jurado y la entidad convocante en aquellos aspectos no previstos inicialmente. Más información en <https://www.ai4es.com/ai4es-datathon/>