JPL SIRI INTERNSHIP ANNOUNCEMENTS OF OPPORTUNITY
Spring 2021
(subject to change without notice)

AO# **11597**
Project: **Machine Learning for Galaxy Surveys**

Background
We are a group of cosmologists working at Caltech and Jet Propulsion Laboratory (JPL) developing data analysis techniques for the next generation of galaxy surveys. These surveys will map hundreds of millions of galaxies and can be used to study the initial conditions, energy content and evolution of the Universe.

Description
We will develop machine learning techniques to improve the speed or/and accuracy of extracting information from these large data sets compared to current standard techniques. We will be developing these techniques, testing them on a small set of simulations, and compare them to standard results. Examples of problems are: 1) redshift measurements and validation with machine learning; 2) compression techniques for large data using neural networks; 3) efficient methods for sampling cosmological parameter posteriors using machine learning. The project may also be adjusted to suit the student's background and experience.

Background, Skills, Courses
Coding skills: familiarity with python, git, C/C++/fortran strongly preferred; experience in research/toy projects/tutorials in neural network or general machine learning techniques preferred. Courses: strength in mathematics and physics required; courses in astronomy/cosmology/statistics are a plus but not required.

Major(s): Astronomy/Astrophysics, Computer Science, Physics

AO# **11596**
Project: **NISAR Operations training and tools**

Background
The NASA-ISRO Synthetic Aperture Radar, or NISAR, satellite is designed to observe and take measurements of some of the planet's most complex processes. These include ecosystem disturbances, ice-sheet collapse, and natural hazards such as earthquakes, tsunamis, volcanoes and landslides. Data collected from NISAR will reveal information about the evolution and state of Earth's crust, help scientists better understand our planet's processes and changing climate, and aid future resource and hazard management. The mission is a partnership between NASA and the Indian Space Research Organization. It is scheduled to launch in 2022.

Description
NISAR uses the HySDS framework for its science data processing system. It was developed at JPL and open sourced in 2017. It is under active development with many new features designed to support upcoming NASA missions like SWOT and NISAR. HySDS scalably powers the ingestion, metadata extraction, cataloging, high-volume data processing, provenance management, and publication of data products for various projects at JPL. Candidates would work on one or more of the following tasks: 1) Work on training materials for HySDS and how we use Amazon Web Services and the ELK Stack 2) Develop scripts, reports, and dashboards for the operations team 3) Work with other interns on further development of the new SDS Watch feature, that was developed as a proof of concept in early 2020. It is designed to stream insights and metrics from the data processing pipeline and science algorithm software to a centralized place. The candidates would need to become familiar with HySDS and SDS Watch, analyze data to determine what features or metrics to filter on, and work on the concept or prototype to leverage SDS Watch for operational monitoring. One of the major ways to use SDS Watch is to build Kibana dashboards to visualize data in real time. This will require looking at the data, possibly normalizing the data, and then designing the dashboards.

Background, Skills, Courses
ELK Stack (Elasticsearch, Logstash, Kibana) Linux Python Docker

Major(s): Computer Science, Information Systems/Technology

AO# **11593**
Project: **SMAP Operations scripts and documentation**

Background
The SMAP (Soil Moisture Active Passive) mission measures soil moisture over a three-year period, every 2-3 days. This permits changes, around the world, to be observed over time scales ranging from major storms to repeated measurements of changes over the seasons. The Science Data System is responsible for producing archive quality Science Data Products used by science teams and researchers.

Description
On SMAP, we are migrating to a new Science Data System based on the Hybrid Cloud Science Data System Framework. This requires updating our science data system operational processes and tools and user interfaces. Working with mentors and fellow interns, the interns will update the current operational procedure documents and build and update operational tools to support this transition. In particular, the interns will be responsible for refactoring the operational tools and updating them to Python3.

