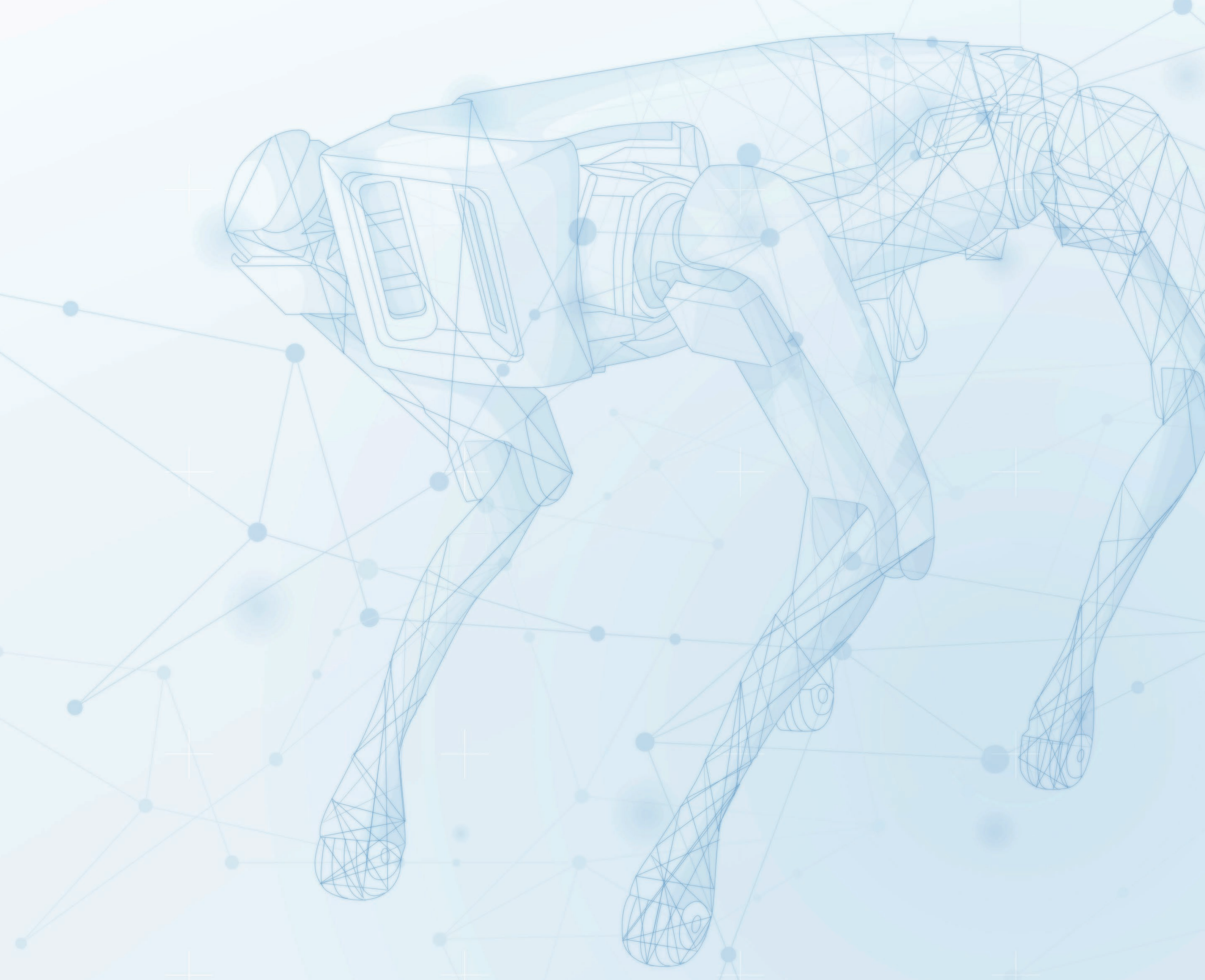
OLTER Shared Analytics White Paper

NET ZERO TECHNOLOGY CENTRE



Net Zero Technology Transition Programme

Net Zero Technology Centre
Technology Driving Transition

OLTER NZTTP Programme

CONTENTS

EXECUTIVE SUMMARY

Artificial intelligence (AI) and Robotics and Autonomous Systems (RAS) are revolutionising offshore energy inspection and maintenance, with the potential to improve safety, reduce cost and carbon footprint, and provide greater resilience to exogenous shocks such as a pandemic.

With RAS becoming a more viable option for offshore operations, there needs to be a place where the offshore industry, supply chain, academia, developers, and any other sectors can connect in deploying RAS for the UK Continental Shelf (UKCS). As such, the Offshore Low Touch Energy Robotics and Autonomous Systems (OLTER) project aims to provide the benchmark for development and use of reliable, on-demand, standardised autonomous systems. With this, it can ultimately deliver a RAS industrial service which supports the offshore energy industry and supply chain to scale and commercialise robotics as a service[1].

Within OLTER, the Data Hub aims to set up the following[1]:

Data Architecture	Shared Analytics Platform	Data Trust
Develop a preliminary "Industry Best Practice" Tech Stack architecture, on which to collect data and grow machine learning competencies.	Create a digital marketplace where developers can access data to create intelligent products, and end users can select/purchase these digital products to meet their operational requirements.	Establish how collected industry data will be governed.

Table 0.1: OLTER Data Hub

This White Paper focuses on the Shared Analytics concept, the second thematic within the Data Hub, which is designed to facilitate a more cohesive digital culture, working across the sector's landscape to improve the development of innovative machine learning and machine vision techniques. It will investigate industry trends and challenges, evaluate potential approaches, and suggest key activities that could uniquely position OLTER's Data Hub.

Currently, there is no existing end-to-end solution where offshore energy players can collaboratively develop machine learning techniques with one another and data science specialists to enhance their operations. Furthermore, the existence of data silos has both reduced data accessibility to train such algorithms and limited the versatility of such machine learning models that are created due to the differences in data formats across data repositories. Data silos are becoming an increasing barrier to sector innovation, and as a result the industry's uptake for machine learning development has been slower when compared to other industries. As the world transitions into a new digital era, offshore energy players risk being left out of the current digital revolution.

A Shared Analytics concept with an incorporated data standards and governance model would better enable the creation of new products and or services employing machine learning techniques through industry collaboration, thereby enhancing productivity and efficiency savings as well as seeding new innovative startup opportunities. The Shared Analytics concept intends to be a home for hosting data and analytics models as well as a marketplace for machine learning algorithms. The concept's inhabitance of machine learning talent allows companies to accelerate their development of machine learning techniques as well as ensure such techniques are built to a sufficient standard.

[1] *OLTER NZTTP Programme – OLTER, www.olter.co.uk*

The data standards and governance model aim to standardise RAS data, remove data silos to improve data accessibility for the training of machine learning algorithms and to make such models more widely available both within and outside of an organisation. This framework enables interoperability between in-house and external development of machine learning models as it allows industry-standard products to be built based on industry-agreed principles. This is crucial in giving companies the option to develop their machine learning techniques in-house should they wish to keep their data private. Therefore, it is suggested that this approach would be recommended for an industry-wide initiative.

As shown below in Figure 0.1, the OLTER Forum would create data standards, architecture, and governance principles. These standards can be interpreted as fundamental axioms upon which machine learning models are built and tested within the platform. Should companies choose to build their machine learning models in-house, they would need to follow these standards to ensure such models are compatible with industry-agreed data principles. If companies choose to upload their data into the platform, this would be used to train machine learning models. Evaluation of models would also be conducted through real-time scoring (Kubernetes). Once machine learning models pass the criteria for deployment, they would be made available in the marketplace. Users can access the analytical workload of models which provide metrics around their results through front-end data visualisation.

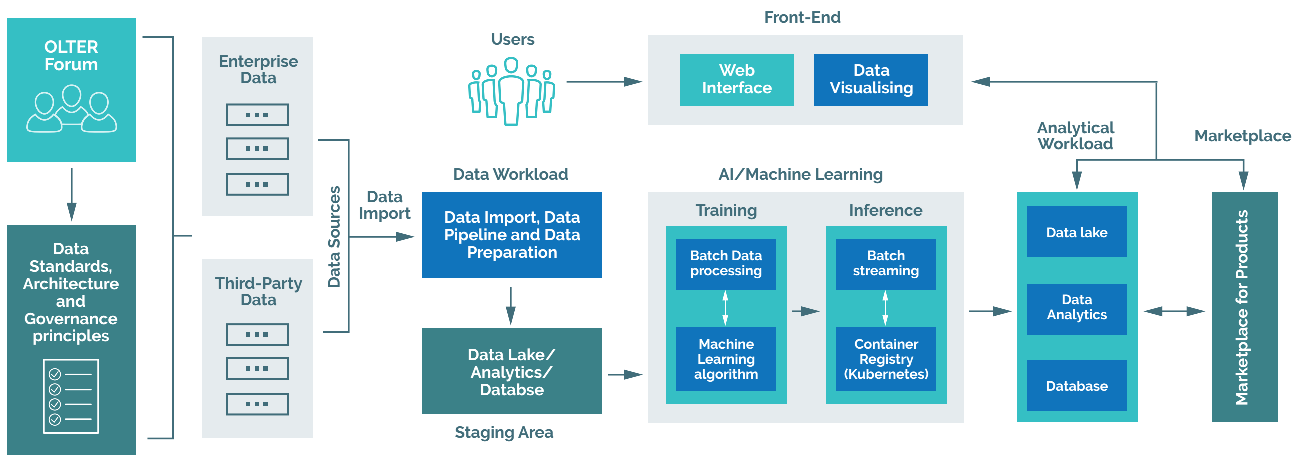


Figure 0.1: Shared Analytics Approach

This approach could bring a variety of benefits for the offshore energy industry, and can:

- Unlock open innovation, allowing a variety of organisations to better work together to support a transition to a low-carbon energy system
- Create new approaches to optimise an increasingly complex network
- Create better outcomes for business users by encouraging innovation in the marketplace, through delivery of new products and services
- Greater facilitate the development of RAS technology

The potential participants of this concept and the value to them are as shown in Table 0.2:

Participant	Value
Integrated Energy Companies and Minor Oil and Gas Companies	• Increase in safety (e.g., early corrosion detection of assets can lead to faster maintenance action and reduce frequency of equipment failure) • Demonstratable CO2 reduction efforts and progress to the regulator • Reduce data silos and enhance data accessibility • Improve versatility of machine learning models through data standardisation • Access to new data science talent. The industry has struggled to attract data science talent for a variety of reasons including poor public perception • Direct access to developers through the platform, enabling faster and more efficient problem resolution • Identify previously unknown similarities/synergies • Increase in transparency into service company performance and information • Faster reaction to operational challenges and solutioning • Service providers will be able to undertake actionable outcomes based off model insights (e.g., a service engineer can visit a failing valve and fix it)
Offshore Renewable Energy Companies	Value would be similar to that of Integrated Energy Companies and Minor Oil and Gas Companies, but they would also be able to: • Gain further insight into impactful RAS machine learning use cases that would significantly improve operations • Prevent the formation of data silos and reduce the re-work/additional effort required in addressing these in the future
Analytics Service Companies	• Ability to build services and new capabilities on the platform and to deliver these to the market • Tailor and improve services through collaboration with end-users • Insight to build out new services to their clients and develop novel business lines that were previously inaccessible or unknown
North Sea Transition Authority/ Regulatory Bodies	• Greater oversight of the offshore industry players and activities • Ability to hold all parties in the offshore energy industry to account for their actions through data analysis • Identify improvement opportunities and drive investment opportunities
Academic Institutions	• Strengthen relationships with offshore energy players through industry partnerships • Tailor research programs to the most relevant industry challenges • Collaborate and share data with other academic institutions • Improve expertise, technology, and R&D development programs (e.g., through PhD programmes) • Demonstrate talent and capabilities directly to the offshore energy industry • De-risk longer term research for both universities/research institution and their funding partner/company

Table 0.2: Value for potential participants within the Shared Analytics concept

1. INTRODUCTION

In June 2019, the UK became the first major economy in the world to pass laws to end its contribution to global warming by 2050 (and Scotland by 2045)[2]. The target, seen as one of the most ambitious in the world, was recommended by the Committee on Climate Change, the UK's independent advisory body on building a low-carbon economy.

Since then, the UK has managed to reduce its greenhouse gas emissions in 2021 by 47% compared to 1990 levels[3]. Its most recent Net Zero Strategy, unveiled in October 2021, sets outs to secure 440,000 well-paid jobs and unlock £90 billion in investment in 2030 on its path to decarbonisation[4].

Digitalisation has become a key driver within the industry as operators are currently facing the need to improve operational efficiency and secure critical data. Beyond an operational standpoint, digitalisation is also important for organisations to support decarbonisation efforts, as well as comply with health, safety, and environmental policies[5].

The key to this digital transformation lies with data. Data has become a new form of capital, vital for organisations to survive and succeed in today's fast-paced business environment. Every interaction in the digital world generates data: from humans, machines, and Internet of Things (IoT) devices to edge systems and beyond[6]. In 2022, it is estimated that 2.5 quintillion bytes of data is generated each day, and this is expected to increase to 175 zettabytes (10^{21} bytes) by 2025[7].

As a result, organisations are now becoming more data-driven in their approach, with expectations for data to be[8]:

- Delivered in real time for quick decision making
- Able to generate intelligence in the form of predictive and prescriptive models for optimised operating models
- Available for self-service with improved quality and security

By being more data-driven, companies can respond to market changes with greater agility, be more confident in their decision making, make more accurate predictions and increase operational efficiencies. As such, organisations are beginning to define and develop their data strategies and data management plans to support their digital transformation.

More specifically within the energy industry, companies are looking to accelerate this journey through applications of artificial intelligence (AI), namely the development of new products and or services that employ machine learning techniques. By leveraging their vast amounts of data to develop machine learning models, energy companies can begin to automate and improve operations. Paired with machine vision - methods used to provide imaging-based automatic inspection and analysis – energy companies can deliver actionable insights to extend asset life – these can be existing offshore oil and gas assets or future offshore renewables assets.

[2] *UK becomes first major economy to pass net zero emissions law – Department for BEIS, www.gov.uk/government/news/uk-becomes-first-major-economy-to-pass-net-zero-emissions-law, 2019.*

[3] *2022 Progress Report to Parliament – Climate Change Committee, www.theccc.org.uk/wp-content/uploads/2022/06/Progress-in-reducing-emissions-2022-Report-to-Parliament.pdf, 2022.*

[4] *UK Net Zero Strategy – Department for BEIS, www.gov.uk/government/news/uks-path-to-net-zero-set-out-in-landmark-strategy, 2021.*

[5] *Digitalization Becoming Increasingly Crucial for the Oil and Gas Industry – GlobalData, www.globenewswire.com/en/news-release/2022/06/14/2462460/0/en/Digitalization-Becoming-Increasingly-Crucial-for-the-Oil-and-Gas-Industry-GlobalData-Plc.html, 2022.*

[6] *Building data products as a competitive differentiator – Accenture, www.accenture.com/gb-en/insights/technology/data-products, 2022.*

[7] *How much data is created each day – DevShed, wpdevshed.com/how-much-data-is-created-every-day/, 2022.*

[8] *Data is Essential to Digital Transformation – Forbes, www.forbes.com/sites/forbestechcouncil/2020/12/03/data-is-essential-to-digital-transformation/?sh=a3c120626c99, 2020.*

As most energy companies are focusing their immediate data science efforts on higher priority areas such as production optimisation, the utilisation of Robotics and Autonomous Systems (RAS) data has remained immature, with huge potential for enabling innovation. This White Paper will therefore refer to RAS payload data specifically, with the potential to be scaled for other data types.

Offshore environments are taking advantage of RAS to accelerate the path to net zero through improving operational efficiencies and reducing resource usage and waste[9]. Beyond this, RAS technology is beginning to replace the need for humans to perform manual inspections, thereby enhancing worker safety. RAS usage over the last several years has increased due to advancements in drones and other similar technologies[10] and the large amount of data sets they collect offshore can facilitate data contextualisation. To maximise their use and potential in the offshore energy industry, there are key challenges and barriers that need to be addressed, particularly in the development of machine learning solutions.

These include the:

- Existence of technical and data silos
- Difficulty accessing specialist talent
- Lack of industry-agreed best practices

This has led to poor data accessibility and less collaboration amongst the offshore energy sector relative to other industries such as Big Tech, which has slowed the progress towards digital transformation.

As such, the Offshore Low Touch Energy Robotics and Autonomous Systems (OLTER) project aims to provide the benchmark for development and use of reliable, on-demand, standardised autonomous systems. With this, it can ultimately deliver a RAS industrial service which supports the offshore energy industry and supply chain to scale and commercialise robotics as a service[1].

Within OLTER, the Data Hub aims to set up the following[1]:

Data Architecture	Shared Analytics Platform	Data Trust
Develop a preliminary "Industry Best Practice" Tech Stack architecture, on which to collect data and grow machine learning competencies.	Create a digital marketplace where developers can access data to create intelligent products, and end users can select/purchase these digital products to meet their operational requirements.	Establish how collected industry data will be governed.

Table 1: OLTER Data Hub

This White Paper focuses on the Shared Analytics concept, the second thematic within the Data Hub. It explores the viability of an analytics platform and marketplace that can be shared and industrialised to remove these barriers and enable the widespread development of data analytics. The term 'Shared Analytics Platform' will be referred to as a 'Shared Analytics concept' in this White Paper, which encompasses a platform element as well as a marketplace.

It is designed to facilitate a more cohesive digital culture, working across the sector landscape to define the expectation of what for many, now sits upon the immediate horizon, machine learning and machine vision. The desired goal is to create an environment where the offshore energy ecosystem (energy companies, RAS providers, academia) and other parties (public developers etc) can come together to share challenges and create new machine learning solutions through access to sufficient training data sets.

Therefore, this White Paper aims to:

- Investigate why the offshore industry has faced challenges in developing machine learning models compared to other industries such as Big Tech
- Investigate industry views on a Shared Analytics concept
- Identify key challenges and barriers related to a Shared Analytics concept
- Understand key industry trends and drivers for data-driven transformation
- Evaluate potential approaches for Shared Analytics

Contribution from Industry

To inform this White Paper, interviews were conducted with members of the OLTER consortium as well as other organisations (Analytics Service Company and IECs). The OLTER consortium members who participated are listed below:

[9] *Overcoming the barriers to robotics in offshore renewables – ORE Catapult, ore.catapult.org.uk/blog/overcoming-barriers-robotics-offshore-renewables/, 2021.*

[10] *Use of Robotics in the Oil and Gas Industry – COPAS, copas.org/use-of-robotics-in-the-oil-and-gas-industry/#:~:text=Robotics%20usage%20has%20increased%20 significantly,gas%20sector%20is%20no%20exception, 2021.*

2. PREMISE FOR A SHARED ANALYTICS PLATFORM AND MARKETPLACE

Machine learning techniques can be seen to benefit the offshore energy industry through:

- Increased production through optimising drilling operations
- Reduced operational and maintenance costs through predictive maintenance of equipment
- Improved predictability of power generation through forecasting wind energy

However, while the benefits of machine learning techniques are well-established, developments in this space have not been as straightforward compared to the other industries such as Big Tech.

2.1 Lessons from Big Tech

There are three key lessons the offshore energy industry can learn from Big Tech companies (Google, Amazon, Apple, Meta and Microsoft) to improve the development of machine learning techniques, as these are:

1. Access to large training data sets to adequately train machine learning algorithms
2. An open-source environment which allow companies to collaborative and solve common industry problems
3. Receptiveness to leverage the vast amounts of data they generate to break into new markets, granting them access to more data and further insight into developing more machine learning models

To create new products or services that employ machine learning techniques, organisations require large amounts of data for training their machine learning algorithms. Big Tech companies have abundant access to training data, and therefore, can sufficiently train their machine learning algorithms such that these algorithms provide precise, reliable, and accurate results when used in practice.

While training data sets are usually proprietary and confidential, algorithm development in Big Tech occurs within an open-source environment. This allows companies to collaborate and learn from each other as well as build innovative solutions to solve common problems experienced within industries. With this, Big Tech companies can leverage their vast amounts of data and talent to develop state of the art machine learning tools. Notable examples of this within Big Tech include TensorFlow[11] and PyTorch[12] and Microsoft Cognitive Toolkit[13], founded by Google, Meta and Microsoft respectively.

In many other industries, the big internet players already have an inbuilt set of advantages, this includes their:

- Existing consumer and business relationships to help increase the use of their products and services aswell as increase customer retention
- Sizeable cloud infrastructure and networks to facilitate the planning, building, and deployment of new technologies
- Access to a vast amount of content in the cloud which allows them to quickly train any new machine learning models they build
- Oversight of usage data across many platforms, products as well as services to gain deeper insights into product/service trends and customer needs

[11] TensorFlow – Google, https://www.tensorflow.org/
[12] PyTorch – Meta, https://pytorch.org/
[13] Microsoft Cognitive Toolkit – Microsoft, https://learn.microsoft.com/en-us/cognitive-toolkit/

This allows Big Tech companies to use their dominant data position to enter and disrupt new industries . Breaking into new industries allows these companies to gather even more data, thus amplifying the ability to overcome additional barriers of entry through further strengthening their data position.

