*Corresponding Email: aga1@hw.ac.uk

A Data Integration Framework for Offshore Decommissioning Waste Management

[a]A.G. Akinyemi*, [a]M. Sun and [b]A.J.G. Gray
[a]School of Energy, Geoscience, Infrastructure and Society
[b]School of Mathematical & Computer Sciences
Heriot-Watt University
Edinburgh, EH14 4AS

Abstract

Offshore decommissioning represents significant business opportunities for oil and gas service companies. However, for owners of offshore assets and regulators, it is a liability because of the associated costs. One way of mitigating decommissioning costs is through the sales and reuse of decommissioned items. To achieve this effectively, reliability assessment of decommissioned items is required. Such an assessment relies on data collected on the various items over the lifecycle of an engineering asset. Considering that offshore platforms have a design life of about 25 years and data management techniques and tools are constantly evolving, data captured about items to be decommissioned will be in varying forms. In addition, considering the many stakeholders involved with a facility over its lifecycle, information representation of the items will have variations. These challenges make data integration difficult. As a result, this research developed a data integration framework that makes use of Semantic Web technologies and ISO 15926 - a standard for process plant data integration - for rapid assessment of decommissioned items. The proposed solution helps in determining the reuse potential of decommissioned items, which can save on cost and benefit the environment.

Keywords: data integration, offshore decommissioning, ISO 15926, semantic web, reuse

1 Introduction

Waste management for a sustainable process industry requires a guarantee of limited energy consumption while producing limited amount of waste, and consuming a reduced amount of natural resources [1]. Critical to achieving low waste production are data management systems that can access heterogeneous data sources, manage large quantities of data in a variety of formats, and carry out data analytics. These systems are required to support data integration to increase availability and traceability of information for decision support [2]. Over the years, such data integration technology applications in the oil and gas industry have focused on meeting the needs of asset life cycle stages other than decommissioning [3] . This is the case because most assets are just reaching the end of their design lives [4]. Currently, limited work has been done on applying the concept of data integration to legacy oil and gas data for offshore decommissioning waste management purpose [5]; and a report by industry stakeholders underscores this need [6]. Consequently, this paper proposes a method for the integration of the common data types relevant in oil and gas asset decommissioning. The framework developed is generic and can be applied whenever there is a need to integrate and query disparate data involving oil and gas assets. The key features include a data integrator that uses ISO 15926 as conceived by Fiatech [3] for reconciling heterogeneous oil and gas domain data, and an interface that makes distributed data available for querying. The prototype presented in this paper shows how the data management challenges associated with assessing the suitability of decommissioned offshore facility items for reuse can be addressed. According to DNS et al. [7], the ability to do this effectively and efficiently will advance the oil and gas industry's transition toward a circular economy.

2 Research Method

In the oil and gas industry a lot of importance is attached to proprietary information. So, data required for this research is not easily accessible. The domain concerned in this study is also large meaning that it is only practicable to address case studies. Consequently, abductive reasoning [8] is used for this research. The research approach addresses the shortcomings of deductive and inductive research approaches [9] using the pragmatist philosophical stance [10] which seeks to solve research problems in a practical way. The approach assumes that with some

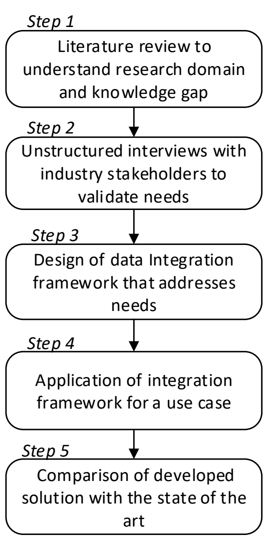
'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

research data it is possible to make logical inferences. Following this approach, this research identified a knowledge gap, obtained synthetic data, and tested a solution developed for the domain of interest using a case study.

Figure 1 details the design for this research. As shown in Figure 1(a), step 1 reviews literature on offshore decommissioning in the UK (Section 3), process plant data integration (Section 4), and related solutions for decommissioning waste management (Section 5) to understand the research domain and identify the research knowledge gap. In step 2, unstructured interviews with industry stakeholders were conducted to validate the needs of a solution (Section 6.1). These interviews were targeted at industry stakeholders attending the meetings listed in Figure 1(b). From the analysis of the findings, a data integration technology framework is conceptualised in step 3 (Sections 6.2) and developed using the Rapid Application Development (RAD) technique of software development [11]. RAD is a minimalist and agile software development technique with limited focus on planning and instead placing an emphasis on sourcing reusable components for prototypes. Consequently, the proposed solution was implemented using open source software programmes that can be easily accessed. The value of the proposed technology framework is then demonstrated for a waste assessment case study in step 4 (Sections 7.3) by using data for a hypothetical asset, based on a real offshore asset, supplied by an oil and gas engineering service provider, Ariosh [12]. Thereafter, step 5 evaluates the proposed technology framework implementation by comparing value attributes between it and the state-of-the-art tool for waste assessment in the UK offshore decommissioning industry (Section 7.4).

(a) Flow chart



(b) Overview of unstructured interviews in step 2

S/N	Conference/Workshop	Year	Location	Participants	Years of Experience
1	Decommissioning and Wreck removal Workshop	2015	Glasgow, UK	5	10 - 20
2	Offshore Decommissioning Conference	2015	St Andrews, UK	3	15 - 20
3	Decommissioning and Wreck removal Workshop	2016	Glasgow, UK	5	10 - 20
4	Offshore Decommissioning Conference	2016	Aberdeen, UK	6	8 - 25
5	Offshore Technology Conference	2016	Houston, USA	2	20 - 30
			Total	21	

Figure 1: Research Design

3 Offshore Decommissioning in the UK

Over the next 30 to 40 years, more than 470 offshore installations in the North Sea's UK Continental Shelf will need to be decommissioned [13]. According to the Royal Academy of Engineering [4] determining the cost of this is extremely difficult because of the volatility in estimated risks, material condition and oil price. In addition to the cost issues, DNS et al. [7] identified limited industry experience and inadequate human resources as challenges. While offshore decommissioning presents a significant business opportunity to UK companies, the owners of the assets and the UK Government are liable for the cost as the UK Government will be granting tax relief to the asset owners [14]. As a result, the UK tax payer is a very important stakeholder. It is therefore very important that the costs are

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

kept as low as possible. The main drivers, issues and opportunities associated with offshore decommissioning in the UK North Sea are highlighted as follows.

3.1 Drivers

- **a) End of Life:** oil and gas production in the UK Continental Shelf (UKCS) started in the 1970s [4]. Considering that a significant number of the facilities have a design life of 25 years [15], then facilities installed prior to the 1990s are likely due for decommissioning. Life extension programmes are currently employed to keep facilities in operation beyond their design lives [16]. However, with continued usage of the assets, Stacey et al. [15] states that challenges with their integrity increases and influences the decision to decommission.
- **b) Return on Investment:** with the drop in crude oil price [17], decommissioning is expected to pick up. This is because asset owners are unwilling to invest in life extension programmes for old assets - as they are unlikely to recover the cost [18]. In addition, when it is not viable to produce oil for reasons other than market price, e.g. low oil yields, production assets are prioritized for decommissioning [19].

3.2 Issues

- **a) Asset Information:** as described in Yang [20] and DNS et al. [7], from the conception of an asset to the point where the decommissioning decision is made, various stakeholders collect significant amounts of disparate data about the asset. As shown in Figure 2, design, construction, operations, maintenance, revamp and as-built data are collected over the lifecycle of an asset and need to be used for decommissioning. For example, design and construction sources include requirement documents, Front End Engineering Design (FEED) database, configuration models, and procurement contract documents. During operations, typical sources include data on compliance, safety, and costs. Maintenance sources include data on maintenance activities, schedule, and costs. Revamp projects have lives of their own, starting with new sets of design and construction data, and proceeding through another data generation cycle. Examples of as-built data include 3D point cloud, photos and videos. The data assembled from these sources, and others, are required for materials management, energy estimation, cost analysis, schedule development, etc. during decommissioning [21]. To determine the condition of a decommissioned item for example, information on design, performance in operation, and maintenance records will be required to make a judgement. However, at decommissioning, some of the required information may be missing and the available data may be in disparate formats coming from a wide range of stakeholders. This makes the assessment process very difficult. Also considering that there are large numbers of similar items on oil and gas facilities, the process is prone to error. The assessors will need to carefully string the available information together to make a determination on reuse or recycle.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

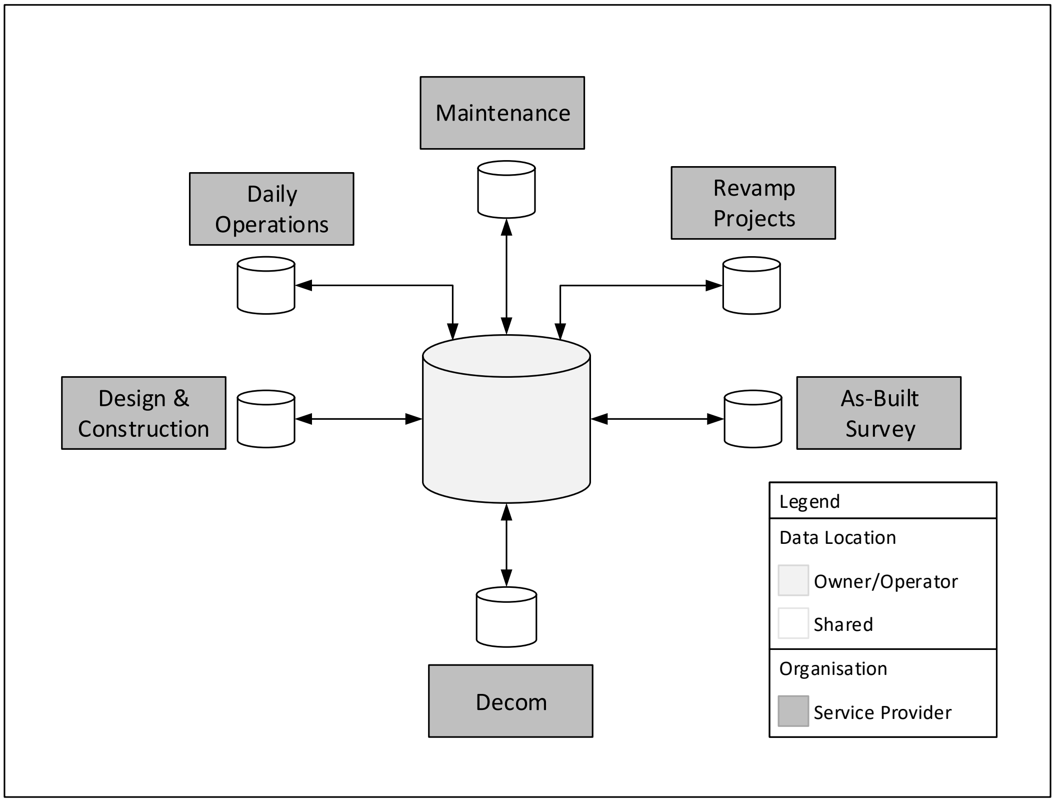


Figure 2: Typical data flow for offshore rigs

b) **Safety:** offshore decommissioning involves the removal of heavy structures from the world's most inhospitable environment - the sea. The process poses significant health and safety challenges considering that typical operations in that environment are associated with hazards of fatal consequences [22]. As a result, decommissioning programmes require strict health and safety considerations [16] and full compliance with applicable safety regulations [23]. Estimates from previous decommissioning projects, made on the statistical probability of serious and fatal accidents occurring during the decommissioning process, e.g. the 'Ekofisk I' (13 jackets and a tank) indicated that the Potential Loss of Life (PLL) during decommissioning of the facility was at least 8% [24]. This is high, usually of concern, and the basis why offshore personnel mobilization and lodging are typically recommended to be as low as practicably possible.

c) **Regulations:** the key consideration of this research with regard to regulations is stated in the guidance by DBEIS [21] and it relates to the requirement that all topsides of all offshore oil and gas installations must be returned to shore for reuse, recycling, or final disposal on land. This is to ensure they are managed appropriately in an environmentally responsible way. Complying with regulations where facilities have significantly deteriorated structurally may require single lift removal by special heavy-lift vessels. As these vessels are few, hiring costs are high and engaging them will require long lead times [25]. These waiting periods can impact the decommissioning schedule and increase the preservation cost of the facilities where production has ceased.

