Map the Subsurface: A Challenge in Multi-scale Representations

AI Co-Innovation Program برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Subsurface data are complex and span multiple scales – from fine well details to large-scale regional trends. This challenge invites you to build a compact, unified representation in a compact structure, token space or otherwise that captures meaningful relations across scales and support interpretation or prediction, even with sparse or misaligned input. The goal is to unlock deeper insight through representation, not brute force.

Expected Outcomes:

- *A lightweight representation that encodes spatial and across-scale continuity.*
- *Clustering/prediction enabled by the structure itself*
- *Prototype that scales and adapts with visuals*

Potential Impact:

- *Enables dimensionality reduction for smarter modeling*
- *Fuels low-label or zero-label workflows*
- *Scales to sustainability subsurface solutions*
- *Accelerates transition to AI-native, energy conscious geoscience.*

Data Preparation

Participants may use the public dataset, which includes 3D seismic angle stacks, well logs, and well metadata. Classical log types (e.g., GR, density, sonic), and synthetic FWS or VSP data may be used for extension.

Any special instructions:

The workflow should be containerized and modular to support reuse and adaptability by users from different technical backgrounds. Flexibility and clarity in deployment will be considered a strength.

☑ **Geoscience** ☐ **Reservoir Engineering** ☐ **Petroleum Engineering** ☐ **Production** ☐ **Drilling**

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

Rock Properties Prediction from Seismic

Challenge Description

The ability of predicting rock properties from seismic with accuracy will have dramatic impact on oil and gas exploration and production

Expected Outcomes:

Develop an AI-based method that can accurately predict rock characteristics such as density, porosity, and velocity using seismic data. This method should require only a limited amount of training data, ideally just a small number of wells (e.g., fewer than five wells).

Potential Impact:

The developed solution is expected to significantly enhance prediction accuracy, achieving at least 90% accuracy.

Data Preparation

Participants may use the public dataset, which includes 3D seismic angle stacks, well logs, and well metadata. Classical log types (e.g., GR, density, sonic), and synthetic FWS or VSP data may be used for extension.

Data Source:

TerraNubis - Data Info of F3 Demo 2020

Any special instructions:

☑ Geoscience ☐ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☐ Drilling

Realistic Reservoir Simulations from Sparse Subsurface Data Using AI

Challenge Description

Accurate and realistic reservoir data plays a critical role in developing robust AI models for the subsurface. However, acquiring sufficient real-world subsurface data is often highly expensive and logistically complex. While synthetic datasets provide a cost-effective alternative, they typically fail to capture the intricate complexities of actual geological conditions, resulting in suboptimal predictive performance and increased operational risks.

Expected Outcomes:

Develop an innovative AI-driven approach capable of generating realistic, accurate, and robust reservoir models using extremely limited real-world data and/or simplified synthetic data. The approach should minimize reliance on extensive domain-specific simulations, effectively bridging the complexity gap to improve the accuracy and reliability of reservoir predictions.

Potential Impact:

The output should enable the training of complex deep learning models effectively which was not possible.

- *Achieve a minimum 20% improvement in predictive accuracy or reduce modeling uncertainty by at least 15% compared to current standard methods.*
- *Demonstrate significant reductions (at least 30%) in dependency on complex, domain-intensive simulation resources.*

Data Preparation

Publicly available sparse geological datasets and synthetic subsurface datasets

Any special instructions:

Participants are encouraged to present highly original ideas and demonstrate how their approaches innovatively solve the stated challenge, emphasizing realistic complexity from minimal data without prescribing specific technical methods.

☑ Geoscience ☑ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☐ Drilling

Optimizing ESPs Operations for Oil and Gas Fields

AI Co-Innovation Program
برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Electrical Submersible Pumps (ESPs) are a crucial component in increasing production from oil and gas fields. However, they require significant amounts of energy to operate and are prone to failures, which can lead to decreased production, and increased maintenance costs. With the growing demand for energy and the need to reduce carbon footprint, it is essential to develop a smart system that optimizes ESP operations to maintain production targets while minimizing energy consumption and pump failures.

Expected Outcomes:

Design a smart system to operate a field with a large number of submersible pumps, ensuring:

- *Maintaining production targets*
- *Minimizing energy consumption*
- *Minimizing pump failures*

Potential Impact:

- *Reduce energy cost needed to operate the field. As a consequence reduce carbon emissions related to the field.*
- *Lower maintenance cost of electrical submersible pumps by increasing their lifespan.*
- *Possibility of increasing hydrocarbon recovery.*

Data Preparation

Publicly available tools/data:

[1] General purpose reservoir simulation (software).

[2] Pure data (e.g., reservoir model)

[3] ESP calculator (software).

[4] ESP maintenance (data and some code).

Any special instructions:

Participants are encouraged to use any data sources, software, and libraries available to show the viability of their proposed solution. The above data sources are provided as a guidance.

☐ Geoscience ☑ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☐ Drilling

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

Drilling parameters multi-modal foundation model

EXPEC ARC DEPT

AI Co-Innovation Program
برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Foundation machine learning models are transforming the way of digitalization in different industries including large language models, remote sensing and single cell biochemistry. Currently there is no relevant foundation model for supporting well drilling.

Expected Outcomes:

Drilling process is described by a number of specific time series such as weight on bit, rate of penetration, torque and others. Fine-tuning of time series foundation models based on drilling parameters is non-trivial task due to uneven patterns and relationships. Properly fine-tuned model would allow to efficiently reconstruct drilling parameters and therefore optimize well planning process.

