A CREWES Data Science / Machine Learning initiative

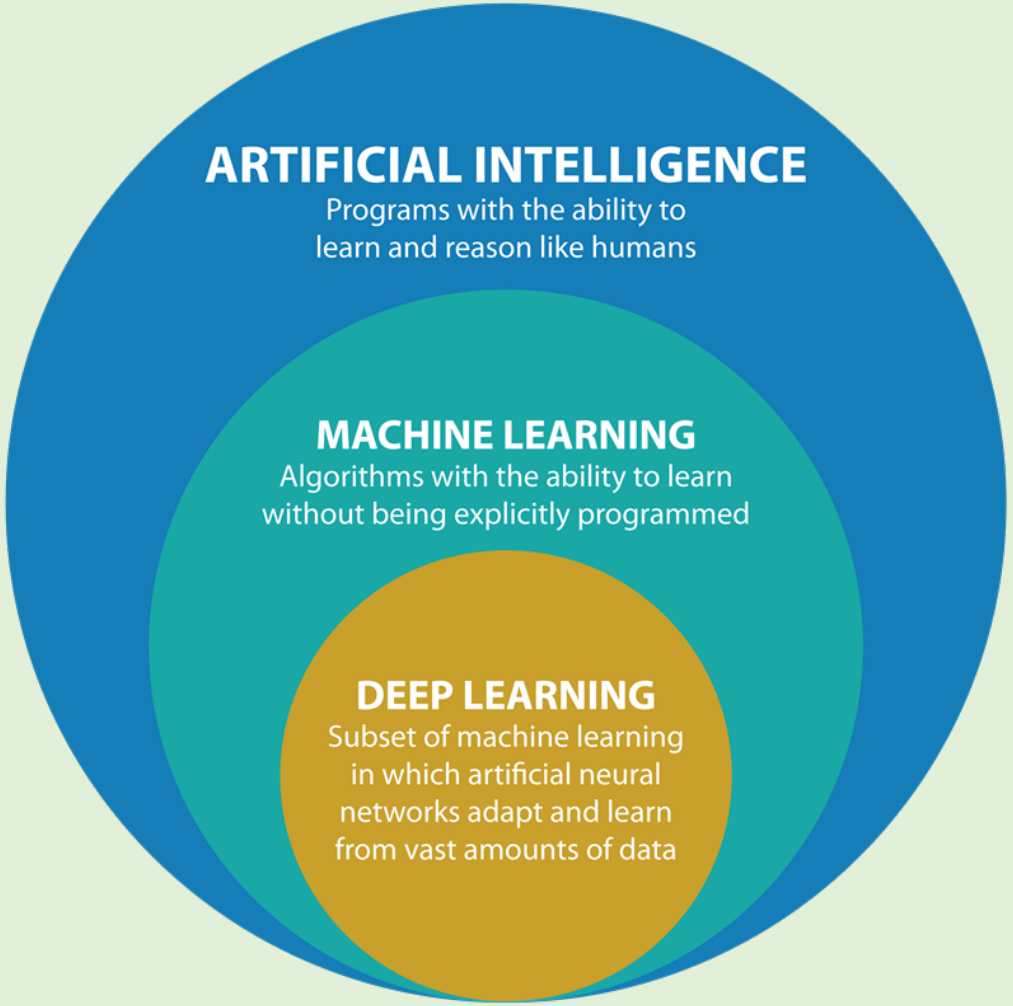
Marcelo Guarido, Daniel Trad, and Kris Innanen

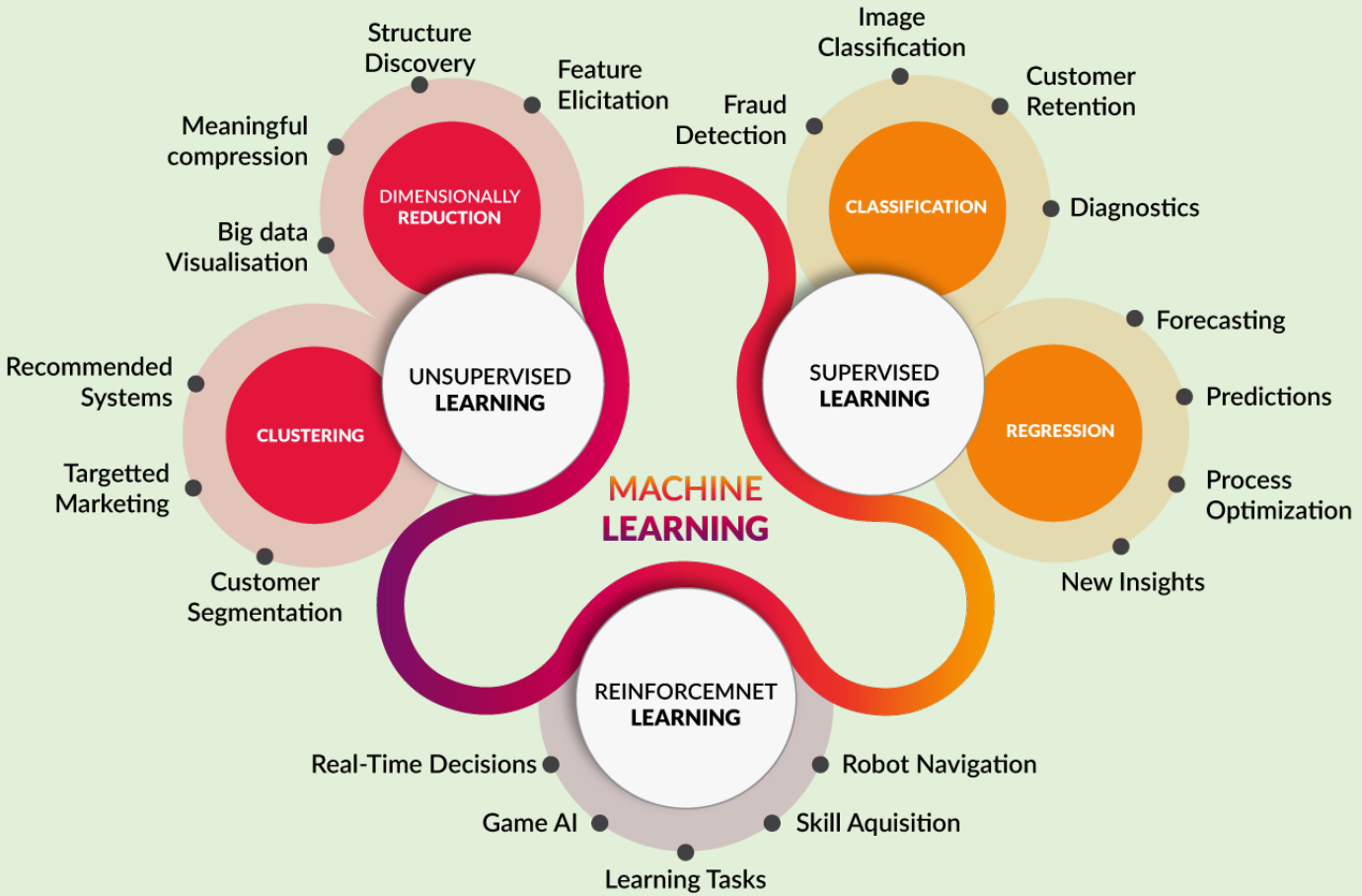
Sponsors Meeting
Banff, December 11th, 2019