3.3 Opportunities

In order to sustain the oil and gas industry in the UK North Sea, the UK Government is strategically supporting the development of the decommissioning industry by advising on (i) development of late-life business models and (ii) ways to eliminate the barriers to cost effective decommissioning [26]. This strategic support, in part, is to leverage the job creation opportunities inherent in decommissioning in order to recover some of the jobs that may have been lost in the exploration and production business areas due to investment challenges. Estimates of decommissioning

'Declarations of interest: none'

jobs across the North Sea till 2040 is valued at £46bn [27] and the main business areas can be grouped into (i) services for removal of a facility from its offshore location and (ii) waste management on land [28]. With regard to the removal process, a study by DNS et al. [7] indicates that several of the existing service providers can adapt their tools and techniques to effectively support the required functions. However, RSA and ZWS [27] and DNS et al. [6] indicate that waste management of decommissioned offshore facilities under prevailing regulations requires innovative ways to exploit the inherent benefits.

According to DNS et al. [7], current waste management initiatives in decommissioning programmes employ the circular economy concept because it provides an avenue for cost reduction. The objective of a circular economy is to keep a product in use for as long as possible and extract maximum value from it in the process [29]. This allows items from a decommissioned facility to be used in other systems and create value by either saving on cost, when applied directly by an owner, or earning cash when sold. In the UK offshore decommissioning industry, the following waste management techniques are often in consideration:

a) **Material recycling**: this process involves removing the materials and components of the decommissioned assets and processing them to change their forms for use in a new entity. Crushing concrete to aggregate and smelting of steel materials are good examples. These conversion processes are carbon intensive, so not a favorite with respect to environmental consideration. Currently, the absence of local capacity for steel smelting in the UK highlighted in RSA and ZWS [27] and DNS et al. [6] means the value created from the process is lost to other countries. The opportunity that exists here is the need for local capacity for steel recycling in the UK.
b) **Component re-use:** cleaned Items from a decommissioned facility can be reused directly in similar or totally different functions without having to rework them at all. It could be that the items are brand new spares, have been maintained properly, or have just been recently replaced. Examples identified by industry professionals in a workshop described in RSA and ZWS [27] include structural steel sections, electric motors, and valves. Also, DNS et al. [7] indicates the possibility of reusing complete modules, e.g. complete helideck and test separator systems.
c) **Repair and re-use:** after assessment of equipment and components from a decommissioned facility, some of them will require some repair, sub component replacement, or sometimes recertification before they can be reused. Examples identified in RSA and ZWS [27] include winches, steel sections, and water treatment filtration units.

By observing the percentages of recycled decommissioned materials in DNS et al. [7] and DNS et al. [6], it can be concluded that asset owners seeking to meet their environmental obligations often adopt this method. However, estimates indicated in RSA and ZWS [27] show that increasing re-use over recycling can increase the value of items by between five and seven times which roughly adds up to a third of the total value of an installation at the decommissioning stage. According to the guidance by DBEIS [21], reuse is also a good strategy in gaining government approval for decommissioning programmes and can help to save on overall decommissioning cost.

4 Process Plant Data Integration

Process plants are used for processing bulk resources into products. Examples of process plants include chemical plants, oil refineries, oil and gas platforms, and gas plants [20]. For a process plant to run properly, supporting infrastructure such as electricity, water supply and sewage connections must be provided. Given the scale and complexity of process plant projects, their design and procurement can be global with several teams and systems exchanging and combining information [30]. For example, in the engineering phase, requirements analysis to fine-tune project goals needs data to be integrated from requirement documents, schedule tags, process documents, cost spreadsheets, etc. During bid management, data from the design team and vendors, e.g. items list, price catalogues, design criteria, etc., are integrated in generating requests for proposals. And during construction, the planning effort involves constructability reviews which integrate 3D model data and schedule tags. The operational life of process plants also benefits from data integration [31]. An example is the integration of operational data sets for maintenance purposes.

4.1 Data Integration Challenges

The data integration challenges associated with process plants are twofold. One part concerns the legacy data of facilities, and the other syntactic and semantic heterogeneity. The common formats of process plant legacy data include technical drawings, data sheets, and manuals - which cannot be applied directly for data integration tasks; and are error prone and time consuming to extract information from [32]. This may be because they are paper-based, unstructured, or versions of the authoring tools used in generating them no longer exist. Parts of typical facility information upgrade programmes focus on abstracting useful data from legacy sources and generating them as required using specialised tools [31].

The second part, which is the focus of this research, deals with the issues hindering integration of heterogeneous data. This is premised on the fact that over the lifecycle of an asset, multiple stakeholders generate information about it, representing information about its objects and activities differently. This is because the data models of authoring tools are usually different [33], [34], and the domain understanding of some terms, according to Fiatech [3], may be unique for different groups. As a result, attempts to combine data from multiple heterogeneous sources is not a straight forward process [35]. Meanings of terms must be resolved and efficient ways to handle large volumes of data in this form, effectively and cheaply, are required.

For decommissioning waste management, the assessor of a type or group of items needs the data associated with them for assessment. These sources are usually heterogeneous. Figure 3 illustrates an example of this situation where a steel inventory is being developed from multiple data sources. From the CAD model and material take off spreadsheets, geometric attributes such as cut length and steel section specifications are extracted. From the maintenance database, information about repairs and replacements are extracted. Where 3D laser scan model of the asset exists, scan positions are also extracted. To carry out data extraction, these sources must be searched and navigated. Thereafter the values extracted are qualified and populated in another spreadsheet or database before any queries are applied for analytics. In a scenario where there are hundreds of similar items, e.g. steel sections, the assessor will need to go through the sources for each item - a process that is inefficient and error-prone; and likely to frustrate any motivation for decommissioning waste reuse.

*Corresponding Email: aga1@hw.ac.uk

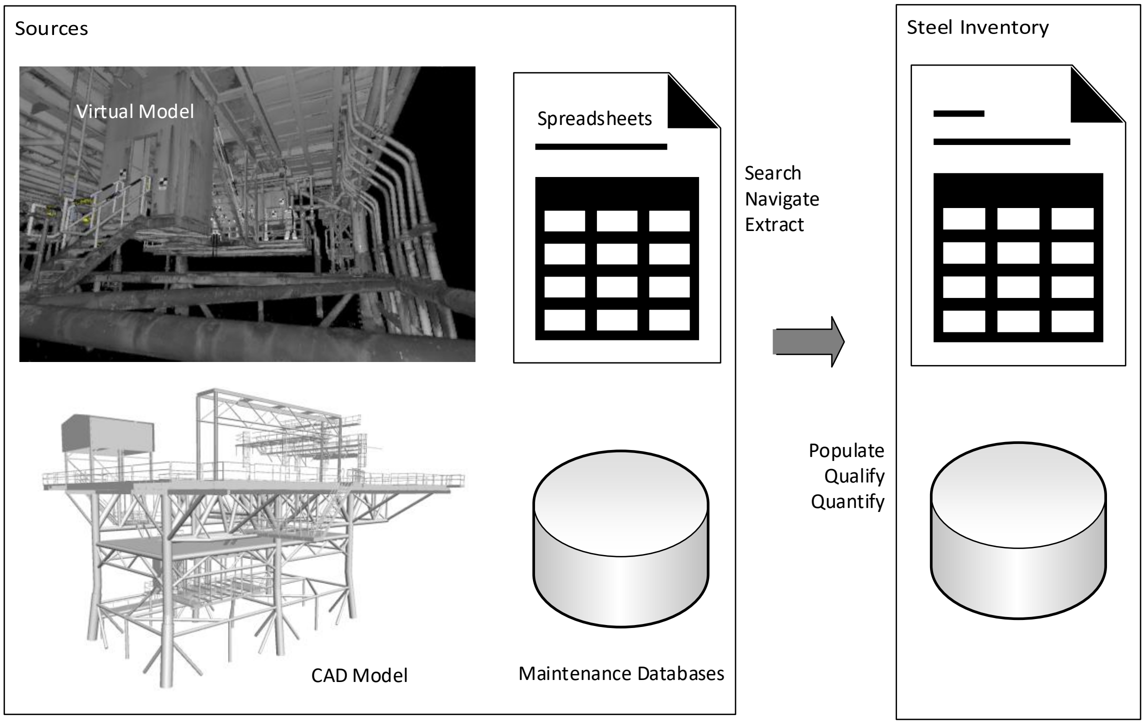


Figure 3: An example of traditional approach for managing disparate data on projects

4.2 Data Integration Standard - ISO 15926

ISO 15926 is detailed in Fiatech [3] as the standard for integration of life-cycle data for process plants including oil and gas production facilities. It addresses the data integration challenges that are the focus of this research. However, the implementation of the standard is difficult because of the effort and skills required in using it [30], [36]. Currently it is made up of twelve parts with varying levels of development. The standard is very important for oil and gas data integration because it provides an ontology - a shared conceptualization [37] - for the process industry domain using its Parts 2, 3 and 4. Part 2 documents the generic data model for representing technical information about process plant projects; Part 3 documents reference data for 2D and 3D geometric properties; and Part 4 documents the initial set of reference data for use with the Part 2 data model. ISO 15926 Parts 7 and 8 provide specifications for data exchange and life-cycle data integration. Part 7 specifies this using ISO 15926 templates based on the Part 2 data model; and Part 8 does the same using Semantic Web technologies. The Semantic Web is an extension of the World Wide Web that enhances information retrieval [38]. Fiatech et al. [30] describes ISO 15926 templates as standard addressable set of specific relationships between known things that together represent some information. Other parts of the standard include: Part 1 which introduces the standard; Part 9 (database storage), Part 10 (conformance), Part 12 (lifecycle integration ontology) and Part 13 (integrated lifecycle asset planning) which are under development; and Part 11 which demonstrates the use of the published parts [39].

