Information System Artefact Development for Evaluating New Business Opportunities

Completed research paper

Shastri L Nimmagadda
School of Management (BIS)
Curtin University
Perth, WA, Australia
shastri.nimmagadda@curtin.edu.au

Gustavo Cardenas Barragan
School of Management, Curtin University
Perth, WA, Australia
g.cardenas@postgrad.curtin.edu.au

Torsten Reiners
School of Management
Curtin University
Perth, WA, Australia
T.Reiners@curtin.edu.au

Lincoln C Wood
Department of Management
Otago University, Dunedin
New Zealand
lincoln.wood@otago.ac.nz

Abstract

The data heterogeneity and multidimensionality, including the exponential growth of upstream businesses, complicate the construct-design and model development, precluding data integration process. Besides, an uncertainty arises with the effectiveness of business models, making prototype evaluation uncertain, risking business prospects. The significance of Information System (IS) artefacts is realized in upstream business contexts, in which unstructured volumes and varieties of data attributes and instances may have complicated the artefact designs, especially in geographically distributed petroleum systems. We consider a petroleum-prospecting company, a case study with a focus on IS development for upstream businesses and their growth with new reserves. As a part of business IS artefact development, we study a multidimensional modelling approach, in which structure and reservoir models are integrated into logical metadata to extract valuable meta-knowledge. The contribution of research lies with facts of designing new IS artefacts and evaluates their efficacy in new knowledge domains using artefact-utility properties.

Keywords: Upstream Business, Information System, Design Science and Knowledge Management

1 Introduction

The current research is intended to explore the scope of design science research in business information system development and bring out useful information through artefact design and its evaluation in industry scenarios. The designs deal with creating new model articulations that are non-existent in petroleum industry contexts. Baskerville et al. (2018) elucidate the creation of an artefact in different knowledge domains. We have found similar research gaps in upstream businesses in which we intend to investigate the connectivity between domains and systems in diverse geographic contexts. In the present study, we propose knowledge-based innovative artefacts through various elements and processes and analyse them through effective use of artefacts in industrial applications. Viable IS artefacts are in good demand in the form of multidimensional models and evaluate their utility properties in addition to assessing their qualities, envisioning the upstream business IS with sustainable architecture. The artefacts, we focused in the present study include design considerations and integrated architectures that accommodate different models, deduced from multiple domains of upstream businesses. In addition, we intend to analyse the connectivity between various petroleum ecosystems and evaluate the use, the reuse and effectiveness of artefacts in diverse spatial-temporal contexts for further knowledge discovery and management. The knowledge depends on evaluable properties of IS artefacts that ensure connectivity between various petroleum systems and upstream businesses. The petroleum systems that constitute different elements, such as structures, reservoirs, seals, source and processes, migration and timing of occurrence and energy accumulations are interpreted with diverse entities and dimensions in a basin-scale (Figure 1a). The conceptualization and contextualization attributes are interpreted to interconnect petroleum systems in a digital ecosystem. At sedimentary-basin scale in between geographically varying attribute dimensions can make connections into innovative multidimensional repositories (Jing et al. 2012). The fine-grained multidimensional data structuring is vital with added domain ontology descriptions to make the IS approach, effective in upstream business scenarios.

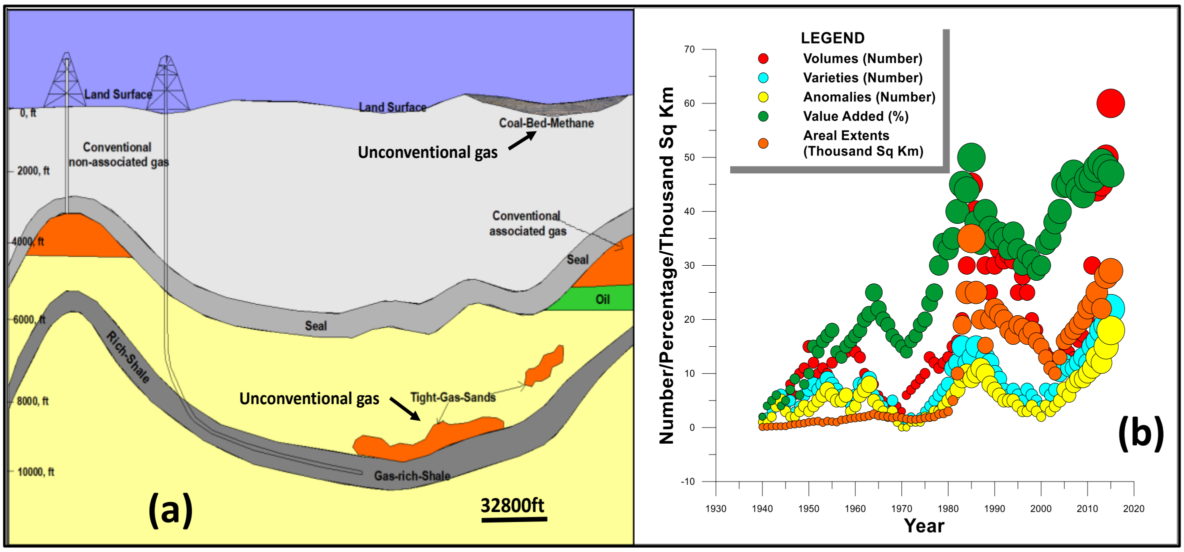

Figure 1: (a) A schematic view of digital ecosystem (b) upstream petroleum resources