Agenda

- What data science and machine learning are?
  - Introduction
  - Domain knowledge
- CREWES data science initiative
  - Types of projects
  - People
  - Courses
- 2019 projects
  - Machine Learning as a Tool to Predict the Mass of Oil from Well Logs
  - Using Natural Language Processing and Machine Learning to Predict Severe Injuries Classification in the Oil and Gas Industry

What Data Science and Machine Learning are?



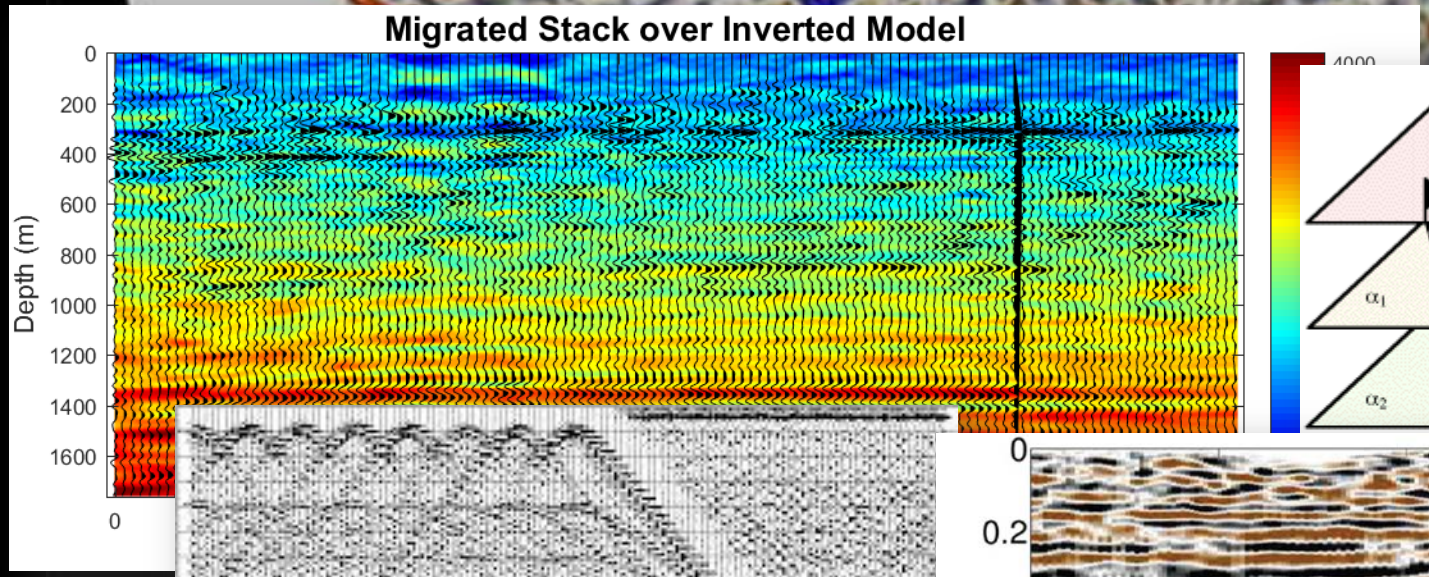


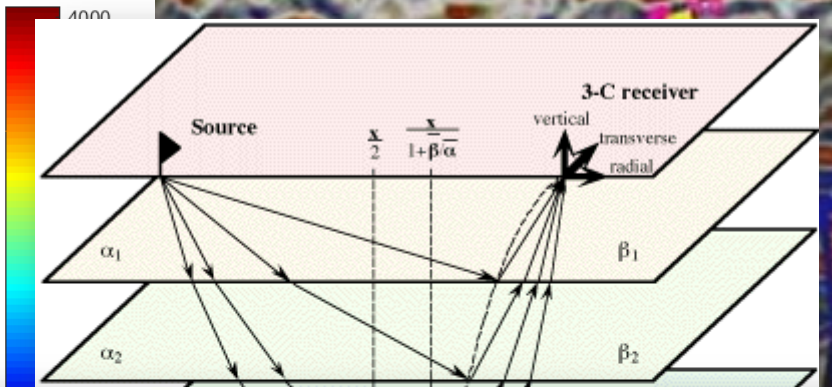
Figures from Argility website

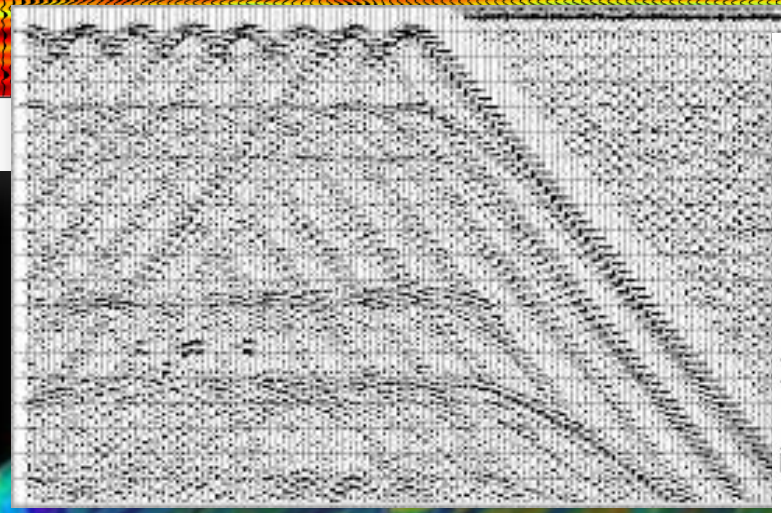
Computer Science
Machine Learning
Math & Statistics
Data Science
Traditional Software
Traditional Research
Subject Matter Expertise