4.3 Semantic Web Application

Semantic Web techniques have been used successfully for integrating disparate data in the design and operations lifecycle stages of oil and gas plants [40]–[42]. The Semantic Web uses the Resource Description Framework (RDF) [43] as a standard graph-based data model with globally defined identifiers and terminologies. This approach supports bringing together disparate data sources and quickly highlights inconsistent terms. To aggregate data, Doan et al. [44] describes that the contents of the sources are associated with semantic markups to form identifiers which facilitate easier integration and more accurate search results. RDF, RDF Schema [45] W3C Recommendation and the Web Ontology Language (OWL) [46] W3C Recommendation are used for the semantic markup. The Semantic Web uses SPARQL [47] for querying RDF data. SPARQL enables expressing queries across multiple sources

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

(Prud'hommeaux et al. 2013 W3C Recommendation) and the querying is done by matching the graph patterns between the queries and RDF data. Curé and Blin [38] describes this graph pattern as a set of triple patterns, and a triple pattern as a three-part statement containing a subject, predicate and object. Dealing with decommissioning waste requires an ability to integrate life cycle information for aging assets. This type of data integration effort requires ingestion of data in a variety of structures, resolution of overlapping terminologies and the ability to query data across sources. These functionalities are capabilities covered by Semantic Web technologies [38] and specified in the Part 8 of ISO 15926.

5 Related waste management resources

A review of the literature on built asset waste management data tools (see Table 1) provides evidence that the ontology-based method can address the challenges associated with built asset data integration [49]. However, no evidence exists on its application for the process plant decommissioning process. Proximally situated to the latter solution is the evidence of the application of technologies including the World Wide Web, relational database (RDB), Big Data, and ifcOWL ontology [49]–[52] for waste management in the building industry (see Table 1). Nevertheless, none of these resources demonstrate use cases and implementation plans that can be easily adapted. The only tool that fits the bill and recommended for decommissioning waste management in the UK offshore decommissioning industry has a manual approach [5]. Given that ISO 15926 is the recommended standard for process plant data integration, and Semantic Web technologies have been recommended for its implementation, this research focuses on the application of these technologies for data integration in offshore decommissioning waste management process in light of the knowledge gap.

'Declarations of interest: none'

Table 1: Survey of waste management data tools for built assets

			Integration Technology			Industry			Lifecycle Stage						use case
Source	Summary	Data Integration Resource	Ontology	Semantic Web	Others	Oil & Gas	Other Process Industries	Building	Engineering	Procurement	Construction	Operations	Maintenance	Decommissioning	Waste Management
(McGrath, 2001)	SMARTWaste is a proprietary tool that enables pre-demolition audit. The tool is a software application targeted toward the construction industry but which can be applied in other domains. To use the tool, waste data needs to be pre-prepared in a recommended template and then imported into the software. Otherwise, attributes of each waste element will need to be entered manually using a form within the software.	Proprietary tool			✔			✔			✔			✔	✔
(Banias et al., 2011)	This study presented a construction and demolition waste management decision support system that makes use of a relational database management system for storing waste data; the World Wide Web (WWW) technologies for delivering web pages to user requests; an algorithmical model for the estimation of generated wastes; and mathematical programming for specifying optimal waste management strategies.	WWW RDB			✔			✔			✔			✔	✔
(Li and Zhang, 2013)	This study proposed a construction waste management system that makes use of a web browser for user interfacing; an application logic for integrating the accounts, data input and analytical modules; a web server for delivering internet web pages; and a relational database management system for storing waste data.	WWW RDB			✔			✔			✔				✔
(Bilal et al., 2016)	This study proposed a Big Data architecture for construction waste analytics. Data from heterogenous sources are migrated to a central repository using an extract, transform and load (ETL) tool. Applied Big Data technologies include: (i) Spark and Hadoop for efficient data processing and (ii) Neo4J for data storage. Also ifcOWL, a standard ontology, is used for knowledge represenetation. Using tools based on this architecture allows integrated data to be applied for resource optimisation in the engineering and procurement life cycle stages of an asset.	Neo4J Spark Hadoop ifcOWL	✔	✔	✔			✔	✔	✔					✔
(D3 Consulting Limited, 2017)	This resource is a freely available tool contributed by an offshore decommissioning waste management service provider for decommissioning waste audit. The tool is a basic spreadsheet that needs to be populated with waste data before use. The tool is a spreadsheet and has no special functions programmed into it.	Manual			✔	✔								✔	✔
		Research Focus	●	●		●								●	●

6 Proposed Data Integration Framework

6.1 Requirements

Table 2 summarises the requirements identified from interviews with offshore decommissioning professionals. The result shows that in a generally sense interviewees believe that a data integration solution for offshore decommissioning waste management should make use of ISO 15926, the process industry standard for integrating life cycle data described in Section 4.2. This seems logical because, at decommissioning, data collected over the lifecycle of a facility will be available for waste management. Considering that the data involved would have been independently generated by multiple stakeholders, interviewees also underscored the need for ways to resolve the meanings of terms and any data representation issues. Another requirement relates to methods for handling data conversion, given that data in the sources are usually in formats other than RDF, which is the recommendation for ISO 15926. A scenario that was highlighted is one in which a data integration system could connect to live relational databases, where access is granted, to use the most up-to-date data e.g. maintenance records. Lastly, Fiatech et al.

*Corresponding Email: aga1@hw.ac.uk

[30] identified that the difficulty with using the ISO 15926, for most engineers, is identifying how to couple the available tools with the requisite technologies for its proper use. This is backed up by the responses of interviewees indicating the lack of simple examples on the use of ISO 15926 discourages potential users.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

Table 2: Summary of unstructured interview responses on requirements for process plant data integration

S/N	Questions	Comment Summaries	Solution Requirements
1	Use of ISO 15925	✔ ISO 15926 covers the lifecycle of a process plant, so will be appropriate for integrating data at decommissioning ✔ Stakeholders and application developers seek to comply with ISO 15926 because it is a global standard ✔ Engineers are not very familiar with the technologies required for the proper use of ISO 15926 ✔ Lack of examples on the use of ISO 15926 for data integration is a challenge for potential users	✔ Appplication of ISO 15926 ✔ Tools for using ISO 15926
2	Systems architecture	✔ Solution should be able to connect to live autonomous sources, particularly relational database ✔ Solution should be able to handle multiple sources simultaneously	✔ Querying technique
3	Inventory development	✔ Need for ways to reduce the effort required for extracting data from different sources ✔ Need for ways to eliminate costly human errors inherent in manual approaches ✔ Need for ways to resolve meaning of terms ✔ Need for ways to resolve data level representation	✔ Smooth data migration ✔ Methods for handling of common data types ✔ Syntatic and semantic homogeneity
4	Data extraction	✔ Data output should be in the order desired ✔ Need for capability to query across multiple sources ✔ Data output should indicate where missing values exist	✔ Querying technique
5	Desirable attributes	✔ Solution should require minimal effort and return very accurate result ✔ Solution should result in reduced task duration ✔ The IT artifacts generated by the solution should be reusable	

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

6.2 Design

Figure 4 depicts the proposed data integration framework for oil and gas asset data. The framework uses ISO 15926 as the common data model for the sources. To integrate data – resolving the semantic and syntactic heterogeneity issues, information patterns made from generic templates based on Parts 2, 4, 7 and 8 of ISO 15926 standard are used to represent the ontology of the domain of interest. Thereafter, data from the sources are mapped to the ontology to generate RDF data. The RDF data is then deposited in a triplestore – a database management system for RDF data - where users and applications can interface with it via SPARQL queries, the standard query language for RDF data. SPARQL has the capabilities specified for querying in Table 2. The generated queries, which answer questions relevant to the purpose for which the framework is deployed, can be documented as result templates that are reusable. The generic steps of the work process for the proposed framework are described below:

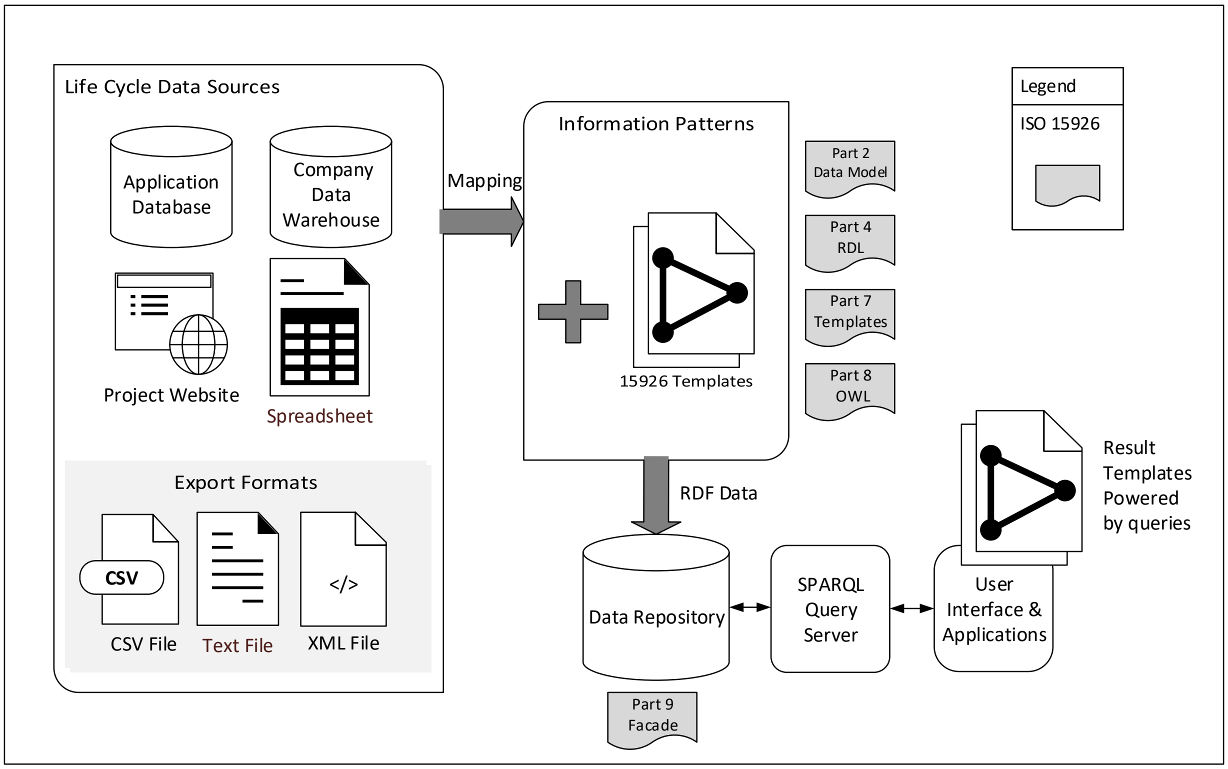


Figure 4: Data integration using ISO 15926 ontology

a) **Pattern Definition:** This involves two steps. The first is identifying the right information templates for representing the entities in the sources. The templates required can be selected from those defined in Parts 7 and 8 of ISO 15926 or new ones formulated based on Parts 2 and 4 of the standard. The second step is the generation of information patterns that represent the ontology of the domain of interest. The information patterns can be created using open source tools like dot15926Editor [53] and Protégé [54].
b) **Data mapping:** This involves organising the sources and generating the mapping files for adding the data in the sources to the information patterns. This research is premised on the assumption that data can be converted to some type of structured formats. Mapping of data from the sources can be done using open source tools like Cellfie [55] and Ontop [56]. Mapping of data can also be done using an open-standard file format like JavaScript Object Notation (JSON) as in TechInvestLab.ru [53] or by using mapping language like 'Relational Database to RDF Mapping Language' (R2RML) [57].
c) **Data Storage:** This is about the architectural decision on where to locate the prepared data for the querying process. This could be a file located on the local machine, a triplestore, or a relational database. Although ISO 15926 Part 9 is yet to be released, the idea around the use of triplestores for storing RDF data is already

well established [58]. The use of a relational database requires an Ontology-Based Data Access (OBDA) model [59]. This helps to relate the data in the relational database to the appropriate ontology entities.

d) Information Extraction: SPARQL is used for information extraction because the output data is in the RDF format. When prepared data is in a single location, normal SPARQL queries can be used for information extraction. However, when querying across multiple sources, federated SPARQL queries [48] are used. Where relational databases are maintained as sources, the OBDA approach can be used to on-the-fly translate the relevant parts of the SPARQL query into a local query. Calvanese [56] describes Ontop as one such tool.

