



# Energy A.I. 2024 4<sup>th</sup> Annual Hackathon Introduction

Dr. Michael Pyrcz and Dr. John Foster  
Energy A.I. 2021-2024 Hackathon Hosts  
Hildebrand Department of Petroleum and Geosystems Engineering

Dr. Matt Balhoff  
Sponsor and Advisor  
Chair of the Hildebrand Department of Petroleum and Geosystems Engineering

Elnara Rustamzade and Fehmi Özbayrak  
Hackathon Architects and Mentors, Graduate students in PGE

Rowan Halliday, Heba Abdel-Rahim, Gabby Banales, Coordinators of Chaos



## Appreciation

*Appreciation to the student participants, the hackers!*

*Thank you for your enthusiasm!*



**None of this would be possible without our sponsors. Thank you for supporting Energy Data Science Education!**

## Thank You to Our Sponsors

---

### Platinum



### Silver

**Elizabeth Huth Coates  
Charitable Foundation  
of 1992**



### Bronze





## Who's Running this Show?

Professor Michael Pyrcz (aka GeostatsGuy)  
Hackathon Host



**Michael Pyrcz**

@GeostatsGuy

Assoc. #Prof @UTAustin @CockrellSchool @txgeosciences @daytum\_io |  
#geostatistics #DataAnalytics #DataScience #MachineLearning #author #dad  
#github #YouTube



**John Foster**

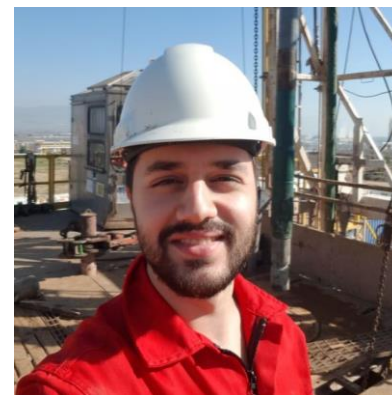
@johntfoster



Professor John T. Foster  
Hackathon Host



Elnara Rustamzade, PhD Candidate PGE  
Hackathon Architect, Volunteer



Fehmi Özbayrak, PhD Candidate PGE  
Hackathon Architect



## Appreciation

Rowan Halliday  
Coordination



Stacia Miller  
Communication, Promotion,  
Photography



Samantha Rabinowitz  
Purchasing

Professor Matt Balhoff  
PGE Chair  
Strong support and engagement



Heba Abdel-Rahim  
Business Development, Student Support



Gabby Banales  
Organizing, Student Engagement





## Welcome Message



*Professor Matt Balhoff*  
*Chair of the Hildebrand Department of Petroleum and Geosystems Engineering*





## Petroleum / Mining / Spatial Engineering and Science Leadership in the Fourth Paradigm

*'We are the original data-driven science, we have been big data long before tech learned about big data!'*

1930-1940s

1950-1960s

1980-1990s

>1990s

Probability and  
Stationarity  
Kolmogorov

Volume  
Variance in  
Mining  
Krige

Geostatistics  
Mathematical  
Morphology  
Matheron

Applications in Oil  
and Gas,  
Environmental  
Journel, Verly, Deutsch

Spatial Statistics, Big  
Data Analytics and  
Machine Learning

*'Complicated, heterogeneous, sparsely sampled, vast systems with complicated physics and high value decisions.'*



## What is a Hackathon?

*'an event in which a large number of people meet to engage in collaborative computer programming.'*

Dictionary.com

*'The goal of a hackathon is to create functioning software or hardware by the end of the event'*