Background, Skills, Courses
Linux Relevant software: Amazon Web Services Elasticsearch Kibana Logstash Redis GitHub Celery RabbitMQ Flask Programming skills: Python JavaScript Bash Java JSON/XML

Major(s): Computer Science, Information Systems/Technology

AO# **11591**
Project: **Orbiting Carbon Observatory 3 (OCO-3)**

Background
The Orbiting Carbon Observatory 3 (OCO3) is an Earth Observing, Earth orbiting satellite mission measuring the levels of carbon dioxide (CO2) in the atmosphere, the leading human-made driver of climate change. The Science Data Operations System (SDOS) is a key part the data return process. The SDOS extracts science data from the returned data stream and prepares higher-level products for the public. With a million observations coming in each day, the SDOS operations cannot do adequate quality control without helper tools.

Description
This project will augment the OCO-3 Science Data Operations System with new tools and processes. The System operators perform various quality checks for each data downlink. The student will modernize & improve web reporting tools which communicate quality checks to operators, as well as codify some of these checks and automate other processes in order to triage the daily products generated by the SDOS. Most products will fall into a nominal range, but the set which don't will be presented to the operator along with initial findings, error messages and log entries. The ability to automate checks or otherwise streamline operations will reduce the daily routine workload of Operators so they may solve real issues more rapidly. The student will work with programmers, testers and systems engineers in order to learn the workflows of the data system and the processes of the operators. The task will involve Python scripting and, perhaps, higher-level language programming.

Background, Skills, Courses
Python scripting, web reporting tools, linux and shell scripting, experience with data analysis and data plotting, good communication skills, experience with customer focused use. Desired: familiarity with AWS.

Major(s): Information Systems/Technology, Computer Science, Data Analysis

AO# **11589**
Project: **Enhancing Technical documentation for Hybrid Cloud Science Data Systems**

Background
The Hybrid Cloud Science Data System (HySDS) is a science data system framework that can leverage a heterogeneous set of worker nodes from private and public clouds as well as virtual and bare-metal machines to perform every aspect of the traditional science data system. The HySDS framework was developed at the NASA Jet Propulsion Laboratory (JPL) over a number of years, before being released as open source software on GitHub.com in October 2017 under the Apache Software license, Version 2.0 HySDS consists of numerous software packages, which can be found under our GitHub org: https://github.com/hysds It is currently used to power the ingestion, metadata extraction, cataloging, high-volume data processing, and publication of the geodetic data products for the Advanced Rapid Imaging & Analysis (ARIA), Getting Ready for NISAR (GRFN), and Water Vapor Cloud Climatology (WVCC) projects at JPL. It is being developed for use in science data systems for the Soil Moisture Active-Passive (SMAP), Surface Ocean Water Topography (SWOT), and NASA-ISRO SAR (NISAR) missions.

Description
Continue the work of previous interns to improve the public-facing documentation and user guides for the HySDS open source project. Continue with a cohesive website to link to wiki, GitHub and other resources. Continue building a framework and sample documents for a contributor guide. Create an automatable changelog/release documentation to track changes and impacts on end-users.

Background, Skills, Courses
Required: Self-starter, communicator, Linux or technical background, website development experience, Python literacy Additional: GitHub, Python, JavaScript, JSON/XML, Amazon Web Services, HTML, CSS, Elasticsearch, Docker, Terraform, Confluence wiki

Major(s): Information Systems/Technology, Computer Science, Earth Science

AO# **11588**
Project: **Genetic inventory of microbes present on spacecraft and spacecraft associated surfaces**

Background
In compliance with the Committee on Space Research Planetary Protection Policy, NASA monitors the total microbial burden of spacecraft as a means for minimizing forward contamination, which is the inadvertent transfer of viable contaminant microorganisms, to extraterrestrial environments. The PP Group at JPL is responsible for evaluating spacecraft microbial contamination, a.k.a. bioburden, for interplanetary missions. Development of modern genomic techniques in the past decade has made it possible to perform bioburden estimations using DNA collected from spacecraft hardware. A metagenomics approach provides a comprehensive taxonomic profile of a microbial community in a given habitat. For the first time, with Mars 2020, metagenomics has been used by PP to evaluate microbial contamination on a flight project. This project has generated large volumes of DNA sequencing data and will apply cutting-edge bioinformatic tools to identify, document, and archive the broadest spectrum possible of potential terrestrial contaminants from critical spacecraft hardware during the M2020 spacecraft assembly process.

