D1.1 : User Requirement Analysis Report

Deliverable No	D1.1 – User Analysis Report
Work package No. and Title	WP1 : End User Analyses, KPI Definitions and Adaptions
Version - Status	Final
Date of Issue	7th May 2021
Dissemination Level	PUBLIC
Filename	OptiDrill_D1.1 User Analysis Report

This project has received funding from the European Union's Horizon 2020 research and innovation action under grant agreement No 101006964

DOCUMENT INFO

Authors

Author	Organization	e-mail
Kevin Mallin	Geolorn (GEO)	kevinmallin@geolorn.co.uk

Document History

Date	Version	Editor	Change	Status
07.05.2021	V.01	Kevin Mallin		Final darft

TABLE OF CONTENTS

ACRONYMS

AI	Artificial Intelligence
BHA	Bottom Hole Assembly – The section of the drill string from the bit to the drill pipe. The BHA, is designed to include any specialised tools, technologies etc., that allow the well to be drilled efficiently.
BUR	The rate that a well can be safely radiused from it's previous path (typically vertical)
DA	Data Analysis/Data Analytics – used to understand and model hidden patterns/relationships in complex data sets.
DDR	Daily Drilling Report – a record of all operations over a 12 or 24 hour period
EOWR	End of Well Report – a full and final report on all activities associated with a well.
LCZ	Lost Circulation Zone – zones where the drilling flush is lost to the formation.
LTI	Lost Time Incident – loss of operational time (rig shut-down), usually due to personnel having an accident or a dangerous occurrence takes place that requires investigating.
LWD	Logging Whilst Drilling – in-hole equipment that logs formations being drilled
MD	Measured Depth – the total length of the well from surface to bottom
ML	Machine Learning
MWD	Measurement Whilst Drilling – records downhole data, azimuth, position etc as drilling progresses.
NPT	Non Productive Time – all the time when the bit is not on the bottom of the well and drilling (May also include LTI).
PP	Pore Pressure – Pressure due to compaction in the interstices of formations (Either fluids or gases). PP above hydrostatic may cause issues with drilling.
ROP	Rate of Pentetration / Rate of Progess (This may either be "instantaneous drilling rate" or overall Time (Time/Depth plots)
RPM	Rotating speed of drill bit
TVD	True Vertical Depth – the depth difference from the top of the well to the base of the well, vertically, in order that hydrostatic pressure can be calculated for well control.
UCS	Unconfined Compressive Strength – Defines rock strength as derived from laboratory tests
WOB	Weight on Bit – the load applied to any drill bit to effect drilling

EXECUTIVE SUMMARY

Data driven approaches to solving problems is becoming the norm for all walks of life and drilling is no exception. It is reasonable to assume that given all the wells that have been drilled around the world, there will be some identifiable trends and patterns developing, that will assist future projects to firstly plan and subsequently execute more efficiently. However, being wholly dependent upon such data, may result in less advances than first expected.

There was a recently published Society for Petroleum Engineers article (1st March 2021) entitled; ***Robust Data-Driven Machine-Learning Models for Subsurface Applications: Are We There Yet****?* The article raises and highlights many issues that need to be considered when relying on ML models for subsurface applications and there needs to be an appropriate mindset that benefits from the extraction from data, to make better decisions.

Information without knowledge has limitations, which is what underpins the ethos of the OptiDrill project. Gathering together drilling (training) data, both digital and analogue, and develop ML algorithms to analyse and present the data in order to assist the drilling team, make better and more informed decisions, with a view achieve better value from drill programmes. It will also contribute to future project planning with better collated data sets and an understanding where efforts to reduce NPT should be focused.

Importantly, the drill string data sensor system being developed within the the OptiDrill project, should additionally provide robust, real-time data, that complements other (offset) data sets gathered together for each well drilled and made available across an open access platform to assist other sub-surface industry sectors, both industrial and research based. With a particular emphasis on deep geothermal resources, it is envisioned that the OptiDrill Drilling Advisory system, will greatly help the transition away from fossil energy sources, including Geological Carbon Capture & Storage, green hydrogen storage and many other green house gas emission mitigating strategies.

1. INTRODUCTION

The Optidrill project sets out to combine numerous technologies, both developed and devoloping, that gather and transmit data, in real-time, from the drill bit/formation interface and translate the streams of information, by means of "Machine Learning and Artificial Intelligence" into a user friendly interface that can assist and advice the drilling team make better decisions and therefore, **OPTI**mise the **DRILL**ing process(es).