Company	Industry	Project/Subsidiary	Description
Amazon	Automotive	Zoox[14]	A self-driving car subsidiary
	Gaming	Amazon Games[15]	Gaming subscription service
	Consumer Discre-tionary	Amazon Smart Home[16]	A virtual assistant technology for home appliances
Apple	Healthcare	Apple Health[17]	A mobile health informatics application
	Consumer Discretionary	HomeKit[18]	A virtual assistant technology for home appliances
Google[19]	Automotive	Waymo	A self-driving car initiative
	Life Sciences	Verily	Life sciences research organisation
	Telecommunications and Energy	Alphabet Access & Energy	Houses the company's telecommunications projects and energy initiatives

Table 2.1: Examples of Big Tech companies disrupting new industries

As a result, Big Tech has historically seen a lot of success with respect to the development of machine learning techniques and have started to disrupt other industry with new business models. This is something the offshore energy industry has yet to achieve due to the existence of certain challenges and barriers which will be discussed in the next section.

2.2 Challenges within the Offshore Energy Industry

Unlike Big Tech, the offshore energy industry faces different challenges in digital transformation, particularly in sharing of data both internally and externally, and as a result has yet to make comparable progress in developing and utilising machine learning techniques.

The limited availability of sufficient training data sets needed to enable machine learning is one of the bottlenecks for innovation in this space, along with access to specialist talent, and cultural barriers. The causes behind these are discussed on the following page.

[14] Alexa Smart Home – Amazon, https://www.amazon.com/alexa-smart-home/b?ie=UTF8&node=21442899011
[15] Amazon Games – Amazon, https://www.amazongames.com/en-us
[16] Alexa Smart Home – Amazon, https://www.amazon.com/alexa-smart-home/b?ie=UTF8&node=21442899011
[17] Apple Health – Apple, https://www.apple.com/uk/ios/health/
[18] Apple Home – Apple, https://www.apple.com/uk/home-app/
[19] Google Strategy Teardown: Google Is Turning Itself Into An AI Company As It Seeks To Win New Markets Like Cloud And Transportation – CB Insights, www.cbinsights.com/research/report/google-strategy-teardown/, 2018.

Data Silos

Operationally, disparate legacy systems, business units and interactions with contracted parties cause great difficulty in consolidating data, resulting in data silos, which inhibit data accessibility for the testing of machine learning algorithms. Additionally, in cases where machine learning models are created and trained using data sets within a specific repository, it is unlikely these models will be compatible with data sets elsewhere as data standards and formats differ. Therefore, unstandardised data arising from data silos often leads to poor versatility of machine learning models created within industry.

Referring to RAS payload data as an example: there are multiple data silos at different points in the supply chain. An energy company can have multiple assets/ business units, each with its own data owners and processes in place for sharing data. There could also be associated legal implications with sharing data externally for initiatives similar to Shared Analytics if they conflict with internal activities the company may be pursuing. In instances where an energy company can share their RAS payload data for Shared Analytics, there is another complication if their RAS data is stored externally. Some energy companies have in-house RAS which may make accessing this data simpler, but the majority subcontract RAS activities to robotics companies or offshore energy service companies. The parties involved contractually agree that the energy company will own the data but often due to a lack of efficiency, visibility, or clear process in place for accessing this data, it will sit with the service provider unused, thereby creating another silo.

Data sensitivity is another important factor to consider when referring to the sharing of energy companies' data. Most energy companies will be hesitant to share their RAS images and videos externally due to these being safety critical (e.g., inspection images of assets) and/or difficult to anonymise as they have embedded metadata that cannot be removed (e.g. location data). This means that any industry approach that is suggested must consider that real data contributions to an external source may not be possible or will only occur on a small scale and will potentially be unsustainable over time. Further investigation is needed to better understand the cases where data can be gathered and shared externally for the benefit of participants.

Specialist Talent

Another key barrier for successfully creating machine learning models for offshore energy is the industry's access to specialist talent. The lack of access to talent with both machine learning skills and domain energy knowledge relating to specific industry problems has resulted in greater costs and slower development of machine learning techniques.

Cultural Barriers

From a cultural perspective, most energy companies prefer to develop their own tools for their specific needs; they primarily extract value from developing machine learning algorithms in the form of intellectual property (IP), rather than to improve and build on top of existing technology. Therefore, energy companies tend to develop their algorithms in house rather than in an open-source environment, and because of this, are not able to collaborate with other companies in solving common engineering problems experienced within industry.

As a result of these barriers and challenges, offshore energy players are less equipped to effectively utilise the vast amounts of data they generate and are less able to collaborate with one another. Consequently, this has caused the development of machine learning techniques to be inefficient and slow. Hence, pressures arise for companies to become more receptive to alternate business models and strategies such that the industry can begin to fully reap the benefits of machine learning tools, keep up with the pace of digitalisation, and transition towards net zero.

2.3 Use Cases within Industry

Two categories of use cases are identified within this section, data science and data standards community approaches, and machine learning model use cases.

Data science and data standards community approaches are examples of communities and platforms that facilitate the sharing of data and/or data standards to create new products and services. Machine learning uses cases are examples of machine learning models which can be developed to solve specific, topical challenges that the offshore energy industry currently faces.

2.3.1 Data Science and Data Standards Community Approaches

Kaggle

Kaggle is a crowdsourcing data science platform with cash prizes for winning entries. Kaggle aims to attract, nurture, and challenge data scientists from all around the world to solve data science, machine learning and predictive analytics problems. It also offers a public data platform, a cloud-based workbench for data science, and artificial education[20]. The platform started in Melbourne Australia, and moved to Silicon Valley in 2011, which was later acquired by Google in March of 2017. As of November 2022, Kaggle has over 11 million registered users[21], and the community spans over 194 countries.

Kaggle competitions occur typically through the following steps[22]:

1. The competition host prepares the data and a detailed description of the problem. The hosts' posted competition data is completely open to the public with only a Kaggle login required to set up competitions, access or download data
2. The competition participants submit their solution to the problem in the form of machine learning models. It is expected that for a given competition, different models will contain different solution approaches. All the work is shared on the platform with the intention of inspiring new ideas to achieve better benchmarks. In most Kaggle competitions, submissions are scored immediately by a set of evaluation criteria which use statistical analysis to determine model accuracy and reliability. They are subsequently summarised and made public on the live leader board
3. When competition reaches its deadline, the host pays the prize money to the winner. Hosts have the sole ownership and royalty-free license to use the winning entry as well as the IP

Alongside its public competitions, Kaggle also offers private competitions which give hosts the ability to limit access to selected participants should they wish to restrict the release of information to the general public. Relating to offshore energy, competitions include classification of oil and gas fields to determine location of mining sites[23], and prediction of wind power production from turbines using sensor data[24].

Through Kaggle, companies can gain access to machine learning techniques and data science talent at low cost using competitions and prize funds. Public developers can access training data sets to create tailored machine learning products to help overcome industry challenges. Furthermore, as all the work is shared on the platform, this encourages collaboration between participants and allows for faster development of new innovative solutions.

[20] *Kaggle – Kaggle, www.kaggle.com/*
[21] *Unique Kaggle Users – Bojan Tunguz, www.kaggle.com/code/tunguz/unique-kaggle-users/notebook, 2022.*
[22] *Kaggle Terms of Use – Kaggle, www.kaggle.com/terms, 2022.*
[23] *Classification of oil and gas fields – Kaggle, www.kaggle.com/competitions/profitrain1/overview/description, 2021.*
[24] *2022 Challenge Data Science Wind Turbine – Kaggle, www.kaggle.com/competitions/challenge-datascience-wind-turbine-cs, 2022.*

Open Subsurface Data Universe (OSDU)

The Open Group OSDU Forum is developing an open source, standards-based, technology-agnostic data platform for the energy industry, facilitated through the collaboration between operators, developers, and academic institutions - the OSDU Data Platform[25]. With this, OSDU aims to reduce data silos through data standardisation and governance, allowing for greater industry innovation and data management as well as accelerating the deployment of emerging digital solutions. In addition to the data platform and forum, the OSDU provides application development frameworks, an OSDU Catalog and an innovation marketplace.

Established in 2018, the OSDU Forum has accumulated over 185 member organisations, including consultancies, oil and gas companies and software companies who are collaborating to accelerate innovation and reduce costs in the energy sector[26]. The Open Group has different membership options with associated yearly fees, and its members range from major corporations, government organisations and consortia to universities and private individuals.

The OSDU Application Development framework serves as a guide for developers to better understand OSDU's ways of working, frameworks, application programming interfaces (APIs), data principles and data model, therefore ensuring all applications conform to OSDU Data Platform's standards. It also encourages developers to collaborate when creating software products and services for the OSDU Data Platform[27].

1. Once an application has been developed and is ready for deployment, it must first be certified through the Standard and Certification Program
2. After certification, the developer registers and lists the application on the OSDU Catalog
3. Developers may also wish to increase availability and easy installation of their application; they would work with one of OSDU cloud service providers (CSP) in this case to have their application integrated, validated, and listed on their cloud marketplace. Integration of their applications with CSP marketplaces will give customers the ability to 'one stop shop' search for their application and improve accessibility to various CSPs[28]

The OSDU Catalog allows users to market and purchase OSDU solutions and services. These can be categorised as[29]:

Category	Description
Training	An up-to-date list of available training courses and services (e.g., application development)
Models	A list of all available services related to the use of AI, machine learning models and other algorithms against data sets

Table 2.2A: Types of products and services within the OSDU Catalog, Part 1

[25] OSDU Mission & Vision – OSDU Forum, osduforum.org/about-us/who-we-are/osdu-mission-vision/

[26] The Open Group OSDU™ Forum Launches the OSDU Data Platform Mercury Release, www.opengroup.org/open-group-osdu-forum-launches-osdu-data-platform-mercury-release, 2021.

[27] OSDU Application Developers – OSDU, osduforum.org/about-us/communities/application-developer-service-provider-cloud-providers/

[28] How can I start developing on the OSDU Data Platform? – OSDU, osduforum.org/application-development/how-can-i-start-developing-on-osdu/

[29] The OSDU Catalog – OSDU Forum, osduforum.org/catalog/the-osdu-catalog/

Category	Description
Data	A list of available data sets that can be loaded into the OSDU Data Platform to run or test applications against
Consulting	A list of consulting companies, system integrators, and applications service providers for support
Applications	A list of applications and services that is provided by Independent Software Vendors (ISV) or as open source and is compatible with the OSDU Data Platform
Platforms	A list of currently available OSDU Data Platforms in the market, ranging from large, dedicated Cloud Service Provider solutions to ready to use SaaS imple-mentations of the OSDU platform.