Description
The Mars 2020 Genetic Inventory team is looking for a motivated undergraduate student intern to assist with computational biology work related to this project. Students will work on automating an optimized metagenomics pipeline for high throughput data processing of genomic sequencing data. Ideal candidates should have at least intermediate level experience with Linux (Shell scripting) and Python programming. Experience with pipeline automation and cloud computing is a plus.

Background, Skills, Courses

Major(s): Computer Science, Biology/Bioengineering

AO# **11586**

Project: **Post-Processing for Exoplanet Imaging with the Nancy Grace Roman Space Telescope Coronagraph Instrument**

Background

The Nancy Grace Roman Space Telescope (formerly WFIRST) will be launched in 2026 with an onboard technology demonstrator for exoplanet direct imaging. The Roman Coronagraph Instrument (CGI) will be capable of detecting and characterizing exoplanets and circumstellar disks in visible light at the unprecedented contrast level of ~10^{-9}. Such a contrast level, which is several magnitudes better than state-of-the-art coronagraphs, raises entirely new challenges that will be overcome using a combination of hardware, calibration and data processing. Detecting a Jupiter-like planet in reflected light requires observations of tens to hundreds of hours, during which the instrumental aberrations may evolve significantly. Therefore, innovative observing strategies and post-processing techniques are being developed, accounting for this very unique instrument and its requirements.

Description

This project is part of an on-going effort to prepare direct imaging observations of exoplanets with the Roman CGI. The student will work with Dr. Marie Ygouf, Dr. Vanessa Bailey and the Roman CGI Project Science Team to develop post-processing tools in python and to perform tests on the most recent Roman CGI simulated data.

Background, Skills, Courses

Excellent coding skills, knowledge of python is preferred Strong willingness to learn

Major(s): Computer Science, Biology/Bioengineering

AO# **11582**
Project: **Machine Learning and Instrument Autonomy Group Capabilities Development Effort**

Background
The roots of the Machine Learning and Instrument Autonomy (MLIA) group at JPL go back over 3 decades. During these years the group has been responsible for some of the earliest instrument autonomy in a spacecraft setting, as well as numerous successful supporting endeavors with terrestrial science applications. This project will involve supporting the aggregation of these accomplishments, as part of an effort to determine future strategic efforts for the group.

Description
This project will involve reading historical papers written by the MLIA group, collating information about these projects and reaching out to group research leads to fill in missing information. The motivation of this work is to ensure group capabilities are well documented, characterized and accessible to future generations of group members. Successful culmination of the task will involve generating contents for the group website, and internal databases, about past group projects. This insight can additionally be distilled and analyzed by the student to help support key group decisions on which future proposals and capabilities the group would like to grow. Day to day activities will involve reading group documentation and reaching out to lead researchers on those tasks with targeted questions to make sure group resources such as the website and Wiki spaces are accurate and complete. As part of this effort, the chosen student will be exposed to a wide variety of machine learning research capabilities and researchers to support their understanding of the academic field and career networking.

Background, Skills, Courses
An interest in reading scientific papers and collecting supplemental information from authors is desired. Website development skills are a significant plus.

Major(s): Computer Science, Information Systems/Technology, Library Science

AO# **11580**
Project: **Data Analysis of Batteries for Deep Space Missions**

Background
Future missions to the outer planetary moons will likely require energy provided by non-rechargeable batteries. These batteries must survive a long cruise time followed by exposure to radiation prior battery usage. Our group is studying the effects that long-term storage and radiation exposure have on battery performance. We are generating large raw data sets from several different experimental test protocols that require efficient analysis to reveal trends in the performance of the battery.

Description
The student will work with the JPL mentors to develop data analysis techniques including writing Excel macros to help the research team manage the data sets, and to effectively search for trends in the experimental data.

Background, Skills, Courses
Desire students that have taken at least one term of calculus.