Bubble plots have definite advantages with visual articulations, in comparison with line graphs, because each bubble describes the size and magnitude of the data point. Bubble plot can better visualize multiple attribute variables on a scaler plot in comparison with a line plot. In the bubble plot view, the diameter of each bubble can vary in size, providing additional dimensions of data. Traditional line or scatter plot views reveal the only number of volumes or varieties in digits. Figure 1b also supports data science with diverse response characteristics of Big Data features. In the bubble plot view, additional dimensions of data can be visualized through response diversity, which describes the extent of heterogeneity of Big Data systems. For example, the data-size increase and decrease observed, in which year are graphically easy to interpret, compared with line graphs. We appraise the existence of Big Data in the upstream petroleum businesses, as it is illustrated and clarified in Figure 1b, including its significance and motivation in the current study area. The petroleum ecosystems are spatial-temporal with the communion of multiple domains and systems. They are geographically inhabited with large size Big Data, which interpret various events, functions and processes that may have ensued and embedded inherently. In the current study, we consider the upstream business characteristics as attribute dimensions in a typical dimensional modelling process. All the fact instances of elements and processes of petroleum systems are identifiable in a manner that users can interpret, perceive and extract useful knowledge in diverse events of petroleum geology and their linked geophysical inputs (Catchpole and Robins, 2013). As discussed in Meena and Meena (2016), innovative data management practices, such as data warehouse repositories and digital petroleum ecosystems though have become popular, lack of engagement of IS artefacts in the data integration process, interoperability and the connectivity between systems are yet to be achieved. In addition, resolving the associativity between conventional and unconventional petroleum systems is challenging. Further to manifest the concept of petroleum ecosystem as shown in Figure 1a, we interpret multiple systems and their data sources within a broad geosyncline in a sedimentary basin setting (Syed et al. 2013). As shown in Figure 1a, many oil and gas fields can exist within a single petroleum system in Big Data scale. The Big Data characteristics and their spatial-temporal factual attribute instances are examined (Figure 1b) to ascertain the modelling requirements. It can minimize the ambiguity involved in the interpretation of models pertained to oil and gas prospects and their analysis in the investigating oil field areas. Use of the Big Data tools and digital ecosystem technologies can change upstream business values, based on the size of volumes and variety of exploration data. Appropriate design and development of IS artefacts may be prerequisites for evaluable interpretation and implementation in upstream business research. In addition, the assessment of utility properties such as use, reuse, efficiency, effectiveness and quality of IS artefacts need attention, with implementable integrated architectures, compatible with upstream petroleum industries.

2 Literature Survey and Motivation

The research gaps are discussed in the section. The authors engage the design science discussion and its analysis in industry scenarios with practical applications. The viability of IS research and practice depends on feasible field application with an explicitly defined sequence of steps and evaluable outcomes. The goal of the systematic approach is to ascertain the most efficient means of generating reliable and decisive outcomes through a variety of field designs. Baskerville et al. (2011) provide an information-system design theory as a prescriptive notion, integrating the normative and descriptive theories to enable decision-maker with desired outcomes. Jing et al. (2012) discuss practical aspects of IT related design tools, such as data-warehouse design requirements and challenges for streamlining the upstream field design with drilling strategies. Still, the study lacks a systematic Business Information System approach that can articulate a robust data modelling methodology, with improved connectivity between systems and interoperable utility properties. Though Abou-Sayed (2012) has described various database and mining applications, the data integration, use and reuse of models, including the effectiveness and efficiency of modelling practices have not been fully explored. Briscoe (2010) have discussed digital ecosystem approach with service-oriented architectures, but without adequate application and feasibility of models resolved in Big Data business scenarios. Syed et al. (2013); Topp et al. (2008) discuss various production challenges in the petroleum industries without discussing the remedies, in particular, the reservoir quality and distribution. The study further needs a holistic IS approach that can offer new artefacts and integrate various business activities and functions with their evaluable utility properties. Downes et al. (2014) discuss various attribute dimensions that contributed to the boom of the Australian resources industry, but the study is in short of a systematic attribute modelling approach. Mallory and Wright (2017); Meena and Meena (2016) describe various concepts of data- knowledge-lakes, but the study is inadequate enough to implement and evaluate of artefact designs, how upstream business IS can accept these concepts to resolve the E&P issues and challenges is not explored. Keeping in view the existing data management (Nimmagadda et al. 2017) and associated application challenges, the authors explore new research in modelling multiple attribute dimensions and their data schemas to accommodate in an integrated architecture. We need a new direction in the Information System (IS) development process, based on which the upstream business contexts are adaptable with their requirements (Catchpole and Robins 2013; Nimmagadda and Dreher 2012). In addition, the domain ontologies are articulated envisaging the integrated approach and exploring the systems' connectivity with which construction of multidimensional warehouse repositories are articulated (Yu 2019). To ascertain the knowledge-lake building process using IS artefacts, we need an integrated architecture with Big Data analytics and automated knowledge management (Agarwal and Dhar 2014).