Potential Impact:

Reducing of non-productive time by 15%.

☐ Geoscience ☑ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☑ Drilling

Data Preparation

[1] General-purpose time series foundation model

[2] Drilling parameters for historical wells

[3] Fine-tuning of large language foundation model for drilling purposes

Any special instructions:

Any required instructions to use the data

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

AI Proxy for Large-Scale Reservoir Simulation

Challenge Description

Building full-physics reservoir simulation models is computationally expensive and time-consuming, often requiring days to weeks for a single forecast. This limits the ability to perform uncertainty quantification, optimization, or real-time decision-making. Developing accurate AI-based proxy models can enable faster predictions without sacrificing reliability

Expected Outcomes:

Develop an AI proxy model that can accurately predict reservoir performance (e.g., pressure, rates, …) and reduce the computational time, across a wide range of geological and operational scenarios, enabling rapid forecasts for production and reservoir performance.

Potential Impact:

Reduce simulation turnaround time by >80% while maintaining forecast accuracy within ±5% of full-physics model results.

Data Preparation

Open-source datasets such as:

[1] SPE10 Benchmark Model: A widely used synthetic model for testing reservoir simulation methods, introduced by Christie & Blunt (2001).

[2] Norne Field Case: A real North Sea reservoir dataset released for research purposes, widely used in history matching and uncertainty studies.

Any special instructions:

☐ Geoscience ☑ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☐ Drilling

Intelligent Geosteering Solution for Maximizing Reservoir Contact

AI Co-Innovation Program
برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Geosteering aims to maximize reservoir contact, a critical factor in optimizing oil and gas well production. Even a 1% increase in reservoir contact can translate to millions of dollars in added value. Effective geosteering is essential for unlocking significant economic benefits.

Expected Outcomes:
The goal is to develop a solution that focuses specifically on the geosteering section of wells, where real-time monitoring and early alerts can be crucial. The ideal solution should provide geosteering operators with sufficient warning time to take corrective action and avoid exiting the reservoir, thereby maximizing reservoir contact and optimizing well production. This requires a system that can quickly detect potential exit points and alert operators in a timely manner, allowing for prompt adjustments to be made to stay within the reservoir boundaries.

Potential Impact:
At least 1% increase in reservoir contact.

Data Preparation

Public datasets such as: Smeaheia [1] and Volve [2] datasets

The data contains:

Horizons

Seismic

Well logs

[1] https://co2datashare.org/

[2] Volve field data set download – Equinor

☑ Geoscience ☐ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☐ Drilling

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

LLM-Driven Visual Alerts from Drilling Reports

AI Co-Innovation Program برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Daily drilling reports contain critical textual information about operations, incidents, and anomalies, such as lost circulation or stuck pipe. These reports are dense and require significant time to interpret. AI-driven visualizations can convert this textual data into intuitive graphics, helping rig engineers and supervisors quickly understand well status and respond effectively.

Expected Outcomes:

Develop an on-premises AI or leverage a large language model (LLM) solution capable of analyzing daily drilling reports and automatically generating clear visual summaries and alerts. These visuals should highlight operational anomalies such as lost circulation, stuck pipe, or borehole instability using timelines, 2D plots, and schematic diagrams.

Potential Impact:

Improve operational awareness and reduce response time to drilling anomalies by enabling engineers to visually assess conditions and events directly from textual reports.

☐ Geoscience ☐ Reservoir Engineering ☐ Petroleum Engineering ☐ Production ☑ Drilling

Data Preparation

Publicly available drilling reports.

Any special instructions:

Focus on operationally relevant events like lost circulation, stuck pipe, and well deviation.

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

Vision-Based Data Extraction from Petroleum Engineering Reports

AI Co-Innovation Program برنامج الابتكار المشترك للذكاء الاصطناعي

Challenge Description

Petroleum engineers frequently rely on historical scanned documents, including PVT and PLT reports, inspection logs, well logs, and well cross-sections. These documents often contain unclear text, handwriting, complex graphical data, and poor image quality, posing significant challenges to traditional parsing methods and limiting engineers' efficiency and accuracy in decision-making processes.

Expected Outcomes:

Develop on-premises computer vision or vision-enhanced large language model (LLM-Vision) solution to accurately extract, classify, and structure data from scanned petroleum engineering documents (such as PVT and PLT reports, inspection logs, seismic images, well logs, and well cross-sections) without relying on external cloud services, due to confidentiality constraints.

Potential Impact:

Significantly reduce manual data processing time; substantially improve data accuracy and usability compared to existing traditional parsing methods. Enable the processing and automation of legacy document management and analysis workflows

Data Preparation

Publicly available synthetic dataset, providing representative samples similar to real-world reports

Any special instructions:

Participants should consider variability in document quality, text clarity, handwriting, graphical complexity, and image resolution when developing extraction algorithms.

☐ Geoscience ☐ Reservoir Engineering ☑ Petroleum Engineering ☑ Production ☑ Drilling

هيئة تنمية البحث والتطوير والابتكار
Research, Development & Innovation Authority

aramco

AI Co-Innovation Program برنامــج
الابتكـار المشتـرك
للذكاء الاصطنـاعي

RDI
هيــئة تنــمية البــحث
والتـــــطوير والابتـــكار
Research, Development
& Innovation Authority

aramco