



# PISCIS: Platform for Interactive Search and Citizen Science

I. Daza<sup>1,2</sup>, I.G. Alfaro<sup>1,2</sup>, J.A. Benavides<sup>3,2</sup>, M. Lares<sup>1,3</sup>, M.V. Santucho<sup>1,2</sup>, J.B. Cabral<sup>4,1</sup>,  
A.L. O'Mill<sup>1,3</sup>, F. Rodriguez<sup>1,3</sup> & M. Koraj<sup>5</sup>

<sup>1</sup> *Instituto de Astronomía Teórica y Experimental, CONICET-UNC, Argentina*

<sup>2</sup> *Facultad de Matemática, Astronomía, Física y Computación, UNC, Argentina*

<sup>3</sup> *Observatorio Astronómico de Córdoba, UNC, Argentina*

<sup>4</sup> *Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas, CONICET-UNR, Argentina*

<sup>5</sup> *Liricus SRL, Argentina*

Contacto / [vanessa.daza@unc.edu.ar](mailto:vanessa.daza@unc.edu.ar)

**Resumen** / Muchos temas de la Astronomía moderna se caracterizan por la identificación de características en imágenes. Si bien esta es una tarea fácil para un ojo entrenado, es difícil de obtener la misma calidad al realizarse mediante modelos o métodos numéricos. Aquí presentamos la experiencia y el desarrollo de una plataforma web de ciencia ciudadana (PISCIS, Platform for Interactive Search and Citizen Science), plataforma que tiene como objetivo la generación de catálogos de valor agregado a partir de datos que comprenden un conjunto de imágenes. Además presentamos un proyecto piloto para la clasificación de pares de galaxias interactuantes usando una plataforma web abierta.

**Abstract** / Many topics in modern astronomy are characterised by the identification of features in images. While this is an easy task for a trained eye, it is difficult to obtain the same quality when performed by models or numerical methods. Here we present the experience and development of a citizen science web platform (PISCIS, Platform for Interactive Search and Citizen Science), that aims at generating value-added catalogues from data comprising a set of images. We also show a pilot project for the classification of interactions in galaxy pairs using an open web platform.

*Keywords* / methods: numerical — methods: observational — methods: statistical

## 1. Introducción

La Astronomía moderna se caracteriza por un crecimiento exponencial en volumen y complejidad de los datos observacionales y teóricos que se van generando en todo el mundo (Szalay et al., 2002). Esto resta considerablemente la eficiencia del análisis manual de datos, volviéndolo en muchos casos una opción inviable.

Con el objetivo de favorecer la ciencia de datos y aprender sobre el desarrollo de un programa desde el diseño del problema hasta la obtención y análisis de los datos, construimos una aplicación informática flexible para ser aplicada en problemas de clasificación visual o similares.

Esta herramienta se denominó PISCIS: *Plataform for Interactive Search and Citizen Science*, puede ser descargada desde el repositorio de GitHub\* y posee documentación disponible en la plataforma "ReadtheDocs".\*\*.

La misma está destinada a aprovechar el interés ciudadano para facilitar la recolección y el análisis de datos, o bien para facilitar el procesamiento o clasificación de datos privados en proyectos de investigación mediante

una interface conveniente.

Esta aplicación además de brindarnos habilidades en el desarrollo de programas de computadora, necesarias en la actualidad para el análisis de grandes volúmenes de datos astronómicos, también le ofrece a la comunidad una herramienta fácilmente adaptable para la recolección de información en sus propias investigaciones a través de la web. La aplicación básica de PISCIS es una encuesta conformada por una sola pregunta relacionada con una imagen que se relaciona con una pregunta de opciones múltiples preestablecidas.

Una diferencia crucial frente a otro tipo de aplicaciones (como Zooniverse), es que éste último tiene opciones preestablecidas de como realizar la investigación, mientras que con PISCIS el usuario cuenta con un entorno base que se puede adaptar según la necesidad. Con conocimientos básicos de python, se espera que PISCIS permita desarrollar un proyecto en un lapso de tiempo más corto y sin la necesidad de tener un equipo de expertos en programación y desarrollo de software. Reportamos además la experiencia en el desarrollo y los resultados científicos obtenidos hasta el momento haciendo uso de PISCIS, además de posibles aplicaciones que puede tener para la recolección de información astronómica.

\*<https://github.com/vanedaza/piscis>

\*\*<https://piscis.readthedocs.io/en/latest/>

## 2. Desarrollo

En la descripción de las herramientas desarrolladas es importante distinguir dos tipos de usuario, a saber, el usuario de PISCIS (UP), que descarga nuestra aplicación para desarrollar parte de su investigación y el usuario del producto (UC), que interactúa con el producto hecho por el investigador a través de la web o de una intranet.

