Data challenges and opportunities for environmental management of North Sea oil and gas decommissioning in an era of blue growth

Fiona Murray\textsuperscript{b},{*}, Katherine Needham\textsuperscript{b,1}, Kate Gormley\textsuperscript{c}, Sally Rouse\textsuperscript{d}, Joop W.P. Coolen\textsuperscript{e,f}, David Billett\textsuperscript{g}, Jennifer Dannheim\textsuperscript{h,i}, Silvana N.R. Birchenough\textsuperscript{j}, Kieran Hyder\textsuperscript{k}, Richard Heard\textsuperscript{l}, Joseph S. Ferris\textsuperscript{l}, Jan M. Holstein\textsuperscript{m}, Lea-Anne Henry\textsuperscript{n}, Oonagh McMeel\textsuperscript{m}, Jan-Bart Calewaert\textsuperscript{m}, J. Murray Roberts\textsuperscript{b,1}

\textsuperscript{a} School of GeoSciences, The Grant Institute, University of Edinburgh, King’s Buildings, James Hutton Road, Edinburgh EH9 3FE, UK
\textsuperscript{b} Institute of Biodiversity, Animal Health and Comparative Medicine, Graham Kerr Building, University of Glasgow, Glasgow G12 8QQ, UK
\textsuperscript{c} School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, UK
\textsuperscript{d} Scottish Association for Marine Science, Kirk Road, Oban, Argyll PA37 1QA, UK
\textsuperscript{e} Wageningen Marine Research, P.O. Box 57, 1780 AB Den Helder, The Netherlands
\textsuperscript{f} Wageningen University, Chair group Aquatic Ecology and Water Quality Management, Droevendaalsesteeg 3a, 6708 PD Wageningen, The Netherlands
\textsuperscript{g} Deep Seas Environmental Solutions Ltd, Ashurst, Southampton SO40 7AP, UK
\textsuperscript{h} Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, P.O. Box 120161, D-27515 Bremerhaven, Germany
\textsuperscript{i} Helmholtz Institute for Functional Marine Biodiversity, University of Oldenburg, Ammerländer Heerstraße 231, 23129 Oldenburg, Germany
\textsuperscript{j} Cefas Lowestoft Laboratory, Parkfield Road, Lowestoft, Suffolk NR33 0HT, UK
\textsuperscript{k} INSITE Programme, c/o Oil & Gas UK, 3rd Floor, The Exchange 2, 62 Market Street, Aberdeen AB11 5PJ, UK
\textsuperscript{l} ECAP Consultancy Group, Spean Bridge, Scotland PH34 4EG, UK
\textsuperscript{m} Seascape Belgium, Secretariat of the European Marine Observation and Data Network (EMODnet), Wandelstraat 7, 8400 Oostende, Belgium

A R T I C L E   I N F O

Keywords: Decommissioning, Offshore energy, Environmental assessment, Blue economy, Open access, ROV survey

A B S T R A C T

Maritime industries routinely collect critical environmental data needed for sustainable management of marine ecosystems, supporting both the blue economy and future growth. Collating this information would provide a valuable resource for all stakeholders. For the North Sea, the oil and gas industry has been a dominant presence for over 50 years that has contributed to a wealth of knowledge about the environment. As the industry begins to decommission its offshore structures, this information will be critical for avoiding duplication of effort in data collection and ensuring best environmental management of offshore activities. This paper summarises the outcomes of a Blue Growth Data Challenge Workshop held in 2017 with participants from: the oil and gas industry; the key UK regulatory and management bodies for oil and gas decommissioning; open access data facilitators; and academic and research institutes. Here, environmental data collection and archiving by oil and gas operators in the North Sea are described, alongside how this compares to other offshore industries; what the barriers and opportunities surrounding environmental data sharing are; and how wider data sharing from offshore industries could be achieved. Five primary barriers to data sharing were identified: 1) Incentives, 2) Risk Perception, 3) Working Cultures, 4) Financial Models, and 5) Data Ownership. Active and transparent communication and collaboration between stakeholders including industry, regulatory bodies, data portals and academic institutions will be key to unlocking the data that will be critical to informing responsible decommissioning decisions for offshore oil and gas structures in the North Sea.

1. Introduction

In an era of rapid global change and increasing international regulations, comprehensive environmental datasets are required to manage marine economic activities and facilitate blue growth (i.e. the long-term strategy to support sustainable growth in the marine and maritime sectors, [12], https://ec.europa.eu/maritimaffairs/policy/blue_growth_en). For the North Sea, a basin with a long history of economic exploitation now entering a period of transition, this is a particularly pertinent issue. Whilst the shipping, fishing, and oil and gas industries have dominated the basin historically, the use of, and stakeholders in, the North Sea are diversifying [39]. For example, offshore
renewables, marine biotechnology, cruise tourism, maritime surveillance, aggregates dredging and large-scale offshore aquaculture [39] have emerged in recent decades, with new activities under development. The management and decommissioning of redundant oil and gas infrastructure provides an excellent case study to explore how environmental considerations are balanced against economic realities and whether society has access to adequate scientific data to underpin management decisions [29]. To achieve the best outcomes, decisions on the timescale, methodology and monitoring of decommissioning must be supported by sound environmental data. The big challenge is to collect, collate and interpret the data required at appropriate spatial and temporal scales across the North Sea to make informed decisions.