Syed et al. (2013); Topp et al. (2008) reiterate that the explorers need long-term exploration plans with strategic commitments for discovering new oil and gas prospects. For sustainable exploration and production in the upstream businesses, the explorers are constrained with strategies and new exploration plans. We need viable data artefacts, adaptable in IS architecture development, to ensure successful implementation at prospect evaluation and drilling campaign phases. In the data modelling perspective, sustainability is a composite attribute, comprising of several related data attributes, based on which the upstream businesses and petroleum systems can be made evaluable. For sustainable business systems, we need quality IS artefacts with evaluable utility properties. In addition to sustainability, the supply and demand of petroleum resources are other key dimensions in the upstream business environment, including pricing of the oil and gas resources (OGERE, 2012). Sustainability depends on the periodic attribute, chosen in the dimensional modelling that can leverage its viable influence on the life span of petroleum prospect. Each petroleum province can produce oil and gas without interruption, keeping in view the size, quality of reservoirs and computed-life span of petroleum occurrences with net present value (NPV) of the market (Catchpole and Robins, 2013). We further investigate the existing historical data that involve several data attribute dimensions and volumes of factual instances associated with multiple oil and gas fields to build empirical models using innovative articulations. The introduction and literature surveys thus motivate us to draw the research purpose and objectives, as discussed in the following sections.

3 Problem Statement and Challenges

The heterogeneity and multidimensionality characteristics pose business-data modelling and integration challenges, including artefact-design requirement analysis. The transformation of voluminous upstream data into valuable new business knowledge is an intricate task. In certain situations, a single database cannot provide answers to the complex problem solutions to G&G domains of upstream petroleum companies. Integrating or assembling information from several databases is a significant challenge of the upstream business informatics while discovering new knowledge of business scenarios. Data integration by traditional approaches is challenging and incompatible because of scalability, format differences, including inflexible *survey-wells-permits* data structures. The schematic, semantic, syntactic and system heterogeneities may have created inconsistencies while structuring dimensional models with their attributes, adding ambiguities in the meaning, naming conventions, terminologies including language rules (Bishop et al. 2011). Managing interoperability, use and reuse of the data dimensions and their structures are added significant features, despite data instances that may have fallen in poor data quality areas, at times they are difficult to interpret. The issues can even affect the IS artefact design considerations and the development process in the proposed integrated architecture. We further investigate the scope of upstream business information systems, where Big Data play a vital role. In such contexts, innovative artefacts are in good demand with a new direction in the evaluation and implementation processes. As described in Figure 1b, we associate the Big Data characteristics with periodic attribute dimensions to move forward with the development of upstream business information systems with multidimensional artefacts in the spatial domain (Nimmagadda and Dreher 2012). The proposed architecture is expected to integrate multidisciplinary data and generate new upstream business opportunities in the investigating areas.

4 Research Methodologies

We deploy mixed method of research, for which data acquired through multidimensional field sensors, their seismic data observations are used to process and interpret the seismic instances by both qualitative and quantitative approaches (Cooksey 2020). In our research, the phenomenon is to explore the connectivity between geology and geophysics entities and their attribute dimensions in the investigating areas and investigate prospective locales. In addition, we need to ascertain more precisely the fact of connectivity between the geological structures, reservoir settings, and matured source and seal rock, as elements including migration pathways, hydrocarbon generation and accumulations as processes. We have chosen to model both elements and processes of petroleum system as attribute dimensions in the IS guided dimensional modelling approach. Using geology and geophysical data attribute dimensions, we build various logical and physical data schemas. These artefacts are guided by design science architecture from modelling to implementation stages, including testing models through instantiation process (Peffers et al. 2007 and Venable et al. 2016).

5 Architecture Development and Modelling Methodology

The conceptual, logical and physical entity-relationship (ER) models, dimensional models, different types of schemas are appropriate to contextualize IS artefacts in a variety of business perspectives. Ontologies that describe the data relationships through a variety of models are considered as artefact development procedures, including achievable knowledge-based connectivity, for example in between multiple petroleum systems. We consider them as IS artefacts, models, and methods in the design-science architecture development (Hevner et al. 2004). The data acquisition, processing and interpretation, are artefact features of IS articulations from which the upstream business research outcomes are deliverable. Even the types of data operations, which we have used in our business research, are no different from designability of IS architectures, rather they represent IS artefact application development procedure (Lee et al. 2015). The design science guided IS design and development is an established analytical approach with purposeful artefacts that address ecosystem challenges and evaluate artefacts through a variety of utility properties (Venable et al. 2016). The purposeful artefact describes features of system development, including overall reasoning of the integrated architecture in upstream business contexts. Peffers et al. (2007) provide a commonly accepted framework for design science research with guidelines and evaluable utility properties. We follow up the design science approach and its guidelines as an acceptable research paradigm in the current IS architecture development. In addition, the methodology presents a rigor on development of design science-articulations, with a business scope and opportunity in oil & gas industries. We take advantage of logical data storage systems and data dictionaries, interlinking millions of geological and navigational fact instances, geographically coexistent with petroleum systems or groups of systems (Nimmagadda and Dreher 2012). Hundreds of attribute dimensions are identified, at places conceptualized, to build ontology models for diverse business contexts.