Wikipedia





## Mentors

Eric Qian

Amy Rueve



Julio De La Colina

Shane Prochnow

Guillaume Dulac

Eduardo

Maldonado-Cruz

Julian Salazar

Alena

Grechishnikova

Yogashri Pradhan



Preetika Srivastava



Matthias Imhof

Marcelo Jimenez

Hector Martinez



Tyrel Krohn





## Judges

**Matt Duke**

General Manager  
Geology Department  
Chevron Technology



**Susan Howe**

President



**Andy Flowers**

Director of Advanced  
Analytics



**Kumar Lakshmipathi**

Principal Solutions Architect



**Doug  
McMaster**

VP Product



**Kathryn Briggs**

Completions Engineering  
Coordinator



**Manuel E Rosales**

Data Analytics Supervisor





## Who is Here to Build?

TBD Teams, from UT Austin Cockrell, Jackson and Natural Sciences  
Including:

ChatPGE

Rubber Duckies

Cool Cat Fracking Party

Bitumen Bandits

HB Analytics

Banana Ducks

Desi Drillers

Wild Cats

Flat Earth Society

6

Petroleum and Geosystems Engineering, Operations Research, Mechanical Engineering, Geological Sciences, Materials Science Engineering, Electrical and Computer Engineering, Data Science, Computer Science, Aerospace Engineering, Mathematics, McCombs, Math, Physics, Economics, Neuroscience,



## The Hackathon Rules

**Submit to GitHub by 12:00 noon, January 21<sup>st</sup>:**

1. Well-documented Python workflow in a Jupyter Notebook. **See template in resources folder.**
2. Results as a .csv DataFrame ESPs (fail in 30 days = 1, otherwise = 0).  
Named: solution.csv, **use the file in data folder.**
3. Short presentation with executive summary, goals workflow choices and defense, results and discussion. Every team member participates in the presentation. **Use template in resources.**

**Participation:** All team members contribute to the above products. There are various roles! Participate in sessions.

**Coding:** Use only open source and methods / workflows developed during the hackathon. Provide code for testing and scoring. All code submitted in Jupyter Notebook. **Readable code!**



## The Hackathon Rules

**Our academic staff have volunteered to assist over the weekend.**

- **Please let them know that we appreciate.**
- **Let's all do house keeping, clean up and disposal of recycling and waste in your work area and general areas, as we go. Take out the trash.**
- **Share contacts in case you get locked out.**



## The Hackathon Rules

**Participate in the Scheduled Workshops and Working Sessions**

**Treat All other Hackers, Hosts, Mentors, Judges, Coordinators with the utmost respect.**

**Cite all code used from other sources in your workflows.**

Pyrcz, M.J. (2020) GeostatsPy 0.0.19 [Source code]. <https://github.com/GeostatsGuy/GeostatsPy>  
Foster, J.T., (2015) 1DPDpy 1.0 [source code]. <http://dx.doi.org/10.5281/zenodo.15795>

**Work Hard, Learn and Have Fun!**



## The Hackathon Rules

**The data has been sanitized. Do not attempt to hack the source!**

But, here's a hint...