Table 2.2B: Types of products and services within the OSDU Catalog, Part 2

While the existing OSDU Catalog offers an area where vendors can display their existing products and services, it lacks functionality for consumers to request solutions and products from vendors. To close this gap, an Innovation Marketplace was created to support upfront demand specification such that the solution developer ecosystem can build solutions for identified consumer requirements. The Innovation Marketplace hopes to broaden the types of OSDU-aligned solutions created and facilitate entry for smaller players, resulting in improved and faster solutions[30].

2.3.2 Machine Learning Model Use Cases for Offshore Energy

The industry research conducted for this White Paper highlighted the need for two types of machine learning solutions:

- Machine learning models that offshore energy companies can integrate and utilise within their in-house data science teams to automate certain human tasks
- Machine learning models that can be integrated with robotics programming to enhance RAS

The first type focuses solely on offshore energy companies and their internal processes, whereas the second type highlights how these companies in collaboration with RAS providers, other robotics experts, and analytics service companies can continuously improve the functionality of RAS through time. Further discovery work is needed to understand the ecosystem for enhancement of RAS functionalities.

Identifying use cases for machine learning with RAS data can be challenging as the technology is not yet fully developed and may need an iterative approach whereby the offshore energy ecosystem collectively decides on high priority issues. Despite this, there are some existing processes which could benefit from machine learning and could form initial use cases:

Corrosion Prediction

Corrosion is a major challenge within offshore energy production, its damage to energy assets poses various threats to the environment and humans in terms of both contamination and accidents. The resulting increased risk of failure and reduced asset life leads to more frequent chemical treatment, repair, inspection as well as unplanned shutdowns.

[30] Announcing the Innovation Marketplace – OSDU, osduforum.org/innovation-marketplace-blog/, 2022.

Automating corrosion prediction can extend asset life, reduce resource usage, minimise contamination, and lower the risk of asset failure. As a result, offshore energy companies would see improvements in sustainability and human safety as well as capital, operating and maintenance expenditure.

This can be achieved through pairing machine learning models with RAS technology to analyse imagery data of offshore equipment. Currently, this has only extended as far as academia, which determine various elements of corrosion through:

- Using image texture analysis and meta-heuristic optimized machine learning approaches to separate images into segments and classify these segments as corrosion or non-corrosion[31]
- Using Bayesian deep learning models to provide confidence estimates at the pixel level for corrosion detection[32]
- Using k-means clustering algorithm to identify centres of corrosion clusters[33]

The importance of corrosion prediction is well-understood between offshore energy companies. However, the automation of this capability is still very much in its infancy, as it has yet to be sufficiently developed enough for industry use. Therefore, the need to develop this machine learning technique is extremely important for the industry in its endeavour to minimise corrosion costs.

Gas Bubble Formations

Multiphase flow is a common yet extremely complex phenomenon within offshore energy operations and has been difficult to measure and simulate numerically. Dysfunctions in multiphase flow can lead to production inefficiencies due to poor heat and mass transfer as well as safety implications.

Measuring gas release from the formation of bubbles within offshore energy production is important for operators to better understand and optimise fluid dynamic behaviour in their processes as well as prevent any potential safety hazards associated with this phenomenon. The quantification of gas release through bubbles can be very numerically demanding as it requires deciphering from bubble sizes in images as well as consideration of multiple flow factors (e.g., flow rate, pressure drops, fluid composition, fluid properties etc.). Furthermore, as some bubbles can be extremely small, precision becomes much more important to minimise the relative error within measurements. As such, relying on humans to quantify gas release through bubble formation can be strenuous, inefficient and error-prone.

Using machine learning and vision can be much more effective in understanding multiphase flow and mitigating potential dysfunctions surrounding this phenomenon through the development of deep learning frameworks to estimate gas release and formation from images[34]. The applicability of these machine learning algorithms can extend through multiple different operations within energy production due to the common occurrence of multiphase flow, examples of this include heat exchangers, separation processes, reaction processes and fluid flow through pipes.

[31] *Image Processing-Based Detection of Pipe Corrosion Using Texture Analysis and Metaheuristic-Optimized Machine Learning Approach – Nhat-Duc Hoang and Van-Duc Tan, 2019.*

[32] *Deep learning corrosion detection with confidence – Will Nash, Liang Zhang, and Nick Birbilis, 2022.*

[33] *Detection and quantitative assessment of corrosion on pipelines through image analysis – Venkatasainath Bondada, Dilip Kumar, Pratihara Cheruvu, and Siva Kumara, 2018.*

[34] *On the Application of Image Processing Methods for Bubble Recognition to the Study of Subcooled Flow Boiling of Water in Rectangular Channels - Concepción Paz, Marcos Conde, Jacobo Porteiro, and Miguel Concheiro, 2017.*

3. PROPOSED APPROACHES FOR SHARED ANALYTICS

Based on the aims of a Shared Analytics concept for the offshore energy ecosystem, this section will evaluate three potential approaches for how this might function.
The intention for the Shared Analytics concept is to create an environment (platform) that enables the development and training of machine learning models, as well as a marketplace in which these can be purchased or sold, and consumer requests can be submitted.

Therefore, the approaches will all include an analytics platform element as well as a marketplace. This section will describe the assumptions made, advantages and risks associated with each approach, potential impacts of AI legislation, AI assurance frameworks, and suggestions on key activities that OLTER's Data Hub can drive.

3.1 High-Level Assumptions

- Common goals exist between end users of the Shared Analytics concept
- The approaches outlined are not mutually exclusive and components from each can be interchanged as needed
- It is accepted that the current scope for this is high-level and ambiguous in nature, and as such, further discovery and refinement processes may need to be undertaken, after which the approach designs are subject to change to meet the more detailed requirements
- RAS payload data will be the current focus of the approaches described in this White Paper, however there is also the potential to incorporate other data types
- Static data has been assumed for all approaches; no real time considerations have been included although the architectural principles used in the designs follow the SOLID[35] foundations so can be extended if required
- User access: Authentication and Authorisation will need to be managed using multi factor authentication (MFA)

3.2 Shared Analytics Approaches

3.2.1 Generic Components

In the three approaches discussed below, there are components which will remain generic and the same throughout. These are the marketplace, third party accreditation, and user interface requirements:

Marketplace

The principal goal of the marketplace is to provide an environment where machine learning as a service (MLaaS[36]) can be vended to energy companies or other end users. Existing machine learning models/products or those created through the analytics platform can be published/advertised and purchased by end users. Energy companies can send requests within the marketplace for solutions they require, and either purchase these as existing products in a catalogue or wait for industry/public developers to create new products. This can create an industry backlog that defines the functional requirements and use cases, and companies can then expand their product and service offerings by building out these capabilities.

[35] *Software interoperability: Principles and practice. In Proceedings of the 21st international conference on Software engineering – J.C. Wileden and A. Kaplan, 1999.*

[36] *MLaaS: Machine Learning as a Service, 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA) – M. Ribeiro, K. Grolinger and M. A. M. Capretz, 2015.*

Third Party Accreditation

There is a potential question around the credibility of vendors and users who are submitting machine learning solutions to the platform for purchase by offshore energy players. A mechanism for building trust around these submissions may be required to accredit the third parties participating. Most of the methods for achieving this involve the ability to ensure that the companies and individuals who propose 'products' for submission are vetted and credentialed. This would need to be investigated further based on industry requirements.

User Interface

Using a web interface, users can log on to a portal environment and search through data and a catalogue of machine learning algorithms as well explore test data sets. The interface will include the functional capability to connect to a data source and run a machine learning algorithm from the platform as well as the option to upload their own data sets and execute machine learning in situ on the actual platform itself.

3.2.2 Approach 1: Data Hub's Current Architecture Integrated with Kaggle and Custom Marketplace

This approach evaluates a business-as-usual model, whereby the OLTER Data Hub goes ahead with a short-term plan to integrate their existing architecture with Kaggle21 or a similar vetted third party to fulfil the analytics platform and public competitions criteria. A custom marketplace has also been incorporated as an enhancement. The key components that constitute this approach are shown in Figure 3.2A, and each component will be discussed below, along with the advantages, assumptions, and risks.

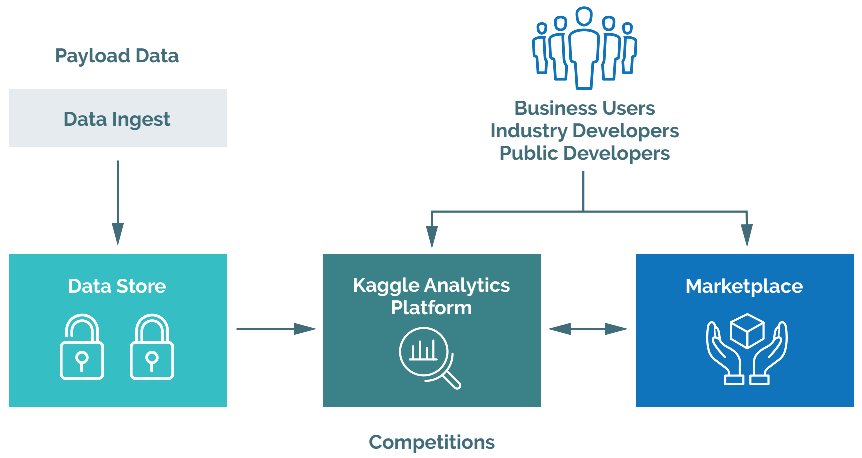


Figure 3.2A: Shared Analytics Approach 1

Data Store

The database element of the whole system will be provided by the OLTER Data Hub's Data Stores. The sensitivity of the data sets (in this example: payload data) will be defined and agreed between OLTER and the energy companies providing data and stored accordingly in the Shared or Open Data Stores. To enable the solution to be stood up quickly, the initial datasets will be a manual load of data. Over time and depending on the frequency required for the refresh of the data, there would be an option for those data producers to have an end point to which they would be able to automatically push data to and thus minimising the overhead of this load.

Kaggle Analytics Platform

Kaggle's Public API or featured competitions packages with custom pricing can be used for integration. An API and command line interface (CLI) tool provide easy ways for users (e.g., industry or public developers) to create, interact with, and download datasets[37]. Users can explore and run machine learning code via the platform's cloud computational environment, Kaggle Notebooks[38].

