

[Inicio](#) > Nuevo consorcio internacional para crear modelos generativos de IA fiables y de confianza para la ciencia

[Nuevo consorcio internacional para crear modelos generativos de IA fiables y de confianza para la ciencia](#)

El Trillion Parameter Consortium se pone en marcha con docenas de socios fundadores de todo el mundo, entre ellos el BSC.



Se ha constituido un consorcio mundial formado por equipos científicos de laboratorios federales, institutos de investigación, universidades y empresas para abordar los retos que plantea la creación de sistemas de inteligencia artificial (IA) a gran escala y el avance de una IA fiable para el descubrimiento científico.

El Trillion Parameter Consortium (TPC) reúne a equipos de investigadores dedicados a la creación de modelos de IA generativa a gran escala con el fin de abordar los principales retos que plantea el avance de la IA para la ciencia. Estos retos incluyen el desarrollo de arquitecturas de modelos escalables y estrategias de entrenamiento, la organización y conservación de datos científicos para modelos de entrenamiento, la optimización de bibliotecas de IA para plataformas de computación exaescala actuales y futuras, y el desarrollo de plataformas de evaluación profunda para evaluar el progreso en el aprendizaje de tareas científicas y la fiabilidad y confianza.

Para ello, TPC

- Construirá una comunidad abierta de investigadores interesados en crear modelos de IA generativa a gran escala de última generación con el objetivo general de avanzar en los problemas científicos y de ingeniería compartiendo métodos, enfoques, herramientas, conocimientos y métodos de trabajo.
- Incubar, lanzar y coordinar proyectos de forma voluntaria para evitar la duplicación de esfuerzos y maximizar el impacto de los proyectos en la comunidad científica y de IA en general.
- Crear una red mundial de recursos y conocimientos para facilitar la próxima generación de IA y reunir a los investigadores interesados en desarrollar y utilizar la IA a gran escala para la ciencia y la ingeniería.

El consorcio ha formado un conjunto dinámico de áreas de trabajo fundamentales que abordan tres facetas de las complejidades de construir modelos de IA a gran escala:

- Identificación y preparación de datos de entrenamiento de alta calidad, con equipos organizados en torno a las complejidades únicas de diversos dominios científicos y fuentes de datos.
- Diseño y evaluación de arquitecturas de modelos, rendimiento, formación y aplicaciones posteriores.
- Desarrollo de capacidades transversales y fundacionales, como innovaciones en las estrategias de evaluación de modelos con respecto al sesgo, la fiabilidad y la alineación de objetivos, entre otras.

El objetivo del TPC es proporcionar a la comunidad un lugar en el que múltiples iniciativas de creación de grandes modelos puedan colaborar para aprovechar los esfuerzos globales, con flexibilidad para adaptarse a los diversos objetivos de las iniciativas individuales. TPC incluye equipos que están llevando a cabo iniciativas para aprovechar las plataformas emergentes de computación a exaescala para entrenar LLMs -o arquitecturas de modelos alternativos- en investigación científica, incluyendo artículos, códigos científicos y datos observacionales y experimentales para avanzar en innovación y descubrimientos.

Los modelos de billones de parámetros representan la frontera de la IA a gran escala y sólo los mayores sistemas comerciales de IA se acercan actualmente a esta escala.

Para entrenar LLM con esta cantidad de parámetros se necesitan recursos informáticos de clase exascale, como los que están desplegando varios laboratorios nacionales del Departamento de Energía de Estados Unidos (DOE) y múltiples socios fundadores de TPC en Japón, Europa y otros lugares. Incluso con estos recursos, el entrenamiento de un modelo de última generación con un billón de parámetros requerirá meses de tiempo dedicado, algo imposible de conseguir en todos los sistemas excepto en los más grandes. Por consiguiente, en estos proyectos participarán grandes equipos multidisciplinares e interinstitucionales. El TPC se concibe como un vehículo de colaboración y cooperación entre estos equipos y dentro de ellos.