5.1 Discussions on Model Articulations

The architecture discussed in Figure 2, describes various data acquisition, processing and analysis tasks, pertained to seismic prospecting with volumes of multidimensional sensor data and their linked drilled-well data -attributes within available permit-licenses of the investigating areas. We have reviewed the research frameworks, existing in the upstream business contexts and accordingly cited them in the literature review section. The novelty of our framework described in Figure 2 lies with facts of managing multidimensional and heterogeneous data sources, with descriptive and exploratory investigations that are lacking in the current research framework articulations. The artefacts used in the integrated articulations are knowledge-based and innovative.

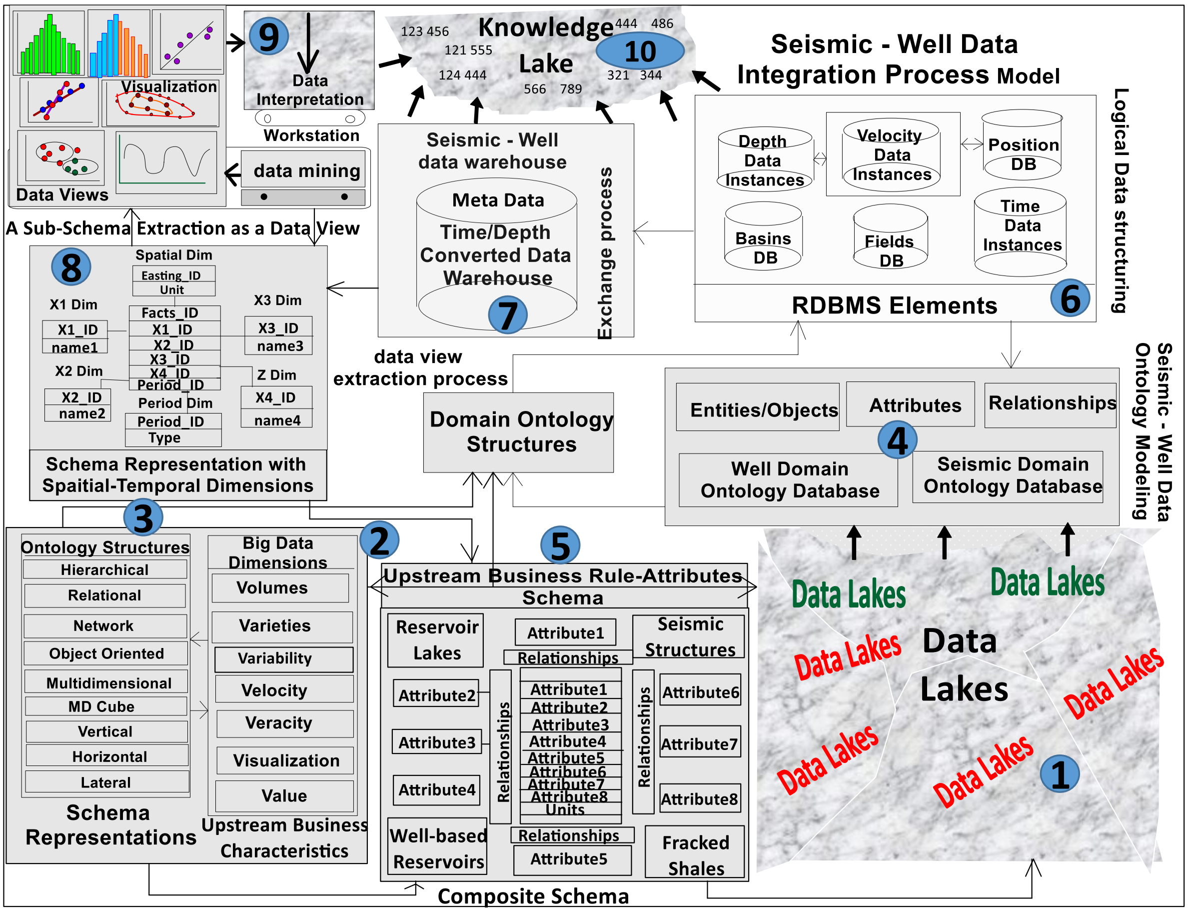

Figure 2: A conceptual architecture development – integrating data sources from multiple domains in a warehouse environment

Before structuring constructs and models, the issues associated with schematic, semantic and syntactic heterogeneities, including naming conventions, vocabularies and terminologies are resolved between variable attributes. In particular, ecosystem contexts, which may have characterized within Big Data descriptions, maybe at times difficult to reconcile due to data quality challenges, when visualization and interpretation are the focus of study. With tacit knowledge of the complexity and schematic visualization of workflow, it is important to describe a comprehensive sequence of events and tasks occurring in the workflow. Item Number 8 is a schematic data visualization extracted from the data warehouse. Item number 8 is itself a sub-schema view, drawn from warehouse metadata, used for data mining, visualization and interpretation, as described in item 9. Figure 2 describes the connectivity between 7, 8 and 9 items. Ultimate role-play is described in item number 10; the data and information acquired from knowledge lakes (1) are translated into knowledge-lakes (10).

