# NCSA Brown Dog and Box Skills Speed up Astronomical Research

http://www.ncsa.illinois.edu/news/story/ncsa_brown_dog_and_box_skills_speed_up_astronomical_research
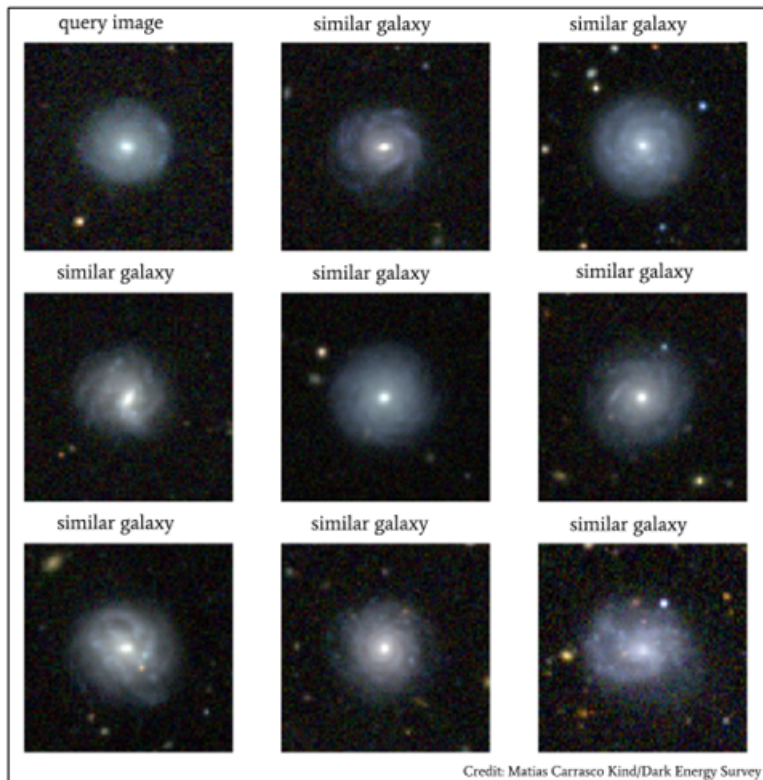
Brown Dog, a prototype Data Transformation Service (DTS) developed at the National Center for Supercomputing Applications (NCSA) with support from the National Science Foundation (NSF) is partnering with Box to bring intelligence framework capabilities to users' content by leveraging its recently introduced Box Skills. With the release of one of our transformations as a preliminary Box Skill, Box users and scientists will soon be able to find images of similar galaxies based on a query galaxy image. Box Skills are not automatically enabled for all Box instances and the Galaxy Skill will have a limited initial rollout. The University of Illinois is planning to enable and configure the Galaxy Skill on the Campus instance. Other Universities or businesses who want to use this Skill will need to create their own Skill application. Anyone interested in having this Box Skill on their enterprise can reach out to the NCSA Brown Dog team at browndog-support@ncsa.illinois.edu for assistance in creating an instance of the Galaxy Skill which can be authorized on their instance of Box.
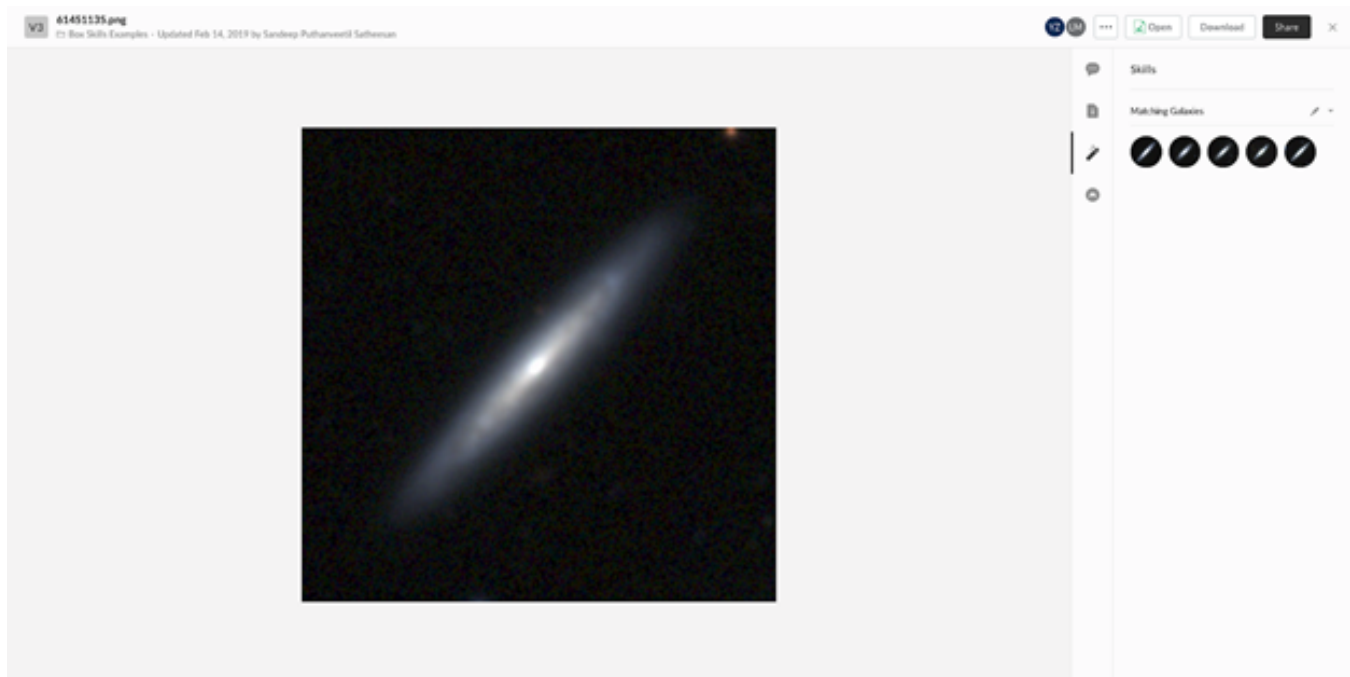
## What is a Box Skill?

