ChiVO:
Observatorio Virtual Chileno

El data center de la Universidad Técnica Federico Santa María (USM) tiene una capacidad de almacenamiento de 1PB para enfrentar el Big Data, siendo el más grande de Latinoamérica.

Dr.Mauricio Solar.
mauricio.solar@usm.cl

ChiVO es un Observatorio Virtual (VO) desarrollado en Chile que nace en torno a la necesidad de almacenar datos que requieren grandes capacidades de almacenamiento, de desarrollar nuevas herramientas para analizar grandes volúmenes de datos y mejores algoritmos para el procesamiento inteligente de datos astronómicos en Chile, principalmente del proyecto ALMA, el radiotelescopio más grande del mundo.

ChiVO es una plataforma construida por profesionales de la Ingeniería Civil Informática, ayudantes de investigación, memoristas de Ingeniería Civil Informática, tesistas de Magíster y Doctorado. Además, participan diversas instituciones entre ellas la Universidad de Chile y la Universidad de Valparaíso, así como también las instituciones asociadas ALMA, REUNA, la Chinesse Academy of Sciences South America Center for Astronomy (CASSACA) y REUNA, Ciencia y Educación en Red.

El trabajo realizado en el VO permite almacenar los datos en el territorio local para compartir los datos de observación de los astrónomos en Chile. Esto constituye un material de uso tanto para los astrónomos como también para los estudiantes de distintas escuelas. Ello, pues el trabajo realizado por los investigadores tiene como propósito facilitar la investigación de la astronomía que utiliza múltiples fuentes de datos.

El costo de realizar una hora de observación en un observatorio moderno fluctúa entre los 20 a 30 mil dólares, explica el Director de ChiVO, Dr.Mauricio Solar. Estos datos obtenidos, agrega, demoran en ser de uso público ya que por una parte los astrónomos que son propietarios de los datos por un año deben publicar en revistas de alto impacto y por otra, se deben estandarizar para que otros investigadores puedan descargar la información.

"En este contexto, ChiVO ha logrado crear un proceso de ingestión de datos para que esos datos de alto costo que se van transformando en públicos, sean alcanzables para el resto de la comunidad que ahora los puede usar libremente para hacer más ciencia sobre ellos a costo cero y en forma instantánea. Es decir, no es necesario presentar una propuesta de observación, y esperar el día y hora de observación para obtener el dato", acota Solar.

La Alianza Internacional de Observatorios Virtuales (IVOA), de la cual ChiVO es miembro, proporciona a los centros de datos un marco estándar para publicar y entregar servicios utilizando sus datos mediante la estandarización de datos y metadatos, mediante protocolos de intercambio de datos y el uso de un registro que enumera los servicios disponibles y qué se puede hacer con ellos. Estos datos están disponibles para todos los investigadores, independiente de su afiliación o acceso a las instalaciones de observación.

Radiotelescopio ALMA(ESO/NAOJ/NRAO), USM/ José Rivas.

La disposición pública de este material, especialmente de ALMA, es un aspecto que la comunidad científica estaba esperando, según Solar. "Para los propios radioastrónomos que trabajan en Chile es muy incómodo tener que bajar sus propios datos desde la NRAO en Estados Unidos, porque son archivos muy grandes y requieren un largo tiempo de comunicación para descargarlos, a veces superando los dos días. Al disponibilizar estos datos a toda la comunidad internacional, se pueden realizar búsquedas de las cosas que existen, y entonces bajar sólo lo que se requiere una vez que se encuentra, optimizando mucho tiempo".

Dentro de los logros más importantes de ChiVO se puede mencionar el paso de un prototipo a nivel de servidor local (computador personal amateur) hacia una plataforma de alto desempeño (HPC) con una capacidad de almacenamiento de 1 PB, lo que según Solar, garantiza la capacidad de almacenar 5 años de datos de ALMA. "Esto se logró gracias a la colaboración con la Academia China de Ciencias y la empresa Huawei que se han aliado con nosotros para implementar el data center más potente en Latinoamérica en términos de capacidad de almacenamiento para datos astronómicos". Mientras que en materia de desarrollo se están realizando pruebas de concepto con datos reales en la detección y clasificación de líneas espectrales.

El impacto que genera el actual trabajo de ChiVO es según Solar, un significativo aporte para la sociedad ya que el trabajo realizado con el procesamiento de grandes volúmenes de datos, el desarrollo de nuevas técnicas para acelerar los tiempos de procesamiento y el desarrollo de nuevos algoritmos para hacer minería de datos, "algún día serán utilizados en el día a día en otras aplicaciones de otros ámbitos como por ejemplo procesamiento de imágenes médicas, procesamiento de imágenes satelitales, en geología y prospección minera, etc".