Public or private competitions can be hosted: the competition host (energy company) prepares the data and a detailed description of the problem. For public competitions, the hosts' posted competition data is completely open to the public with only a Kaggle login required. For private competitions, the hosts can restrict access to participants. The competition participants (industry or public developers) submit their machine learning models as the solutions. These are scored using a set of evaluation criteria which use statistical analysis to determine model accuracy and reliability. Once the competition deadline has passed, the host pays the prize money to the winner. The host will have the sole ownership and royalty-free license to use the winning entry as well as the IP23.

Marketplace

A provision for a marketplace as outlined in Section 3.2.1 that coexists with the Kaggle platform is included in this approach.

Key Advantages

- Compatible with the existing OLTER Data Hub architecture
- Leverages an existing data science community and platform and can utilise Kaggle's front-end
- Potential to demonstrate value early in the form of machine learning products as the Data Hub's architecture can be integrated relatively quickly via Kaggle's API

Assumptions and Risks

Assumptions	Risks
Energy companies will provide the required data sets needed for machine learning development and that this will be stored in the OLTER Data Hub	High-risk assumption as energy companies are hesitant to share such data externally and the process for accessing and sharing this data is challenging
There are shared data standards and governance across the industry data that bring parity to the results derived from ML analysis	High-risk assumption due to lack of industry insight. Data formats can vary between companies for the same data types, result-ing in the development of ML models that may not be usable for other companies

Table 3.1: Assumptions and risks associated with Shared Analytics Approach 1

[37] *Public API – Kaggle, www.kaggle.com/docs/api*
[38] *Notebooks – Kaggle, www.kaggle.com/docs/notebooks*

3.2.3 Approach 2: Custom Built Analytics Platform and Marketplace

This approach evaluates the use of a standardised method for building an analytics platform along with a marketplace. Using standard cloud services and building functionality on top of these following the SOLID[35] recommended software development lifecycle, a web application and supporting software stack can be orchestrated that will be easy to set up from an infrastructure perspective as well as to further develop. This design pattern is recognised as the most efficient and follows most mainstream developments. The key components that constitute this approach are shown in Figure 3.2B, and each component will be discussed below, along with the advantages, assumptions, and risks.

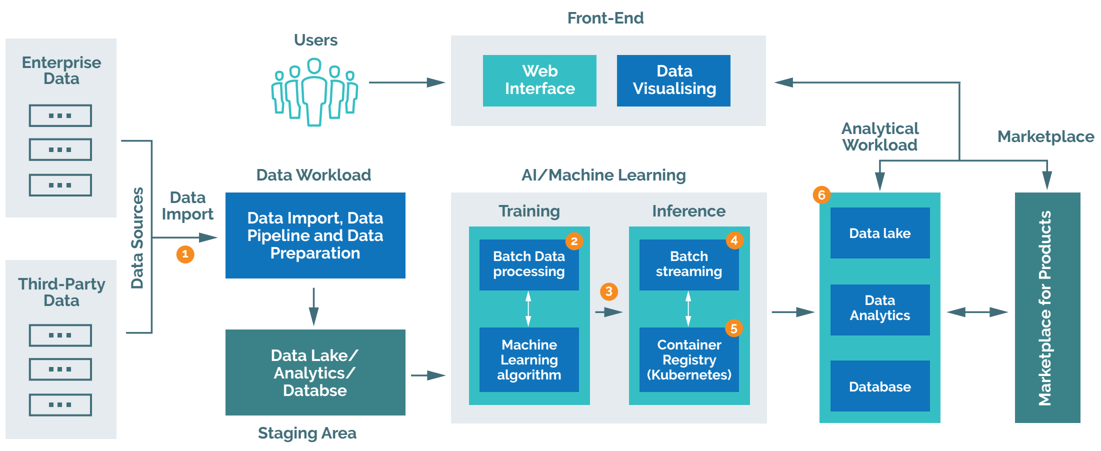


Figure 3.2B: Shared Analytics Approach 2

1. **Data Ingestion**
   - Data is pulled from a source database (OLTER Data Hub Shared/ Open Data Store) and copies it to a cloud-based data lake/database.
2. **Model-training pipeline**
   - **Prepare data:** The training pipeline pulls the data from the data lake and uses a processing option to group it into datasets for training the models.
   - **Train models:** The pipeline trains models for all the datasets that were created during data preparation. It uses the function API to train multiple models in parallel. After a model is trained, the pipeline registers it into the machine learning component along with the testing metrics.
3. **Model-promotion pipeline**
   - **Evaluate models:** The promotion pipeline evaluates the trained models before moving them to production. A DevOps pipeline applies business logic to determine whether a model meets the criteria for deployment. For example, it might check that the accuracy of the testing data is over 80 percent.
   - **Register models:** The promotion pipeline registers the models that qualify to the production Machine Learning workspace.
4. **Model batch-scoring pipeline**
   - **Prepare data:** The batch-scoring pipeline pulls data from data lake and uses a processing option to group it into data sets for scoring.
   - **Score models:** The pipeline uses the function API to score multiple datasets in parallel. It finds the appropriate model for each dataset in the Machine Learning component by searching the model tags. Then it downloads
     the model and uses it to score the dataset. It uses the connector to the analytics/ database component to retain the results.
5. **Real-time scoring**
   - Kubernetes can do real-time scoring if needed. Because of the large number of models, they should be loaded on demand, not pre-loaded.
6. **Results**
   - **Predictions:** The batch-scoring pipeline saves predictions to the analytics component.
   - **Metrics:** A data visualisation tool connects to the model predictions to retrieve and aggregate results for presentation.

Marketplace

A marketplace component as outlined in section 3.2.1 is included in this approach for machine learning solutions to be sold as products or as a service, and to enable industry requests to be shared.

Key Advantages

- Clear, defined process for standing up an analytics platform using standard technologies and methods that can be iteratively built on as business needs change with time
- Provides required functionalities for machine learning model development, training, and refinement
- Compatible with the existing OLTER Data Hub architecture

Assumptions and Risks

Assumptions	Risks
Energy companies will provide the required data sets as the data sources needed for machine learning development and that this will be stored in the OLTER Data Hub	High-risk assumption as energy companies are hesitant to share such data externally and the process for accessing and sharing this data is challenging
There are shared data standards and gov-ernance across the industry data that bring parity to the results derived from ML analysis	High-risk assumption due to lack of industry in-sight. Data formats can vary between companies for the same data types, resulting in the devel-opment of ML models that may not be usable for other companies

Table 3.2: Assumptions and risks associated with Shared Analytics Approach 2

3.2.4 Approach 3: Data Governance, Custom Built Analytics Platform and Marketplace

While Approach 2 outlines the required analytics services and infrastructure needed for the Shared Analytics concept, it still shares a high-risk assumption that the data sources used for importing the data will have the same data standards and formats applied. This limits the future scalability and usability of any machine learning products created via this approach across the industry. To maximise the value and efficiency of this concept, the products that are built should be interoperable within industry. This approach evaluates how this can be achieved through establishing an OLTER Forum in addition to the analytics platform and marketplace outlined in Approach 2. This is displayed in Figure 3.2C, building on Figure 3.2B:

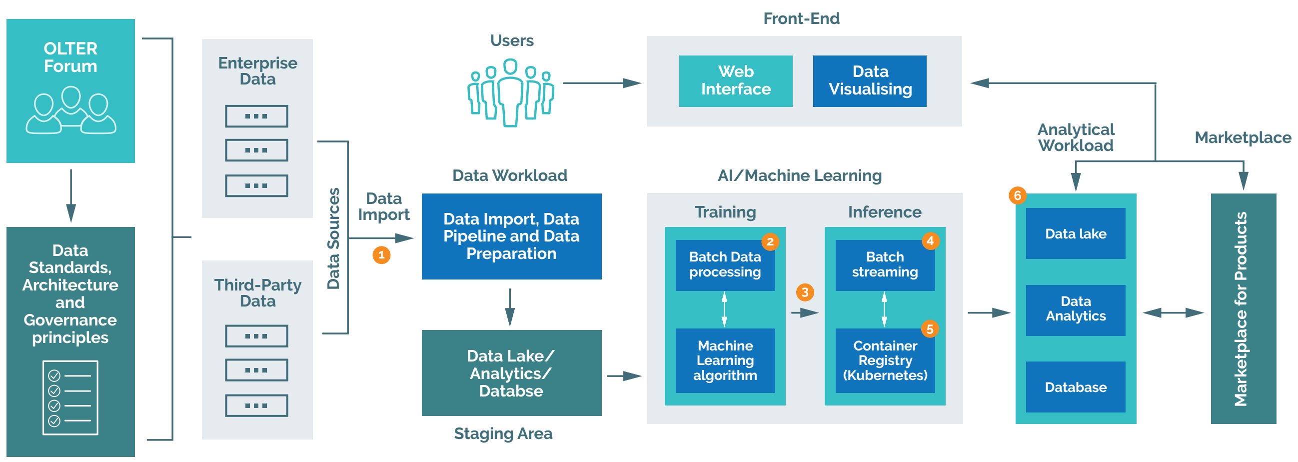


Figure 3.C: Shared Analytics Approach 3

OLTER Forum

A forum consisting of energy companies, RAS companies, academia, and analytics service companies at an industry-wide scale can collaborate to create RAS data standards. To do this, the forum may need to interact with existing open standards forums to ensure an industry approach is followed. As the group size increases, a management committee can be formed to shape the future path of the forum, defining new goals and initiatives to continue the development of data frameworks and new solutions. Costs for maintaining working groups, the analytics platform and marketplace can be shared between members through a yearly membership fee.

The forum can identify high priority use cases for machine learning solutions, discuss the associated data types and formats that exist in industry, and create industry standards for them.

Data Sets

This approach still includes the creation of an analytics platform for machine learning development, and it operates in the same way as described in section 3.2.3. However, because of the forum's work on data standards, architecture and governance principles, the processes followed and data that is loaded into the platform will be aligned with an industry-agreed framework. This mitigates the risk of creating siloed products that cannot be marketed or utilised widely. Additionally, if the data required for the platform's training data sets are limited or unattainable, the forum can provide guidance on the necessary data formats and analytics service companies can use this to create synthetic data instead. Depending on the use case or data types involved, investigation would be needed to confirm the feasibility of this method. It can be assumed that synthetic data can use publicly available data and alter it to match the guidelines, or it can create these data sets from scratch to meet specified requirements. This can potentially mitigate the challenge around energy companies contributing their data to the Shared Analytics concept.