Box Skills are bits of code that operate on folders in Box. Apply them to a folder, and a Box Skill automatically analyzes each file placed into that folder using a machine learning algorithm, and then writes the output of its analysis as metadata on the file. For example, apply an audio Skill to a folder, place an audio file in that folder, and the Skill processes the audio file using a machine learning algorithm and then adds a transcript to that audio file. Users can then view the transcript when viewing the file in Box. The metadata information can drive other Box functionality, such as search. With the Galaxy Skill, Box scientists can upload an image of an unknown galaxy and have a set of possible galaxy matches returned.

## The Galaxy Skill

The original tool which this was based upon was developed by the DES Labs for the Dark Energy Survey (DES) at NCSA. Brown Dog and Box are partnering to bring this technology, once reserved for mission-scale science, to Box users everywhere. This will create new opportunities for scientists and Box users to widely utilize machine learning tools to extract actionable data from images. An example of this skill in operation: Take the following images of galaxies captured by DES selected by a query image on the top left. This is just an example of the capabilities of this skill which is able to search among thousands of galaxy images.



Credit: Matias Carrasco Kind/Dark Energy Survey

## How it works:

Using the above example, Box users can upload an image of a galaxy to Box, which is then sent to the Autoencoder Deep Learning model, which is pre-trained and running in the back end. The model then creates a one-dimensional compressed representation in a so-called "latent space" of reference for this new image and compares that to thousands of galaxy images in a database in that same space. Each one of the newly created floating points represents a distinct characteristic about the submitted image, such as roundness, brightness, size, orientation, etc.

The autoencoder also compresses these images from roughly 40,000 pixels (in three channels) down to about fifty floating points (roughly a 2000x compression), which vastly expedites the search process by using smaller sizes. Judging by similarities in the submitted galaxy image, the autoencoder will reveal its 'Top 5 most similar galaxies.' The images of the five similar galaxies are then displayed in the sidebar when previewing the galaxy image in Box.

## What does this mean for the Box community?

The future goal for the next version of this skill is for users to be able to apply this same model to any set of images - plants, faces, galaxies, etc.- to quickly find similar images throughout a database.

## Challenge:

To maintain a reliable and robust user technology that will produce an accurate 'Top 5' result, there is a continuous effort to update and maintain the deep learning model.

If you want to develop a Box Skill yourself, you can do so by building with the Box Skills Kit, a developer toolkit that makes it easy to create and configure Box Skills. Using the Box Skills Kit, you can apply the AI/ML technology of your choice to enhance data in Box. You can use any applicable third-party AI /ML service or your own internally-built AI/ML service. You can learn more about the Box Skills by visiting box.com/skills. Learn more about Brown Dog, Box Skills, and DES.