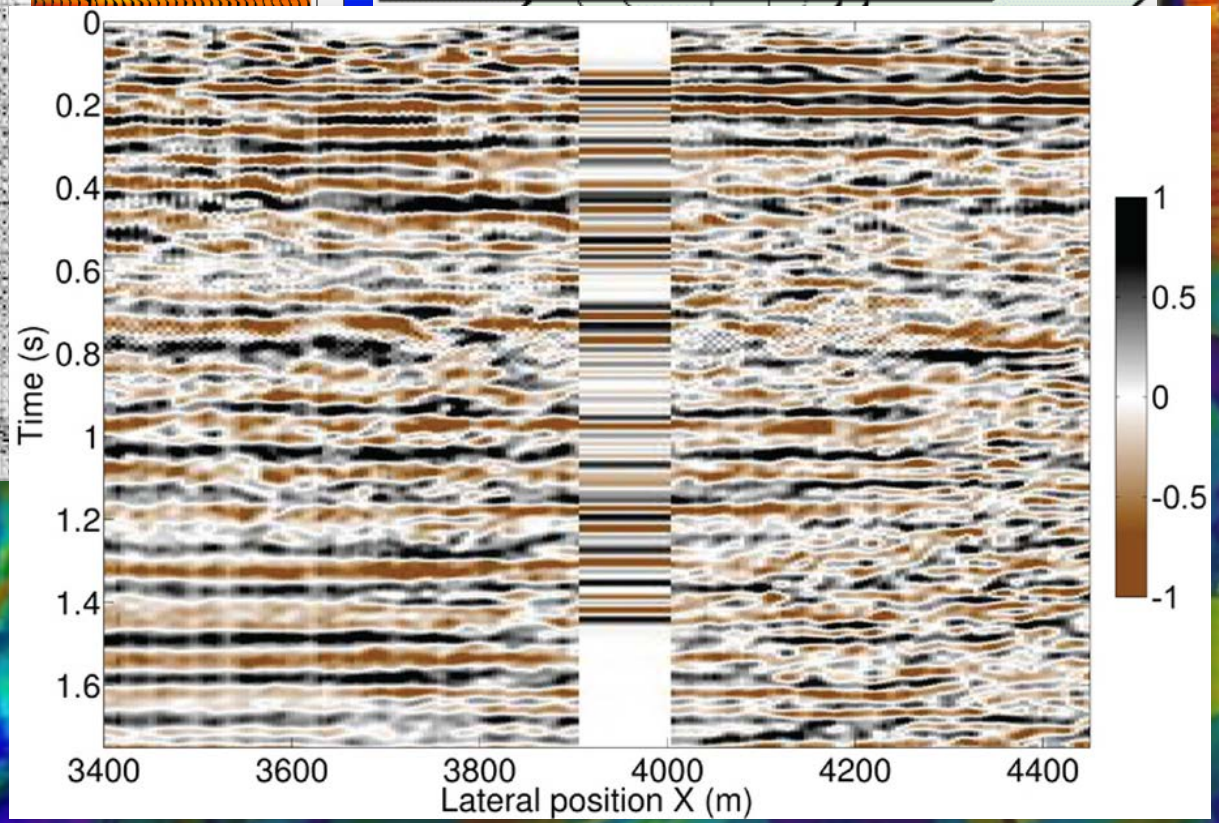
Source

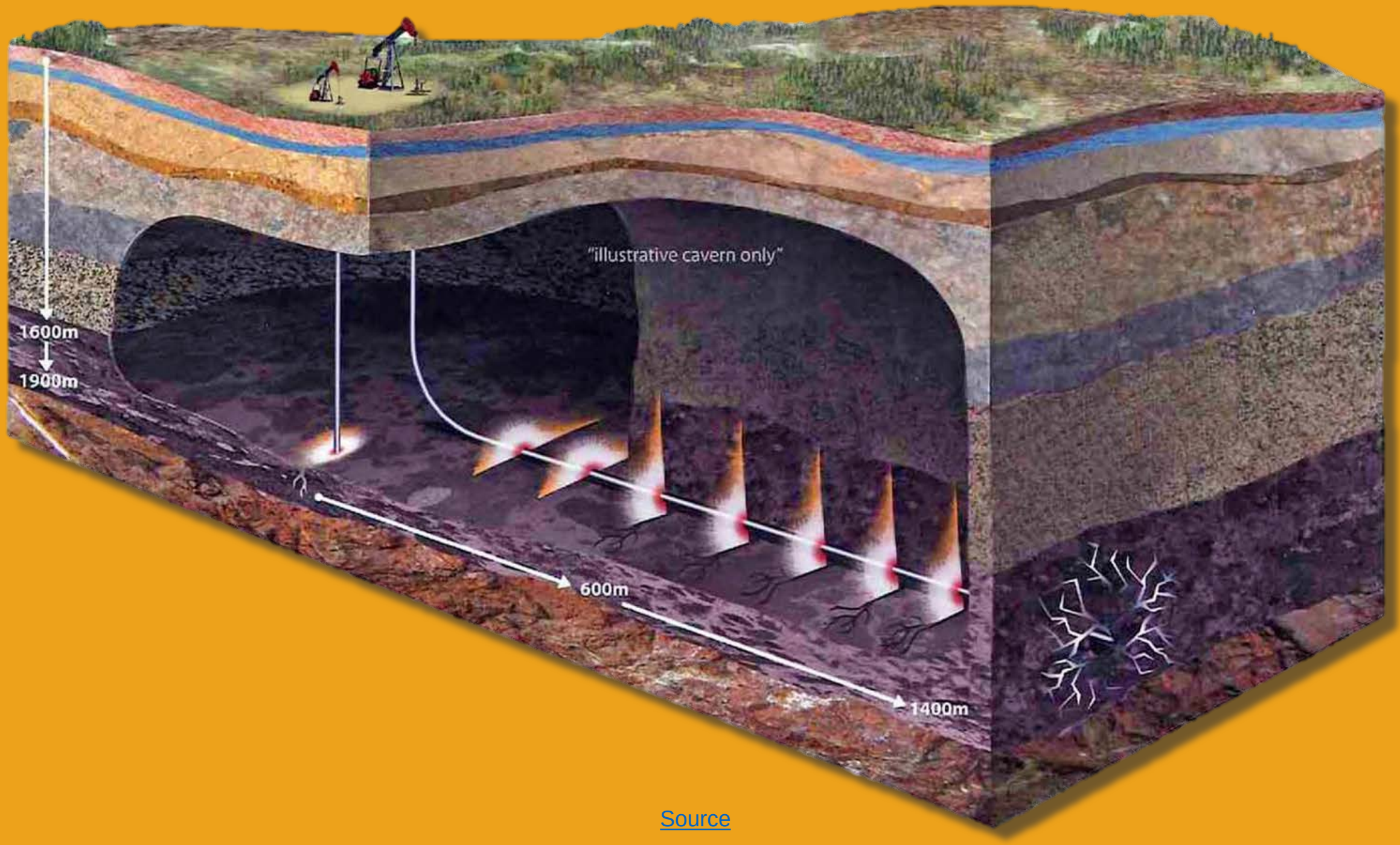
Source


Migrated Stack over Inverted Model
Depth (m)
0
200
400
600
800
1000
1200
1400
1600


Source
3-C receiver
vertical
transverse
radial