In addition, the Big Data guided data-lakes of upstream E & P business information solutions are adaptable to manage, govern, access and explore the data-driven decisions, based on which the reservoir lakes are interpreted (OGERE 2012). Details of Big Data and knowledge lakes given in Meena and Meena (2016); Mallory and Wright (2017) have motivated us to extend the current contextual application in knowledge-based upstream business-information system development scenarios. As demonstrated in Figure 2, we articulate IS artefacts that are compatible with upstream business contexts. The E & P data are interlinked with multiple reservoir lakes, which are controlled by spatial attribute dimensions. As described in Figure 2, the itemized numbers 1, 2, 3, 4, 5, 6, 7 and 8 characterize different artefacts that are integrated through domain ontology procedures, ascertain the connectivity between attributes and unifying the overall architecture. Ontologies are descriptions of data relationships, attributed in between attributes and their models to deduce knowledge of upstream businesses. The multidimensional data structures, simulated as IS artefacts, are shown in Figure 2. Numbers 4, 5 and 6 represent the data artefacts respectively needed in the integration process with associated data mining, data visualization and interpretation activities, undertaken in the architecture development. Numbers 9 and 10, as shown in Figure 2, represent the data mining, visualization and interpretation artefacts. The knowledge-lake (item number 10) is developed by means of exploration activities carried out at stage 9. When the E&P is associated with the Big Data tools, the E&P is made compatible with multiple schemas. To be able making connections in diverse upstream business contexts, petroleum and artefact-evaluation systems and finally arrive at unified metadata, implementable within the architectural framework. The method ensures business connections among various petroleum systems in E&P scenarios (Catchpole and Robins, 2013; Downs et al. 2014). We further add values to the artefacts relevant to upstream BIS through data- and knowledge-lakes in support of novel integrated architectures. In pursuance of schemas designed in the E & P scenarios, we explain the compatibility or the congruence between dimensions of E & P entities, including their multidimensional representations. The e-clusters and their associated file systems, including applications of HDFS are installed in our labs. We use Apache Hadoop tools to store large volumes of data using MapReduce for data processing. Item Number 8 in Figure 2 is a schematic view of data categorization to mine data views in spatial-temporal attribute dimensions and process them for visualization and interpretation. (https://www.geeksforgeeks.org/introduction-to-hadoop-distributed-file-systemhdfs/?ref=lbp).

5.2 Business Knowledge Lake, a Conceptualized IS artefact

The data lakes may possess several conceptualized, contextualized attributes with unstructured information (Meena and Meena 2016). The knowledge may be ambiguous that cannot interpret the interconnectivity between elements and processes of petroleum systems. However, the exploration data-science emerges as a dominant knowledge-lake builder in the oil and gas industry, predominantly for exploration businesses. Knowledge Lake may be a pathway to interpret big data contexts, architecting and preparing data lake for queries using analytics solutions. Big data guided knowledge lakes derive new insights from large size upstream business metadata structures. In data repository environment, several data marts can cater to knowledge lakes with relational data subsets for specific applications. In addition to big data analytic solutions, for specific applications, we need data mining, visualization and interpretation artefacts to add value to various products of knowledge lakes (Beheshti et al. 2018). Additional values can come from IS articulations, in which we integrate several logical and physical schemas from hundreds of knowledge-based attributes, interpreted in G & G and petroleum system scenarios. The attribute dimensions and their instances are connectable to fact tables through one-to-many and many-to-many data relationships, as corroborated in Nimmagadda et al. (2017). Big Data is an added advantage, supporting the type of data relationships in warehouse architectures, adaptable in upstream businesses. Applications of Data Warehouse (DW) architectures are discussed in Abou-Sayed (2012); Jing et al. (2012). Using Oracle-driven warehouse structures, metadata cubes are generated to process the data views by OLAP engines.

Logical data relationships are explored through data modelling of elements and processes of several petroleum systems; the connections grow from the structure and reservoir-based domain ontologies through links of common attributes of multidimensional artefacts. The cuboid cells that hold the petroleum data instances are connectable to the characteristics of the Big Data and their tools (Chen et al. 2012). It is one of the innovative upstream business IS solutions for exploring the connectivity between elements and processes of petroleum systems and minimizing the risk of E&P interpretation. In addition, for systems dealing with the Big Data in the upstream petroleum exploration, the field of geo-informatics plays an inclusive role in the study of fundamental geological problems owing to the exponential explosion of sequence and structural information with time and geography (Nimmagadda and Dreher 2012). The process articulates a geo-informatics solution with integrated IS articulations. For extracting knowledge on favorable geological-structures that hold commercial oil and gas, unified petroleum systems' approach need to collaborate with applications of Big Data. As demonstrated in Figure 2, data in multiple domains connect with petroleum systems and their related upstream businesses to assess the economic viability through evaluable artefacts in sustainable oil and gas development scenarios. Sustainability may depend on the production of oil and gas for longer periods of time (Downes et al. 2014). For implementing the integrated architecture in the current contexts, we make sure that the IS artefacts are articulable in systems' integration. Like any structured data, for upstream exploration businesses, in a geological sense, the data are presented in different structures from data-lakes (Mallory and Wright 2017). However, they are usually deposited in multidimensional repositories. Data views are extracted using structured query languages and other data mining algorithms (Chen et al. 2012). The artefacts are evaluated using various utility properties, as discussed in Venable et al. (2016). Analyzing the effectiveness of artefacts is achievable through interpretability, including the use and reuse properties. The exchange of data structures in different applications, even meta-knowledge extractable from metadata with large-size knowledge lakes are attainable.