"En nuestro laboratorio y en un número cada vez mayor de instituciones asociadas de todo el mundo, los equipos están empezando a desarrollar modelos de IA de vanguardia para uso científico y están preparando enormes colecciones de datos científicos sin explotar previamente para el entrenamiento", dijo Rick Stevens, director asociado del laboratorio de informática, medio ambiente y ciencias de la vida en el Laboratorio Nacional Argonne del DOE y profesor de informática en la Universidad de Chicago. "Hemos creado TPC en colaboración para acelerar estas iniciativas y crear rápidamente los conocimientos y herramientas necesarios para crear modelos de IA con capacidad no sólo para responder a preguntas específicas de un dominio, sino también para sintetizar conocimientos de todas las disciplinas científicas."

Los socios fundadores de TPC pertenecen a las siguientes organizaciones (enumeradas por orden alfabético organizativo, con un punto de contacto):

AI Singapore: Leslie Teo

Allen Institute For AI: Noah Smith

AMD: Michael Schulte

Argonne National Laboratory: Ian Foster

Barcelona Supercomputing Center: Mateo Valero

Brookhaven National Laboratory: Shantenu Jha

CalTech: Anima Anandkumar

CEA: Christoph Calvin

Cerebras Systems: Andy Hock

CINECA: Laura Morselli

CSC - IT Center for Science: Per Öster

CSIRO: Aaron Quigley

ETH Zürich: Torsten Hoefer

Fermilab National Accelerator Laboratory: Jim Amundson

Flinders University: Rob Edwards

Fujitsu: Koichi Shirahata HPE: Nic Dube

Intel: Koichi Yamada

Jülich Supercomputing Center: Jenia Jitsev

Kotoba Technologies, Inc.: Jungo Kasai

LAION: Jenia Jitsev

Lawrence Berkeley National Laboratory: Stefan Wild

Lawrence Livermore National Laboratory: Brian Van Essen

Leibniz Supercomputing Centre: Dieter Kranzlmüller

Los Alamos National Laboratory: Jason Pruet

Microsoft: Shuaiwen Leon Song

National Center for Supercomputing Applications: Bill Gropp

National Renewable Energy Laboratory: Juliane Mueller

National Supercomputing Centre, Singapore: Tin Wee Tan

NCI Australia: Jingbo Wang

New Zealand eScience Infrastructure: Nick Jones

Northwestern University: Pete Beckman

NVIDIA: Giri Chukkapalli

Oak Ridge National Laboratory: Prasanna Balaprakash

Pacific Northwest National Laboratory: Neeraj Kumar

Pawsey Institute: Mark Stickells

Princeton Plasma Physics Laboratory: William Tang

RIKEN Center for Biosystems Dynamics Research: Makoto Taiji

Rutgers University: Shantenu Jha

SambaNova: Marshall Choy

Sandia National Laboratories: John Feddema Seoul

National University, South Korea: Jiook Cha

SLAC National Accelerator Laboratory: Daniel Ratner

Stanford University: Sanmi Koyejo

STFC Rutherford Appleton Laboratory, UKRI: Jeyan Thiyagalingam

Texas Advanced Computing Center: Dan Stanzione

Thomas Jefferson National Accelerator Facility: David Dean

Together AI: Ce Zhang

Tokyo Institute of Technology: Rio Yokota

Université de Montréal: Irina Rish

University of Chicago: Rick Stevens

University of Delaware: Ilya Safro

University of Illinois Chicago: Michael Papka

University of Illinois Urbana-Champaign: Lav Varshney

University of New South Wales: Tong Xie

University of Tokyo: Kengo Nakajima

University of Utah: Manish Parashar

Source URL (retrieved on 10 Oct 2024 - 07:02): <https://www.bsc.es/es/noticias/noticias-del-bsc/nuevo-consorcio-internacional-para-crear-modelos-generativos-de-ia-fiables-y-de-confianza-para-la>