



¿CÓMO ESTUDIAR MILES DE MILLONES DE DATOS Y NO MORIR EN EL INTENTO?

El Centro de Astrobiología (CSIC-INTA) desarrolla herramientas de análisis para la explotación científica de la misión Gaia

21-09-2016

El pasado 14 de septiembre, la Agencia Espacial Europea (ESA) presentó los primeros resultados de la misión Gaia, un proyecto que tiene por objetivo cartografiar nuestra galaxia con una precisión sin precedentes. El catálogo publicado representa el mayor mapa estelar construido hasta la fecha y proporciona posiciones y brillo para más de mil millones de estrellas, de las que 400 millones no habían sido catalogadas previamente. Además, el catálogo ofrece información precisa de la dirección y la velocidad a la que se desplazan en el espacio dos millones de estas estrellas. Y esto sólo es el primer paso. En los próximos años, nueva información astrométrica, fotométrica y espectroscópica hará del catálogo Gaia una herramienta imprescindible para los astrónomos y permitirá una revolución en el conocimiento que actualmente tenemos de la Vía Láctea.

Esta ingente cantidad de datos, unida a la de otros archivos astronómicos que proporcionarán información adicional de gran interés, en particular en aquellos casos en los que el objeto sea muy débil o muy brillante para Gaia o que emita gran parte de su energía fuera del rango visible, plantea un reto interesante para la comunidad astronómica: ¿seremos sepultados por esta enorme avalancha de datos o, por el contrario, contamos con los mecanismos necesarios para extraer el conocimiento científico que dichos datos encierran? Dos son las claves para dar una respuesta positiva a la pregunta anterior: “*Big Data*” y Observatorio Virtual.

Big Data es uno de los conceptos de moda en el mundo informático. Se habla de *Big Data* cuando, como ocurre con Gaia, el volumen de datos a tratar supera la capacidad del *software* habitual para su manejo. Es impensable imaginarse un batallón de expertos humanos intentando clasificar cada uno de los millones de objetos observados por Gaia. En estos casos es necesario olvidarse de metodologías clásicas y aplicar una batería de técnicas distintas entre las que destacan la minería de datos y los desarrollos astroestadísticos que, por ejemplo, van a permitir acometer análisis multidimensionales mucho más profundos que los estudios tridimensionales a los que estamos limitados por nuestra visión espacial.

Por su parte, el Observatorio Virtual (VO, según sus siglas en inglés) es una iniciativa internacional que nació hacia el año 2000 con el objetivo de crear una federación de archivos de datos astronómicos a nivel mundial. Esta federación de archivos actúa como una única base de datos virtual y permite que, dado un objeto astronómico, cualquier persona pueda conocer de manera fácil y rápida todo lo relacionado con el mismo (dónde, cómo y cuándo fue observado y qué se conoce de dicho objeto). El Observatorio Virtual es mucho más que un “Google” astronómico, ya que proporciona

información de manera estandarizada, lo que permite que todos los archivos se “entiendan” al compartir una misma lengua. Esta normalización en la forma de comunicarse los archivos astronómicos permite abordar retos impensable hace tan sólo unos años.

Herramientas de análisis: VOSA y Clusterix

INTA, a través del Centro de Astrobiología (CAB, CSIC-INTA), lidera y coordina, desde 2004, las actividades de Observatorio Virtual en España. En el marco del proyecto Gaia, el grupo de VO se ha centrado en el desarrollo de dos herramientas de análisis que permitan a los astrónomos analizar de manera rápida y precisa toda la información que, sobre los objetos Gaia, existe en el universo del Observatorio Virtual: VOSA y Clusterix.

En esta primera versión del catálogo, Gaia proporciona información del brillo a una única longitud de onda. Para poder conocer parámetros físicos fundamentales de las estrellas, como su temperatura, radio, masa o edad, es necesario complementar la información de Gaia con medidas de brillo a otras longitudes de onda, desde el rango ultravioleta hasta el infrarrojo y analizar estos datos comparándolos con modelos teóricos. VOSA, desarrollada en el CAB en colaboración con científicos de la Universidad de Valparaíso (Chile), realiza todo este proceso de manera automática para miles de objetos a la vez (figura 1).