$\frac{x}{2}$
$\frac{x}{1+\beta/\alpha}$
α_1
β_1
α_2
β_2

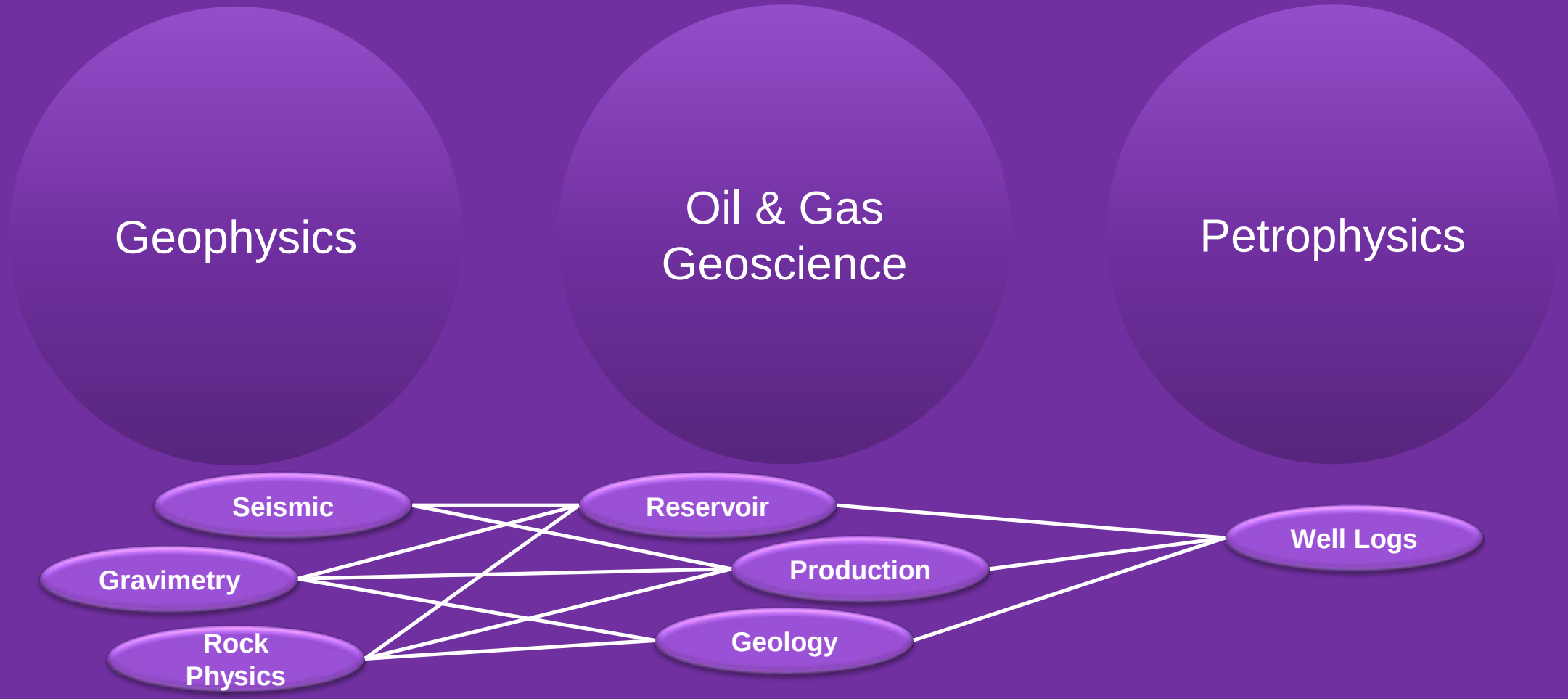



Time (s)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
3400
3600
3800
4000
4200
4400
Lateral position X (m)
1
0.5
0
-0.5
-1



