# 2020 Workshop: Data Science in CEDAR

Long title
Data Science in CEDAR: Progress, Capacity-Building, and Traversing Disciplines
Conveners
Ryan McGranaghan
Asti Bhatt
Bharat Kunduri
James Ahrens
Enrico Camporeale
Marcin Pilinsk
Description

Characterizing the geospace environment requires measurements from several regions within the geospace. Fortunately, data to advance the scientific understanding of the geospace environment are growing across the four V's of 'big data': 1) Volume; 2) Variety; 3) Veracity (i.e., uncertainty); and 4) Velocity. This growth represents both a challenge, to efficiently and comprehensively utilize these data, and an opportunity for new discovery by embracing new technologies and analysis capabilities that scale well to the geospace environment. These developments have revolutionized the creation of new scientific insights from data through the union of statistics, computer science, applied mathematics, and visualization, i.e., data science.

Numerous previous efforts to identify, understand, and progress data science in the CEDAR community set the stage for a new session that will not only share the latest progress, but will also create a more firm and lasting structure on which CEDAR data science will flourish.

The proposed workshop is a timely effort to sustain and amplify momentum from several previous workshops with a data science focus (see a list here).

Our specific objectives will be to:

1. Identify problems and challenges that can immediately be addressed using data science tools (i.e., the compelling and transformational 'use cases');

2. Promote interaction and collaboration between the CEDAR community and related disciplines (e.g., Earth Science);
3. Improve agility and capability within CEDAR science through embracing newer technologies and sound digital data scholarship; and
4. Grow methodology transfer to enhance CEDAR science.

Agenda

**First hour (June 23; 11 AM-12 PM MT)**

(11:00-11:10) Introduction - Ryan McGranaghan: CEDAR Data Science Evolution and Community-Building & Innovation to virtual interactions - new modes of collaboration (pdf)

(11:10-11:18) Janet Kozyra “Using Amazon Web Services to understand and predict space weather superstorms: A pilot study”

(11:18-11:26) Hyunju Connor “Prediction of global geomagnetic field disturbances during the 9-March-2012 geomagnetic storm using Recurrent Neural Network" (pdf)

(11:26-11:34) Yang Pan: “Machine learning method to recover the missing data of the global total electron content maps” (pdf)

(11:34-11:42) Jenny Yang: “New discoveries about time lags for solar wind-geospace connections in Global Navigation Satellite Signal (GNSS) data”

(11:42-11:50) Garima Malhotra (FDL GNSS 2019): “Breakthroughs in predicting Global Navigation Satellite Signal (GNSS) scintillation” (pdf)

(11:50-11:58) Roxana Bujack - “Los Alamos National Laboratories (LANL) Data Science at Scale plans for space science”

(12:35-1:00) Panelists introduce each theme

1. Handling large volumes of data: (pdf) Shea Hess-Weber (Stanford)

2. The data analysis methods that are changing (CEDAR) science: (pdf) Karthik Venkataramani (ASTRA)

3. Creating platforms for visualizing and interacting with data: (pdf) anet Green (Space Hazards Application)

4. Increasing the value of data:  (pdf) Adam Kellerman (UCLA)

5. Broadening participation/Using data science to improve collaboration in the virtual setting: Abigail Azari (SSL)

(1:00-1:25) Breakouts

Handling large volumes of data

Convener: Marcin Pilinski

Student moderator: Deepthi Ayyagari

The data analysis methods that are changing (CEDAR) science

Convener: Enrico Camporeale

Student moderator: Md Nurul Huda & Brian Swiger

Creating platforms for visualizing and interacting with data

Convener: Asti Bhatt & Jim Ahrens

Student moderator: Diaby Abdel Aziz & Komal Kumari

Increasing the value of data (e.g., using data science to formalize relationships between data)

Convener: Ryan McGranaghan

Student moderator: Elizabeth Hernandez

Broadening participation/Using data science to improve collaboration in the virtual setting

Convener: Bharat Kunduri

Student moderator: Sarah Luettgen and Sophie Phillips

View PDF