Los cúmulos estelares son estructuras básicas para entender el nacimiento y evolución de las estrellas. Las estrellas de un cúmulo son parte de una misma “familia” y, como tales, comparten rasgos de identidad propios –todas tienen la misma edad, se mueven en la misma dirección y su composición química es muy similar–. Se estima que Gaia va a incrementar significativamente el número de cúmulos conocidos, pero para ello es necesario determinar y comparar las características de dichos cúmulos, comprobando que son diferentes a las de sus vecinos. Para facilitar esta ingente tarea, el grupo de VO, en colaboración con científicos de la Universidad de Barcelona y del Centro Astronómico Hispano-Alemán (CAHA), ha desarrollado Clusterix, una herramienta que permite, de manera transparente para el astrónomo, recopilar toda la información existente sobre edades, movimientos y distancias, visualizarla y asignar la probabilidad de pertenencia de un determinado objeto a un cúmulo (figura 2).

El reto es enorme y la carrera no ha hecho más que comenzar. En el marco del proyecto Gaia, el grupo de Observatorio Virtual continuará dedicando recursos y ofreciendo servicios a la comunidad astronómica para avanzar en el conocimiento del universo.

Sobre el CAB

El Centro de Astrobiología (CAB) es un centro de investigación mixto del Consejo Superior de Investigaciones Científicas (CSIC) y del Instituto Nacional de Técnica Aeroespacial (INTA). Creado en 1999, y asociado al NASA Astrobiology Institute (NAI), fue el primer centro del mundo dedicado específicamente a la investigación astrobiológica. Su objetivo es estudiar, desde una perspectiva transdisciplinar, el origen, presencia e influencia de la vida en el universo.

En el centro trabajan biólogos, químicos, geólogos, astrofísicos, planetólogos, ingenieros, informáticos, físicos y matemáticos, entre otros. Además de todo lo que tiene que ver con la comprensión del fenómeno de la vida tal y como lo conocemos (su emergencia, condiciones de desarrollo, adaptabilidad a ambientes extremos, etc.), también involucra la búsqueda de vida fuera de la Tierra (exobiología) y sus derivaciones, como son la exploración espacial (planetología) y la habitabilidad. El desarrollo de instrumentación avanzada es también uno de sus objetivos fundamentales.

Actualmente, más de 150 investigadores y técnicos trabajan en el CAB en diferentes proyectos científicos tanto nacionales como internacionales. En el CAB se ha desarrollado el instrumento REMS (Rover Environmental Monitoring Station), una estación medioambiental a bordo de la misión Mars Science Laboratory (MSL) de la NASA que explora actualmente Marte. También participa en las próximas misiones a Marte tanto de la NASA (instrumentos TWINS para InSight y MEDA para Mars2020) como de la ESA (instrumento RLS para ExoMars2020).