Source

Our domain of knowledge



CREWES
Data Science Initiative

Types of Projects

Short-term:
1 to 4
months

POC (proof of concept)
Student/industry training

Mid-term:
5 to 12
months

Thesis chapter(s)
Publications

Long-term:
1+ years

Thesis
Publications

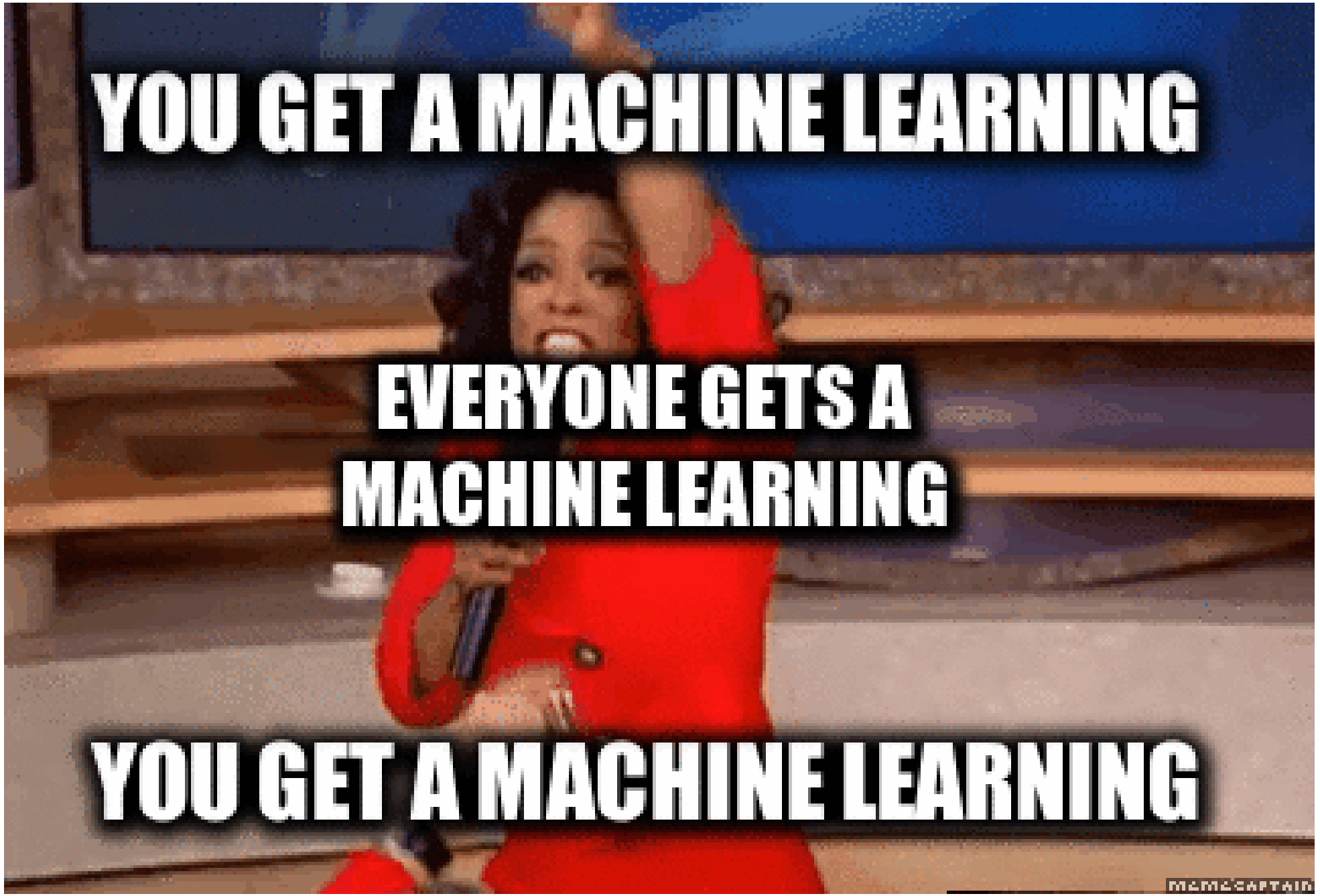
People

All CREWES grad students and staff*

* subject to supervisor/manager approval…

Courses

- Basic Programming
  - ✓ Python
  - ✓ R
  - ✓ MatLab
- ML Programming
  - ✓ General cases
- Custom ML Programming
  - ✓ Sponsor specific
- Mentorship