7 Case Study Evaluation of Proposed Solution

RSA and ZWS [27] reported that a typical North Sea platform is about 97% steel, and an estimated 500,000 tonnes of the over 5 million tonnes of steel - including platform topsides, jackets and subsea components - presently installed will be decommissioned before 2023. These assertions indicate that steel is going to be a large portion of the waste from offshore decommissioning. According to DNS et al. [6], decommissioning projects often report up to 98% recycle and reuse (Section 3.3 description) of waste materials. However, only about 2-3% is actually reused. Consequently, the benefits of the current practice are significantly reduced because of the energy consumed in the recycling process. A survey in 2008 discovered that 0.2% of 9 million tonnes of steel scrapped was reused [60]. This indicates that steel reuse is very limited. However, an earlier survey in 1998 found that the amount of steel reused in the construction sector was 31% of 129,000 tonnes sent for recycling [61]. This indicates the potential reuse opportunity that exists with structural steel in the construction industry.

For decommissioned steel sections to have application in construction, they need to have traceability and documentation. DNS et al. [6] reported that specialist asset re-sale companies specified information from certifications, warranties, material inventory, fabrication history record, condition report, 2D drawings and 3D models for assessment purpose. These may also need to be corroborated with virtual 3D images from laser scanning according to DNS et al. [7]. Information obtained from the sources is used to determine the quality, condition, age and coating of fabricated steel. However, considering the heterogeneity of the sources - as discussed Section 4.1 - building an asset inventory to efficiently deliver the required assessment is always a challenge. Consequently, this research proposes a data integration framework that can mitigate the challenge compared to the state-of-the-art tool for the same purpose.

7.1 Data Sources

A demonstration based on the case study described in the latter paragraph makes use of synthetic data generated based on a real offshore platform documented by Ariosh [12]. The data sources are three of the commonly used sources for steel assessment and are shown in Figure 5. The sources include:

- a structural report from a 3D application (S1)
- a maintenance database on the steel sections in S1 (S2) and
- a summary of 3D point cloud data scan positions (S3)

S1 contains the description, specification, geometric properties, and location of reported steel members. S2 contains the details of maintenance service, the contact of service provider and specific remark for instances of maintenance activity carried out. S3 lists the links of the scan positions that are closest to the zones of a platform to which a steel member belongs. From these scan positions, an assessor can see adjacent scan positions and very quickly peruse the zone of interest. The case study has limited instances in the sources because it is only for demonstration purpose. For context, the items listed in the report have an approximate weight of 113 tonnes compared to single North Sea platform that DNS et al. [7] describe as weighing up to 48,000 tonnes and over. From the sources, this evaluation seeks to deduce all the relevant data required for the assessment of a steel asset.

S1: Structural report from facility model

ID	Description	Spec	Cut Length	Area	Zone
B.01	Circular Hollow Section	CHS 168.3x5.6	14 m		Boat Deck
B.02	Circular Hollow Section	CHS 168.3x5.6	14 m		Boat Deck
B.03	Circular Hollow Section	CHS 219.1x5.0	18 m		Boat Deck
B.04	Circular Hollow Section	CHS 219.1x5.0	17 m		Boat Deck
B.05	Circular Hollow Section	CHS 323.9x5.6	18 m		Boat Deck
B.06	Circular Hollow Section	CHS 323.9x5.6	17 m		Boat Deck
B.07	Circular Hollow Section	CHS 355.69x8.0	20 m		Boat Deck
S.01	Universal Beam	UB 762x267x226	10 m		Boat Deck
S.02	Parrallel Flange Channel	PFC 260x75x28	15 m		Boat Deck
S.03	Grating	38x5 SG		98 sq m	Boat Deck
C.01	Circular Hollow Section	CHS 168.3x5.6	14 m		Cellar Deck
C.02	Circular Hollow Section	CHS 219.1x5.0	18 m		Cellar Deck
C.03	Circular Hollow Section	CHS 323.9x5.6	18 m		Cellar Deck
C.04	Circular Hollow Section	CHS 355.69x8.0	20 m		Cellar Deck
C.05	Universal Beam	UB 762x267x226	15 m		Cellar Deck
C.06	Parrallel Flange Channel	PFC 260x75x28	14 m		Cellar Deck
C.08	Grating	38x5 SG		150 sq m	Cellar Deck
M.01	Universal Beam	UB 762x267x226	15 m		Main Deck
M.02	Parrallel Flange Channel	PFC 260x75x28	15 m		Main Deck
M.03	Grating	38x5 SG		70 sq m	Main Deck

S2: Structural maintenance database

maintenance

id	structure	structure Id	serviceDate	serviceId	remarks
M001	Boat Deck	X.BD	9/5/2010	S003	Repairs done
M002	Grating	S.03	7/5/2011	S001	Replaced
M003	Cellar Deck	X.CD	7/6/2013	S001	Repairs done
M004	Main Deck	X.MD	7/27/2013	S001	Repairs done
M005	Boat Deck	B.07	11/15/2015	S002	Fatigued
M006	Grating	S.03	9/18/2017	S002	Corroded and needs replacement

contact

id	company	phone	website
C0001	UTEC	+441224 812020	http://utecsurvey.com
C0002	Siem	+44 1224 945500	https://siemoffshorecontractors.com/

services

id	contactId	label
S001	C0001	Topside inspection and maintenance
S002	C0001	Structural inspections and cleaning
S003	C0002	Splash zone inspection and maintenance

S3: 3D scan summary of facility laser model

Platform Level	Virtual 3D
Boat Deck	C:\Navisworks Data\Delta PP\Revised Truview with Model Superimposed\DATA\BOAT & SUBCELLAR DECK.htm
Cellar	C:\Navisworks Data\Delta PP\Revised Truview with Model Superimposed\DATA\CELLAR DECK.htm
Main Deck	C:\Navisworks Data\Delta PP\Revised Truview with Model Superimposed\DATA\MAIN & UPPER DECK.htm

Figure 5: Sample data on design specification, maintenance and 3D virtual image of structural steel in a facility

7.2 State-of-the-art Tool

The Late Life Planning Portal [62] is a repository for knowledge sharing on decommissioning projects. The platform documents best practices and tools for decommissioning practitioners. The recommended tool for decommissioning

*Corresponding Email: aga1@hw.ac.uk

waste management is a tool by D3 Consulting Limited [5], a specialist decommissioning waste management organisation. The tool, shown in Figure 6 is a spreadsheet template with attributes associated with items registered as column headers. To use the tool, all the attributes related to an item have to be observed from the relevant sources and logged in the correct column cells corresponding to the row of the item. After the tool has been populated with the required data, spreadsheet functions are then used to manipulate the data to sort, count etc. There is no real integration of the data and useful results depend on careful manipulation of the spreadsheet functions.

Asset	Topside/ Jacket/ Subsea	Functional Category	Description	EWC	EWC Description	EWC Classification	Net (Te)	Units	Source

Figure 6: Recommended spreadsheet template for offshore decommissioning waste management

7.3 Proposed Solution

7.3.1 System Architecture

To realise the proposed data integration framework for decommissioning waste management, open source applications are coupled for its implementation. Figure 7 shows the plan for the data sources. S1 and S3 are provided as exported data files while S2 is assumed to be residing in an active relational database. For the sources that are data files, data in the sources are mapped to information patterns that are based on ISO 15926 templates and standard entities. The resulting RDF data is then deposited in a triplestore. For the source that is a relational database, the source is sustained as-is. The data from the relational database is made available as RDF data in a virtual RDF triplestore using an OBDA model which links the relational data to appropriate ontology entities. SPARQL queries are then used to extract information of interest from the repositories. Queries over the ontological view of the virtual RDF triplestore are translated into Structured Query Language (SQL) queries that extract the required data from the relational database. The results are then mapped with RDF and combined with the results obtained over the other sources. Several permutations of these RDF data generation methods are possible, and any combination of data files, relational databases and triplestores, as sources, will work well.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

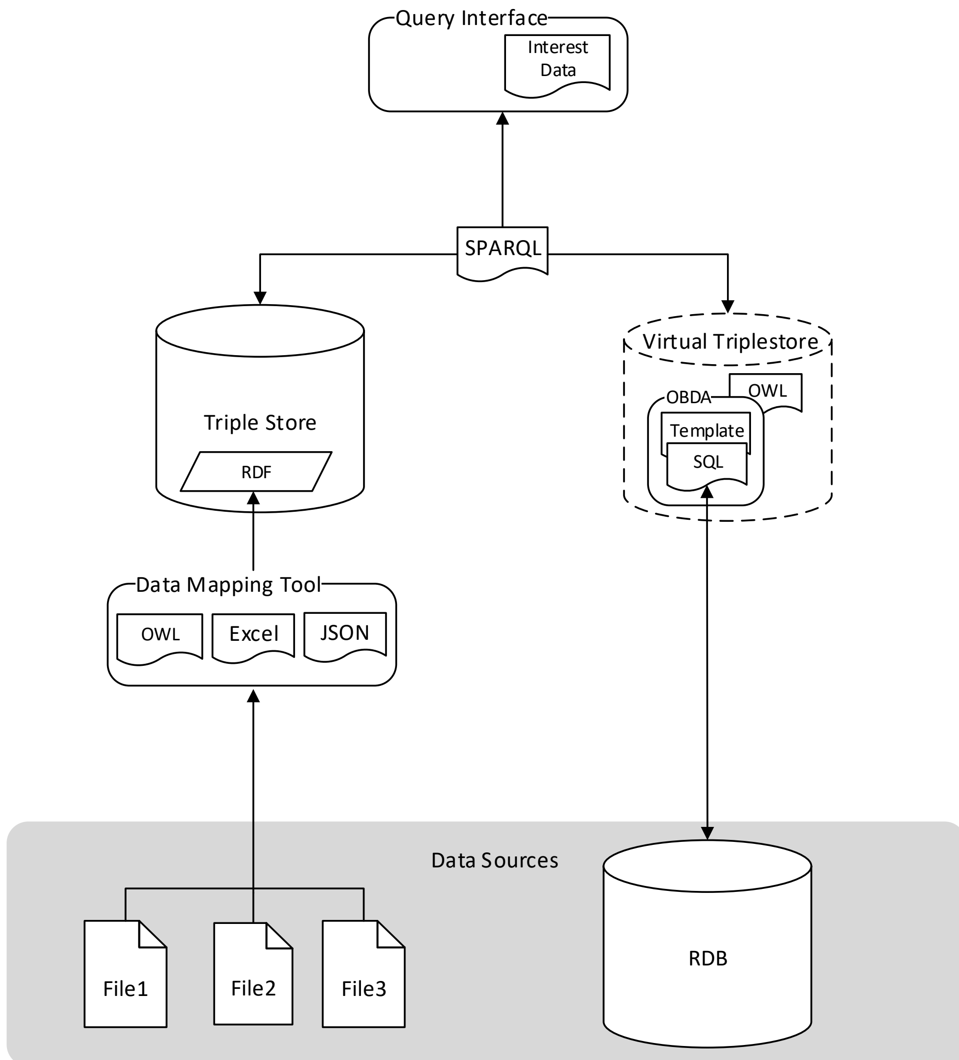


Figure 7: Implementation architecture for case study

The open source tools used for the implementation of the proposed framework are Ontop and Eclipse RDF4J described in Knap et al. [58]. Eclipse RDF4J is a Java framework for storing, reasoning, querying, etc., RDF data. It was used as the RDF data repository and SPARQL query server. Calvanese et al. [56] describes Ontop as an OBDA system that allows querying relational data sources provided in terms of an ontology to which the data sources are mapped. The version of RDF4J server that has Ontop prepacked into is used for the implementation of the proposed framework. Guidance for replicating the implementation of the framework is provided in [63].