Marketplace

The marketplace component in this approach follows that of the marketplace in Approaches 1 and 2, with additional governance in that the products that are listed in this case will have to be aligned to the industry standards established via the forum. This may add further reassurance and confidence in the products or services that are marketed.

Key Advantages

- Forum approach enables offshore energy players to engage in areas where no one competes (data standards and governance) to create uniformity across RAS data types
- Forum allows initiatives to be executed at an accelerated pace to create new technological developments for the entire industry
- Iterative approach enables members to prioritise the development of solutions based on industry needs
- Agreed data standards and governance enable scalability of both the analytics platform and the products created within it, providing the most efficient and sustainable long-term approach

Assumptions and Risks

Assumptions	Risks
Energy companies can provide the re-quired data sets as the data sources needed for machine learning develop-ment and that this will be stored in the OLTER Data Hub	Low-risk assumption as if energy companies cannot contribute this data, other methods such as creation of synthetic data can be used with their guidance and expertise
OLTER will establish a forum and receive participation from key offshore industry players	Medium-risk assumption as OLTER currently has existing SteerCo and Design Authority members from industry, however more companies will need to be involved in order to agree on industry standards
The process for creating industry-accepted data standards for RAS data can be achieved within OLTER	High-risk as further discovery work is needed to understand existing data standards and open forums. The timeframe required for this could be significant depending on industry participation

Table 3.3: Assumptions and risks associated with Shared Analytics Approach 3

3.3 AI

3.3.1 Impacts of Legislation

All AI applications created or adopted by an operator will need to be internally reviewed to ensure they meet with the organisation's safety guidelines. However, additional legislature is projected to have an impact on the operation of the analytics platform. The EU AI Act[39] will impose that owners of 'high-risk' applications, which encompass any AI system which have some level of interaction with humans (directly or indirectly), will be subject to a conformity assessment on their system before it can be placed in the EU market. In the context of OLTER's Shared Analytics Platform, algorithms that are implemented in the proposed marketplace that have an element of human interaction would therefore be considered high risk and would need to undergo conformity assessments. How conformity assessments would explicitly impact on the logistical operation of deploying algorithms in a marketplace is to be further determined; for example, whether having certification of a conformity assessment would be required or whether all high-risk applications are banned outright.

The AI Liability Act[40] is another proposed EU law that will place a "presumption of causality" on the AI systems. This means that the burden of proof for demonstrating how an AI application has made an impact on an individual's physical or mental wellbeing will be lowered, and owners of AI systems may be required to provide compensation for adverse impacts. This law may significantly increase the legal risk of hosting high-risk algorithms on the Analytics Platform marketplace for OLTER. Therefore, the development of AI should follow assurance frameworks to mitigate such risks and ensure human safety guidelines are met.

The UK has its own directive known as the "National AI Strategy"[41]. While the National Strategy does not currently look to enforce restrictions on the use of AI, one of the key pillars of the Strategy is to 'Govern AI Effectively', which includes the establishment of a governance framework, an AI Standards hub and a Centre for Data Ethics and Innovation (CDEI) assurance roadmap, all of which look to establish cross-industry standards in the application of AI. The National AI Strategy further aims to collaborate with the EU AI Act to harmonise on the construction of safe AI frameworks.

3.3.2 Development of Frameworks for AI Assurance

In response to the propagation of AI in industrial applications and the responsive legislation, several frameworks have been developed to inform operators on how to deploy their AI systems in a safe and secure manner. Four key frameworks have been identified as being highly appropriate to the development and use of AI within OLTER's Shared Analytics concept, in brief:

Framework	Author	Description
capAI[42]	University of Oxford	Gives a list of requirements on how to create an EU-Artificial Intelligence Act-compliant application and a short description on how to meet the requirements. It is the one of the few frameworks created with specific (upcoming) legislation in mind and provides a robust framework to establish basic data assurance principles.
Safety Assurance Objectives for Autonomous Systems[43]	Safety Critical Systems Club (SCSC)	Gives guidance for owners of autonomous systems to en-sure that their systems have safe AI practices from an architectural, computational and the "platform" (in the context of OLTER, this would be a RAS vehicle) perspective. This framework is particularly critical in ensuring that the RAS vehicles conforms to its underlying AI behaviour.
Securing Machine Learning Algorithms[44]	European Un-ion Agency for Cybersecurity (ENISA)	Considers various vulnerabilities an AI application can have throughout its development cycle and the corre-sponding tools that can be implemented to counteract any cybersecurity threats. The report gives multiple struc-tured perspectives on how to securely interact with all the main threats that an AI system can face.
DNV-RP-0510[45]	DNV	Gives a series of claims that can be made about an AI ap-plication's safety assurance and provides guidance on how to make each claim. It is structurally similar to the ENISA and SCSC reports, but with additional guidelines which focus on safety assurances from business management and documentation perspectives.

Table 3.4: AI assurance frameworks applicable to OLTER's Shared Analytics concept

To progress with the Shared Analytics concept, the aforementioned frameworks should be used as a baseline to establish a forum discussion, such that a bespoke framework of safety guidelines for algorithms can be created.

3.4 OLTER Data Hub

In addition to the Data Hub's ownership of the activities proposed above, there are key areas where the Data Hub can position itself to lead and drive forward an industry-accepted Shared Analytics approach:

- Defining industry-agreed data acquisition recommendations and guidelines e.g., providing the ideal set of conditions RAS images and videos should meet to be compatible for machine learning model development. This can help inform robotics developers and energy companies on the specifications needed when selecting RAS vehicles
- Establishing data standards, architecture, and governance principles to create a RAS data framework accepted by the offshore energy industry. This can be achieved using a similar approach to existing open data forums: creating a management committee and working group/s that are representative of the offshore energy industry to discuss, define and agree on these standards for any machine learning use cases that will utilise the Shared Analytics platform. Further discovery work is required to better define whether this can be achieved solely within OLTER, or if OLTER's Data Hub should position itself as a conduit between OLTER's forum and existing open standards committees (e.g., the British Standards Institution)

[39] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts – European Union, 2021.*

[40] *Proposal for a Directive on adapting non contractual civil liability rules to artificial intelligence – European Union, 2022.*

[41] *National AI Strategy – HM Government, 2021.*

[42] *capAI - A Procedure for Conducting Conformity Assessment of AI Systems in Line with the EU Artificial Intelligence Act – L. Floridi, M. Howleg, M. Taddeo, J. A. Silva, J. Mokander and Y. Wen, 2022.*

[43] *Safety Assura nce Objectives for Autonomous Systems – The Safety of Autonomous Systems Working Group, 2022.*

[44] *Securing Machine Learning Algorithms – European Union Agency for Cybersecurity, 2021.*

[45] *Framework for assurance of data-driven algorithms and models – DNV, 2020.*

4. PARTICIPANTS AND END USERS

This section gives a brief introduction relating to each participant and end user and discusses how they may contribute and find value with respect to the third solution approach of Shared Analytics concept. This solution approach takes the holistic view through considering both the data governance and shared analytics aspect of the platform.

It is important to note that, the contributions and value outlined for each participant group are propositions and not definite as more work is needed to better understand this. Furthermore, as the Shared Analytics concept is intended to support activities within the North Sea, this section takes an offshore perspective regarding participant and end user context, contributions, and value.

4.1 Integrated Energy Companies (IECs) and Minor Oil and Gas Companies

4.1.1 Context

IECs are characterised by their highly variable asset base and have historically been the key developer, enabler and financier of the UK's oil and gas reserves in the North Sea. They have operations that play across the entire value chain from exploration to downstream refining and as such have access to a wealth of data and experience. Since the 2014 oil price crash, many IECs continue to divest marginal assets and scale back on large scale development and investment in the basin. As of recently, many IECs are beginning to transform their ways of working to become more digitally as well as energy focused and are making significant progress in their plans to increase renewable and low carbon investments by 2030.

Minor oil and gas companies share many similarities with IECs, albeit at a smaller scale. They typically are not involved in activities across the oil and gas value chain (like their larger competitors) and are not involved in other forms of energy production (e.g., renewables). Minor oil and gas companies tend to be focused only on upstream activities, normally exploration and production.

As IECs have begun to divest their assets, minor oil and gas companies have risen to prominence in the North Sea through mergers and acquisitions (M&A), helped by an influx of private equity funding. The COVID-19 pandemic has accelerated this trend, as a period of low oil prices caused some assets devalue sharply, particularly in mature basins such as the North Sea. Some private equity funds have used this as an opportunity to buy assets at knocked down prices – in 2021, private equity funds bought £11.9 billion of European oil and gas business, as UK businesses made up five of these 12 deals, totaling £2.4 billion[46]. This trend started more than a decade ago as the proportion of private exploration and production within the UK increased from 8% in 2010 to 30% by 2020[47].

[46] *Private equity funds accelerate acquisitions of oil and gas assets – Mayer Brown, www.mayerbrown.com/en/news/2022/private-equity-funds-accelerate-acquisitions-of-oil-and-gas-assets, 2022.*

[47] *The new North Sea players riding the wake of the retreating majors – Rystad Energy and FT, www.ft.com/content/93d5f778-833c-4553-ae29-785e3aa3d4d3, 2022.*

4.1.2 How can IECs and Minor Oil and Gas Companies Contribute?

IECs and minor oil and gas companies will play an integral role within the Shared Analytics concept. They are key in both the development of an open standards governance model and the Shared Analytics platform itself.

The potential ways IECs and minor oil and gas companies can contribute to the Shared Analytics concept are:

- Provide valuable information around the necessary data format requirements and offer important feedback through the development of data standards
- Leverage engineering expertise to contextualise common problems faced within the industry and the development of machine learning algorithms
- Identify high value use cases for the Shared Analytics concept
- Where available, provide data sets for the training of machine learning algorithms

4.1.3 Value for IECs

Through the development of an open standards data governance model, IECs will be able to reduce existence of data silos within industry, this will improve data accessibility as more data sets become available for the training of machine learning models. Furthermore, the standardisation of data types/formats across the industry will enhance the versatility of these algorithms, allowing for the development of industry-standard machine learning models that can be used by any offshore energy company subscribed and aligned to the data governance model. This will allow for greater efficiency in the creation of machine learning techniques.

From a Shared Analytics perspective, an IEC will see the value of pooling resources with other IECs to tackle common or recurring problems and issues. This would optimise problem solving within the industry, as IECs can share expertise around solutions they have developed to overcome problems experienced in the past. For example, if an operator has collected inspection photos of a pressure vessel and developed an algorithm to assess the condition of the vessel, to determine whether corrosion exists and if maintenance is required, the IEC can then choose to share this with other operators. If other operators upload additional pressure vessel images to the data set it will further improve the accuracy of the algorithm. Furthermore, a Shared Analytics environment will provide IECs with better access to data scientists and machine learning talent. This will therefore accelerate the development of machine learning techniques as well as ensure such techniques are built to a sufficient standard.