The North Sea is a shared resource between Belgium, Denmark, France, Germany, the Netherlands, Norway and the United Kingdom. These states include part of the North Sea in their territorial waters and other EU Member States have claims to its resources (e.g. through fishing quotas). Oil and gas reserves have been exploited near the coast since the 1850s [7] and commercial offshore exploration began in earnest in the 1960s. North Sea oil and gas exploration and production is ongoing, but after over 40 years of production the more mature fields are now being decommissioned. To date, the oil and gas industry has brought economic growth to the region, for example to the UK through an estimated 375,000 jobs and over £100 billion in tax revenue between 1998 and 2014 ([115], other estimates vary). By contrast, decommissioning is expected to cost the UK state and private companies £40 billion by 2040 [35], before opening areas of the North Sea that have been exclusively exploited by the oil and gas sector for 50 years to new economic activities. Decommissioning of offshore structures will be an important environmental and economic issue with competing priorities from stakeholders. The options for decommissioning each rig in the North Sea will vary by installation as each has individual characteristics; however, each decision needs to consider the wider ecological context. By collating the environmental data the industry has amassed since the 1970s, there is the potential to reduce the costs of assessing decommissioning scenarios for individual installations, more readily inform stakeholders of the environmental costs and benefits of each approach and build public trust that wider environmental considerations are incorporated into the decision-making process.

This paper aims to highlight the emerging themes and issues associated with collection and access to environmental data for the North Sea. It focuses on the long-term management of environmental data collected by the oil and gas industry, whilst drawing on examples from other industries. The discussions are the result of a North Sea Blue Growth Data Challenge workshop held in February 2017 and organised by the University of Edinburgh through the INSITE research programme (see Acknowledgements). Discussion workshops have proven value when promoting the integration of monitoring activities across North Sea countries [46]. The overall objective of the workshop was to bring together members of the scientific community, government regulators and the oil and gas industry to discuss the challenges and opportunities surrounding data relevant to decommissioning; and the issues surrounding sourcing and maintaining environmental data for the offshore energy industries. Over 30 representatives from industry, academia, online data portals, government science and regulators attended with discussion including the current availability of environmental data for the North Sea, example case studies from academia using industry data, potential quick wins and long term aims. Each presentation block was followed by a semi-structured breakout discussion giving the three themes addressed in this paper: (1) the present situation and challenges; (2) the opportunities; (3) the best way forward to promote and enable sharing environmental data.

2. Present situation and challenges

The North Sea requires comprehensive temporal and spatial data to differentiate between human-driven influences on the ecosystem and natural variation. The position of the North Sea as the mixing point between water masses makes it a region of high climatic variation. It is strongly influenced by the North Atlantic Oscillation (NAO, [41]) which drives changes in species distributions and environmental conditions over relatively short time frames [111,17]. Long-term and broad scale trends in ecological variability, together with an understanding of causes of variation, are necessary to assess environmental quality [16]. It is also likely that the environmental baseline is shifting due to climate change [22]. Deciphering human impacts in this dynamic ecosystem is further confounded by a long history of human activity (notably fishing, see [28]), but a relatively short history of environmental monitoring. Only with the data to provide a holistic overview of North Sea ecology can environmental surveys and assessments be compared to an appropriate baseline.

Extensive efforts have been made in recent years to collect and collate marine environmental data at local, national and international levels. The most relevant of these projects for the North Sea is EMODnet: the European Marine Observation and Data Network [5]. EMODnet is a network of organisations funded by the European Maritime and Fisheries Fund (EMFF) as part of the European Commission's Marine Knowledge 2020 strategy to support the implementation of EU's Integrated Maritime Policy (IMP). The overarching aim of the network is to convert Europe's otherwise fragmented marine data landscape into an interoperable data sharing framework, adopting the “collect once, use many times” data philosophy (INSPIRE Directive 2007/2/EC). This is achieved by pulling together Europe's many data sources [45], processing and harmonising the data in accordance with international standards, and making the information freely available through online data portals. EMODnet categorises data by scientific discipline (geology, biology, chemistry, physics) and by general thematic categories such as seabed habitats, bathymetry and human activities. These thematic data portals are supplemented by an overarching central portal (www.emodnet.eu) which acts as an information hub and gateway to the network's data and information resources. It also provides a number of additional data services to browse, visualise and retrieve data layers from various disciplines and themes simultaneously. The work is achieved by more than 150 organisations including national/regional data centres, hydrographic offices, geological services and thematic data aggregators that collate, process and host data that is then made available through both their own websites and the EMODnet portals. For example, MEDIN (Marine Environmental Data Information Network, http://www.oceannet.org/) facilitates and coordinates UK marine environmental data sharing. At present, the data standards employed by MEDIN for UK datasets exceed those set by EMODnet and, for now, it is understood that MEDIN will remain part of the EMODnet network after the UK leaves the EU [34].

EMODnet has amassed datasets for a wide range of environmental parameters. The North Sea basin is now one of the best-covered marine areas in Europe for environmental monitoring. However, a recent “stress test” of available datasets to assess a range of environmental scenarios identified key data gaps and questions over data quality and comparability [23]. Geological, chemical and physical datasets were, for the most part, available and compatible across the basin, however, habitat information and biological and ecological parameters were found to be much more difficult to obtain, collate and standardise. Time-series datasets monitoring species distributions and populations over time are rare, and the occurrence of unidentifiable organisms, which are not recorded in a standardised way, in datasets can limit their use [50]. Qualitative data can be difficult to standardise and quality control (including species identification, [48]), and these datasets often tend to be more problematic to identify and access than quantitative data (e.g. weather measurements).

