

Geoscience Meets Data Science:

How to blend domain knowledge with learning from data

Researchers Links Workshop



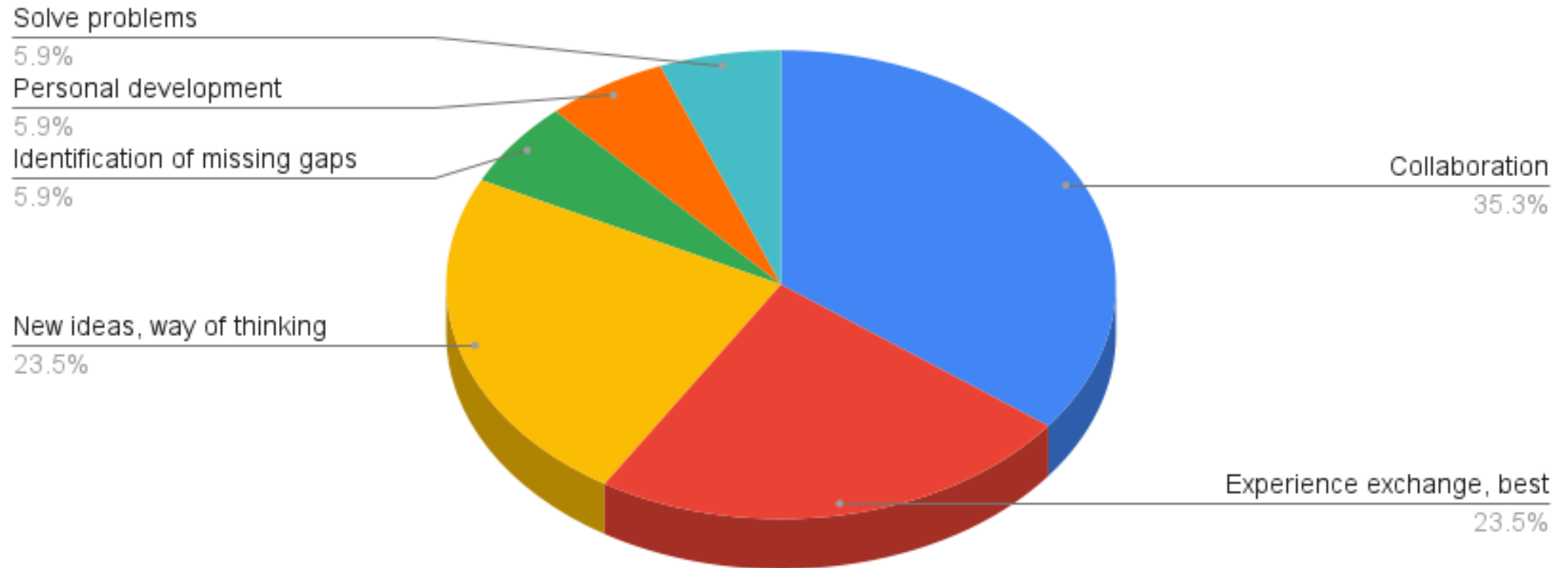
<https://geodatascience.hw.ac.uk/geoscience-meets-data-science>

14 December 2021

Aims

- Bring together researchers across different disciplines
- Develop a new vision on tackling geoscience problems with data science optics
- Establish collaborative networking environment
- Promote active interaction between academics and industry and public sector.

Expected Outcome



BroadenViews
LearnfromExperts
Gain^{New}Understanding
CuttingEdge
knowledge Solve problems
Discussions
Feedback
NewIdeas

Plan

- Warm-up 1h sessions:
Geoscience challenges, Data science challenges
 - 7 December
 - 15 December
 - January TBC
- February hybrid workshop:
 - 4 half-day networking sessions

Let's shape it together

Today's outline

- Break out room 1 – Introductions 15 mins
 - Who you are and what interested you to come
- Live talk – prof Ivan I Priezzhev 20 mins
- Break out room 2 – Reflection/exchange 15 mins
 - What stood out?, What was interesting?, Any ideas?
 - Share summary in the chat
- Wrap up 5 mins

Breakout room 1

Greetings and introductions
in groups of threes:

- Who you are
- What interested you to come

Design of a system to build a lithofacies model based on artificial intelligence elements

by Dr Ivan Priezzhev
Gubkin State Oil and Gas University, Moscow



Breakout room 2

Reflect on the topic:

- What stood out?
- What was interesting?
- Any ideas?

Share summary in the chat

Q&A – Prof. Ivan Priezzhev

1. What are the primary goals of AI applications in geoscience?

- We believe that the main task of AI in geosciences is to create an intelligent assistant with the following goals
- maximum automation of routine procedures for the analysis of initial data (seismic, borehole, etc.).
 - Assistance in complex analysis of multi-scale data from various sources.
 - minimization of the human factor (multi opinion factor) (how many geologists – so many opinions)

2. What are the remaining challenges in building lithofacies models or it is just about getting a ML software license from your favourite vendor?

Ambiguity in obtaining a model of lithofacies is not solved by simple changing the vendor

3. How do you see the middle ground between data AI and geoscience?

I think that when applying AI, the goals of geoscience should be set.

4. How to overcome the black-box developing awareness and literacy of numerical methods in geosciences for those novice to data science?

Only deep knowledge of geoscience tasks can help with the application of AI. I believe that only special (geoscience oriented) methods should be used to solve the problems of geosciences.

5. Where are the main challenges of integrating all the different data types with each other? How far can we go with data only and will we eventually have to fall back to geoscientists?

Integration of different-scale and variously accurate data from various sources is the main problem of geosciences and it is impossible to do without a lot of experience of geologists

Q&A – Prof. Ivan Priezzhev

6. How to establish a 2-way communication between geos and ML scientists, learning what each other requires?
More collaborative than just a pipeline – Kirstie Wright

Yes, it should be a very close collaborations in working joint groups

7. Where are the main challenges of integrating all the different data types with each other? How far can we go with data only and will we eventually have to fall back to geoscientists? – Bastian Steffens

The correct use of heterogeneous data types with different accuracy and details requires special approaches and experience of geoscientists. A simple naive use of machine learning algorithms for such data can give a negative result and a deep disappointment in the capabilities of AI.

8. Should we delve deeper into the methods of machine learning with transfer learning, where the algorithm will be trained on the data of a large number of analogous fields, and calibrated for the studied field? As far as I know, learning algorithms require a huge amount of data, and in the fields there is usually a lack of data. At the same time, for us, the most valuable are the cases of ML application at greenfields, about which we know very little.

Of course, application of modern learning algorithms using the trained neural networks at analogical fields is a very promising direction, especially for greenfield

9. How important is a set of open data for the industry on which you can train your approaches or algorithms and compare your results with published results? A good example on the F3 cube (free data on SEG wiki). Sabine Klarner raised this issue at the EAGE Data Science conference (Novosibirsk online 2020) (more about seismic), but the audience was not quite prepared for such discussion, so they have come of it then . - Sasha Volkova:

Application of open datasets to test new algorithms proved to be very important.

Next session

15 December , 10am GMT

Data, data everywhere!

Data driven challenges in Geoscience

by Prof Jennifer McKinley,
Queen's University, Belfast