IECs will also see the value in increasing operational efficiency. Using logistics information as an example, if a helicopter or supply vessel is transiting to Asset X and Asset Y of another operator is close by, the helicopter could deliver required equipment or additional supplies to both sites in one journey. As a result, IECs can reduce costs, improve sustainability as well as miminise worker safety through reducing the number of trips needed for transportation.

As automation becomes more prevalent in the oil and gas industry, a Shared Analytics Platform could further facilitate the development of autonomous systems both on the surface or subsea. Other values to IECs are listed below:

- Demonstratable CO_2 reduction efforts and progress to the regulator
- Identification of previously unknown similarities/synergies
- Increase in transparency into service company performance and information
- Faster reaction to operational challenges and solutioning
- Service providers will be able to undertake actionable outcomes based off model insights. (e.g., a service engineer can visit a failing valve and fix it)

4.1.4 Value for Minor Oil and Gas Companies

The value for minor oil and gas companies would be like that of IECs. However, due to their lean operating models, the value may be harder to articulate and subsequently realised for minor oil and gas companies.

There is potential for these companies to see the investment and level of effort as a greater barrier compared to IECs, particularly in the earlier stages of development the Shared Analytics concept. That said, some examples of value creation that could be envisaged specifically to minor oil and gas companies on top of those already stated above for IECs are discussed below.

Leverage of Data and Competency

Minor oil and gas companies may not have access to the same level of historical data sets as the IECs nor the same size of internal analytics teams. By participating in the Shared Analytics concept, they would be able to equally compete pulling data and leveraging from the wider data science landscape.

Operational Efficiency Increase

Many minor oil and gas companies' portfolios are acquired from an IEC divestment program and are characterised by ageing assets. These assets are commonly reaching the end of their functional design lives and therefore suffer from a range of maintenance related challenges. Placing payload data into a Shared Analytics platform would enable the sharing of learnings on how best to manage respective maintenance programs. For example, if images were taken by a drone of the external visage or cladding of an offshore asset paired with the specification information, then it would enable the development of machine learning algorithms (by third parties) that could identify which areas were most at risk of failure or needed replacement. This could then help optimise the maintenance program and related logistics required to carry out the operation, enabling further synergies with other assets activities.

4.2 Offshore Renewable Companies

4.2.1 Context

Currently, the offshore wind is the only form of the UK's offshore renewable energy production, and therefore this section will focus on the offshore wind. However, as decarbonisation efforts continue, it is anticipated that offshore renewable production will diversify beyond offshore wind, and such this section will refer to participants as "offshore renewable companies" unless points made are otherwise specific to the offshore wind industry.

Since its inception in 2000, the offshore wind industry has risen to become an integral part of the UK's energy system[48] – by 2021, offshore wind energy makes up almost one third of total renewable electricity generation and as of June 2022, total offshore wind power capacity installed in the UK reached 13.1 GW[49], in which the Government plans to grow this to 50 GW by 2030[50]. With the need to decarbonise the economy, offshore wind has become pivotal in UK's energy transition. As of recently, IECs have made significant strides to break into the industry through joint ventures and M&A.

[48] *UK offshore wind industry – ORE Catapult, https://guidetoanoffshorewindfarm.com/offshore-wind-history#:~:text=2000%3A%20First%20offshore%20project,less%20than%202km%20from%20shore.*

[49] *Renewable electricity capacity and generation – HM Government, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1107457/ET_6.1_SEP_22.xlsx, 2022*

[50] *Offshore wind – HM Government, https://www.great.gov.uk/international/content/investment/sectors/offshore-wind/#:~:text=This%20ambition%20was%20increased%20through,this%20should%20utilise%20floating%20technology., 2022*

Offshore wind energy is a less of a developed technology compared to oil and gas, as the industry has yet to optimise the parts of their infrastructure, such as asset integrity. Furthermore, offshore wind energy companies face technical difficulties of turbine construction and connection to the National Grid[51]. With this, players are beginning to recognise the need to invest in digitalisation to aid with the design, construction, operation, and maintenance (O&M) of their assets, as this is a key driver for O&M cost reduction.

4.2.2 How can Offshore Renewable Companies Contribute?

Offshore renewable companies can contribute in very similar ways to IECs through providing their information around data formats, engineering context of problems, and where possible, the data sets themselves. Although contributions may be similar to IECs, they will mostly exist in a different domain within the Shared Analytics platform, which caters more towards machine learning models for offshore renewable operations (i.e., maintenance of wind farm assets).

4.2.3 Value for Offshore Renewable Companies

The value for offshore renewable companies derived from the Shared Analytics concept are also very similar to that of IECs as they are both operators within the offshore environment. However, these companies may benefit in unique ways due to being in different point of their life cycle.

As offshore wind companies look to integrate digitalisation into their business model at an earlier stage of the industry's life cycle, development of digital technologies in this space may be slower due to the lack of impactful RAS machine learning model use cases for renewables players. Through collaboration between operators and developers within a Shared Analytics environment, insights into useful RAS use cases can be realised and subsequently lead to manifestation of useful machine learning tools for offshore wind players.

Furthermore, a well-established data governance framework will help prevent formation of data silos for RAS data. As seen for IECs, who are currently facing problems with respect to working around data silos; this complication will be evaded for offshore renewable companies as they will not need to exert a considerable amount of effort in dealing with siloed RAS data in the future.

4.3 Oilfield Service (OFS) Companies and RAS Providers

4.3.1 Context

OFS companies deliver products and services associated with oil and gas exploration, development, construction, production, and maintenance. OFS companies are involved with manufacturing, repair and maintenance of equipment and assets used in the extraction and transportation process.

OFS companies range in capabilities and levels of expertise and can provide a range of services such as, drilling rigs, offshore scaffolding, and logistics. They enable IECs to deliver and manage complex operations that would not otherwise be possible.

[51] *Onshore vs Offshore Wind Energy – Boythorpe Wind Energy, https://www.boythorpewindenergy.co.uk/wind-turbine-advice/onshore-vs-offshore-wind-energy/*

As offshore energy production now involves renewable energy and moves towards digitalisation as well as automation, many OFS companies have begun to invest heavily in the use of RAS technology and are now becoming robotics providers. RAS technology is very much still in its infancy as the first deployment of an offshore autonomous robot was in 2018[52]. However, the use cases for robotics within an offshore environment have become apparent in improving inspection, monitoring and maintenance efficiencies as well as worker safety.

Several RAS providers are now beginning to adopt Robotics as a Service (RaaS) as a business strategy in which offshore energy companies can avoid the cost of inventory and obtain robotic services when required. This could lead to market standardisation in robotics depending on the service rendered. As such, RAS vendors are looking to customise robots for several different applications, giving both IECs and offshore renewable companies an expanded pool to choose from for their specific needs.

4.3.2 How can OFS Companies and RAS Providers Contribute?

Data ownership is dictated through contractual agreements between OFS companies and RAS providers and offshore operators and generally, OFS companies and RAS providers do not legally own the data captured from their services as this is usually owned by the offshore energy companies. Therefore, OFS companies and RAS providers are restricted from supplying data sets to help train machine learning models on the Shared Analytics platform, and their contributions may be limited.

Despite this, OFS companies and RAS providers can provide insightful knowledge and information around the development of certain machine learning models, specifically related to engineering activities, timings, trips, condition of equipment and materials and duration related to construction, production, and maintenance. This information can be combined with other factors such as: weather, time of year, production rates etc., to improve any planning and resource management machine learning models that are developed within the Shared Analytics platform.

4.3.3 Value for OFS Companies and RAS Providers

As the Shared Analytics platform matures OFS companies and RAS providers will have the potential to build applications based off the data, statistical models, and APIs exposed through the platform through which they can start to offer new services. The APIs would be built on the third party's infrastructure and then published to a catalogue on the platform to which end users can subscribe to. This will enable OFS companies and RAS providers to refine their global services, equipment, and asset allocations to the parts of the basin that are more likely to require them, enabling faster response times to clients, shorter resolution times and a better overall service experience.

OFS companies and RAS providers could also see value through having greater ability to:

- Tailor and improve their existing or potential services, equipment, and assets through additional data sources
- Build out new services to their clients and develop new novel business lines that were previously inaccessible
- Conduct more extensive qualification testing on new solutions, in a more cost-effective manner. Strong potential for conducting extensive modelling and testing programs on new solutions given the access to the new data sources

[52] *World's First Autonomous Offshore Robot – Anybotics, https://www.anybotics.com/worlds-first-autonomous-offshore-robot/, 2018.*

4.4 Analytics Service Companies

4.4.1 Context

Analytic service companies are focused on delivering IT, professional, data and security focused services to offshore energy company and across the whole energy value chain to help them extract additional value. Analytic service companies have been at the forefront of helping the industry transform their analogue operations to digital ways of doing things. These can be large scale organisations who leverage their experiences and best practices from across multiple industries to provide new insights and ways of working.

4.4.2 How can an Analytics Service Company Contribute?

Service providers who have pre-existing machine learning models can load these into the Shared Analytics platform for offshore energy company to use. The capability will exist for service providers to also create/publish new models that are relevant to the industry.

Service providers will be involved in analysing the data to improve predictions and enable applications to take place based on the analytics. Specifically, the service providers involvement could contribute:

- Algorithms
- Bespoke APIs
- Data science capabilities
- Design experience for a suitable platform for the open data to reside
- Experience from other industries. Open data is not a new concept, being far more developed in the public sector as well as other industries

4.4.3 Value for Analytics Service Companies?

The Shared Analytics concept allows analytics service companies the opportunity to improve their expertise within the offshore energy industry. This enables them to provide better services to offshore energy companies and grow their presence within this space. Furthermore, they can take learnings from certain machine learning techniques that are developed within the Shared Analytics environment and apply them to the work they do in other industries.

Analytics service companies could also see value through a greater ability to:

- Build services and new capabilities on the platform and to deliver these to the market
- Tailor and improve services through collaboration with end-users
- Insight to build out new services to their clients and develop novel business lines that were previously inaccessible or unknown

4.5 North Sea Transition Authority (NSTA) and Other Regulatory Bodies

4.5.1 Context

The NSTA, previously known as the Oil and Gas Authority (OGA) until March 2022, regulates the licensing of exploration and development of the UK's offshore and onshore oil and gas resources, carbon storage, gas storage and unloading activities. The NSTA's role is to take the steps necessary to secure maximise the economic recovery of UK oil and gas and support the UK government is its drive to reach net zero[53].