Gaps in available marine environmental data arise for two main reasons. One is a genuine monitoring gap, i.e. the data have never been collected. Whilst this may be the primary cause of data gaps in other seas, the North Sea has seen many studies, monitoring surveys and
environmental assessments, but often the data cannot be accessed, are difficult to collate (for example due to missing metadata), are archived in an inappropriate format, or are excessively expensive to obtain for most users [23]. Progress has been made by academia and government bodies in identifying and providing open access data, including industry data (Table 1). Considerable financial, technical and human resources have been dedicated to creating data portals and generating the capacity to process, store and manage large volumes of environmental data. This has been a costly, time consuming and technically challenging process requiring government investment in public sector organisations and their infrastructures.

In some instances, newer industrial users of the marine environment have been provided with repositories and are encouraged to submit data to these repositories. For instance, the Marine Data Exchange (www.marinedataexchange.co.uk) was established to facilitate data sharing for the UK renewable energy and aggregate industries. The Marine Data Exchange provides a central platform for these industries to submit and display areas where there are developments and their associated data sets. Use of the Marine Data Exchange is encouraged by government and regulatory bodies, but is voluntary.

Older industries, such as oil and gas, have developed without such arrangements to support environmental data collation, but voluntary data pooling has taken place in the past (Box 1). Whilst there are regulatory requirements to collect and submit environmental data to the relevant regulatory body for all offshore industries, the UK for example does not require these data to be made publicly accessible or deposited in data repositories (e.g. MEDIN, see [1]). Such obligations are unusual for North Sea industries, but there are exceptions where such agreements to submit data to open access repositories have been successfully implemented (e.g. German offshore wind industry, Box 2).

EMODnet is also working towards making it as easy as possible to share industry data whilst addressing confidentiality concerns (see Section 4.2.1) through its data ingestion portal (www.emodnet-ingestion.eu), as well as developing incentives to use these resources. Such incentives include offering data management services and comparing data provided to existing records. This can build confidence in the datasets provided because data collation quickly identifies anomalies (i.e. datasets or data points that do not conform to the overall pattern of the collated data). This may indicate either that there is something unusual at the location where the data were collected or that there is a problem with the collection methodology. In turn this can reduce data collection costs through reducing the number of replicate observations required to have confidence in the data, in comparison to datasets gathered and considered only in isolation, and allows a broader understanding of the wider impacts of industries on the basin (e.g. see Box 2).

Collecting environmental data is a hugely costly activity, especially

Table 1
Information on marine environmental data collection and availability for selected offshore industries, summarised from [1]. ✓ indicates that data are collected but not generally made available to third parties. ✗ indicates that data are both collected and generally shared upon request (but not necessarily deposited in open access data repositories). Blank cells indicates that data are not routinely collected.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Parameter</th>
<th>Physical</th>
<th>Biological</th>
<th>Primary data sharing barriers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bathymetry</td>
<td>Water/</td>
<td>Sediment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sediment Quality</td>
<td></td>
</tr>
<tr>
<td>Oil &amp; Gas</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Marine Aggregates</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Subsea Cables</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Wave &amp; Tidal</td>
<td>Renewable Energy</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Offshore Wind</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Box 1
Examples of UK Oil and Gas industry data sharing

- **UK Benthos**
The UK Benthos database (https://oilandgasuk.co.uk/product/ukbenthos/) is a rare example of an open access archive of oil and gas industry environmental surveys containing data from more than 11,000 stations across the North Sea. Originally created at Heriot-Watt University, UK Benthos contains historic quantitative data from 1975 to the present day on biological communities (macrobenthos, > 0.5 mm body size), total oil content, aromatic hydrocarbons, trace metals, and sediment properties. Survey metadata also include information on drilling history, station locations, distance from operations, aspect, species biomass, diversity and composition. Data were originally held in hundreds of separate industry reports in the archives of individual operators and environmental consultancies but long-term efforts to collate these has now resulted in a fully digitised and standardised version which is regularly updated.

- **RigNet**
Real-time weather data are provided to the UK Met Office from automatic weather stations on offshore installations through Nessco, a RigNet company [1]. This improves weather forecasting at sea which directly benefits the industry leading to more reliable weather warning systems and improved safety at sea.