Más información

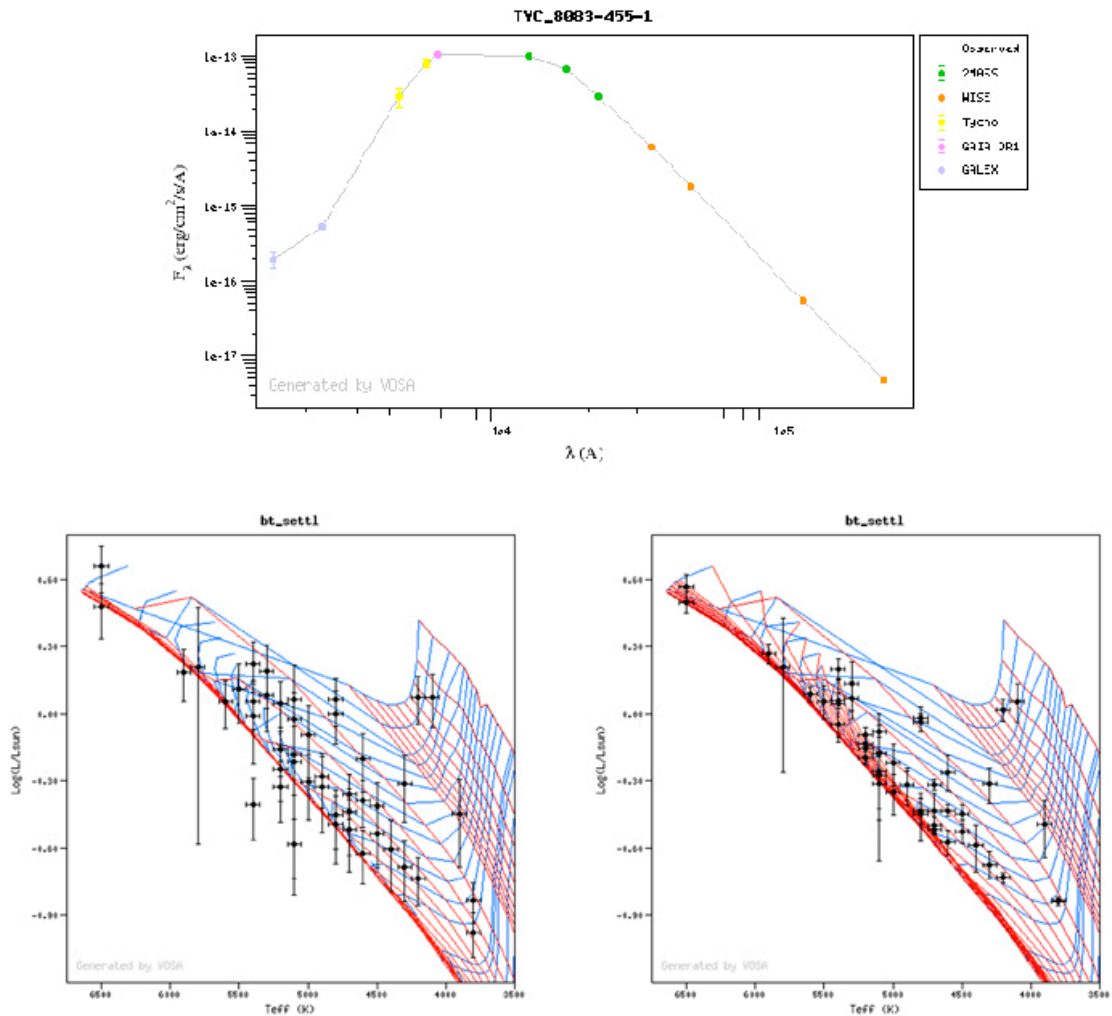


Figura 1. VOSA utiliza la fotometría de Gaia y de otros archivos del Observatorio Virtual (arriba) para construir la distribución espectral de energía (centro) y poder estimar la temperatura, composición química, masa y edad de la estrella (abajo). Los objetos estudiados corresponden a una muestra de estrellas jóvenes cercanas a la Tierra. Se aprecia como el uso de las distancias proporcionadas por Gaia (abajo derecha) permite estimar la luminosidad de objetos que, con determinaciones previas, caían por debajo de la malla de modelos (abajo izquierda). Esta determinación más precisa de los parámetros físicos facilitará la búsqueda de exoplanetas alrededor de estas estrellas.