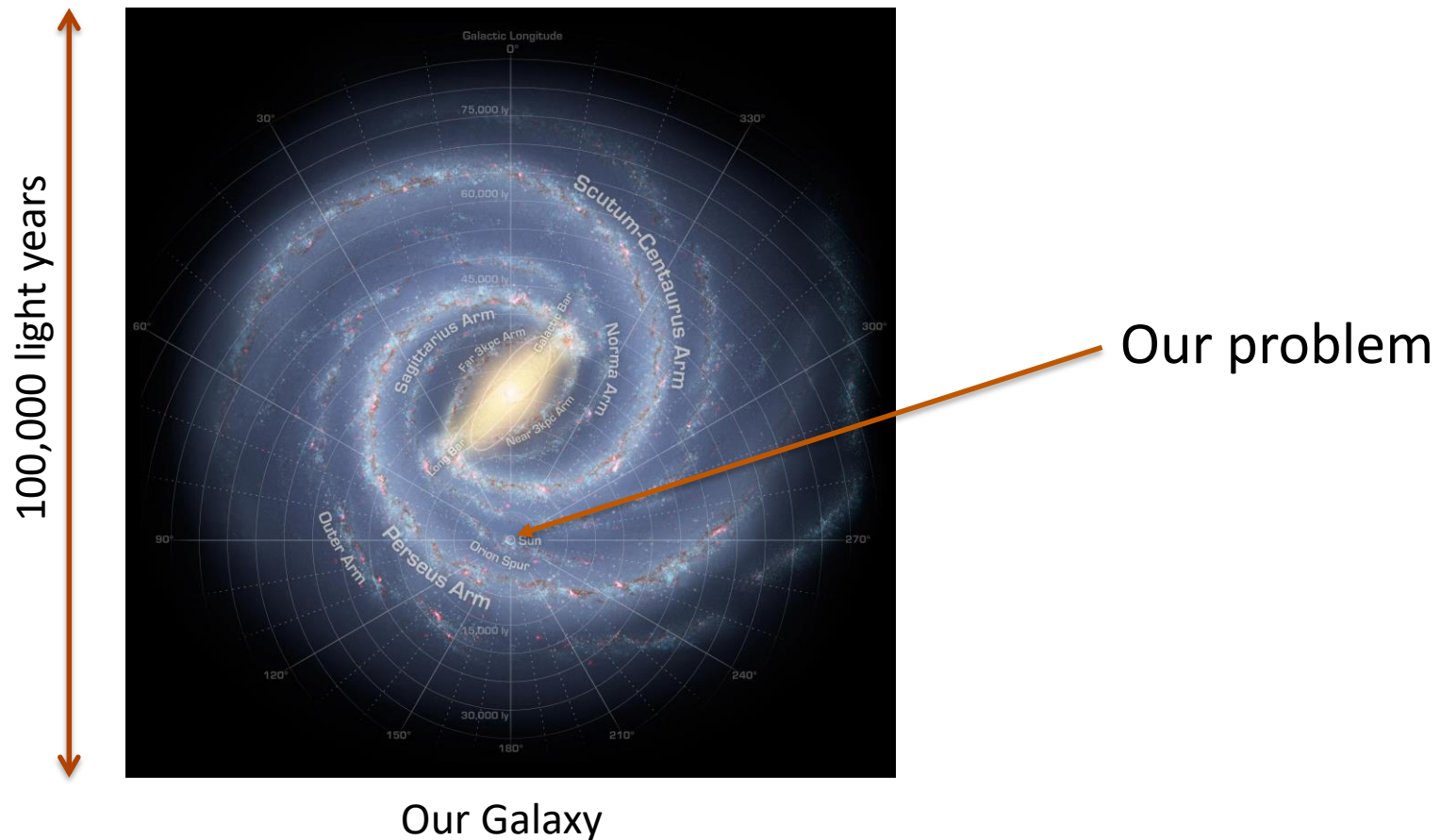




## The Hackathon Rules

**The data has been sanitized. Do not attempt to hack the source!**

We can provide the following general location of the data set.





## Hackathon Team Scoring

### **Results: 75% - Results, Results, Results!**

- Average of rank transform of accuracy error measure and uncertainty model goodness over all groups.

### **Presentation: 20% - and We Must Be Able to Communicate Our Work!**

- Executive summary, project goals, workflow description, results and discussion, finish on time

### **Workflow: 5% - Others Must Understand our Work for Adoption!**

- Scoring metrics: readable code, efficient code, documentation of steps

Use the provided templates for results, workflow and presentation. Follow the submission guidelines and submit on time.