- **System of Industry Metocean Data for the Offshore and Research Communities (SIMORC)**
SIMORC is a data sharing platform, hosted at the British Oceanographic Data Centre (BODC) that collates metocean in situ marine data (e.g. sea level, sea temperature, wind, waves, conductivity) collected globally by the offshore oil and gas industry. The set-up costs were met by the European Commission and the ongoing running is funded by industry. Access to the data is regulated and all users must be registered. The SIMORC database reduces the overall costs of analysing and hosting metocean data for individual operators and creates extra knowledge of local and regional marine systems [1].
In Germany applicants and operators of offshore wind farms have carried out extensive environmental impact assessments (EIA) on benthic invertebrates and demersal fish following a standard investigation concept \[8\] and produced a data coverage never before available in the German Exclusive Economic Zone (EEZ). Operators are legally obliged to submit all raw EIA data to the regulatory body (BSH, Federal Maritime and Hydrographic Office) for official approval. EIA information, and data from research projects (e.g. from the Alfred Wegener Institute) and legal mandatory monitoring programmes (e.g. in relation to the Marine Strategy Framework Directive, Directive 2008/56/EC) are stored in a data information system, standardised, collated, harmonised and quality checked by independent scientific institutions. Analysing tools are developed to extract specific information required by authorities, regulators, industry and researchers. At present, these are provided as static data products, however dynamic tools are under development to allow online analysis of the most up to date information through an open-access portal. Data provided are aggregated by means of (geo-)statistical methods on (a) spatial raster, (b) areas of interest, or (c) different time scales. Through this process, raw data are anonymised and cannot be reconstructed or linked to the original source. This concept enables non-public data (e.g. owing to commercial sensitivity or embargo periods) to become publically available (via www.geoseaportal.de) and to be searchable via EMODnet.

This system has benefitted the offshore wind industry by allowing the German regulator to authorise cluster-investigations as part of an EIA \[8\]. This allows companies applying for neighbouring clusters of offshore wind farm sites to carry out joint EIA monitoring and to share reference sites, both improving data coverage and reducing costs. The first such cluster-monitoring solution was conducted by DONG Energy in 2014 and is estimated to have reduced their monitoring costs by millions of Euros.

**Box 2**

Dual use and benefits of a large information system on offshore wind farm data in the German sector of the North Sea

In Germany applicants and operators of offshore wind farms have carried out extensive environmental impact assessments (EIA) on benthic invertebrates and demersal fish following a standard investigation concept \[8\] and produced a data coverage never before available in the German Exclusive Economic Zone (EEZ). Operators are legally obliged to submit all raw EIA data to the regulatory body (BSH, Federal Maritime and Hydrographic Office) for official approval. EIA information, and data from research projects (e.g. from the Alfred Wegener Institute) and legal mandatory monitoring programmes (e.g. in relation to the Marine Strategy Framework Directive, Directive 2008/56/EC) are stored in a data information system, standardised, collated, harmonised and quality checked by independent scientific institutions. Analysing tools are developed to extract specific information required by authorities, regulators, industry and researchers. At present, these are provided as static data products, however dynamic tools are under development to allow online analysis of the most up to date information through an open-access portal. Data provided are aggregated by means of (geo-)statistical methods on (a) spatial raster, (b) areas of interest, or (c) different time scales. Through this process, raw data are anonymised and cannot be reconstructed or linked to the original source. This concept enables non-public data (e.g. owing to commercial sensitivity or embargo periods) to become publically available (via www.geoseaportal.de) and to be searchable via EMODnet.

This system has benefitted the offshore wind industry by allowing the German regulator to authorise cluster-investigations as part of an EIA \[8\]. This allows companies applying for neighbouring clusters of offshore wind farm sites to carry out joint EIA monitoring and to share reference sites, both improving data coverage and reducing costs. The first such cluster-monitoring solution was conducted by DONG Energy in 2014 and is estimated to have reduced their monitoring costs by millions of Euros.

for the marine environment. The amount of information that can be gleaned remotely from satellites is often limited to the sea surface. Understanding of the water column and seabed still relies on collecting physical samples or deploying instruments, often from an expensive surface vessel, to take images, videos and measurements (e.g. gliders, remotely operated vehicles (ROVs), sidescan sonar, multibeam equipment, CTD casts). Despite technological developments such as automation driving costs down, marine environmental data are more expensive to collect than terrestrial or freshwater measurements. Private industry spends an estimated $3 billion every year on marine data, outsourcing both government agencies and research institutes (including collecting new data, buying access to existing data sets and data processing, [12], Table 1).

For the oil and gas industry, environmental data collection falls into two categories: dedicated environmental surveys (e.g. to fulfil regulatory requirements); and incidental observations from operational surveys. There is a precedent for sharing some information derived from mandatory surveys (Box 1). Such data sharing initiatives have been successful, either because of a clear tangible benefit to industry (improved helicopter safety through sharing metocean data, [20]; improved accuracy of nautical charts, Box 1) or because the data have already been shared outside of the company (e.g. benthic survey data in the UK are submitted to regulatory bodies, and companies must make data available to regulators upon request). In both instances, the data made available are a processed data product rather than raw datasets and this reduces the potential for misinterpretations (see Section 4.2.2). Operational surveys can result in incidental observations from which environmental data could be obtained. For instance, oil and gas operators are required by regulators to conduct surveys to assess the condition of their subsea infrastructures and in order to identify repair work needed to ensure the safety of offshore personnel, and to prevent pollution events. The videos produced from these surveys are used to assess the condition of the structures, identify where repair work is needed, and estimate marine growth coverage (e.g. algae, mussels, anemones, sponges and corals growing on the structures, see Fig. 1 for examples). These videos can also, therefore, be used as a record of how the marine life interacts with structures (e.g. [32,49]) and enhance understanding of the influence these structures have on the wider ecosystem. If this information were to be centralised, this visual survey information would create a valuable new resource to all North Sea stakeholders, helping to streamline environmental assessments, costs of monitoring, and widening the benefits of decommissioning.