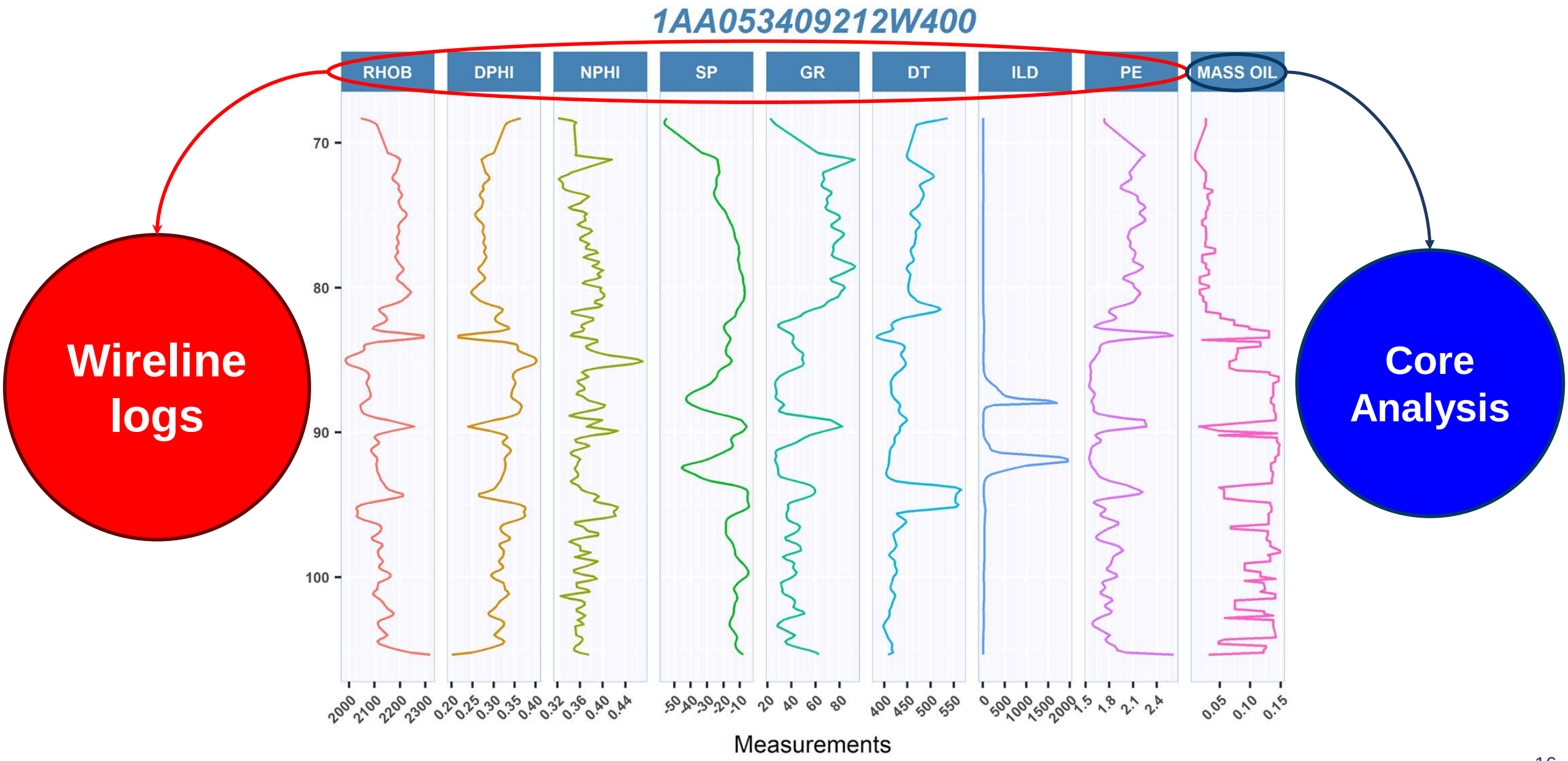

YOU GET A MACHINE LEARNING
EVERYONE GETS A
MACHINE LEARNING
YOU GET A MACHINE LEARNING

2019 Projects

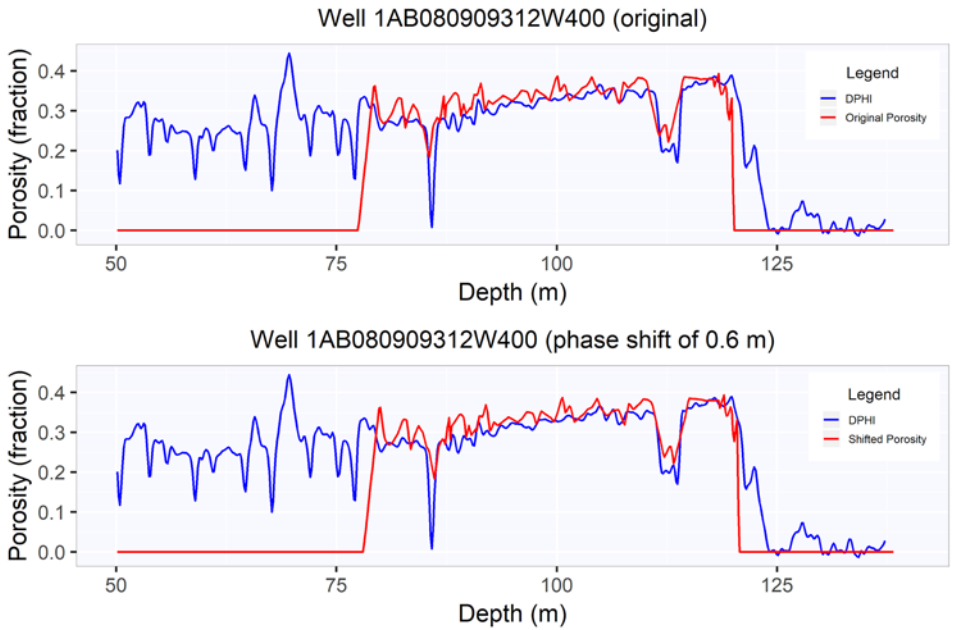
Machine Learning as a Tool to Predict the Mass of Oil from Well Logs