The need to drill deeper and deeper wells, often in complex geological settings, yet reduce the overall cost of drilling operations, pro-rata, requires a multi-faceted approach to reduce Non Productive Time (NPT), caused by equipment failure, lithological problems, overpressure in the formation etc. So it makes very good sense to concentrate on being able to assist the driller to anticipate issues that could lead to avoidable NPT occurences.

It is important to understand that data needs to be analysed for it to be useful, even at a basic level. Sophisticated Data Analysis is the bedrock for Machine Learning, that helps build relationship models for prediction and response, using algorithms to recognise the inputs and outputs from the data. AI then uses the ML data to build a predictive model that requires no human intervention/interaction, possibly updating the model(s) from the feedback.

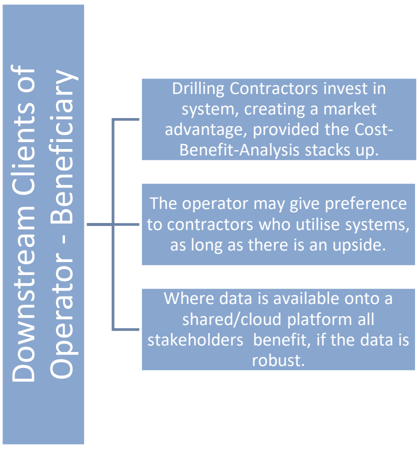
Sankaran et al. 2020 defined DA as a broad framework, whereby it helps determine what happened (descriptive), why it happened (diagnostic), what may/will happen in the future (predictive), or how can we control what will happen (prescriptive). ML is an increasingly important tool in the predictive and prescriptive elements and a core enabler for the AI decision making application (*Donoho – 2017*).

ML models will always require human review, as often the models are based on assumptions, that may not include or do not know all of the pertinent data, and the more ML models that are used the more people become overconfident, even when the assumptions do not match the facts.

Whilst there are many down-hole Data, Measurement-While-Drilling (MWD) and Logging-While-Drilling (LWD) systems, most of them deliver raw data, whereas the Optidrill project aims to compare real-time drilling derived data, with data captured from other wells, sub-surface data libraries and the development of "what-if" scenario planning. Through the use of advanced Machine Learning (ML) and Artificial Intelligence (AI) the real-time data will be rapidly and thoroughly analysed and will be used to predict multiple helpful information sets such as drilled lithology and to also flag-up possible NPT issues, before they occur. The end result should be a powerful drilling advisory system that avoids "digitising" the human input and costly computer driven errors.

2. User Designations

Superficially such a system as the OptiDrill system would be aimed at drilling contractors, looking to optimise the drilling process, minimise their NPT and maximise earnings. The issue here is that those wells drilled on a “day-rate” contract, where the operator (client) takes on all geological and sub-surface risks, the incentives for the contractor to drill more quickly are not in their interests. The advantage to the operator is that the contract should reach a completion subject to not exceeding budgetary constraints.



Are the system advancements most beneficial to the ultimate end-user? If drilling becomes more cost-effective, certainties increase and developers have a wider and more stable market share, that has to be good for everyone!

Drill times may decrease, with potential revenue losses for contractors, but operators will drill more wells, with the proviso that the resource is not over exploited – renewable, does not mean infinite.

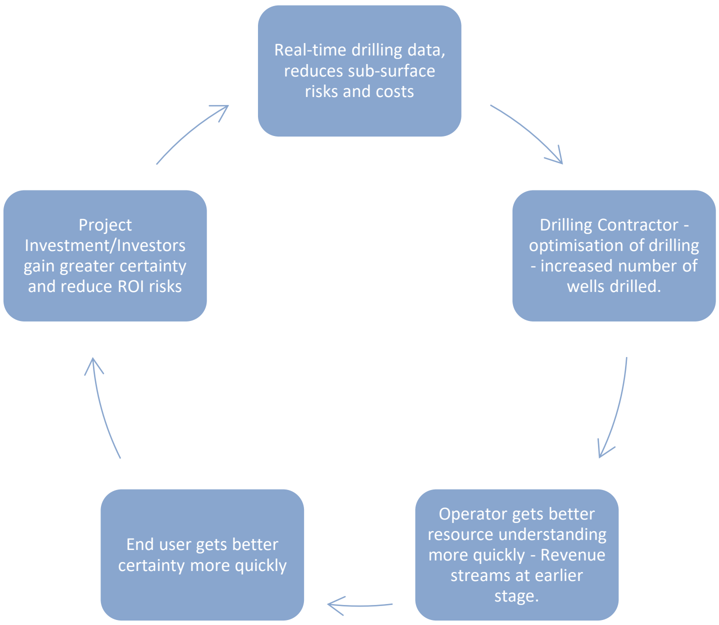


Fig. 1 Data quality and stakeholders relation diagrams

Obviously, the better the range of data generated the greater the efficacy of the system and its uptake amongst stakeholders; the speed of interpretation of the data, the more effective it becomes. Machine Learning technology can accelerate the way in which the drill crew can assimilate and optimise the information streams to optimise the process and minimise drilling problems and increased costs.