### 2.1. Alcances e Infraestructura

El desarrollo de nuestra aplicación informática (APP) se basó principalmente en que cumpla con los siguientes requerimientos:

- que sea una herramienta útil y fácil de usar para la comunidad científica en el área de la Astronomía,
- que facilite la recolección de datos mediante una pregunta simple de opción múltiple junto a una imagen, de tal manera que permita al público general proporcionar una respuesta rápida,
- que posea capacidad administrativa de creación, actualización y eliminación de objetos de contenido, con un registro de todas las acciones realizadas a través de una interface de fácil uso, y
- que los datos recopilados por la misma se guarden de forma ordenada, fácilmente accesibles.

En base a estos requerimientos, desarrollamos PISCIS sobre un entorno web, el cual cuenta con un conjunto estandarizado de conceptos, prácticas y criterios para enfocar un tipo de problemática particular, y sirve como referencia para enfrentar y resolver nuevos problemas de índole similar. En particular, el entorno web que reúne todos los requerimientos necesarios para la fabricación de nuestra app y la distribución de la misma es DJANGO. El mismo sigue el patrón de diseño MVT: Modelo – Vista – Template, y su lenguaje de programación principal es PYTHON.

Para DJANGO, un Modelo es un tipo especial de objeto en donde se define la información de salida que es generada en la APP, y que puede ser visualizada y editada con herramientas como DB BROWSER FOR SQLITE. En cuanto a los patrones de diseño Vista y Template de Django, la Vista describe qué datos serán presentados, mientras que el Template determina cómo se verán los mismos. En resumen, en el diseño MVT primero definimos los Modelos, elegimos qué campos de los mismos serán observables desde las Vistas y como serán presentados a los usuarios de la web mediante los *Templates*.

### 2.2. Desarrollo de la aplicación

El objetivo es el desarrollo de una aplicación que genere y recolecte datos a través de imágenes acompañadas de una encuesta, que son dos elementos fáciles de modificar por cualquier UP. El formato más simple para esto es el de un sitio web de ciencia ciudadana, cuyo diseño lo basamos en el siguiente lineamiento:

Para iniciar, instalamos un entorno virtual con el fin de aislar la configuración de nuestro proyecto de otro que estemos desarrollando, si bien este paso no es ne-

cesario, es altamente recomendable para evitar posibles problemas de compatibilidad con librerías previamente instaladas. Luego, en el entorno virtual, instalamos las librerías necesarias para el desarrollo de PISCIS, DJANGO, PYTZ, SPHINX, DJANGO-NOSE y PILLOW, siendo esta última una librería de PYTHON que permite manipular y guardar diferentes formatos de archivo de imagen, las versiones de cada una de estas librerías se encuentran en el archivo requirements.txt dentro del repositorio de GitHub. Posteriormente, configuramos las opciones generales de DJANGO, como el idioma, la zona horaria, el tipo de base de datos a usar, etc. Manipulamos el patrón de diseño MVT, comenzando con el diseño de los modelos, es decir, definimos la cantidad de bases de datos que queremos y su contenido. Como cada página en Internet necesita su propia URL (*Uniform Resource Locator*), trabajamos paralelamente en las vistas y los *templates*, asociando una dirección URL a cada una de las primeras y un archivo formato .html a los segundos. Como último paso hacemos PISCIS, el UP puede hacer uso de servidores públicos, privados o incluso usar su ordenador como servidor local.

### 2.3. Prueba

El objetivo de las pruebas es mejorar el manejo, rendimiento y optimización del sitio web. Dado que trabajamos sobre un entorno web que incluye sus propias pruebas la verificación del sitio se redujo a:

- Verificar el tiempo que se demora en cargar la página, las imágenes, vídeos, archivos multimedia, etc.
- Mejorar los problemas de codificación del lenguaje JAVASCRIPT, CSS, HTML que son parte de los *templates*.
- Detectar error 404, el cual se refiere a páginas no encontradas en el servidor.
- Detectar enlaces rotos que afectan el posicionamiento de una web.
- Realizar pruebas en estilo de código y funciones programadas.

## 3. Ejemplos del uso de PISCIS

PISCIS es una herramienta que puede ser adaptada para cualquier encuesta basada en una imagen, por ejemplo, la clasificación de raza de gatos, tipos de mariposas, hongos, productos comerciales, etc. Sin embargo, hemos diseñado PISCIS con el propósito de recolectar información astronómica, como lo es, la clasificación de estrellas variables o exoplanetas a partir de series temporales visualizadas en imágenes, la morfología de galaxias y una infinidad de otros estudios, donde la observación detallada de características de un evento u objeto astronómico permite su clasificación.

### 3.1. PISCIS usado en la clasificación de pares de galaxias

Como caso de estudio, PISCIS fue implementada para recolectar datos, presentando al usuario (UC) una

imagen de un par de galaxias cercanas y solicitando seleccionar el tipo de interacción observada en base a las opciones alta, media y baja. El UC vota qué tipo interacción cree observar para el par de galaxias presentado, utilizando como referencia la documentación y ejemplos incluidos en la página web destinada para este producto del uso de PISCIS.