Marcelo Guarido and Daniel Trad

Predicting Mass of Oil

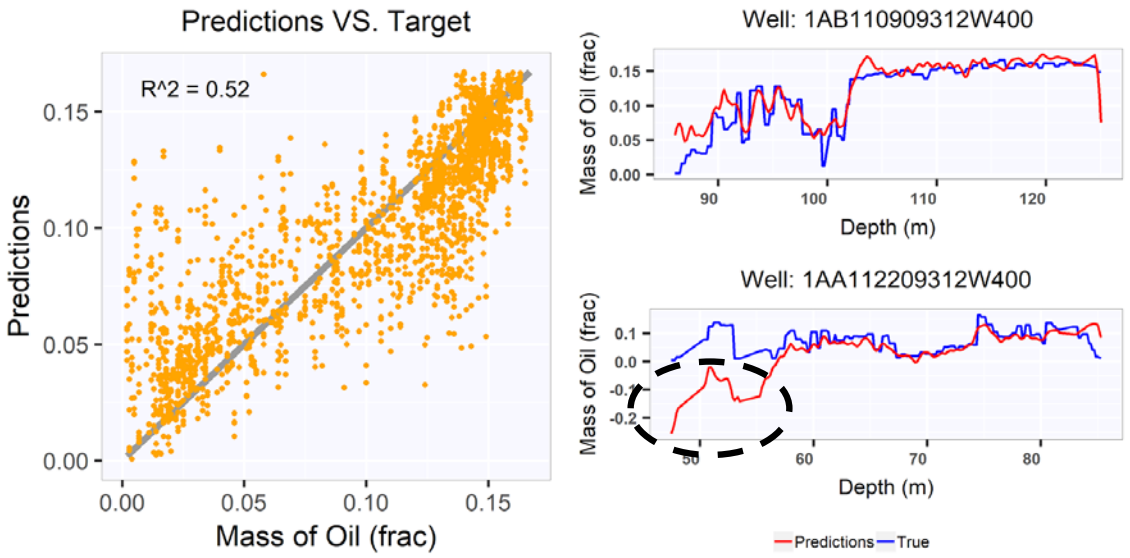


Predicting Mass of Oil



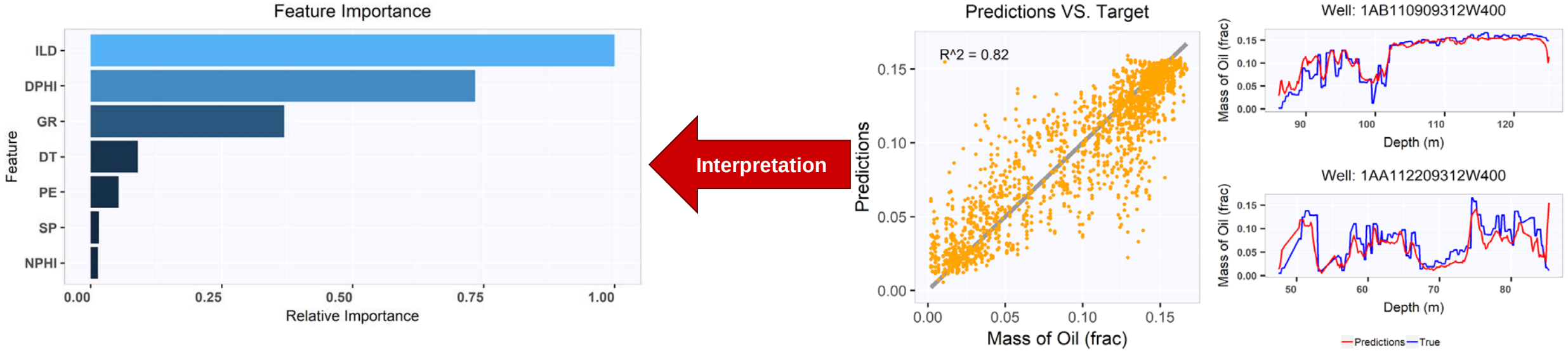
Predictions

Linear Regression



Interpretation

Gradient Boosting Regressor



Using Natural Language Processing and Machine Learning to Predict Severe Injuries Classification in the Oil and Gas Industry

Marcelo Guarido and Daniel Trad

Severe Injuries

Injury	Part of the Body	Description
Fractures	Lower leg(s)	Three correctional facility guards were escorting a restrained federal prison inmate when he became disruptive, requiring the use of force. Two guards and the inmate fell onto the Lieutenant's right leg, fracturing his fibula. He was transported to the hospital and released the following day.
Ampu		ad roller of surgically chine.

TF-IDF (Term Frequency-Inverse Document Frequency)

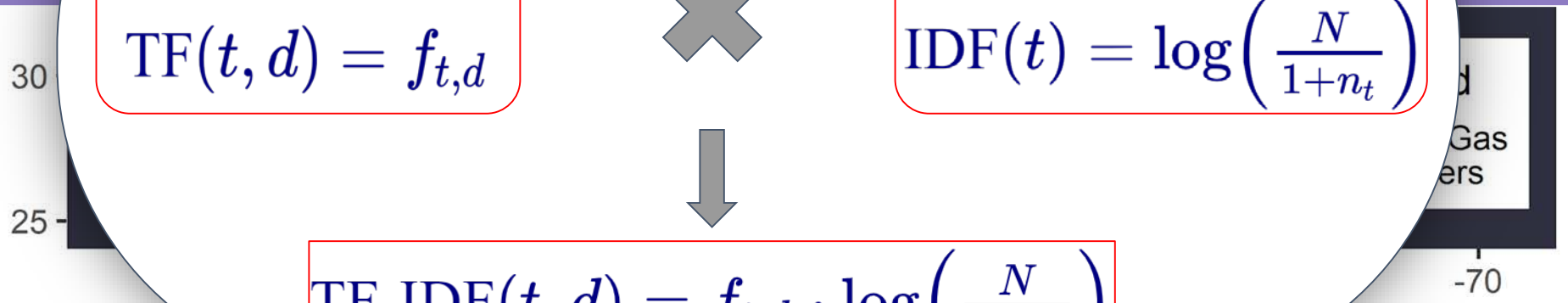
Common locally:

$$\mathrm{TF}(t,d) = f_{t,d}$$

Global rarity:

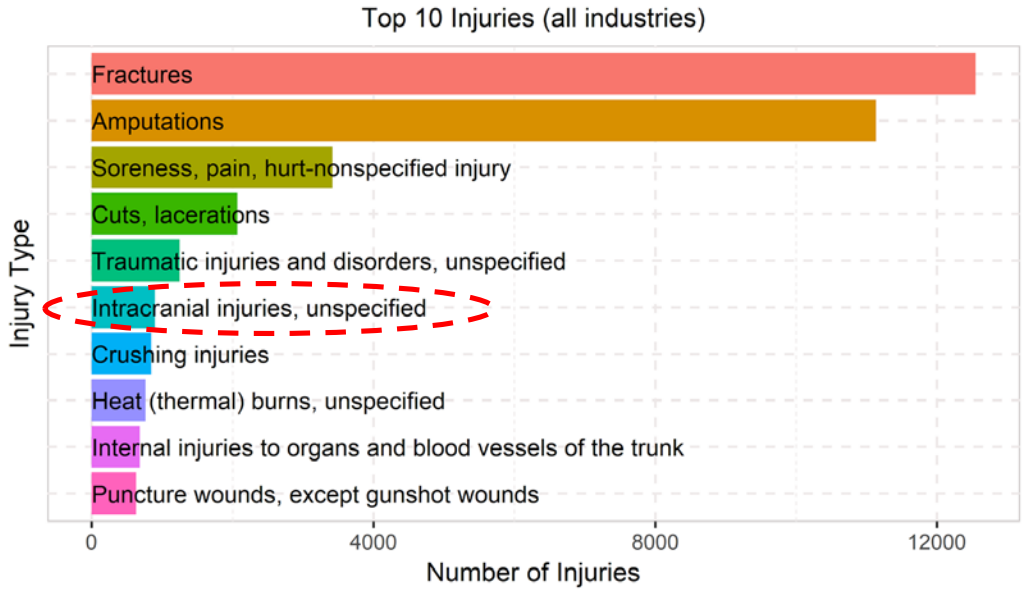
$$\mathrm{IDF}(t) = \log\left(\frac{N}{1+n_t}\right)$$