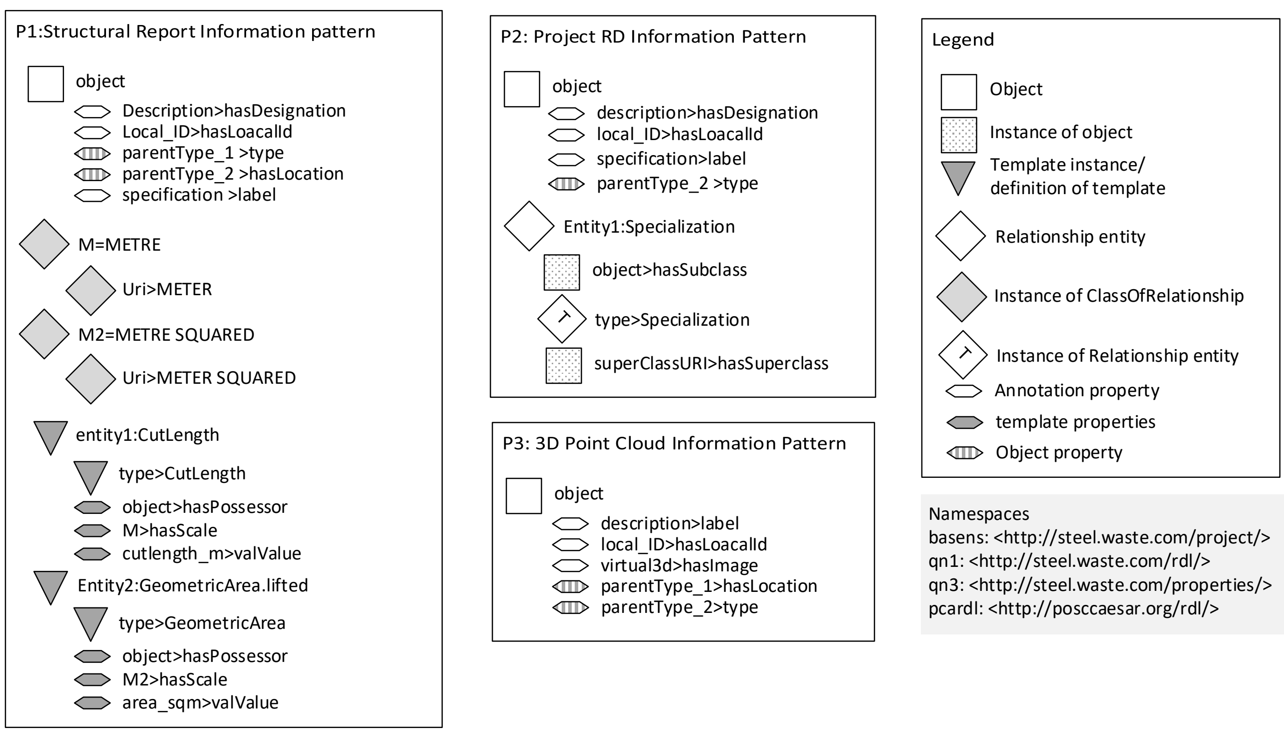
7.3.2 Data Mapping Processing

The open source tools used for data mapping include dot15926 Editor, Protégé, and Ontop. TechInvestLab.ru [53] describes dot15926 as an architecture and set of specific libraries that enable working with ISO 15926. It facilitates reading, visualizing, exploring, searching, reasoning, mapping, etc., of ISO 15926 data. It was used for RDF data generation for the data file sources in the case study. Guidance for replicating the mapping task in this paper is provided in [64]. Musen [54] describe Protégé as an ontology editor that supports the latest OWL standard. Ontop is available as a plugin in Protégé. The two tools help in generating mapping axioms used for extracting data from a relational database and generating RDF data in a virtual triple store. Guidance on using this tools for the task in this case study is provided in [63].

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

To map data from the sources, mappings files and information patterns must be developed. With the use of dot15926 Editor, the *data file* type sources - preprocessed into spreadsheets - are converted to RDF data. The first step in this process is to identify the entities in the sources. By inspection of the sources S1 and S3 in Figure 5, eleven entities including 6 types of steel sections, a grating type, 3 deck types and point cloud representation must be identified in ISO 15926 Reference Data Library (RDL) [3]. This RDL is maintained by POS Caesar Association (PCA), an organization Fiatech [3] describes as responsible for ensuring the integrity of registered Part 4 entities. Identifying them in the RDL will ensure that anyone that interacts with the output data can reference the same meaning and capture the exact description intended. The eleven entities identified are the reference data for the sources. The spreadsheet with header H1 shown in Figure 8 is prepared for the project reference data. The grey cells in the header are added as columns to the base information derived from the sources. A namespace <http://steel.waste.com/rdl/> is formulated for the resources and concatenated with the newly formed local IDs to generated Uniform Resource Identifiers (URIs). This is compliant with best practices for identifiers [65]. Where the URIs are too long, Doan et al. [44] suggests the qualified names of the namespaces may be used. For example, the qualified name for ISO 15926 RDL entities is ‘pcardl’ and the namespace is <http://posccaesar.org/rdl/>. Also for this step, the Part 2 type of the entities in the sources, their Part 4 RDL superclasses and the superclasses’ URIs are logged.



H1: Header for project Reference Data (RD) spreadsheet

Namespace	URI	Desription	Specification	Local ID	Part 2 Type	Part 4 RDL Superclass URI	Part 4 RDL Superclass

H2: Header for structural report spreadsheet

Namespaces	ID	URI	Description	Parent URI_1	ParentID@RDL	Spec	Cut Length	cutlength_m	Area	area_sqm	Zone	ParentID@RDL	Parent URI_2

H3: Header for 3D point cloud spreadsheet

Namespaces	ID	URI	Platform Level	ParentID@RDL	Parent URI_1	Virtual 3D	ParentID@RDL	Parent URI_2

Figure 8: ISO 15926 information patterns for structural report and 3D scans

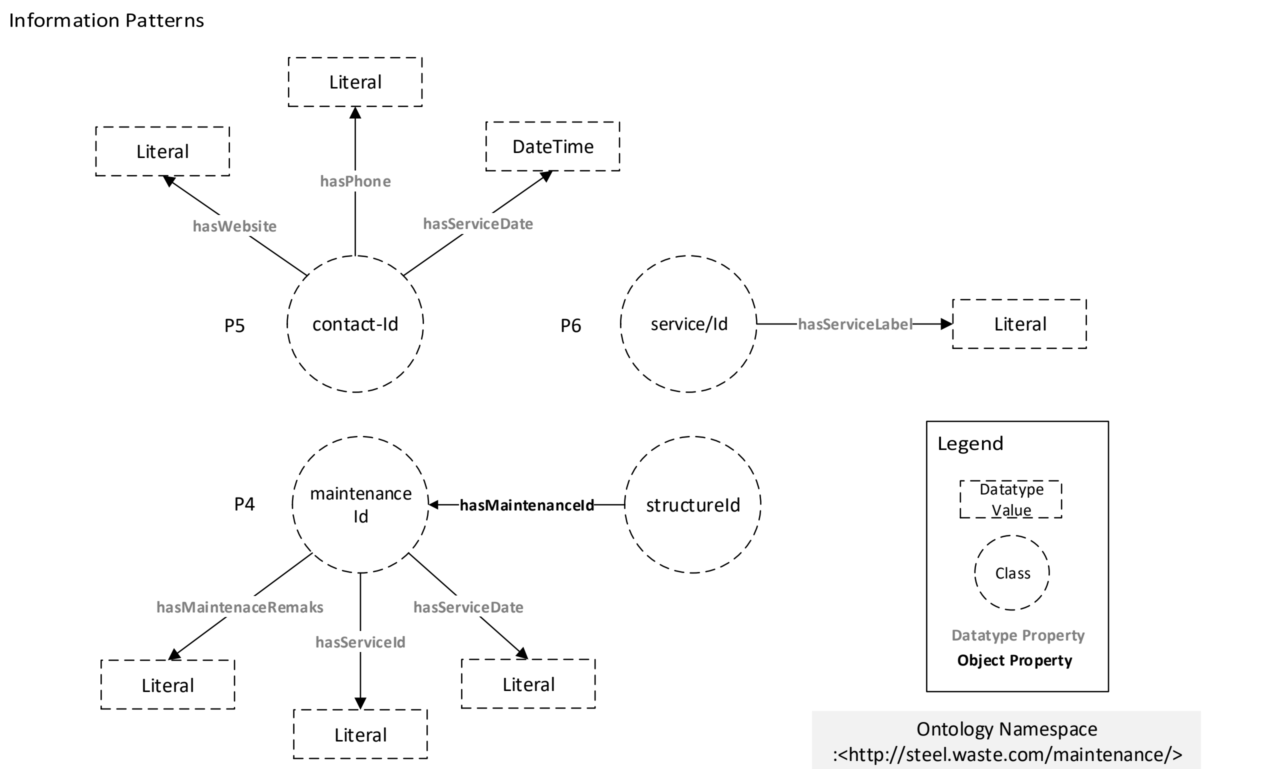
For data in S1, the spreadsheet with header H2 as shown in Figure 8 is prepared. In the spreadsheet, the assigned project namespace <http://steel.waste.com/project/> is concatenated with the ID of the resources in the source to generate URIs for them. For each resource, the corresponding project reference data entities (ParentID@RDL) and their URIs are regenerated in new columns adjacent to them. Lastly, the numerical values of cut length and area for the resources in S1 are extracted into separate columns for detail mapping. For data in Source S3, the spreadsheet with header H3 is prepared. Again, project namespace <http://steel.waste.com/project/> is concatenated with the ID generated for the resources in the source to create URIs for them. Also, the project reference data entities (ParentID@RDL) and their corresponding URIs are regenerated in new columns. Once the data sources are prepared as described, the information patterns are developed based on them. For S1, pattern P1 in Figure 8 is created. It has a signature where its object (URI) has annotation properties 'hasDesignation', 'hasLocalId' and 'label'. Also, the same object URI has object property 'type' and 'hasLocation'. Two Part 7 templates are also in the information pattern. One of them is a template for cut length and the other is for geometric area. These are to capture the records relating to length and area of entities. Each template has properties for the possessor of the record, the record's scale of measurement, and the value of the measurement. Also, the ISO 15926 Part 4 representations of the measurement scales are instantiated in the pattern.