6 Analysis and Discussions

To deliver a sustainable upstream business solution, the artefacts, models and methods, must meet the evaluable utility property criteria (Venable et al. 2016), with measurable and deliverable sustainable oil and gas production. In addition, we need to understand the connectivity between utility properties. The data qualities can affect the key utility properties, affecting business research outcomes. As we construe the complexity of IS articulations in composite domain applications in upstream businesses, we ascertain the use of multiple evaluable utility properties, and assess IS artefacts in spatial-temporal dimensions.

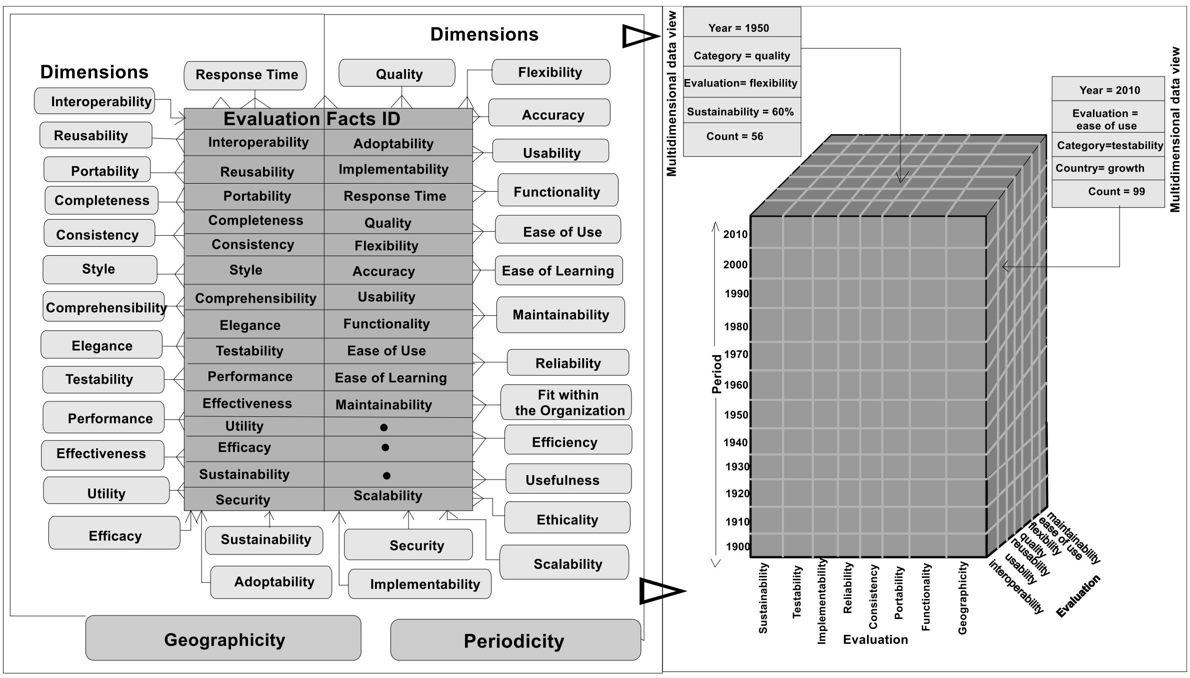

Figure 3: (a) Schema connecting evaluable utility properties (b) Cuboid structure

The added significance of "governance, interpretability, ease of use, usefulness and effectiveness" in the contexts of knowledge-based lakes of upstream business IS makes the schema architecture more robust and holistic (Figure 3a). The connectivity between utility properties in the current contexts is done through data schemas, considering data varieties, multiple data types and their logical relationships. In addition, for standardizing the structuring process, data storage and retrieval methods, we consider the scalability, extensibility and compatibility as criteria to benchmark the utility properties (Venable et al. 2016). Further, we conceptualize a schema comprising of utility properties and facilitating the interconnectivity between features of the articulated evaluation system, as demonstrated in Figures 3a and 3b. Different schemas considered in the study are compatible with cuboid data structures, which are synonymous to multidimensional warehouse structures. The business data attributes represented in attribute dimensions, and their artefacts are evaluable through various utility properties. The method can ascertain the efficacy of the IS architecture as upstream business IS solution.