Clusterix 2.0

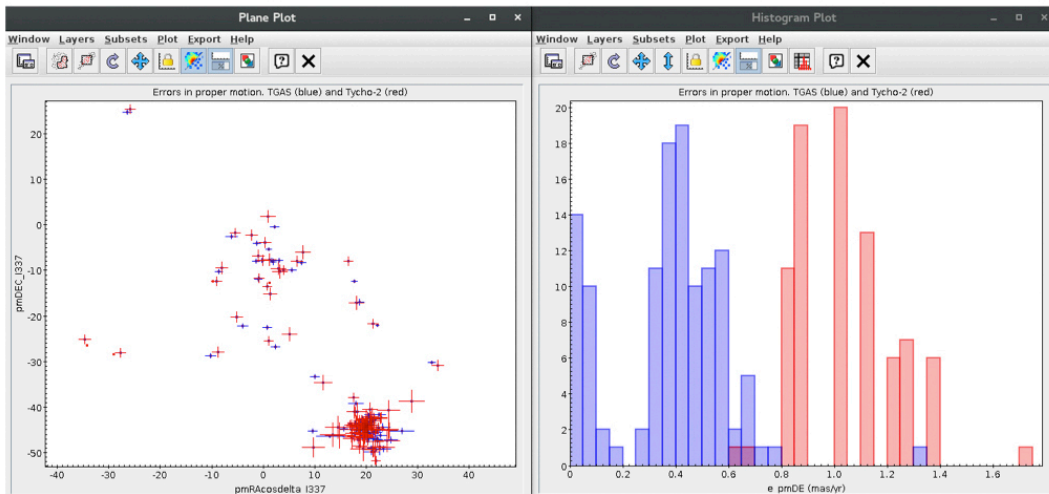


Figura 2. Comparación con Clusterix de movimientos propios (y errores asociados) utilizando el catálogo Tycho-2 (rojo) y el catálogo TGAS generado por la misión Gaia (azul). La agrupación que aparece abajo a la derecha es el cúmulo de las Pléyades.

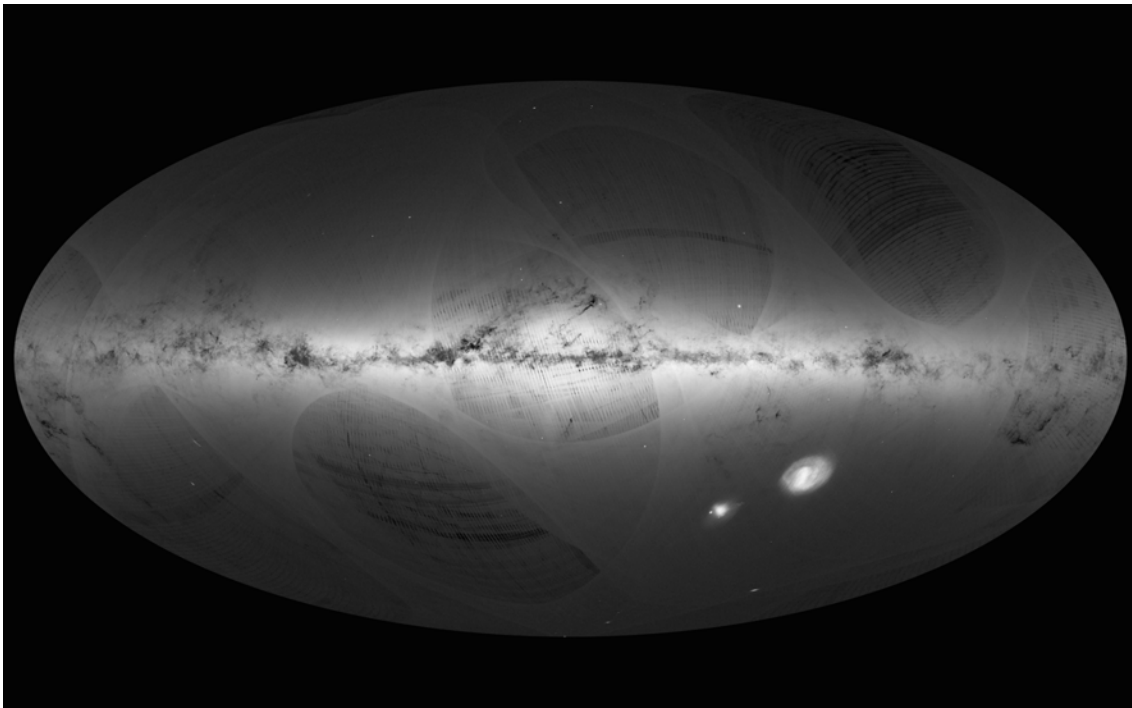


Imagen. Primer mapa del cielo de Gaia. © ESA/Gaia/DPAC

Enlaces de interés:

Misión Gaia

Observatorio Virtual Español

VOSA

Clusterix

Contacto del Investigador Principal del Observatorio Virtual Español

Enrique Solano: esm (+@cab.inta-csic.es); (+34) 918131154

UNIDAD DE CULTURA CIENTÍFICA DEL CAB

Cristina Delgado: cdelgado (+@cab.inta-csic.es)

Juan Ángel Vaquerizo: jvaquerizog (+@cab.inta-csic.es)

(+34) 915206438