For source S3, information pattern P3 in Figure 8 is created. It has a signature where its object (URI) has annotation properties 'label', 'hasLocalId' and 'hasImage'. The same object has object property, 'type' and 'hasLocation'. For project reference data entities, an information pattern P2 is created. It has a signature where its object has annotation properties 'hasDesignation', 'label' and 'hasLocalId'. In addition, a Part 2 specialization entity showing the relationship between the resources and their Part 4 superclasses is added to the pattern. For all the annotation and object properties in S1 and S2, <http://steel.waste.com/properties/> is used for the namespace. Once the information patterns are ready, mapping files are generated to map the data in the sources to them. JSON is used to represent the mappings in this case study and RDF data was generated from the process.

Regarding the mapping of data from the relational database S2, an OBDA model and ontology of the entities in the source are required. The OBDA model contains mapping axioms comprising SQL queries and target triple templates. Calvanese *et al.* [56] states that the mapping axioms can either be generated using R2RML and imported into Ontop, or automatically generated by Ontop itself using its native mapping language. The triple templates of the mapping axioms in this case study are contained in information patterns P4, P5 and P6 generated for the maintenance database and shown in Figure 9. P4 has a signature with an object property 'hasMaintenanceId' linking objects that belong to classes 'structureId' and 'maintenaceId'. Objects that belong to the latter class in addition have datatype property 'hasMaintenanceRemarks', 'hasServiceId' and 'hasServiceDate'. Patterns P5 has a signature where objects in class 'contact-Id' have datatype property 'hasWebsite', 'hasPhone' and 'hasServiceDate'. The last information pattern P6 has a signature where objects that belong to the class 'service/Id' have datatype property 'hasServiceLabel'.

The complementing part of the mapping axioms used for extracting data from source S2 are the corresponding SQL queries for mappings M1, M2, M3 and M4 shown in Figure 9. The SQL query for M1 selects values of 'id' and 'structureId' from the maintenance table in S2. The SQL query for M2 selects values of 'id', 'serviceId', 'serviceDate' and 'remarks' from the same table. The SQL query for M3 selects 'id', 'company', 'phone' and 'website' from the contact table in S2; and the SQL query for M4 selects 'id' and 'label' from the services table in S2. All the resources from source S2 use the ontology namespace <http://steel.waste.com/maintenance/> except for objects of class 'structureId' in M1 that have a namespace <http://steel.waste.com/properties/> that is the same with the resources from sources S1 and S3. It should be noted that resources that share the same URI in a project are one and the same. This is a consequence of representing the integration data model in RDF, which has a globally defined namespace. The generated files from this process are deposited in a custom Ontop triple store in RDF4J framework. This triplestore behaves like a virtual one because the actual RDF data does not reside in it. Instead, it is generated on-the-fly when a database query is executed.

*Corresponding Email: aga1@hw.ac.uk



Mapping	Target TripleTemplate	Source SQL Query
M1	<http://steel.waste.com/project/{structureId}> **:hasMaintenanceId** :{Id} .	SELECT "id", "structureId" FROM "maintenance"
M2	:{Id} :hasServiceDate {serviceDate} ; :hasServiceId {serviceId} ; :hasMaintenanceRemarks {remarks} .	SELECT "id", "serviceId", "serviceDate", "remarks" FROM "maintenance"
M3	:contact-{id} :hasCompanyName {company} ; :hasPhone {phone} ; :hasWebsite {website} .	SELECT "id", "company", "phone", "website" FROM "contact"
M4	:service/{Id} :hasServiceLabel {label} .	SELECT "id", "label" FROM "services"

Figure 9: Information patterns and mapping axioms for maintenance database

7.3.3 Data Extraction

According to DNS et al. [6], an opportunity for the construction sector to interrogate offshore steel inventories to ascertain suitable sections for their projects is a proposition with potential benefits. For example, a sustainable design practitioner interested in using suitable old steel gratings for drainage covers may interrogate a repository with this type of data to find out the type and size of gratings available; any maintenance activities carried out on them; and the dates of the maintenance activities. Also, 3D scan positions close to the shortlisted grating elements can be linked for visual inspection. Figure 10 shows similar sample queries on the case study data. Query Q1 aims to return all the relevant information about steel gratings on the boat deck of the hypothetical facility. The objectives include finding out any maintenance activities carried out on the steel items and the dates the activities were performed. Also, the scan positions that are closest to the grating elements are sought for visual assessment. Q1 answers this question by declaring with the first two conditional triple patterns that grating on the boat deck (BD) is what is required from source S1. It should be noted that grating objects are represented by their reference IDs. Using the common location, BD, the RDF triples from source S3 related to the required images are matched. The corresponding RDF triples from S3 are then matched with the appropriate triples with the correct images. The closing part of the query extracts from source S2 the maintenance remarks and service date information for each steel item selected. It does this by matching the corresponding RDF triple patterns in S2 through the reference IDs of grating objects and maintenance occurrences. The result R1 of this query is presented in Figure 10.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

Query Q2 in Figure 10 aims to list the types of physical objects in source S1 and describe them. This is where the generated reference data becomes useful. Q2 answers the question by using an RDF triple pattern from the reference data that declares that the items sought are of type ‘ClassOfInanimatePhysicalObject’. This pattern is then matched with a triple from S1 that connects the description of the items to the class. The result of this query is presented as R2 in Figure 10. The last query, Q3, demonstrates that it is possible to return answers when parts of the data are missing. Q3 requests cut length and area values for steel items on the main deck of the hypothetical facility. The query extracts the result R3 in Figure 10 by matching an RDF triple pattern from source S1 that declares that the requested items are on the main deck (MD). This triple pattern is then matched to the corresponding triple patterns on cut length and area in the project data. For this query, the area and cut length values are optionally matched (SPARQL query function) through blank nodes in the triple patterns. Cyganiak et al. [43] describes blank nodes as anonymous resources used for creating links between declared data objects. Blank nodes in query Q3 are represented with qualified name qn4 and alphanumeric resources.

Q1

```
#Query to extract maintenance and image data on boat deck gratings
PREFIX : <http://steel.waste.com/maintenance/>
PREFIX basens: <http://steel.waste.com/project/>
PREFIX qn1: <http://steel.waste.com/rdl/>
PREFIX qn3: <http://steel.waste.com/properties/>
PREFIX pcardl: <http://posccaesar.org/rdl/>

SELECT ?steel ?image ?remarks ?date
WHERE {
?steel qn3:hasLocation qn1:BD .
?steel pcardl:hasDesignation "Grating" .
basens:X-BD qn3:hasLocation qn1:BD .
basens:X-BD qn3:hasImage ?image .
SERVICE <http://localhost:8080/openrdf-sesame/repositories/Steel_Waste> {?steel :hasMaintenanceId ?id. ?id :hasMaintenanceRemarks ?remarks; :hasServiceDate ?date .}
}
```

R1

Steel (from S1)	Image (from S3)	Remarks (from S2)	Date (from S2)
S.03	"C:\Navisworks Data\Delta PP\Revised Truview with Model Superimposed\DATA\BOAT & SUBCELLAR DECK.htm"	"Replaced"	2011-07-05T00:00:00.0
S.03	"C:\Navisworks Data\Delta PP\Revised Truview with Model Superimposed\DATA\BOAT & SUBCELLAR DECK.htm"	"Corroded and needs replacement"	2017-09-18T00:00:00.0

Q2

```
# Query extracts the id of physical objects in data and their descriptions
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX pcardl: <http://posccaesar.org/rdl/>
PREFIX dm: <http://rds.posccaesar.org/2008/02/OWL/ISO-15926-2_2003#>

SELECT ?item ?description
WHERE {?item rdf:type dm:ClassOfInanimatePhysicalObject.
?item pcardl:hasDesignation ?description
}
```

R2

Item	Description
CHS168.3-5	"Circular Hollow Section"
CHS219.1-5.0	"Circular Hollow Section"
CHS323.9-5.6	"Circular Hollow Section"
CHS355.69-8.0	"Circular Hollow Section"
UB762-267-197	"Universal Beam"
PFC260-75-28	"Parrallel Flange Channel"
38-5SG	"Grating"

Q3

```
# Query extracts the geometric propertes of items on the main deck of the hypothetical facility
PREFIX qn4: <http://tpl.rdlfacade.org/data#>
PREFIX qn1: <http://steel.waste.com/rdl/>
PREFIX qn3: <http://steel.waste.com/properties/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dm: <http://rds.posccaesar.org/2008/02/OWL/ISO-15926-2_2003#>

SELECT ?itemNo ?cutLength ?area
WHERE {?itemNo qn3:hasLocation qn1:MD.
OPTIONAL {?z qn4:RBC4BE4B0BBAF47E4A0BEFFAE45B5C24B ?itemNo;
qn4:R0854474F25FA4F9699F0DA191B971B58 ?cutLength.}
OPTIONAL {?y qn4:R371BC00E162544848115CB372BB3DEF3 ?itemNo;
qn4:RD097D4A601E34341851FD3F991E93569 ?area.}
}
```

R3

ItemNo	CutLength	Area
M.03		"70 "
M.02	"15"	
M.01	"15"	
X-MD		

Figure 10: Sample queries and results for proposed solution

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

7.4 Comparison

The proposed solution is compared to the state-of-the-art tool for decommissioning waste assessment (Section 7.2) based on value attributes that are essential for developing the required inventory and extracting the correct information efficiently. Table 3 summarises the comparison of the two solutions. The first issue is the development of an inventory that will contain all the required data for the process. This is going to be an arduous process if each line item in the state-of-the-art spreadsheet tool has to be manually populated. The proposed solution simplifies this process by preparing the sources and ingesting the data into a repository using data mapping techniques. This is particularly useful where large volumes of data are involved, e.g. the large quantity of structural steel elements on an oil and gas platform [66]. It will eliminate any error that may be introduced in a manual process. Also, the mappings can be reused for additional data on the same project or for other projects that are similar.

Another advantage of the proposed solution is that query results can be automatically ordered to suit a viewing preference. This is powered by the SPARQL used for querying. On the other hand, the state-of-the-art solution has to be manually ordered. Some columns have to be hidden and may be forgotten for other cases if the user does not pay careful attention. Another issue is that the filtering function in spreadsheets can eliminate rows of data if empty cells exist in the filtering column. Again, if the user does not pay careful attention, that information may be overlooked. In the proposed solution, SPARQL has the capability to query across empty value parts of data and return correct values. An example is demonstrated with query Q3 in Figure 10.