For companies, sharing expensively acquired data can be difficult to justify to managers and shareholders where there is neither a regulatory requirement nor an immediate economic gain (for exceptions see Box 1). Further, there are perceived costs and risks of publishing environmental data. Firstly, where data publication is voluntary, visibility of activities to competitors could have a commercial cost because not all operators share data. This is likely to be less important for data collected with respect to decommissioning than for exploration and production. These risks can be mitigated through embargo periods or through aggregating and anonymising data. For example, fisheries data collated by the International Council for the Exploration of the Sea (ICES) are released in aggregate products and the fishing effort of individual vessels and fleets cannot be identified. It should be noted that the fishing industry does not pay for data collection, archiving or management [1]. However, this limits the potential for timely knowledge to be gleaned from the data to inform management decisions, and indeed on how much new knowledge can be generated as data aggregation can limit future analysis. Where data are aggregated or anonymised at too coarse a geographical resolution (or geo-referencing metadata is removed) it becomes very difficult to place into an appropriate context and to interpret. There are further concerns over data ownership, accreditations, confidentiality and whether there is potential liability where 3rd parties have used the data (see Section 4.2.5). These concerns are largely surmountable and mechanisms to handle these issues already exist e.g. anonymization and embargo methods developed by the EMODnet data ingestion portal.

Concerns over the misinterpretation or misuse of data released are more difficult to address. For instance, where ROV video surveys have a primary focus on examining the structural integrity of assets and determining what maintenance is required, the survey parameters were not designed to answer ecological questions. Secondary data analysis should be interpreted with the original survey aims in mind, but data owners cannot control or influence misinterpretations once the data have been made open access. There are also concerns that data in its rawest form (e.g. video footage) are susceptible to misconception. Cracks in concrete casings around a pipeline, for example, may pose no technical or environmental risk from an engineering perspective, but could be misconstrued to suggest a lack of due diligence on the part of the pipeline owner. There is a reluctance within the oil and gas industry to release truly raw data particularly in the forms of videos and images. In practice, therefore, a considerable amount of time is needed to process and check data before it could be made widely available. In principle, oil and gas operators are not opposed to environmental scientists accessing and using ROV video survey data, indeed the INSITE research program is dependent on access to these data and there are many examples of industry/academic research collaborations that have
been very successful (e.g. [9,18]). For instance, data extracted from ROV video survey footage for some platforms and pipelines have revealed evidence for a range of species colonising or aggregating around these structures [32,43,44]. These collaborations, however, may become more difficult in a rapidly changing research environment where journals require authors to publish the data in support of their manuscript. This has the potential to bring industry and academic partners into conflict (this is not always the case: see [49,9] for examples where operator data has been published in an open access forum). These issues could be addressed by journals offering embargo periods on publishing the datasets.

The final primary barrier from an industry perspective is the accessibility of “open access” data to industry. Current funding models for large-scale data collation efforts often rely on both government investment and on charging some users. In the UK, which has the highest oil production levels of any EU country, relevant public environmental data are often held in NERC (Natural Environment Research Council) Data Centres, for example the British Geological Survey and the British Oceanographic Data Centre (http://www.datacentres.nerc.ac.uk/). Under NERC policy, use of environmental data is often free of charge, although under licence, to non-profit users (e.g. academics, government agencies and NGOs), but costs can be applied where data are to be used for commercial activities [38]. In part, this is because open access data portals are expensive to maintain, but if industry is expected to submit data free of charge and cannot benefit from the resource they have contributed to without paying, it is difficult to incentivise data sharing.

Taken together, solutions to these barriers can be divided into five categories: 1) Incentives, 2) Risk Perception, 3) Working Cultures, 4) Financial Models and 5) Data Ownership (Fig. 2). These five categories are discussed in turn below (Section 4.2).

3. The Opportunities

Recent advances in big data analytics and the technological advances in online data storage have revolutionised the way in which researchers can generate and interact with data, from citizen science projects ([24,6], e.g. recreational divers temperature profiles [52], anglers recreational fishing data [51]) where volunteers contribute to generating more holistic data sets, to specialist highly technical fields like genomics (e.g. Genbank® [3]). From big data projects, it is evident that greater advances in scientific understanding are made by collating and analysing large volumes of data collected by different groups than from each individual effort in isolation. Furthermore, avoiding duplication of effort drives costs down (e.g. Box 2). Progress has been made in Europe to link up national, thematic and regional databases through EMODnet, thus enabling better interpretation of information and the ability to make large scale assessments, for example over the entire North Sea basin. The opportunities this presents vary depending on the stakeholder. Consistent metadata standards are vital to maximising the potential use. This includes measures of uncertainty in the data, which, at present, are rarely available but determine what uses data is appropriate for [23].

There is a window of opportunity now, near the start of the North Sea decommissioning process, to establish best practices and to develop tools and protocols to facilitate standardised, low-cost data collection to provide consistent, standardised datasets endorsed by operators, scientists and regulators. The outcome will be to lower the cost of data collection and analysis. Over time, this would build a comprehensive, regional-based set of environmental data, including from on and around man-made structures, that is available to those operating facilities in the North Sea, regulators and the wider scientific community. This could directly support the concept of regional environmental assessments [21] and hence reduce the need for the extensive asset-specific environmental investigations performed by operators (e.g. Box 2), and to integrate models and monitoring solutions [25]. This type of regional approach has been taken by the International Seabed Authority in creating its own data management procedures and database for subsea mining companies in order to assess wider regional issues, cumulative impacts and natural temporal and spatial variability [27].