2.1 Data Requirements

Do all stakeholders require the same information? Will the data be beneficial to non-immediate stakeholders, such as academia and parallel technology providers?

To the operator the most important data sets will be accumulative drilling cost and the target resource's viability to produce economically for a sufficient period that repays the CAPEX inputs, with a commercial profit.

To the drilling contractor, the most important data sets will be those that help them to optimise the performance of in-hole tools and to reach target depths within budget.

So are these two requirements mutually exclusive or can both primary partners benefit from the technology improvements that systems, such as OptiDrill?

As drilling technology decreases the uncertainties associated with sub-surface risks, who benefits the most and is it necessary to reshape the interaction between operators and contractors that deliver overall improvements to all, especially when associated with decarbonisation of the enrgy sector?

Data Types/Provision

Well Planning Stage – Table 1

1. Well objective: Exploration, appraisal, development, re-injection.
2. Timeline -
3. Data Required ○ **Seismic/Geological profiles** ○ **Lithology/Petrophysical profiles/correlation** ○ **Surface location** ○ **Offset wells** ○ **Risk/Mitigation register** ○ **Well completion details (conceptual design)** ○ **Completion profile** ○ **Completion pressure testing** ○ **Stimulation requirements** ○ **Casing Running** ○ **Production pressures/temperatures** ○ **Produced fluids chemistry** ○ **Well lifecycle plan (future interventions etc.)** ○ **Pore pressure /Fracture Gradient plot** ○ **Possible shallow gas risks (offset wells, shallow seismic, pilot wells)** ○ **Well directional targets (possible downhole constraints – BUR's, Bit-to-Bend ratio etc)** ○ **Required and desired zonal isolations (this will include shallow aquifers)** ○ **Geothermal gradient (temperature profile)** ○ **Handover plan from drilling to production** ○ **Abandonment plan**
4. Specialist Functions (Services) **a. Wireline geophysical logging programme** **b. Coring programme** **c. Mud lugging / Geological logging** **d. Well testing requirements – DST, Flow Tests, Temperature**

The table below is designed to outline all of the data types required or wished for by the various stakeholders and what OptiDrill can offer/include. The list is not exhaustive. For simplicity, it is assumed that the data collection starts from the day of mobilisation. Primarily looking at geothermal resource wells, but can be applied to other resource projects.

OPERATOR	CONTRACTOR/ WELL ENGINEERING	GEOSCIENTISTS	OPTIDRILL
Cost accumulator	Cost accumulator (Daily Audit P&L)		Possible Module?
DDR's and NPT analysis	DDR's and NPT analysis		Possible Module?
Geology (Pre, Daily, Post)	Geology (Pre, Daily)	Geology (Pre, Daily, Post)	Geology (Minute-by-Minute Correlation)
Time/Depth Curves	Time/Depth Curves		Possible Module?
Performance Drilling Analysis	Dependent on contract type?	Passing Interest	Yes, with other software?
Pore pressure (PP) data	Pore pressure data	Pore pressure data	Possibly
MWD/LWD/SWD/ NEUTRON PULSE	MWD/LWD/SWD/ NEUTRON PULSE	MWD/LWD/SWD/ NEUTRON PULSE	Hybrid, tailored to drill programme?
Real Time Data for drilling optimisation[1]	Real Time Data for drilling optimisation[1]	Real Time Data for drilling optimisation[1]	Aim of Project
Drilling Problems that can lead to NPT or LTI[2]	Drilling Problems that can lead to NPT or LTI[2]	Possibly; where they can tie-in geophysical data	Aim of Project
	Annular Pressures		Annular Pressures
Formation Temperature	Formation Temperature	Formation Temperature	Formation Temperature
	Annular Flows		Annular Flows
	Lateral Vibration		Lateral Vibration
	Longitudinal Vibration		Possibly