Major(s): Mathematics, Computer Engineering

AO# **11571**
Project: **Analysis and Archiving of Near- and Mid-Infrared Observations of Jupiter and Saturn**

Background

Images and spectra of Jupiter and Saturn from visible, near- and mid-infrared instruments are sensitive to temperatures, abundances of a major condensate (ammonia) opacity of clouds and the variability of the molecular para vs. ortho-H2 ratio. These define the fundamental state of the atmosphere and constrain its dynamics. This research will focus on observations obtained from a variety of instruments used at large professional telescopes: NASA's Infrared Telescope Facility, Gemini North and South Telescopes, ESO's Very Large Telescope, and the Subaru Telescope, and the Juno mission images of Jupiter in reflected sunlight from the JunoCam instrument. The general objective of the specific tasks below will be to create fully reduced data from unreduced or partially reduced sets. In some cases, our objective is to format the data for input into an atmospheric retrieval code from which atmospheric properties will be derived.

Description

Several specific topics are available: 1. Analyze images made by the JunoCam imaging instrument on the Juno mission. To some extent, each of these tasks may require work with a transformation of the images to a latitude-longitude map using "ISIS3" software supplied by the US Geological Survey. A task associated with any of the following science goals could be automating this process. 1a. Understanding the dynamics of Jupiter's high northern latitudes. Search for sequential JunoCam observations near the north polar region with time dependence in order to determine the wind field of this region and the degree to which Jupiter's winds flow east-west to something different, possibly chaotic, but with a component of east-west prevalent winds. 1b. Search for and measure hazes in Jupiter's atmosphere, i.e. particles lying above Jupiter's main cloud deck. This will include the following, each of which could be a separate research task: (i) Identify clouds near Jupiter's dusk region that appear to have different shapes or positions that are wavelength dependent, creating a "rainbow" appearance. (ii) Identify persistent hazes, their level of transparency and changes in their morphology between successive Juno orbits to determine a velocity. Where possible, verify their appearance in a special filter that is sensitive to high-altitude particles. (iii). Survey JunoCam images of Jupiter's horizons to detect and measure "detached" haze layers. 1c. Continue previous work in the first 20 orbits of the Juno mission to detect and measure small-scale waves in Jupiter's atmosphere to increase statistical evidence for their latitudinal distribution and association with larger features. 2. Archive images of Jupiter made in support of the Juno mission. To fulfill contractual obligations to NASA, we need to archive our infrared imaging of Jupiter from various instruments from the last four years of mid-infrared images and two decades of near-infrared images with NASA's Planetary Data System (PDS). These data must be accompanied by required ancillary files in a specific PDS format. The goal of this work is to collect copies of the images and related files into a single location with a specific nomenclature and submit these to review by the PDS. Aspects of this work could be done concurrently with other student work on the long-term variability of Jupiter. 3. Reduce near-infrared images of Jupiter. We will be acquiring a large volume of observations of Jupiter that are designed to support observations from instruments on the Juno spacecraft. We want to reduce the data and, to the extent possible, analyze the results. There are several objectives in this broad category. 3a. Develop quasi-automated software for reducing near-infrared imaging observations. The basis of this software exists in the Interactive Data Language (IDL), but the order of operations must be reversed at one stage and a subtracted pair of images must be reversed at another stage. 3b. Develop software for combining Mercator maps derived from images taken at different times as the planet rotates; use these to create full maps of Jupiter over all longitudes, as well as to polar project those

maps – for example - to investigate correlations between different phenomena in the neutral atmosphere and the aurora. 3c. Create an absolute calibration of the reflectivity of these images by referencing the flux from measured standards stars. Compare this with a calibration scaled to spacecraft observations of the near-infrared spectrum. 3d. Measure the distribution of cloud properties in the atmosphere with near-infrared reflectivity, including high-resolution adaptive-optics stabilized images. Use these data to characterize the chemistry and dynamics of the atmosphere, associated with specific atmospheric features e.g. polar hazes, the Great Red Spot and its environment, their evolution and their relationship with temperatures and winds. 3e. Reduce scanned spectral observations of Jupiter that create a hypercube of data (two dimensions of imaging and one of wavelength), deriving spectra at each pixel of the slit. Analyze these for properties of clouds and hazes, comparing them with models. 4. Examine the long-term variability of longitudinally averaged temperatures and other properties in Jupiter to create accurate and self-consistent calibrations of all data from a variety of telescopes. Extend a current program to input longitudinally averaged data over Jupiter's full disk to include observations at facilities where only a northern or a southern hemisphere of Jupiter could be captured. Format these data to be an input to an atmospheric retrieval program. Organize the output of this program to enable rapid plotting and correlation with previous studies and between different retrieved atmospheric properties.