4. The way forward

4.1. The short term (5 years)

An emerging theme from this and other recently held data workshops (EMODnet Sea-Basin Checkpoints Stakeholder Conference
Brussels 2017, MEDIN Open Meeting “The Industry Marine Data Revolution” London 2016) has been the lack of industry awareness of large-scale data collation initiatives. This includes EMODnet and the national initiatives that feed into it (e.g. MEDIN for the UK). Indeed, there seems to be confusion over which of the many open access data portals that have emerged in recent years should be used and how they are linked (e.g. that MEDIN feeds into EMODnet). In the short-term, a proactive awareness campaign about these services is important. A lack of knowledge within the oil and gas industry of these data facilities and control procedures, and quality assurance. EMODnet has protocols on metadata standards, estimated error within the datasets, formatting and control procedures, and quality assurance. EMODnet has protocols in place already to address industry concerns over data confidentiality and ownership.

A second emerging theme from regulators and industry was a desire to take stock of existing marine environmental data both in the public domain and held privately, and the advances in marine environmental data sharing that have been made. The EMODnet North Sea basin data “stress test” [23] is the most comprehensive exercise conducted to date in assessing the availability of environmental data for the North Sea. Through the INSITE programme, a dedicated Data Initiative project conducted a more targeted assessment of the availability of data relevant to assessing the role of structures in the North Sea ecosystem producing an INSITE ‘Data Roadmap’ [36]. As shown by the UK Benthos project (Box 1) facilitating the sharing of environmental data submitted to regulators as part of statutory requirements can be highly successful. Indeed, it is now stated in the UK EIA guidelines for the Oil and Gas Industry that “It is also normal procedure to submit all benthic infauna data to UK Benthos” [2]. Investigating other parameters required by regulations, and therefore routinely collected, may be a relatively straightforward starting point to opening up industry environmental data.

In addition to documenting publicly available data (either open access or by paying an access fee) the INSITE Data Initiative surveyed details of the marine growth records in ROV video survey archives for 8 major oil and gas operators (BP, Centrica, CNR International, ExxonMobil, Marathon Oil, Shell, Repsol Sinopec Resources UK and Total) and identified 280 TB of ROV footage suitable for extracting marine growth data from [37]. Operators have made some of this type of data available to researchers on a case-by-case basis (e.g. through the INSITE research programme where it has been used for modelling species connectivity between platforms, www.insitenorthsea.org), but such arrangements are ad hoc and lack standardised data sharing protocols. This can lead to protracted delays in access. In some cases, operators demand signed confidentiality agreements from researchers, which may take several months when multiple organisations are involved. This can cause problems for research projects that are funded over short timeframes and, more significantly, can make it difficult for researchers to work in a transparent and fully impartial manner. The production of standardised contracts, Memorandum of Understanding (MOU), Service Level Agreements (SLA), and Standard Operating Procedure (SOP) documentation for data sharing and research collaboration between oil and gas operators and research institutions could streamline this process.

4.2. The longer term

To achieve open access to environmental data from the oil and gas industry in the longer term the five barriers identified above (Section 2 and Fig. 2) need to be overcome. These barriers are now considered in turn, with suggestions of how solutions may be found.

4.2.1. Incentives

At present there are several incentives for operators to engage in developing data sharing methods for environmental data. Firstly, at a time when the price of oil has slumped, data sharing is an opportunity to reduce the costs of decommissioning. For instance, there is evidence that the benthos in the more environmentally stable, deeper, northern North Sea takes longer to recover from drilling activity than at more dynamic sites in the shallow southern North Sea [21]. Effective post-decommissioning monitoring regimes, therefore, could be tailored in their frequency and extent to be appropriate for specific regions, depths and biological communities [21], and where appropriate companies could share reference sites (Box 2). This insight comes from oil and gas industry data collated in the UK (UK Benthos database, Box 1) and is an example of the type of information that can only be gleaned from combined data and not from individual monitoring efforts. Sharing environmental data can help to establish best practice, refine methodologies across companies throughout North Sea decommissioning, and optimise monitoring and data collection methods.

Secondly, under the OSPAR 98/3 Decision, the central piece of
legislation covering decommissioning in the North Sea (Box 3), there is an allowance for derogations to be sought to leave all or part of very large, heavy concrete or steel structure components in place (OSPAR [40] Annex 1). Where structures are eligible for derogation to leave partly or wholly in place, environmental considerations including “impacts on the marine environment”, “biological impacts arising from physical effects”, “conflicts with the conservation of species”, and “the protection of their habitats” should be taken into account as part of a comparative assessment (OSPAR [40] Annex 2). At present operators cannot apply for a derogation to leave structures that are not listed in Annex 1, nor can they consider a rigs-to-reefs approach to decommissioning as has been used in other areas (e.g. Gulf of Mexico, [31]).