OPERATOR	CONTRACTOR/ WELL ENGINEERING	GEOSCIENTISTS	OPTIDRILL
Permeability and recharge	Permeability, especially where it leads to well control issues	Permeability and recharge	
In-situ stress / stress regime		In-situ stress / stress regime	
Bit position (Azimuth, MD, TVD)	Bit position (Azimuth, MD, TVD)		
Seismic (induced and drilling related)	Seismic (induced and drilling related)	Seismic (induced and drilling related)	

1: Formation/Lithology/Petrophysical, UCS, Fracture(s) network/gradient, Circulating Temperature, Bit Performance, Torque, Annular Velocity, BHHP

2. Issues with formations (PP), LCZ's, Equipment failure, Unplanned deviation,

2.2 Data Flows

Drilling data can take many forms from digital to analogue and this needs to be considered and evaluated throughout the entire well process. Couple to this, data that is not derived directly from drilling related processes, but geophysical properties (seismic profiling for example) can heavily influence how wells get drilled, which can lead to additional costs, when things change from the prognosis.

Often, drilling contracts have to forward purchase bits and specialised tooling, which may prove to be unsuitable to match the geological conditions actually encountered. Not only can this lead to additional financial losses, but will result in additional NPT.

Whilst most geothermal projects (and many other resource/energy projects) refer to “drilling risks” and the “high costs of drilling”, as being considerable obstacles to the development and uptake of geothermal energy as part of the Carbon Zero target, should we also be thinking that the pre-drill data needs to be more robust?

The data that goes down the well, heavily influences the data that comes up the well and may help to optimise the drilling programme. So is this where Machine Learning (ML) can play its part, helping sort and grade information as it is created and compare it rapidly with historic data?

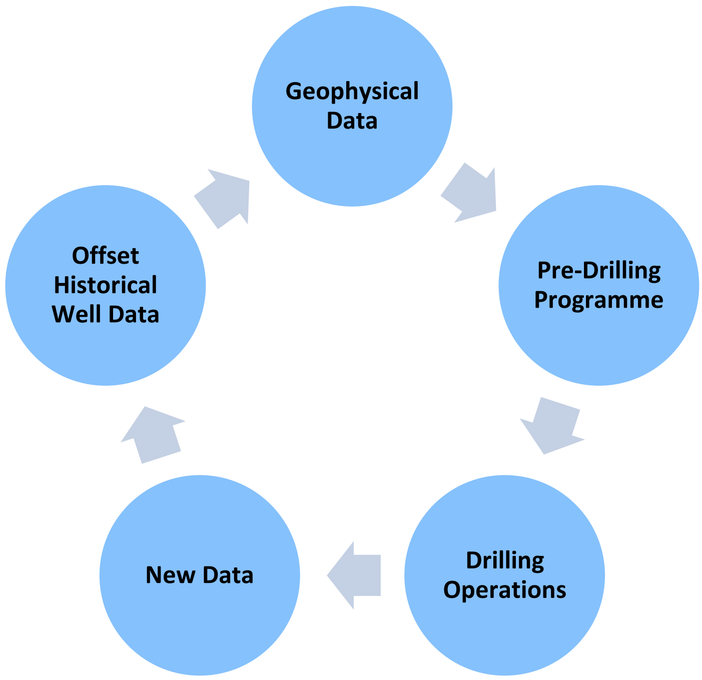


Fig. 3 Drilling process data flow cycle

Each stage assimilates and creates its own set(s) of data, which can both help and/or hinder the next stage and as with all data it requires to be contextualized, interpreted and quantified, so that subsequent wells can benefit from “lessons learned”.

From the diagram (Fig. 3), it can be clearly seen and understood that you can start at any node and use the data to enhance the next node. In wholly new fields, each data lobe is very weak, until it is reinforced by the completion of the next node. Each node will have a particular end-user interest, but will also need to be understood by all stakeholders.

