



# Energy A.I. 2022 2<sup>nd</sup> Annual Hackathon Introduction

Dr. Michael Pyrcz and Dr. John Foster

Energy A.I. 2022 Hackathon Hosts

Hildebrand Department of Petroleum and Geosystems Engineering

Dr. Jon Olson

Sponsor and Advisor

Chair of the Hildebrand Department of Petroleum and Geosystems Engineering

Wen Pan and Elnara Rustamzade

Hackathon Architects and Mentors, Graduate students in PGE

Gabby Banales Sara Hernando and Tracey Wilson

Coordinators



## Appreciation

*Appreciation to the student participants, the hackers!*

*Thank you for your enthusiasm!*



## Appreciation

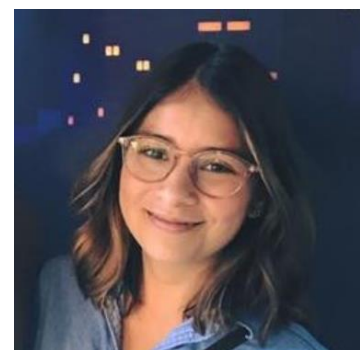


Gabby Banales  
Organizing, Student Engagement



Tracey Wilson  
Organizing, Student Support

Professor Jon Olson  
PGE Chair  
Strong support and engagement



Sara Hernando  
Business Development, Student  
Support



## Appreciation to Our Sponsors

None of this would be possible without our sponsors.

Thank you for supporting Energy Data Science Education!

### THANK YOU TO OUR SPONSORS

Elizabeth Huth Coates  
Charitable Foundation  
of 1992



PIONEER  
NATURAL RESOURCES





## Welcome Message



*Professor Jon Olson*  
*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  
Kriging

Geostatistics  
Mathematical  
Morphology  
Matheron

Applications in Oil  
and Gas,  
Environmental  
Journal, 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



## Who's Running this Show?

Professor Michael Pyrcz (aka GeostatsGuy)  
Hackathon Host



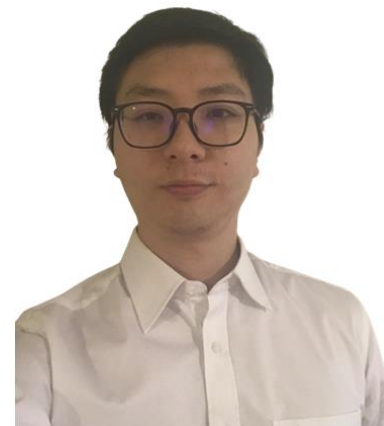
**Michael Pyrcz**  
@GeostatsGuy

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



**John Foster**  
@johntfoster

Professor John T. Foster  
Hackathon Host



Wen Pan, PhD Candidate PGE  
Hackathon Architect



Elnara Rustamzade, PhD Candidate PGE  
Hackathon Architect, Volunteer





## Mentors

Sercan Gul



Ilyas Iyoob



Fabien Laugier



Thatcher Thornberry



Shane Prochnow



Matthias Imhof



Alena Grechishnikova



Nickolas Raterman



Nkem Egboga



Elizaveta Onegova



Allison Gilmore





## Judges

Ben Amaba

CTO Cloud & Cognition



Kumar Lakshmipathi

Principal Architect



David Holmes

CTO Energy



Sarita Salunke

Team Leader



Doug McMaster

VP Product



Vikram Jayaram

Head of Data Science





## Who is Here to Build?

21 Teams, from UT Austin Cockrell, Jackson and Natural Sciences

TBD

Petroleum and Geosystems Engineering, Operations Research, Mechanical Engineering, Geological Sciences, Materials Science Engineering, Electrical and Computer Engineering, Data Science, Computer Science, Aerospace Engineering, Mathematics,

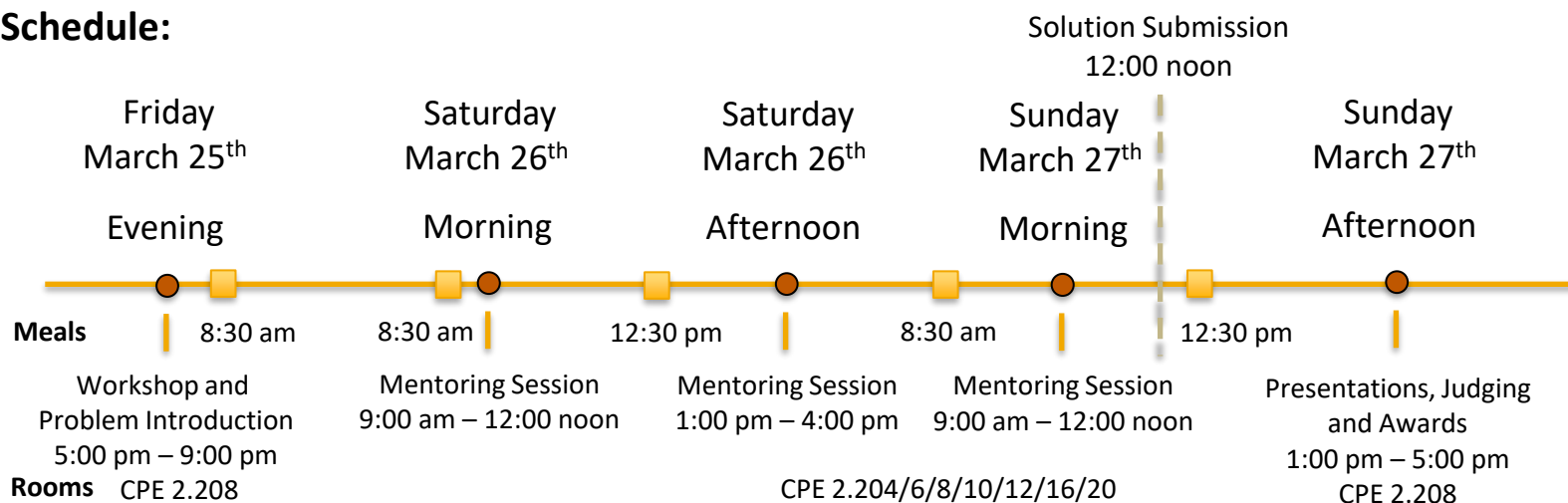


## The Plan at 30,000 ft



## 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. At least 1 undergraduate student per team.

### 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 Hackathon Rules

### **Submit to GitHum by Noon March 27<sup>th</sup>:**

1. Well-documented Python workflow in a Jupyter Notebook
2. Results as a .csv DataFrame with 3 well locations (x, y, unit (upper or lower) and estimated 2 year cumulative production.
3. Short presentation with executive summary, goals workflow choices and defense, results and discussion. Every team member participates in the presentation.

**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 by during the hackathon. Provide code for testing and scoring. All code in Python, Jupyter Notebook. **Readable code!**



## The Hackathon Rules

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

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

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

**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>

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



## Hackathon Team Scoring

### **Results: 60% - Results, Results, Results!**

- Average of rank transform of accuracy error measure and of sum of 2-year cumulative production for 3 proposed wells.

### **Presentation: 25% - and We Must Be Able to Communicate Our Work!**

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

### **Workflow: 15% - 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 for Today's Workshop

### **DAY 1 / March 25th - Energy A.I. Hackathon 2022 Workshop Schedule**

5 pm – 5:15 pm: Hackathon Welcome, Introduction and Review Plan and Rules, Prof. Olson / 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: Spatial Mapping and Spatial Data Analytics – 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





We are looking forward to a great event.