### 3.1.1. Datos

Utilizamos un catálogo de pares de galaxias del Sloan Digital Sky Survey (SDSS, York et al. (2000)), provisto por O'Mill & Rodríguez, el cual cuenta con 8182 sistemas, con mediciones de corrimiento al rojo entre  $z = 0.0105$  y  $z = 1.9967$ , y una media de  $z = 0.0913 \pm 0.0383$ . El mismo está definido a partir de las diferencias de velocidad, distancia y magnitud aparente de las galaxias que conforman el par y teniendo en cuenta criterios de aislamiento de este sistema con otras galaxias (Rodríguez et al., 2020). De este catálogo tomamos una muestra de 399 sistemas, y generamos las imágenes a través de SCORPIO<sup>\*\*\*</sup>. Este programa permite a partir de información de las coordenadas Ascension Recta, Declinación y corrimiento al rojo (AR, DEG,  $z$ ) del par de galaxias generar una imagen con la resolución deseada en alguno de los relevamientos disponibles: SDSS, Wide Field Infrared Survey Explorer (WISE, Wright et al. (2010)) o Two Micron All Sky Survey (2MASS, Skrutskie et al. (2006)).

### 3.1.2. Implementación

El proceso para utilizar PISCIS para configurar una base de datos con una interface interactiva para clasificación de imágenes se puede resumir como sigue:

Luego de descargar PISCIS del repositorio en Github y de instalar las librerías requeridas, manipulamos los valores por defecto de las opciones fundamentales de la app, en el archivo `settings.py` elegimos idioma español y la zona horaria de Argentina.

Seguidamente, agregamos la información contenida en la encuesta, es decir, el tipo interacción. Este paso se realiza desde el archivo `/encuesta/models.py` donde se diseñan las bases de datos de las imágenes y la votación, el diseño de esta base de datos se hace a partir de Programación Orientada a Objetos en PYTHON definiendo dos clases, `Images` y `Choice`. En la Clase `Choice` se encuentran los campos a votar, identificación del usuario y la identificación de la imagen a votar. Con los campos de las bases de datos definidos, queda la carga de datos de las imágenes y de la encuesta, la primera por parte de UP en el proceso de desarrollo y la segunda por los UC quienes son los visitantes de la página web.

Con la encuesta definida la carga de información y control sobre el sitio web queda a cargo de un superusuario creado por el UP.

Con la parte funcional del proyecto finalizada, se rea-

liza el diseño de la apariencia varios *Templates* en HTML, CSS Y JS contenidos en PISCIS, los cuales fácilmente se pueden modificar, desactivar o eliminarlas según se desee. Para el desarrollo de esta encuesta, los *Templates* son usados para que la página web cuente con cinco pestañas, una principal que se conecta con una página de contacto, inicio de sección, encuesta y otra con una descripción sobre la app que se uso para crear la encuesta, es decir, el proyecto PISCIS.

Finalizados los cambios en la parte funcional y estética de la encuesta de pares de galaxias, se configuró la app para ser desplegada en uno de los servidores de la Universidad Nacional de Córdoba. La misma se puede acceder a través de la url del proyecto de clasificación de pares de galaxias <sup>\*\*\*\*</sup>. Este proyecto en particular es de acceso abierto aunque requiere un registro previo.

## 4. Conclusiones

En este trabajo se reporta la experiencia del desarrollo de un programa para ayudar a un usuario en la clasificación visual de imágenes. Hemos adquirido habilidades para la elaboración de un producto que adquiera información astronómica, así como también la experiencia de trabajo interdisciplinaria con el fin de desarrollar una herramienta que puede ser usada para cualquier estudio astronómico en donde se necesite realizar alguna clasificación a partir de un análisis visual. Además, generamos una muestra de pares de galaxias con descripción de su tipo de interacción al realizar una inspección visual. Esta herramienta será publicitada mediante las oficinas de prensa del OAC y del IATE para invitar al público en general a participar del proyecto. Como trabajo a futuro, analizaremos el impacto en la comunidad científica y la eficiencia de usar información provista por la comunidad no experta. Se analizarán además los datos recabados para determinar diferencias en la percepción de usuarios expertos y de usuarios del público en general, y se aprovechará el lanzamiento de la herramienta para trabajar en la difusión de conceptos de astronomía extragaláctica.

*Agradecimientos:* Este trabajo ha sido financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET) y la Secretaría de Ciencia y Tecnología (SeCyT) de la UNC.

## Referencias

- Rodríguez F., et al., 2020, *A&A*, 634, A123
- Skrutskie M.F., et al., 2006, *AJ*, 131, 1163
- Szalay A.S., Gray J., Vandenberg J., 2002, *American Astronomical Society Meeting Abstracts, American Astronomical Society Meeting Abstracts*, vol. 201, 134.06
- Wright E.L., et al., 2010, *AJ*, 140, 1868
- York D.G., et al., 2000, *AJ*, 120, 1579

<sup>\*\*\*</sup>[github.com/josegit88/SCORPIO](https://github.com/josegit88/SCORPIO)

<sup>\*\*\*\*</sup><https://paresdegalaxias.iate.conicet.unc.edu.ar/>