Background, Skills, Courses

The data reduction programs are written in the Interactive Data Language (IDL, which is close to Matlab in format). The analysis code is written in FORTRAN. At least rudimentary knowledge of these (or willingness to learn before the beginning of the research) is highly recommended. Some programming experience is required. With a significant level of contribution, students are welcomed as co-authors on papers emerging from this research.

Major(s): Planetary Science, Astronomy/Astrophysics, Computer Sciences

AO# **11563**
Project: **Creating a data base for photonic devices**

Background
Photonic technologies/devices are fast growing and knowing who are the players with their capabilities and how they test them their devices.

Description
Create a directory/database on the existing foundries and their capabilities and testing, who are the experts, how they do qualification/testing of photonic devices. Such a database includes also photonic components which have been used or qualified for space applications including methods used for their screening and qualification.

Background, Skills, Courses
Understanding and interest in optics, especially the newer photonic Integrated Circuits (PIC)

Major(s): Electrical Engineering, Physics/Applied Physics

AO# **11555**
Project: **Evaluation of MATLAB/Simulink HDL Code Generation for Assurance**

Background
To support a new FPGA Assurance program, we are exploring ways to assure HDL code generated by MATLAB/Simulink.

Description
This project will explore the elements of MATLAB/Simulink models and how they are mapped to HDL language elements, culminating in a roadshow presentation targeted to scientists and algorithm designers.

Background, Skills, Courses
Exposure to MATLAB/Simulink in at least one course and/or on-the-job experience Exposure to VHDL and/or Verilog HDL

Major(s): Computer Engineering, Computer Science

AO# **11554**
Project: **AI Assistant for Driver Safety**

Background

The Intelligent Sensor Processing Object Recognition & Tracking Systems (iSports) Lab in Bio-Inspired Technologies & Systems Group (349B) is conducting research in AI and Deep Learning for automatic object recognition, speech recognition, image understanding and Artificial Intelligence for assisting first responders to avoid accidents. The research is applicable to the assistance of astronauts and autonomous guidance of spacecraft landing/docking/rendezvous and hazard avoidance.

Description

We are interested in sponsoring 6 intern students for the deep learning and AI projects. The candidates will help to train intelligent computer programs to automatically detect, recognize, and track objects from various data sources. The AI system helps the robots and autonomous vehicles to understand the environment, and perform autonomous maneuvers.

Background, Skills, Courses

Critical thinking, creativity, curiosity, good communication skills, C/C++, Python, TensorFlow, Torch, Caffe, etc.; Courses: normal undergraduate math, electrical and computer engineering Useful, but not required: knowledge of image processing, neural networks, computer vision.

Major(s): Electrical Engineering and Computer Science, Computer Engineering, Applied Math

AO# **11534**
Project: **Next generation of miniature high-resolution spectrometers**

Background
Water is important for us! We drink water but also, in general life as we know it, depends on water. Studying the creation of water and its transfer in the Solar System contains information about the evolution of Solar System and how did water end up on the Earth. One of the approaches to study where the water on Earth came from is by studying OD/OH ration on various objects in the Solar System. However, the OD/OH emission in 308nm is very faint and a very difficult measurement to make, often required in situ mass spectrometers to be sent to the astronomical bodies with spacecraft. In my lab at JPL, we have developed a new technology that has can potentially make these measurements remotely using small aperture telescopes.

Description
The student will be involved in developing the mechanical frame and the vibration simulation for flying the instrument on sounding rockets and on future missions. Also depending on the background of the student he/she would be involved in reviewing the mechanical and system requirements and integration strategy.

Background, Skills, Courses
Mechanical engineering and analysis Electrical engineering and analysis Programming 3D printing

Major(s): Mechanical Engineering, Physics/Applied Physics, Engineering