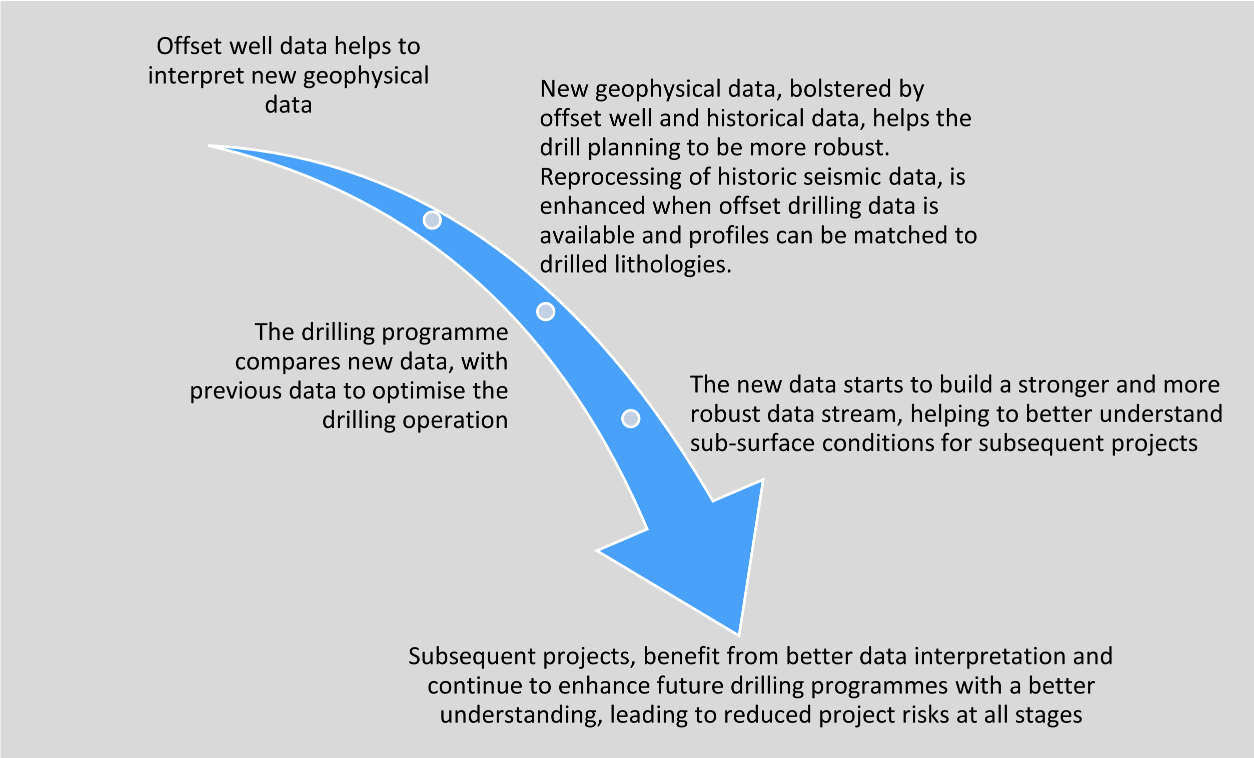


Fig. 4 OptiDrill system position within data flow cycles

The OptiDrill system sits within the New Data section of the above diagram and will clearly benefit from the previous data sets and add to the subsequent sector. The OptiDrill system will also benefit instantaneous operations, through the ML/AI support, being developed within the project, yet will allow the drilling team to make more informed decisions through the ability to predict possible scenarios and lithologies and suggest process parameters to optimise the process. It is very important to understand that the OptiDrill is an advisory system and not a tool for the automation of the sub-surface drilling process.

Given the approximate volume of wells drilled (estimated at 12,000,000 at an average diameter of 0.2m) to depths greater than 3,000m and assuming all the data from every well was available in a format that could be used for ML purposes, the amount of data generated would still be a very, very small percentage (approximately 0.000007%) in terms of the volume of the earth's crust (estimated at 1,083,206,917 km^3). However, through the usage of novel machine learning algorithms, trends and patterns can be recognized, as the variations in drilling tools is quite limited and are normally a "best-fit" approach, with some best practice inputs to optimise the outcomes.

Whilst the OptiDrill project will build a stand-alone data sensor and acquisition system, along with some specialist drilling tools, it will also aim to integrate and accommodate other in-hole and topside systems currently available and those in development. This is important so that the system can be used in as many projects as practical and to continually build the learning processes.

2.2.1 End-Users

Targetting end-user(s) of the OptiDrill Advisory System is both easy and hard to define; primarily it will be utilised by drilling contractors, whose drill teams will benefit from the stream of data being produced on second-by-second basis, analysed by ML algorithms, trained to use the patterns in data sets from previous wells that will both help to predict and respond to the downhole conditions, so Rates of Penetrations (ROP's) can be optimised, matching lithologies with the downhole tools in use.

The OptiDrill system will be based on multiple ML models which have been trained in a way that they will utilise the data from as many other wells as is possible, and will continually enhance its accuracy as it generates its own data and other systems share their data. It is therefore imperative that data is initially shared for training purposes, if we are to drive down the costs of drilling deep wells, by reducing formation issues and in-hole equipment failures. One of the huge advantages of robust ML models is the ability to predict problematic issues and set-out preventative actions and with AI can physically intervene. The issue with deep drilling is that the data may not always be as robust as it could be, for a number of reasons, especially where human intervention is not fully captured and encapsulated – the old adage "they seem to have less problems on the night shift".