OSPAR Decisions are, however, reviewed on a regular basis and, in recent years, there have been calls to revisit the regulations surrounding decommissioning, including OSPAR [40], and consider wider environmental factors as criteria for more structures to be left, whole or in part, in the marine environment beyond those that meet Annex 1 criteria (e.g. [14]). These include limiting drill cuttings piles disturbance [21], reducing the energy and materials required for decommissioning, including the carbon footprint [13], and potential habitat provision through leaving structures in situ [30,33,4]. For platforms in the northern North Sea, for example, the protected cold-water coral Lophelia pertusa grows on several installations [19,43]. Should it become possible for operators to apply for derogations on environmental grounds in the future, they would need to have evidence that they understand the environmental costs and benefits of different decommissioning scenarios for structures not listed in Annex 1 of OSPAR [40]. Such applications to OSPAR would have to be underpinned by comprehensive ecological data of the role of the industry’s infrastructure in the marine ecosystem.

In the longer-term releasing data may reduce overall decommissioning costs. Through publishing data from early feasibility studies and monitoring of decommissioning sites, the overall cost to both the taxpayer and private companies can be brought down by avoiding duplication of effort. Scientific understanding of how these structures interact with the marine ecosystem can be improved and increasing openness and improved communication can lead to better collaboration across business sectors.

4.2.2. Risk perception

There is a perception within industry highlighted at the workshop that there are risks associated with data sharing, therefore it is critical that this perception is addressed. There are two primary concerns, both of which apply more widely than the oil and gas industry (for instance, very similar concerns are expressed by contractors engaged in developing deep-sea mining, D. Billet – International Seabed Authority - personal observation). The first concern is commercial advantage: if one company chooses to release data whilst another one chooses not to; the open access company has incurred two costs: 1) the additional data processing costs and 2) accepted the risk that other companies could capitalise on the data released. Indeed, the strongest signal to emerge from the workshop was the desire for regulators to make depositing environmental data (that is not commercially sensitive) in open access repositories a legal requirement to maximise cooperation and ensure the risks are shared across companies (the UK government is actively pursuing further engagement with industry to collate data through the Marine Science Coordination Committee https://www.gov.uk/government/groups/marine-science-co-ordination-committee). The additional costs are addressed below (Section 4.2.4). Here, where the focus is on data for informing decommissioning choices rather than other oil and gas activities in exploration or extraction, the scope for competitors gaining a commercial advantage from the data is much smaller.

The second perceived risk is a concern that open access data will be misinterpreted, with the potential for reputational damage regarding the environment. This is particularly pertinent for the oil and gas industry where accidental oil spills have often been widely broadcast and are emotive to the public (e.g. the Deepwater Horizon well blow-out in 2010, the Exxon Valdez oil tanker spill in 1989 and the supertanker Torrey Canyon spill in 1967). In the context of oil and gas operations in the North Sea the legacy of the Brent Spar should not be underestimated [42], and highlights the need for public support throughout the decommissioning process. Transparency here is key, and industry stands to gain reputational benefits from research collaborations. For example, in clinical research data sharing has been credited with improving accuracy, accelerating progress and restoring trust - in particular where industry partners are involved [26]. During the decommissioning of the Brent Spar, both Shell and the UK Government were portrayed as secretive by Greenpeace, marring public perception and ultimately leading to boycotts [47]. Public opinion is likely to be an important component in determining acceptable decommissioning scenarios. Open data is one way to create transparency and build trust in the process, and to facilitate discussions on the best practice for decommissioning the North Sea. Publishing data will also increase operators’ societal responsibility output, showing that they take these responsibilities seriously and will go beyond the minimum legal requirement in meeting them. Indeed, some operators have explicitly stated that this is their primary reason for sharing data or sponsoring research (J.W.P. Coolen, personal observation). There is a shared responsibility here for researchers to ensure that their work analysing any shared data is easily accessible and in the public domain, either through open access publications or through depositing pre-publication manuscripts in accessible online portals.

A final risk raised at the workshop was the question of liability. If conclusions are drawn from any data released in relation to questions that the original collection survey was not designed to address (e.g. ecology questions from ROV structure inspection surveys) and these are later shown to be false: who is liable? Responsibility for data interpretation can be transferred through end user agreements whilst retaining data ownership. This may deter wilful misinterpretation of data. It is important to emphasise that data sharing is not a binary issue. Compromises can be reached over publishing data in parts, anonymization and potential embargo periods as are commonplace in the public sector (e.g. individual patients cannot be identified from UK National
Health Service data, ONS https://www.ons.gov.uk/aboutus/transparencyandgovernance/onsdatapolicies/howwekeepdatasecure and in academia (e.g. embargos on publicly funded PhD theses to allow students to publish their work before the data is made public). The frameworks established through EMODnet facilitate this.

4.2.3. Working cultures
Preparing industry datasets for open access and moving towards unrestricted sharing of industry environmental data will require cooperative efforts across business, public sector bodies, and scientific research. This is now incentivised thanks to the increasing use of metrics that capture dataset citations in addition to research paper citation. However, outside academia challenges remain for employees of government regulatory bodies to defend time spent on formatting data for deposition in online repositories where compliance is voluntary. Finally, there would need to be a cultural shift within industry to allow: a) data to be shared without restrictions; b) employees to dedicate time to collating/arranging data products; and c) resources to be allocated. Changes in working practice, cultures and values are particularly difficult to achieve in large organisations (including universities, governments and multinational corporations) where there can be institutional inertia [10], however there are structures in place that can facilitate this. Building on existing frameworks in the UK for example, it would be feasible for the industry through their national trade associations (e.g. Oil and Gas UK) to contribute to funding a focal person within a National Data Centre to facilitate oil and gas data deposition into MEDIN and therefore into EMODnet. This would be a relatively small investment for the industry that all operators would benefit from. This strategy has been successful elsewhere, for instance ISA member states fund a data manager (D. Billet, personal observation). Including training on data management best practices in undergraduate and postgraduate courses for science students, some of whom may go on to work in industry, can initiate and support cultural change.