6.1 IS Artefact Analysis and Interpretation

Large-scale businesses invest in IS designs and architecture development, for which evaluation is becoming critical. The evaluation depends on the application domain and the business value it can fetch to the industry. The upstream petroleum businesses must evaluate IS business values periodically, including spatially evaluable utility measures. Type of design science features can affect IS articulations, including their evaluations as provided in Hevner et al. (2004). An architecture designed for upstream business IS contexts, offers features for systematic data acquisition, processing and interpretation. Various artefacts represented with entities, dimensions, including ontology driven constructs and models are given in Nimmagadda et al. (2017). In the current contexts, the IS artefacts we focused, are from cross-sectional research domains. They are in the form of ontology structures, and their descriptions made from upstream petroleum domains, representative to different geographic and periodic contexts. There are many proven petroleum-fields worldwide, each with thousands of sq. km of areal extents with complex topographies and geographies. We consider secondary data factual instances and their empirical modelling in business research, including utility property analysis. Marthandan and Tang (2010); Venable et al. (2016) analyse IS artefacts through various utility properties, extending their analysis through empirical means in design-science research projects. For onsite workstation processing and interpretation artefacts that offer various computing and data processing facilities, we evaluate their data qualities based on which the artefacts are strengthened in the system integration and evaluation. Different plot views are computed to assess IS articulations of integrated architecture and high-resolution cuboid metadata. We further explore the opportunity of interpreting in several petroleum data views in new E&P knowledge perspectives. The new knowledge depends on the evaluable utility properties of IS artefacts, specifically artefact designs affected by data qualities. The data quality has been a significant concern in addressing the design considerations and their values in IS architecture development (Dan 2011). The data quality influences the existence of evaluable utility properties. As shown in Figure 4a, three distinct peaks of surveys interpreted from upstream business data have brought out matching utility property features, conforming actualization and goal utilities. In spite of data quality challenges, the trends and patterns disclose useful knowledge on "ease of use", "usefulness" and "effectiveness" of the evaluable utility properties of IS artefacts, as shown in Figure 4. In particular, the use and reuse of repository systems are blessings to upstream businesses. Interoperability is an added utility, which provides IS designs and data services to multiple users, systems and/or processes, through "share", "exchange" and "access" amenity attributes. The ease of use, usefulness and effectiveness of IS artefacts, including qualities based on functionality of the upstream business scenarios, are deliverable utility properties of IS artefacts.

In the current application, we analyse metadata associated with upstream businesses. The metadata comprises of surveys-drilled-wells-permits data to analyse hundreds of entities and dimensions with multidimensional constructs and models, including their utility properties and evaluations. As shown in Figure 4a, we interpret three significant features in the form of positive peaks with attribute strengths and their instances. The upstream business attribute strengths match with design-science evaluable utility property strengths as shown in Figures 4a and 4b. In X axis, attribute instance "Year" and "Percentage" on Y-axis. Percentage implies the measurable utility properties, such as "Quality", "Governance", "Interoperability", "Ease of Use", Usefulness and "Effectiveness" attribute strengths and their percentage rates. In Figure 4b, the authors present a regression model for an attribute "Number of Data Sources" along with "Added Value", corroborating with evaluation utility properties, deduced for contextual upstream business scenarios. The attribute "Value" being one of the characteristics of Big Data, appears matching its instance with instances of evaluation utility properties and their upstream business attribute instances. The petroleum data views extracted from warehouse metadata structures are continuously monitored for checking G & G data qualities and their veracity of Big Data in the high-speed computing workstations. It is done for visualization and metadata interpretation, ensuring that the new knowledge obtained in various petroleum ecosystem contexts is up to date and implementable in upstream business information systems. This process further evaluates IS articulations of the architecture, as and when new data arrive and a demand for new artefacts is necessitated in flexible upstream business contexts.