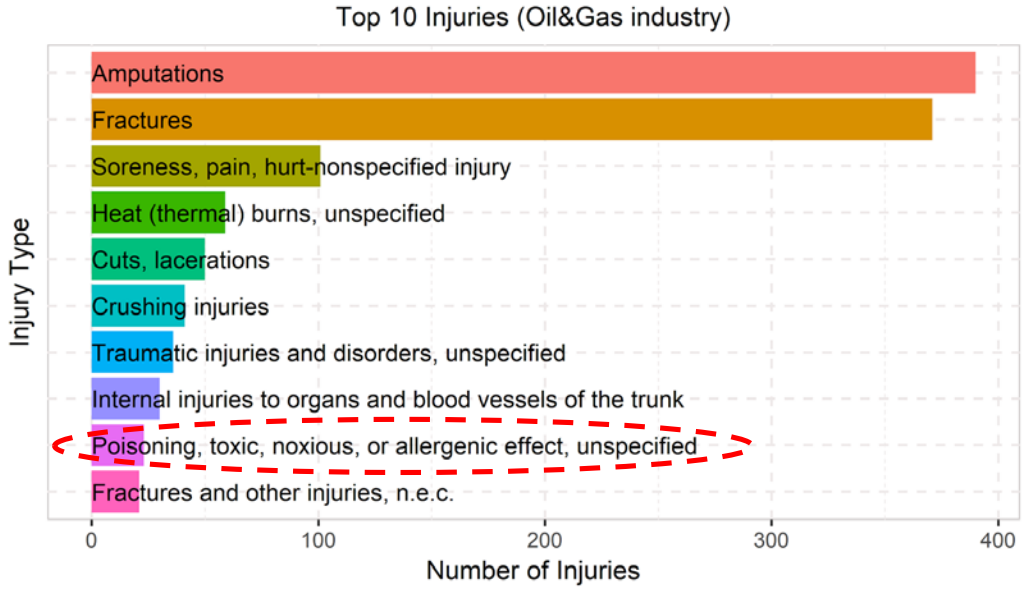
$$\text{TF-IDF}(t,d) = f_{t,d} \cdot \log\left(\frac{N}{1+n_t}\right)$$

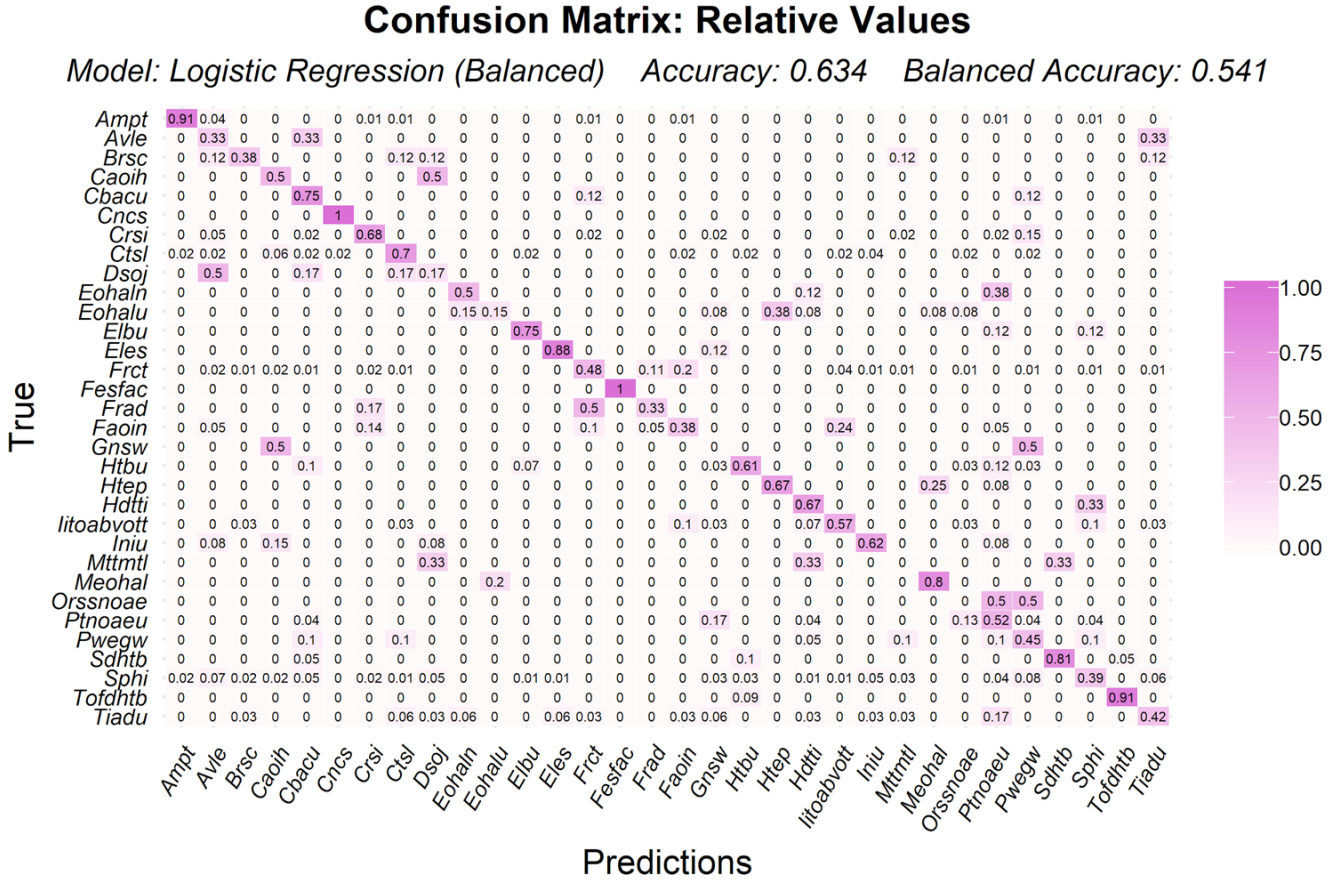


Data from the Occupational Safety and Health Administration of the United States Department of Labor (OSHA)

Severe Injuries







Acknowledgments

- CREWES sponsors
- Natural Science and Engineering Research Council of Canada (NSERC)
- GLJ Petroleum Consultants
  - Bill Spackman
  - Michael Morgan
- CREWES staff and students

Thank you!!!