Therefore, everyone involved with sustainably exploiting the sub-surface, has the potential to be an end user and a contributor, not just to OptiDrill, but to all data ML driven modelling processes.

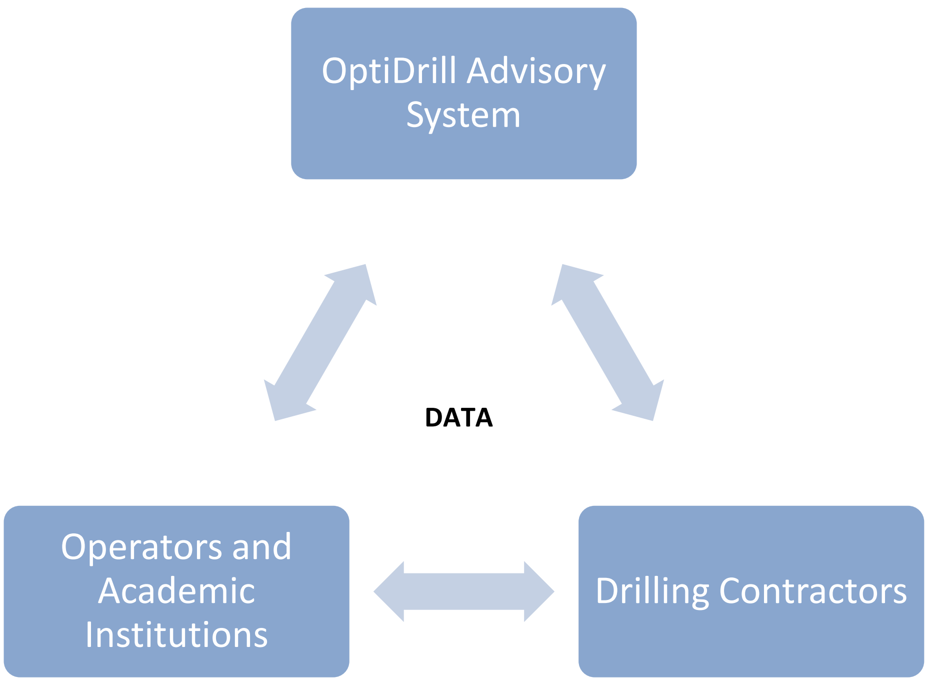


Fig. 5 OptiDrill stakeholders

The above (Fig. 5) clearly demonstrates that all stakeholders are end users, contributors and collaborators and the more data that flows the stronger the system becomes, so it is very important that OptiDrill fully engages across industry and academia to maximise its effectiveness. A break in any of the data-flows adversely affects the robustness of the system.

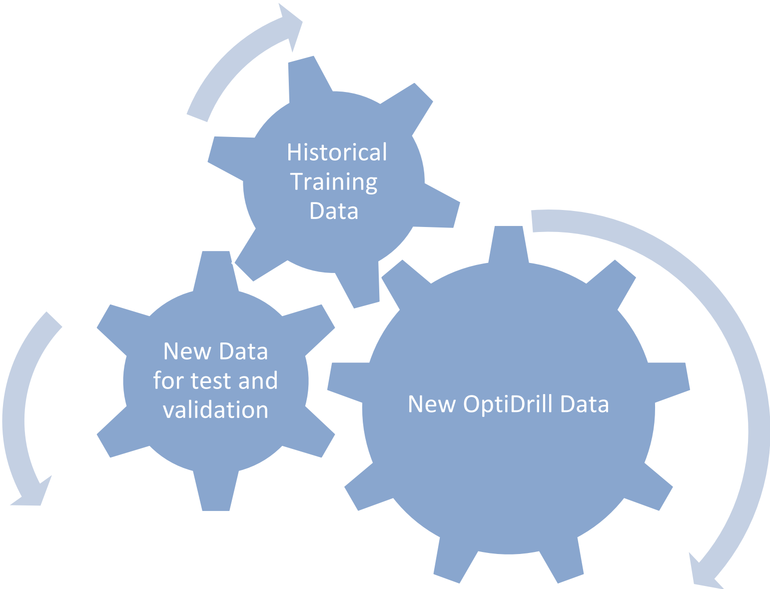


Fig. 6 – *Data Flow Interaction*

The end-users of the OptiDrill system become a very important cog in making the system increasingly effective, coupled with historical data and feeding into validation/test data. Unlike many other ML/AI systems that are aiming toward automation of drilling operations, the OptiDrill system will provide the drilling team with a range of suggestions and scenarios to advise on possible courses of action, it will not instigate any actions under its own volition. The OptiDrill team feel that this is vitally important, as it avoids "skewed" data becoming imprope0rly used. At no point will the OptiDrill system override the driller's operational sequences, although it will assimilate the data generated and validate it against all the drilling parameters.