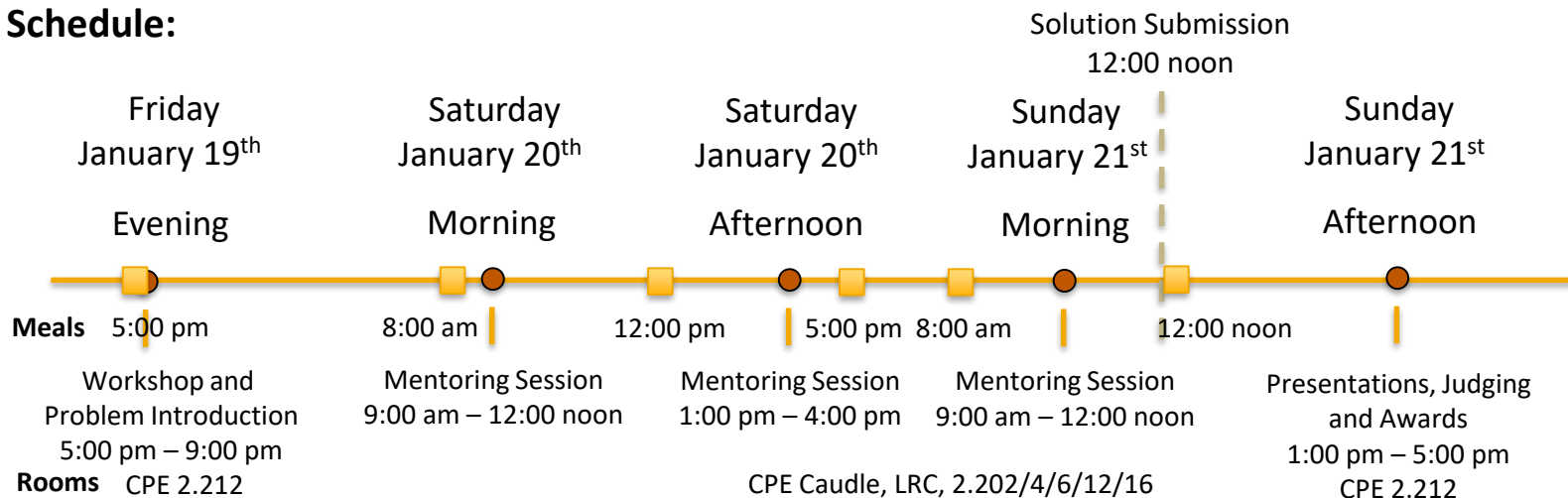


## The Plan at 30,000 ft

Apply Data-driven Solutions with Data Analytics and Machine Learning



### Schedule:



Enhance Learning



Promote collaboration & data-driven science



Inspire Innovation

### Teams:

Register teams of 3-4, which can include students from other departments and schools (UT EID required), but one student must be a UT PGE student and one 1 undergraduate student per team (could be the same student).

### Awards:

Winners of the A.I. Hackathon will be awarded bragging rights and \$5,000 for first place, \$2,500 for second place, \$1,000 for third place and \$500 for fourth place teams.



## The Plan for Today's Workshop

### **DAY 1 / March 25th - Energy A.I. Hackathon 2024 Workshop Schedule**

5 pm – 5:15 pm: Hackathon Welcome, Introduction and Review Plan and Rules, Prof. Balhoff / Prof. Pyrcz

5:15-6:30 pm: Essential Energy Data Science, Numpy, Pandas, Git - Prof. Foster

6:30 – 7:00 pm: Feature Importance, Engineering and Selection, Multivariate Analysis and Shapley Values  
Prof. Pyrcz

7:00 – 7:30 pm: Uncertainty Models – Prof. Pyrcz

7:30 – 8:00 pm: Machine Learning Basics, Train and Tune Overview of Methods - Prof. Pyrcz

8:00 – 8:30 pm: Machine Learning in Python, scikit-learn and TensorFlow Packages – Prof. Foster

8:30 – 9:00 pm: Introduce the Energy A.I. Hackathon Problem and Mystery Data Set – Prof. Pyrcz / Prof. Foster

9:00 pm - : Teams Break-out for Initial Data Review and Planning



## Top Teams from Last Year

### DrillPy, First Place



### Hackelope, 2<sup>nd</sup> Place





## Don't Jump to Complexity



Deep learning generated image of Professor Pyrcz.



Deep learning generated image of Professor Foster.