Other benefits of the proposed solution relate to reasoning. The use of the state-of-the-art tool involves manually resolving meaning between terms to determine if they refer to same thing e.g. 'CHS 168.3X5.6' and 'CHS 219X5.0' in Figure 5 are circular hollow sections. Also, the relationships between items have to be manually established and taken into consideration during the assessment process. The proposed solution on the other hand uses an ontology view of all the sources referenced to ISO 15926 for handling meaning and relationships between terms. For example, query Q2 in Figure 10 selects all types of physical objects in the case study. To carry out the same function with the state-of-the-art tool, the user has to filter through all the columns in the spreadsheet, collate the relevant objects and delete duplicates. The proposed solution also allows the specification of the detail of interest as in Q3 where only the information about the main deck of the hypothetical facility is desired.

Table 3: Comparison of State-of-the-Art Tool and Proposed Solution

S/N	Issues	State of the Art Tool	Proposed Solution
1	Inventory Development	Manual inventory development process that is repetitive and therefore error prone.	Mapping files are used to migrate data from the sources. Method guarantees data in the sources will be ingested as-is. Also mappings created are reusable on similar tasks irrespective of the amount of data in the sources.
2	Concordanced Information	Method does not support automatic concordance table generation.	Supports automatic concordance table generation. The results are presented in the order desired.
3	Missing values	The filtering function in spreadsheets can eliminate rows of data relevant to a result where empty cells exist in the filtering column.	Makes use of SPARQL, a declarative query language that allows users to give broad instructions on data to retrieve. The engine underneath handles the retrieval process and returns the most complete result including where missing values exist.
4	Semantic Heterogeneity	Data elements with same meaning but represented differently have to be manually resolved in the spreadsheet.	Data elements that mean the same thing but represented differently are automatically resolved using ontology developed for the sources.
5	Level of Detail	This solution cannot reason about data with hierarchical relationships. The relationships have to be manually resolved.	The ontology developed in the solution handles the relationships between data entities and SPARQL can be used to extract information at desired level of detail.

'Declarations of interest: none'

7.5 Limitations

One challenge of the proposed solution is that there is limited opportunity to correct any wrong data already existing in the source. This will be mapped as-is and may introduce errors in the integrated data. Given this, integrity of the sources is critical. Another challenge is that some programming effort is required by the users to generate data mapping files and query the integrated data. This can be mitigated by leveraging open source tools that aid mapping and query generation e.g. dot15926 Editor. There is also a challenge with identifying the right entities in ISO 15926 RDL due to the existence of many similar items. The dot15926 Editor helps with limiting the search results but the user still needs to thoroughly check before selecting entities.

Another challenge is the scale of the data. Enterprise grade servers are better suited for this type of technology deployment [49]. This may not be suitable for small businesses as they may have budget constraints [67]. However, the cost should be offset against the current human approach that requires vast amount of costly error-prone human effort, and opportunities with pay-per-use cloud computing [68] considered in mitigating the cost. Lastly, this type of project usually requires expert resources including data engineers, data architects, and data scientists, working together when they are large scale and complex. For small data projects, it is possible for domain engineers with appreciable knowledge of Semantic Web technologies, database technologies, and scripting languages to work alone. However, this is a rarity.

8 Implications for UK Offshore Decommissioning Industry

This study has the potential to impact the following elements related to offshore decommissioning in the UK.

8.1 Supply Chain

Currently, screening out decommissioned items that are useful is a difficult task because information from multiple sources is required for such assessment. As a result, it is difficult to engage the supply chain and potential market outlets about the supply. Application of the solution proposed in this study is valuable because it can be applied for integrating heterogenous data for web platforms that enable searching. This will enable the searching of certified decommissioned items and reduce their time to market.

8.2 Jobs

Over 80% of installations in the UK Continental Shelf are fixed steel structures [28]. Consequently, when these facilities are decommissioned there is a huge opportunity to extend their value based on circular economy principles. However, because of the preference of asset owner for recycling and the limited steel recycling capacity in the UK, most of the materials are sent overseas for processing. The owners' preference for recycling is, in part, due to the challenges associated with assessing life cycle information of steel materials for reuse. Application of the proposed solution in this study will enable assessments for reuse. Such assessments have the potential of changing the course of decommissioned steel by finding applications that will use them in the UK. Examples of reuse applications for steel include for piling and as railway sleepers [6]

8.3 Decommissioning Cost

The solution developed in this research can help to improve time and cost efficiencies of the offshore decommissioning process. It will improve time efficiency by reducing the effort spent in making sense of heterogenous data. It will also reduce costly human errors and save on paid person-hours. These, in turn, will result in compressed work schedules, allowing for quicker progress.

8.4 Small Businesses

This study demonstrated the use of open source tools for data integration. It also provided a data integration architecture that can be adopted. These will be useful resources for small businesses with limited budget for proprietary applications but in need of data integration tools.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

8.5 Environment

As stated earlier (Section 3.2c), decommissioning in the UK must comply with relevant environmental regulations. As a result, decommissioned materials must be managed appropriately in an environmentally responsible way. The three methods stipulated for managing decommissioning waste are reuse, recycling, and disposal on land. Disposal on land (degradability issues) and recycling (carbon footprint) have higher environmental implications than reuse, making it the best option for the environment. However, determining whether decommissioned items can be reused requires reliability assessments which depend on the associated life cycle data of the items. Considering the assessments requirement, the proposed solution in this study is relevant because it can be applied for integrating life cycle data of process plant items. In addition, its querying capability can help provide useful insights into the reliability of decommissioned items.

9 Conclusion

Business processes and methods achieve higher efficiency and improve bottom line when redesigned with the use of information technology. The solution developed in this research can help improve on work efficiency and make savings on the overall cost of offshore decommissioning. It can also change the course of offshore decommissioning wastes by using them for wealth creation in the UK instead of shipping them abroad for recycling. The technology solution demonstrated makes use of ISO 15926 to enable integration of heterogeneous data from multiple sources. The solution ensured semantic homogeneity while allowing the continued use of existing applications. This will ensure a low risk implementation and the least disruption to existing practice. The framework developed in this research is extensible and can be used to support other offshore decommissioning activities like cost calculation, schedule management etc. For next steps, the developed framework should be applied by engineers working on similar tasks in the process industry and evaluations on productivity, scalability, and cost benefits carried out. Future research efforts should also seek to further reduce the setup cost of the proposed framework to lower the barrier for its application. This could mean developing solutions that integrate the mapping process, data storage, and querying interface in one place. Such solutions should also automate all the steps in the process for ease of use by non-programmers.

For decommissioning waste management, the next step should seek to build a graphical user interface that will exploit the querying capabilities of the developed framework to support reliability assessment of specific decommissioned items. Such solutions will be valuable for decommissioning supply chain; enabling the searching of certified decommissioned items that can be reused, thus reducing the time of such items to market. Lastly, a Linked Data application - an application that allows the publishing structured data so that it can be interlinked and become more useful through semantic queries, could be developed for offshore decommissioning waste management. Such an application should mirror an online retail platform but be targeted at providing useful information about items for reuse or resale. The developed platform should be able to synthesise available life cycle data about a decommissioned item, carry out prescriptive analysis, and recommend future uses to which it can be applied. In addition, the platform should allow users to create accounts and add items with their associated life cycle data for appraisal, display and sales.

Acknowledgement

The case study example in this paper is adapted from an offshore facility data produced by Ariosh, an engineering service company that supports oil and gas asset lifecycle information management in West Africa. The opinions, findings and recommendations in this work are those of the authors and do not necessarily reflect the views of the company.

References

[1] W. Terkaj and T. Tolio, “The Italian Flagship Project: Factories of the Future,” in *Factories of the Future*, T. Tolio, G. Copani, and W. Terkaj, Eds. Springer, Cham, 2019.

[2] T. Tolio, G. Copani, and W. Terkaj, “Key Research Priorities for Factories of the Future—Part I: Missions,” in *Factories of the Future*, T. Tolio, G. Copani, and W. Terkaj, Eds. Springer, Cham, 2019.

'Declarations of interest: none'

*Corresponding Email: aga1@hw.ac.uk

[3] Fiatech, “An Introduction to ISO 15926,” 2011.

[4] Royal Academy of Engineering, “Decommissioning in the North Sea: A report of a workshop held to discuss the decommissioning of oil and gas platforms in the North Sea,” 2013. [Online]. Available: https://www.raeng.org.uk/publications/reports/decommissioning-in-the-north-sea. [Accessed: 16-Aug-2017].

[5] D3 Consulting Limited, “Waste Management in Offshore Decommissioning,” *Late Life Planning Portal*, 2017. [Online]. Available: https://l2p2.net/SitePages/Tool-Download.aspx?DocId=445&Version=4.0. [Accessed: 18-Oct-2017].

[6] DNS, AMEC, and ZWS, “Decommissioned Steel Reuse in Construction,” Aberdeen, 2016.

[7] DNS, ZWS, and ABB, “Offshore Oil and Gas Decommissioning,” *Offshore Oil and Gas Decommissioning*, 2015. [Online]. Available: https://library.e.abb.com/public/d689c2f70f0c447586610ac566c9aa7e/ABB-Offshore-Oil-and-Gas-Decommissioning-2015.pdf. [Accessed: 16-Aug-2017].

[8] A. Bryman and E. Bell, *Business Research Methods*, 4th ed. New York, NY: Oxford University Press, 2015.

[9] F. J. Riemer, S. D. Lapan, and M. T. Quartaroli, *Qualitative research: an introduction to methods and designs*. San Francisco: Jossey-Bass, 2012.

[10] A. J. Onwuegbuzie, R. B. Johnson, and K. M. T. Collins, “Call for mixed analysis: A philosophical framework for combining qualitative and quantitative approaches,” *International Journal of Multiple Research Approaches*, vol. 3, no. 2, pp. 114–139, 2009.

[11] M. L. Despa, “Comparative Study on Software Development Methodologies,” *Database Systems Journal*, vol. 5, no. 3, pp. 37–56, 2014.

[12] Ariosh, “Ariosh,” 2018. [Online]. Available: http://www.ariosh.com. [Accessed: 03-May-2018].

[13] Oil and Gas UK, “Economic Report 2014,” 2014. [Online]. Available: http://oilandgasuk.co.uk/wp-content/uploads/2015/05/EC041.pdf. [Accessed: 24-Oct-2019].

[14] HM Revenue & Customs, “Annual Report and Accounts 2011-12,” 2012. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/89198/annual-report-accounts-1112.pdf. [Accessed: 16-Aug-2016].

[15] A. Stacey, J. V Sharp, and M. Birkinshaw, “Life Extension Issues for Ageing Offshore Installations,” *Proceedings of 27th International Conference on Offshore Mechanics and Arctic Engineering*, 2008. [Online]. Available: http://www.hse.gov.uk/offshore/ageing/life-extension-issues.pdf.

[16] Health and Safety Executive, “Offshore Oil & Gas Sector Strategy 2014 to 2017,” 2014. [Online]. Available: http://www.hse.gov.uk/offshore/offshore-strategic-context.pdf. [Accessed: 16-Aug-2017].

[17] Nasdaq, “Crude Oil,” 2017. [Online]. Available: http://www.nasdaq.com/markets/crude-oil.aspx?timeframe=5y. [Accessed: 16-Aug-2017].