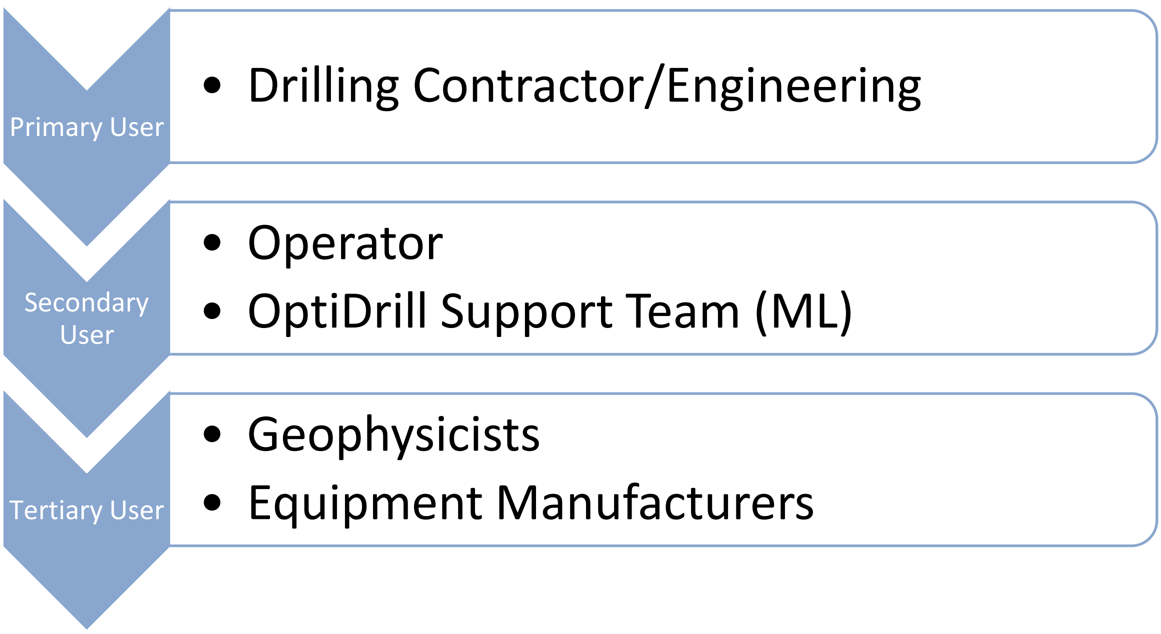


Fig. 7 Users of the OptiDrill System

In order to get the OptiDrill system operational, it is necessary to obtain data sets from previous wells, that show the drilling methods used, the operational parameters (Weight on Bit, RPM, Flushing Rates, ROP's etc), the lithologies drilled and where possible the corresponding Daily Driller's Report, that records all operations and issues during a 12 or 24 hour period. Additionally, MWD and LWD data would be a bonus.

Whilst the initial training data will be "random" in so far as it could be from wells anywhere in the world, there will be trends that can be analysed to start with the ML training work and as more wells are drilled that use the OptiDrill system, the more intuitive the system will become. What will be very important is the fact that the drill team will either act upon or ignore the system to optimise the drilling programme, which in turn will make all subsequent data, far more robust. The information will not override the knowledge and the knowledge will make the information more applicable.

The OptiDrill system will have at its heart a newly developed downhole sensor system, with data transfer through the drill pipe in real-time which in case its used can increase the system's efficiency substantially, although the ML user-interface will operate as a stand-alone.

3. Conclusions

Data driven technologies are the current driving force across all walks of life and the drilling and subsurface energy resource industries are very much part of this trend. However, unlike most other industries, whatever amount of data is available, it is still a tiny fraction of the earth's crust that has been explored, so at best, any data driven system can only be considered effective on a very local scale.

Whilst there will be geological settings that present similarities, there will never be two wells that drill exactly the same, so relying on data alone is a risk in itself. Additionally the choice of drilling methods (tools, BHA etc) will affect the data generated from the drilling, so requires to be adequately assessed and analysed to ensure the system "grows" robustly.