The NSTA seeks to be a progressive and highly effective authority, doing all it can to attract investment and with those jobs, helping to anchor valuable skills and expertise in this country. It looks to the evidence and data provided by industry to enact control and provide oversight of the UK's oil and gas reserves, whilst ensuring it is done in the most efficient and sustainable way possible. The NSTA will be key in facilitating the oil and gas industry position as the UK moves towards the 2050 (2045 in Scotland) Net Zero target.

The NSTA recognises that data is a key asset within the oil and gas industry and is actively trying to encourage data sharing to drive innovation and accelerate the development of digital technologies. In 2019, the NSTA (OGA at the time) launched the UK's first oil and gas National Data Repository (NDR), which is a public data bank that contains petroleum-related information such as carbon storage, exploration, production, well data. The NSTA believes data is invaluable in all the operations on the UKCS from exploration to decommissioning and in assisting the drive to net zero, for example through identifying potential locations for carbon storage sites, and with this, is making more and more data freely available to industry, academia, and others[54]. With the NDR, it aims to enhance industry collaboration, preserve, and protect valuable data, and help create conditions to drive investment and new technologies[55].

4.5.2 How can the NSTA and Other Regulatory Bodies Contribute?

The NSTA would take a regulatory perspective within the Shared Analytics concept, as they can provide expertise around regulatory implications of sharing data as well as using machine learning models for activities within the North Sea.

The NSTA could also contribute by providing:

- Insight into the Shared Analytics concept's impact on other relevant government initiatives e.g., CO2 production
- Benchmarks across basin incumbents using industry performance dashboards
- Relevant API catalogue entries from other government platforms

4.5.3 Value for the NSTA and Other Regulatory Bodies

The Shared Analytics concept will promote the NSTA's remit to help drive innovation and provide the right environment for the creation of additional value and services. The NSTA and other regulatory bodies could also see value through being able to:

- Improve oversight of the industry players, activities, and a real time view of the operations in the basin
- Hold all parties in the industry to account for their actions through data analysis
- Spot improvement opportunities and help drive investment opportunities

[53] *NSTA About Us – https://www.nstauthority.co.uk/about-us/*

[54] *NSTA Data Centre Overview – NSTA, https://www.nstauthority.co.uk/data-centre/overview/*

[55] *UK's first oil and gas National Data Repository receives strong industry backing – NSTA, www.nstauthority.co.uk/news-publications/news/2018/press-release-ndr/, 2018*

4.6 Academic Institutions

4.6.1 Context

Academic institutions, which consist of universities and research institutions, have long been a critical part of developing talent, expertise, and technology for the North Sea basin. Offshore energy companies have long partnered with academic institutions in a mutually beneficial alliance, they give academic funding and scholarships to help them conduct cutting edge research and develop and attract new talent that could enable a competitive advantage.

4.6.2 How can Academic Institutions Contribute?

Academic institutions who have pre-existing machine learning models can load these into the platform for offshore energy companies to use. Specifically, academic institutions could contribute:

- Algorithms
- Data developed from laboratory experiments or pilots
- Data from scientific literature or previous/ongoing research projects
- Data science capabilities
- Simulations and new mathematical methods, formulas, or techniques
- Data related to new materials, or materials in different environments
- New formulae from theoretical or empirical models
- Scientific standards and references
- Robotics experts and SMEs

4.6.3 Value for Academic Institutions

The Shared Analytics concept can lead to industry partnerships and stronger relationships between academic intuitions and offshore energy companies. With more academia-industry collaboration, research programs can be tailored to the most relevant industry challenges across the whole energy value chain. Furthermore, academic institutions could also share data with each other or work collaboratively on the Shared Analytics platform.

Academic institutions could also see value as they would be able to:

- Improve expertise, technology, as well as research and development programmes (e.g., through PhD programmes)
- Demonstrate talent and capabilities directly to the offshore energy industry
- De-risk longer term research for both universities/research institution and their funding partner/company

5. CONCLUSION

The Shared Analytics concept can present an opportunity within offshore energy, to pool resources to achieve scale and momentum in developing machine learning techniques. The idea of Shared Analytics aims to improve data accessibility, provide offshore energy players with better access to machine learning specialists as well as promote collaboration within the industry.

This White Paper considered three solution approaches in facilitating a Shared Analytics environment:

1. Data Hub's Current Architecture Integrated with Kaggle and Custom Marketplace
2. Custom Built Analytics Platform and Marketplace
3. Data Governance, Custom Built Analytics Platform and Marketplace

Both the first and second approaches depend on the offshore energy industry's ability to share RAS data. However as discussed, most offshore energy companies will be hesitant to share these data sets due to these being safety critical (e.g., inspection images of assets) and/or difficult to anonymise as they have embedded metadata that cannot be removed (e.g., location data). Even if offshore energy companies agree to share data, the lack of industry-agreed data standards, architecture and governance models would mean that data formats would differ across companies and there would be a high risk of creating siloed machine learning products. Therefore, the first two solution approaches do not adequately overcome data accessibility challenges faced within the offshore energy industry.

The third solution approach, Data Governance, Custom Built Analytics Platform and Marketplace, includes an industry forum with the purpose of establishing an agreed data standards and governance framework for RAS data, as this would mitigate the data sharing issues. Furthermore, agreed data standards and principles would enable interoperability between in-house and external development of machine learning models/products. This is crucial in giving offshore energy companies the option to develop their machine learning techniques in-house should they wish to, without preventing them from utilising and aligning with RAS industry best practices. Therefore, it is suggested that this approach could be the most efficient and sustainable in the long-term.

Creating greater value from RAS data is possible but requires a change in how data and AI are being used across companies. Leading with data is essential for companies on their digital transformation journeys, and initiatives such as the Shared Analytics concept can provide the opportunity to connect data and people as well as ideas and outcomes[56].

[56] *Data-Led Transformation – Accenture, https://www.accenture.com/gb-en/services/applied-intelligence/data-value*

6. GLOSSARY AND LIST OF ABBREVIATIONS

Term	Abbreviation (if applicable)	Definition
Algorithm		A finite sequence of rigorous instructions, typically used to solve a class of specific problems or to perform a computation. Algorithms are used as specifications for performing calculations and data processing to predict certain characteristics and pa-rameters within machine learning models.
Artificial Intelligence	AI	The theory and development of computer systems able to perform tasks normally requiring human intelligence.
Data Analysis/ Data Analytics		A process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusion, and supporting decision-making.
Internet of Things	IoT	Physical objects (or groups of such objects) with sensors, processing ability, software, and other technologies that connect and exchange data with other devices and systems over the Internet or other communications networks.
Machine Learning		A method within data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention.
Machine Vision		The technology and methods used to provide imaging-based automatic inspection and analysis for such applications as auto-matic inspection, process control, and robot guidance, usually in industry.
Offshore Low Touch Energy Robotics and Autonomous Systems	OLTER	A project which aims to provide the benchmark for develop-ment and use of reliable, on-demand, standardised autonomous systems. The ultimate goal of OLTER is to deliver a RAS industrial service which supports the offshore energy industry and supply chain to scale and commercialise robotics as a service.
Robotics and Autonomous Systems	RAS	Intelligent machines that can perform tasks and operate in an environment independently, with limited/no human control or intervention.

Table 6.1: List containing descriptions of key terms and abbreviations where applicable

Legal Disclaimer

This report has been prepared by Accenture (UK) Ltd ("Accenture"). All copyright and other rights herein vest in Accenture to the exclusion of rights belonging to third party or the Client. No part of this report may be reproduced or transmitted in any form or by any means or stored in any retrieval system without prior written permission of Accenture. Applications for permission for use of copyright material including permission to reproduce extracts in published works including works meant for internal, private and/ or limited circulation shall be made to Accenture. Where such permission is granted full acknowledgement of Accenture and source must be given.

This report has been prepared for the Net Zero Technology Centre Limited ("NZTC") only, and is subject to the terms and conditions of the OLTER Service Agreement dated 6th September 2022 between Accenture and NZTC (together "Agreement") and any use hereof, or reliance hereon shall be subject to such terms and conditions, the terms of any Reliance Letter entered into between Accenture and the relevant addressee of this report, and the various assumptions, caveats and qualifications contained herein. Each addressee of this report understands and accepts that it shall be fully responsible for the consequences of any decisions made by it, and/ or actions taken or not taken by it, based on this report.

Save as provided in the Arrangement Letter, the contents hereof should not in any circumstance, without Accenture written consent, be made available to any third party which has not entered into a Reliance Letter with Accenture in accordance with the Arrangement letter and may not be referred to or used for any purpose whatsoever (including, without limitation, in the preparation of any financial promotion or the provision of investment advice), without Accenture's prior written consent. No person other than Client or a person who has entered into a Reliance Letter with Accenture aforesaid, is entitled to use this report and Accenture disclaims any and all responsibility and liability to any such person.

This report is based on information publicly available and information obtained from various persons including those with whom personal and/ or individual interviews have been conducted. To the extent that any portion of this report refers to a view of a person or attributes information to a particular person, such views and/ or information are the personal views of that person and need not necessarily reflect the views of the organisation with whom he or she is associated or employed. Accenture assumes no responsibility or liability in respect of any such view(s) or information. Further, Accenture has assumed all such views and information to be complete and accurate and has not independently verified such information or views. Where data set out herein has been identified as being from a particular source, Accenture has assumed all such views and information to be complete and accurate and has not independently verified such data.

This report has been prepared by Accenture to provide NZTC with general and limited information regarding the environment in which it operates and may also contain some other general information. This report is not intended to be and should not be construed to be advisory or recommendatory in nature nor does it seek to or provide legal, financial or investment advice. Any person desirous of acquiring the whole issued share capital or the business and assets of, or investing in, or providing finance in relation to any such acquisition or investment or in any business or area of activity associated therewith is hereby advised to undertake his/ its own investigation and study into these matters and not to rely on this report for the same.

Accenture does not undertake any obligation to update this report to reflect events occurring subsequent to the date of the study, new information or data received after the date of the study or to correct any inaccuracies herein. Any person having access to this report agrees that Accenture shall not be responsible or liable as to, or in connection with, the accuracy or completeness of this report (or the lack thereof), any omissions in this report and any such responsibility or liability hereby is expressly disclaimed.

Net Zero
Technology
Centre

Technology Driving Transition

Contact number:
+44(0)1224 063200

Media enquiries:
pressoffice@netzerotc.com

Net Zero Technology Centre
20 Queens Road, Aberdeen AB15 4ZT

www.netzerotc.com