[18] Deloitte, “Insights: The impact of plummeting crude oil prices on company finances,” 2015. [Online]. Available: http://www2.deloitte.com/ng/en/pages/energy-and-resources/articles/crude-awakening-the-impact-of-plummeting-crude-oil-prices-on-company-finances.html#. [Accessed: 16-Aug-2017].

[19] Oil and Gas UK, “Activity Survey 2015,” 2015. [Online]. Available: http://oilandgasuk.co.uk/wp-content/uploads/2015/07/EC044.pdf. [Accessed: 16-Aug-2017].

[20] R. Yang, *Process Plant Lifecycle Information Management*. Bloomington, IN: iUniverse, 2009.

[21] DBEIS, “Guidance Notes: Decommissioning of Offshore Oil and Gas Installations and Pipelines under the Petroleum Act 1998,” 2018. [Online]. Available:

'Declarations of interest: none'

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/760560/Decom_Guidance_Notes_November_2018.pdf. [Accessed: 26-Oct-2019].

[22] M. Christou and M. Konstantinidou, “Safety of offshore oil and gas operations: Lessons from past accident analysis,” 2012. [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC77767/offshore-accident-analysis-draft-final-report-dec-2012-rev6-online.pdf. [Accessed: 16-Aug-2017].

[23] Bureau Veritas, “Decommissioning on the UK Continental Shelf - an overview of regulations,” 2011. [Online]. Available: https://www.bureauveritas.com/white-papers/offshore-decommissioning-guide. [Accessed: 16-Aug-2017].

[24] P. Ekins, R. Vanner, and J. Firebrace, “Decommissionig of offshore oil and gas facilities: Decomissioning Scenarios: A Comparative Assessment Using Flow Analysis,” 2005. [Online]. Available: http://www.psi.org.uk/docs/2005/UKOOA/Decommissioning-Working paper.pdf. [Accessed: 16-Aug-2017].

[25] Y. Dalgic, I. Lazakis, and O. Turan, “Vessel charter rate estimation for offshore wind O&M activities,” in *Developments in Maritime Transportation and Exploitation of Sea Resources*, C. G. Soares and F. L. Peña, Eds. CRC Press, 2013, pp. 899–907.

[26] Oil and Gas UK, “Decommissioning Insight 2015,” 2015. [Online]. Available: https://decomnorthsea.com/uploads/pdfs/projects/Decommissioning-Insight-2015.pdf. [Accessed: 16-Aug-2017].

[27] RSA and ZWS, “The RSA Great Recovery & Zero Waste Scotland Programme: North Sea Oil and Gas Rig Decommissioning & Re-use Opportunity Report,” 2015. [Online]. Available: https://www.thersa.org/globalassets/pdfs/reports/rsa-great-recovery---north-sea-oil-and-gas-report.pdf. [Accessed: 06-Feb-2018].

[28] ARUP, Scottish Enterprise, and DNS, “Decommissioning in the North Sea,” 2014. [Online]. Available: https://www.raeng.org.uk/publications/reports/decommissioning-in-the-north-sea. [Accessed: 16-Aug-2017].

[29] Ellen MacArthur Foundation, “Towards a Circular Economy: Business Rationale for an Accelerated Transition,” 2015. [Online]. Available: https://www.ellenmacarthurfoundation.org/assets/downloads/TCE_Ellen-MacArthur-Foundation_9-Dec-2015.pdf. [Accessed: 06-Feb-2018].

[30] Fiatech, PCA, and USPI, “Building the Semantic Web for the Process Industries using RDF, OWL and SPARQL for the Integration, Sharing, Exchange, and Hand-over of distributed Plant Lifecycle Information on the basis of ISO 15926,” 2015. [Online]. Available: https://web.archive.org/web/20180811061734/http://infowebml.ws/. [Accessed: 06-Feb-2018].

[31] F. Olubunmi and R. Ward, “The End of Handover,” 2017. [Online]. Available: https://www.youtube.com/watch?v=wfVP7snMis8. [Accessed: 06-Feb-2018].

[32] X. Wang, H. Li, J. Wong, and H. Li, “An integration framework of advanced technologies for productivity improvement for LNG mega-project,” *Journal of Information Technology in Construction*, vol. 19, no. 22, pp. 360–382, 2014.

[33] B. Bayer and W. Marquardt, “Towards integrated information models for data,” *Computers and Chemical Engineering*, vol. 28, pp. 1249–1266, 2004.

[34] A. Tolk and S. Y. Diallo, “Model-based data engineering for web services,” *IEEE Internet Computing*, vol. 9, no. 4, pp. 65–70, Jul. 2005.

[35] S. M. Embury, S. M. Brandt, J. S. Robinson, I. Sutherland, F. A. Bisby, and W. A. Gray, “Adapting integrity

enforcement techniques for data reconciliation," *Information Systems*, vol. 26, no. 8, pp. 657–689, 2001.

[36] A. Wiesner, J. Morbach, and W. Marquardt, "Information integration in chemical process engineering based on semantic technologies," *Computers and Chemical Engineering*, vol. 35, no. 4, pp. 692–708, 2011.

[37] T. R. Gruber, "A Translation Approach to Portable Ontologies," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, 1993.

[38] O. Curé and G. Blin, *RDF database systems : triples storage and SPARQL query processing*. Elsevier, 2015.

[39] P. Denno and M. Palmer, "Modeling and Conformance Testing for the Engineering Information Integration Standard ISO 15926," pp. 1–22, 2013.

[40] P. Novák, E. Serral, R. Mordinyi, and R. Šindelář, "Integrating heterogeneous engineering knowledge and tools for efficient industrial simulation model support," *Advanced Engineering Informatics*, vol. 29, no. 3, pp. 575–590, 2015.

[41] R. Šindelář and P. Novák, "Ontology-Based Simulation Design and Integration," in *Semantic Web Technologies for Intelligent Engineering Applications*, S. Biffl and M. Sabou, Eds. Springer, Cham, 2016, pp. 257–277.

[42] E. Kharlamov *et al.*, "Ontology Based Access to Exploration Data at Statoil," *Proc. of the 14th International Semantic Web Conference (ISWC 2015)*, vol. 9367, pp. 93–112, 2015.

[43] R. Cyganiak, D. Wood, and M. Lanthaler, "RDF 1.1 Concepts and Abstract Syntax," *W3C Recommendation*, 2014. [Online]. Available: https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/. [Accessed: 06-Feb-2018].

[44] A. Doan, A. Halevy, and Z. Ives, *Principles of Data Integration*. Elsevier, 2012.

[45] D. Brickley and R. V. Guha, "RDF Schema 1.1," *W3C Recommendation*, 2014. [Online]. Available: https://www.w3.org/TR/rdf-schema/. [Accessed: 06-Feb-2018].

[46] B. Motik, B. C. Grau, I. Horrocks, Z. Wu, A. Fokoue, and C. Lutz, "OWL 2 Web Ontology Language," *W3C Recommendation*, 2012. [Online]. Available: https://www.w3.org/TR/owl2-profiles/. [Accessed: 06-Feb-2018].

[47] The W3C SPARQL Working Group, "SPARQL 1.1 Overview," *W3C Recommendation*, 2013. [Online]. Available: https://www.w3.org/TR/sparql11-overview/. [Accessed: 06-Feb-2018].

[48] E. Prud'hommeaux, C. Buil-Aranda, A. Seaborne, A. Polleres, L. Feigenbaum, and G. T. Williams, "SPARQL 1.1 Federated Query," *W3C Recommendation*, 2013. [Online]. Available: https://www.w3.org/TR/sparql11-federated-query/. [Accessed: 26-Oct-2016].

[49] M. Bilal *et al.*, "Big data architecture for construction waste analytics (CWA): A conceptual framework," *Journal of Building Engineering*, vol. 6, pp. 144–156, 2016.

[50] C. McGrath, "Waste minimisation in practice," *Resources, Conservation and Recycling*, vol. 32, no. 3–4, pp. 227–238, 2001.

[51] G. Banias, C. Achillas, C. Vlachokostas, N. Moussiopoulos, and I. Papaioannou, "A web-based Decision Support System for the optimal management of construction and demolition waste," *Waste Management*, vol. 31, no. 12, pp. 2497–2502, 2011.

[52] Y. Li and X. Zhang, "Web-based construction waste estimation system for building construction projects," *Automation in Construction*, vol. 35, pp. 142–156, 2013.

[53] TechInvestLab.ru, ."15926 Editor," 2013.

*Corresponding Email: aga1@hw.ac.uk

[54] M. A. Musen, “The Protégé project: A look back and a look forward,” *Association of Computing Machinery Specific Interest Group in Artificial Intelligence*, vol. 1, no. 4, 2015.

[55] Protégé Project, “Cellfie,” 2017. [Online]. Available: https://github.com/protegeproject/cellfie-plugin. [Accessed: 20-Aug-2018].

[56] D. Calvanese *et al.*, “Ontop: Answering SPARQL queries over relational databases,” in *Semantic Web*, 2017, vol. 8, no. 3, pp. 471–487.

[57] S. Das, S. Sundara, and R. Cyganiak, “R2RML: RDB to RDF Mapping Language,” *W3C Recommendation*, 2012. [Online]. Available: https://www.w3.org/TR/r2rml/. [Accessed: 06-Feb-2018].

[58] T. Knap, P. Hanecák, J. Klímek, C. Mader, and M. Necaský, “UnifiedViews: An ETL Tool for RDF Data Management,” *Semantic Web*, 2018.

[59] D. Calvanese *et al.*, “Ontologies and databases: The dl-lite approach,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5689 LNCS, pp. 255–356, 2009.

[60] WellMet2050, “Conserving our Metal Energy,” 2010.

[61] Bioregional, “Reclaimed building products guide,” 2008. [Online]. Available: http://www.wrap.org.uk/sites/files/wrap/Reclaimed building products guide.pdf. [Accessed: 16-Aug-2017].

[62] DNS, “Late Life Planning Portal,” 2017. [Online]. Available: https://l2p2.net/SitePages/Home.aspx. [Accessed: 16-Aug-2017].

[63] A. G. Akinyemi, “Ontology-Based Data Access Using Ontop,” *Zenodo*, 2018. [Online]. Available: https://doi.org/10.5281/zenodo.1439101. [Accessed: 30-Sep-2018].

[64] A. G. Akinyemi, “Data Mapping Using dot15926 Editor,” *Zenodo*, 2018. [Online]. Available: https://doi.org/10.5281/zenodo.1405561. [Accessed: 29-Aug-2018].

[65] J. A. McMurry *et al.*, “Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data,” *PLoS Biology*, vol. 15, no. 6, pp. 1–27, 2017.

[66] A. A. L. Abdullah, S. A. Helmi, A. A. Z. Kadir, and M. Hisjam, “Cost Estimation Model of Structural Steel for Super Structure of Wellhead Platform in Oil and Gas Industry,” in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 2018, pp. 2092–2100.

[67] C. Walsh, “Data and Analytics: Open Source Data Integration Tool Comparison,” 2016.

[68] F. Etro, “The Economics of Cloud Computing,” in *Cloud Technology: Concepts, Methodologies, Tools, and Applications*, IGI Global, 2015, pp. 2135–2148.

'Declarations of interest: none'