Analysis of the data will form a framework for the determination of what happened, reasons as to why events happened, what may happen (given what we already know) and how can we best make use of the data to control future outcomes. Manual data analysis will always be the most robust way of interrogation of what is being presented, but this requires a huge effort of personnel, so incorporating ML will help to streamline the processing of raw data, recognising trends and patterns.

It is important to remember that much of the drilling data from other wells may be "coloured" by the fact that Driller A responds differently to Driller B when encountering Formation Y, which may actually have been better addressed by Driller C's response. We also have manufacturers of downhole tools, pushing products, that respond better to some formations, but not others and because tools often have long lead times (particularly drill bits) when downhole conditions change from prognoses, compromises have to be accepted. There may not be a great deal of data available from previous wells, particularly for geothermal reservoirs and what may be lacking is a contextualisation of the data; often wells are drilled in such a way, because that's how these types of wells are drilled. A good example would be the drilling of micro-crystalline rocks with tricone roller bits instead of using a down the hole percussion drilling method, both methods will result in a well, but one method will be substantially quicker than the other and generate a very different set of data. Undertstanding the data is therefore paramount, as the percussion drilling method may not be suitable in pyroclastic formations, although superficially a ML model may interpret that it is. New and emerging drilling technologies will also provide radically different data sets, yet the geology will remain the same, so it is important that all users of the OptiDrill (or any ML derived model) understand what is actually happening downhole and how to best use the data to optimise their outcomes. We must not let Rate of Penetration override Rate of Progress; a well is only a well once it has been successfully completed, as designed.

Amongst the oil and gas companies, data sharing is quite commonplace, across easily accessible platforms, but the uploading of data seems somewhat ad-hoc, from structured to unstructured which requires a large amount of additional effort and analysis to make the data useful, in relation to ML programmes. There is also a huge amount of data, both digital and analogous, stored within servers and basements of buildings, that could prove to be highly valaubale across the sub-surface industries and certainly for the OptiDrill system's efficacy.

Large amounts of data exist in "silo's" and not always filed in such a way that ease of access is guaranteed and day-to-day activities generate more and more data, adding to the problems, where it is not stored in an easily accessible manner.

An additional problem with drilling derived data, is that a great deal of it is in analogue form, with the DDR's being typed into spreadsheet format, although there are a growing number of software programmes that are addressing this issue. This data is extremely valuable, as it encompasses human input/interaction and captures information that may be useful to understand non drilling NPT occurences, but requires an understanding of the drilling industry and how it ties into the digital data, to make full use of it.

Defining who the end-users of the OptiDrill system will be and how they will make use of the system is an extremely difficult task, as the end-users become the inputters into the system as well as using the outputs. The OptiDrill system will also harvest drilling data, in real-time, compare it the ML model training data, with the final evaluation being undertaken by the drill crew.

Getting operators and other stake-holders, willing to share their drilling data, will be hugely beneficial, but the OptiDrill team also believe that extensive "what-if" scenarios, based on the team's previous project experience will also prove to be extremely useful in training the ML system, as hypothetical problems can be introduced, not always seen in actual well data. This will make the system more robust and help develop the data sensors installed in the drill string in terms of the data captured; after all, drilling costs can only be reduced if we minimise NPT through optimising downhole tool selection and predicting issues before they become problems. It should be remembered that a great deal of NPT results from incorrect drill planning assumptions and basic human errors, so careful scrutiny of EOWR's will bolster the efficacy of the OptiDrill or any data driven system approach.

To measure is to manage and without data (or experience) this becomes much more difficult or even worse, far riskier. We could go as far as saying "to measure is to manage and to manage is to mitigate", remembering that at all times, information is only as good as the knowledge that makes use of it.

Undoubtedly, ML based model capabilities will evolve rapidly, as AI enhances predictive modelling and the range of quality of data increases exponentially and the OptiDrill system grows in the types of data it gathers in real-time, whilst drilling and the ML models interpret the data into a friendly user-interface.

References

Donoho, D. 2017. 50 Years of Data Science. Journal of Computational and Graphical Statistics **26** (4): 745– 766.

Sankaran, S., Matringe, S., Sidahmed, M., et al. 2020. Data Analytics in Reservoir Engineering. Richardson, Texas: Society of Petroleum Engineers.

Society for Petroleum Engineers: https://jpt.spe.org/robust-data-driven-machine-learning-model...qL5fI-HuDha5Z1Dxfnky0QWu_otes--LnZFZKN7xHgN_0l1Y8Ugqftz4WKQqw