4.2.4. Financial models
Funding is a twofold problem. The primary issue is paying for the work to create datasets suitable for submission to data portals. The staff time required to process and quality check data, especially where the format and collection methods are likely to have changed over time is a major obstructing factor in releasing environmental data from both industry and other sectors. Progress is being made on automating video analyses, but at present extracting data from images and videos requires the time of skilled staff, and collating information from different surveys and companies into a usable format is labour intensive. Costs could be reduced by standardising reporting across the sector, which would allow automated approaches for extraction and collation of the data. Currently it is unclear how the costs of processing, collating and quality checking data or of developing automated processes could be met.

The second issue is the current funding model for open access data portals. To widen participation, data portals should be free for all users to access. To supplement government investment in data centres such as MEDIN, it may be appropriate for some services to have costs applied to users, but the stored data sets to be freely accessible. For example data files (e.g. spreadsheets) that are currently free for all to view [38] but use may be charged for, could be made free for all to use; but charges could be applied to all users who need support utilising the data. The Norwegian Hydrographic Service, for example, uses this approach, making all data products open access, but applying charges for information products (e.g. maps, figures, tables) that are produced to meet the customers' needs, and for any product support required to help users understand or manipulate the data (www.kartverket.no). Addressing the funding model of open access data portals to incentivise industry data sharing will be important for industries that are not required through legislation to submit their data.

4.2.5. Data ownership
Clear guidelines will need to be established, helping to appropriately credit those who generate raw data, those who process it and those who create new information from it. A standard distinction between data products and information products should be formalised to clearly identify ownership, authorship and intellectual property rights (e.g. through Creative Commons Licences, Digital Object Identifiers (DOI)). This could help promote trust over data use and accreditation. For example, where a map of species distributions is created, all the underlying data are rarely collected by the creator of the map (e.g. bathymetry layers, species observations etc.); sources are therefore often referenced. The map produced, however, is the intellectual property of the map maker, the underlying data are not. At present, appropriately accrediting data and information generated from data is not standardised and is very difficult where metadata are missing (e.g. geographic location, collection dates etc.). EMODnet is well placed to work towards developing suitable guidelines on IP rights and accreditation that could be adopted across sectors.

5. Conclusions
The development of national data centres and their coordination within EMODnet means that the structures are in place to handle and archive industry data for the benefit of all. Many of the issues relating to the archiving of data, safeguarding it and making it available under controlled conditions are in place, but awareness of these facilities within the oil and gas industry is low. The main challenges are: raising awareness of this infrastructure, addressing the perception of reputational risk and establishing how data management will be paid for. Industry trade bodies are well placed to disseminate information on EMODnet and the national data centres. This should be a priority.

Where data are deposited for open access in a transparent way, the opportunity for reputational gain may outweigh the risks from misinterpretation. Trust and transparency issues are currently a greater barrier to data sharing than technological capabilities. Maintaining communication between stakeholders through workshops and working groups can build trust and develop working relationships that will facilitate the development of data sharing protocols and overcoming the barriers identified.

In the case of decommissioning, the costs of data processing and management are likely to be dwarfed by the costs of decommissioning itself. There is potential for reducing these costs through effective communication between operators and collating datasets. The oil and gas industry could fund a focal person within a national data centre would be a cost-effective way forward.

The legacy of the oil and gas industry in the North Sea will include the lessons learned for managing future blue growth. Offshore renewables, for example, are required to plan for decommissioning prior to starting to install offshore structures. The design and operation of future structures in the North Sea will be informed by the experiences of the oil and gas industry and preserving the environmental data that has been collected by the industry will provide an important resource for both users and managers of the basin.

Acknowledgements
The Blue Growth Data Workshop was organised by the University of Edinburgh through the INSITE Data Initiative funded by the INSITE (INfluence of man-made Structures In the Ecosystem, www.insitenorthsea.org) research programme. The authors thank participants from BP, British Geological Survey, Gardline, DeepTek, theDataLab, Hartley Anderson, Marine Scotland Science, Heriot-Watt University, BMT Cordah, Shell, BEIS OPRED and Marathon Oil for their contributions. INSITE follows the 2012 Oil and Gas UK led “Decommissioning Baseline Study” joint industry project that identified data gaps in our understanding in the influence of man-made structures
on the ecology of the North Sea. J. Murray Roberts and Katherine Needham acknowledge further support from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No. 678760 (ATLAS). David Billett was supported by funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 689518 (MERCE: Marine Ecosystem Restoration in Changing European Seas). Lea-Anne Henry was supported by the INSITE Project “Appraisal of Network Connectivity between North Sea subsea oil and gas platforms”. Kieran Hyder was supported by INSITE project “Assessing the ecological connectivity between man-made structures in the North Sea”. Silvana Birchenough was supported by the INSITE project “Understanding the influence of man-made structures on the ecosystem functions of the North Sea and the European Union’s Horizon 2020 Project COLUMBUS (652690) "Knowledge Transfer for Blue Growth". This output reflects only the authors’ views and the European Union cannot be held responsible for any use that may be made of the information contained therein.

References