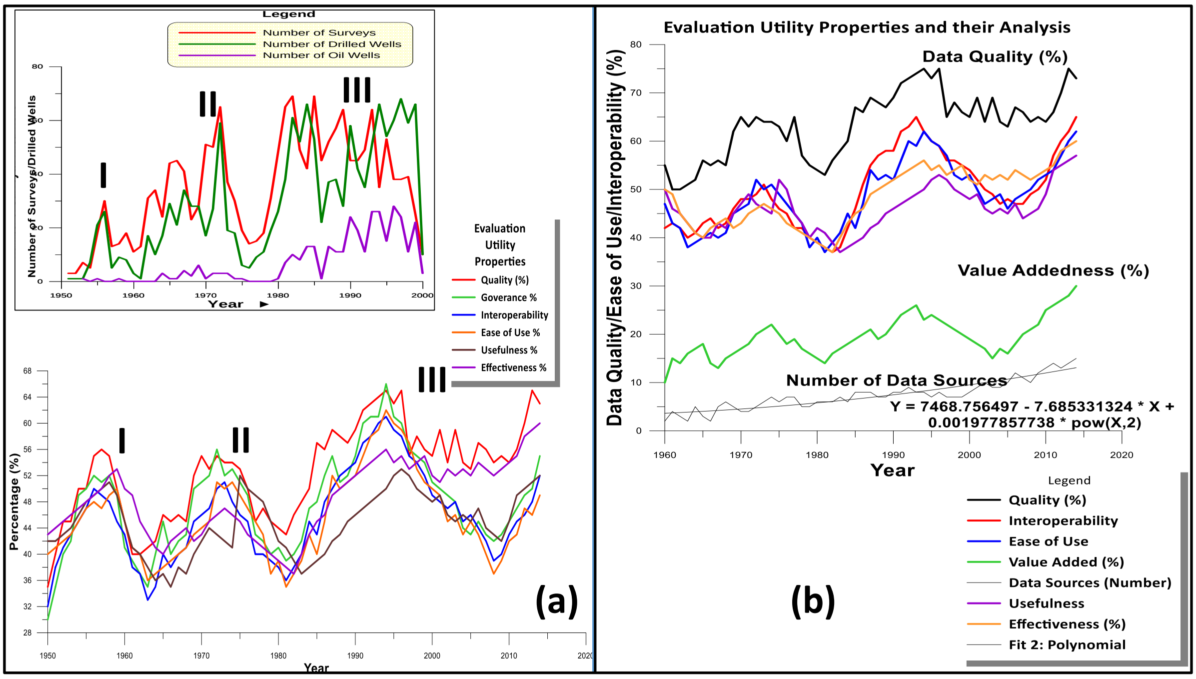

Figure 4: Analyzing evaluation utility properties with "year" dimension

6.2 Business Knowledge Extraction and Evaluation

The discussions explain the research contribution. Implementation of IS artefacts in upstream business depends on the efficacy of integration and validation of the instantiation process as described in Venable et al. (2016). The data cubes explored for upstream business connections through visualization, interpretation artefacts, are evaluable in the form of data views, how successful the IS artefacts and their implementations are in knowledge-based interpretations (Chen et al. 2012). The IS artefacts in the form of data schemas generated for Australian upstream business contexts, establish their implementations in spatial-temporal dimensions. To substantiate the research claim of successful implementation of IS artefacts and their linked schemas of the integrated architecture; we have evaluated attribute models using "structures" and "reservoirs" attributes of upstream businesses as shown in Figures 5a and 5b.

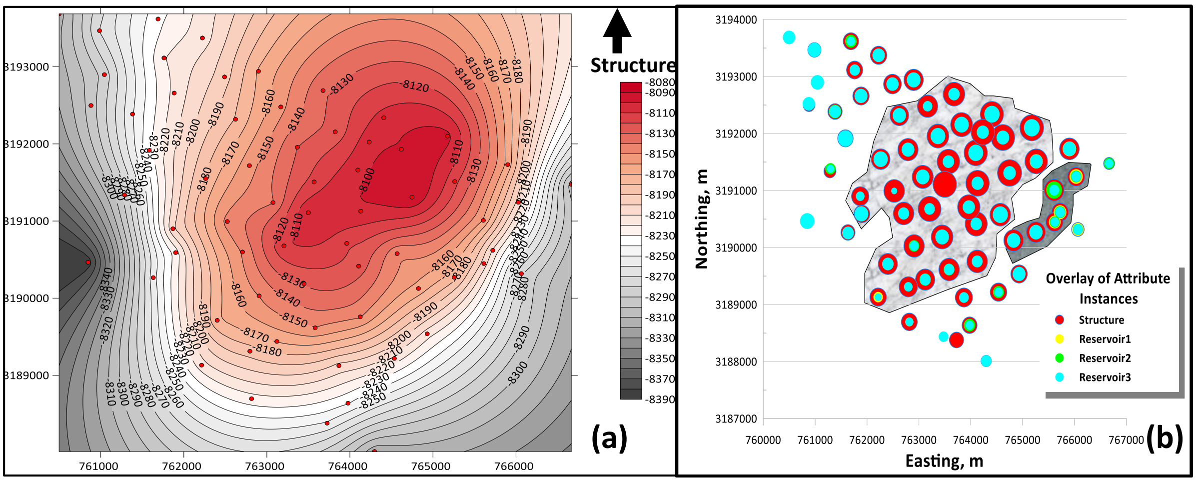

Figure 5: Exploring IS connections from (a) structure and (b) reservoir metadata

We use empirical data instances of upstream petroleum business data for drawing various data views from metadata volumes (Abou-Sayed, 2012; Jing et al. 2012). The map views drawn from metadata cubes are shown in Figures 5a and 5b for visualization and interpretation. Several areas of high porosity petroleum reservoirs that represent large-size bubbles have been interpreted as prospective areas. The bubbles exhibit closely connected clusters, suggesting multi-stacked reservoirs in the investigating area (Figure 5b). This information is immensely useful for petroleum business researchers, oil and gas explorers, and asset managers to further propose drillable campaigns and accrete more oil and gas reserves in the investigating areas. In addition, the map views are interpreted for discovering new knowledge of geological structures and associated oil and gas assets, in which the upstream businesses, linked petroleum systems and their elements and processes have all played major roles in venturing new investigating areas. As an example, Figure 5a shows the interpretation of (geological) structure in the petroleum province in which, we examine a sequence of events occurred while corroborating with IS artefacts, such as multidimensional data modelling, data warehousing and mining and interpretation artefacts. The geological structural feature interpreted in the map view as shown in Figure 5a is visualized as a structural high, in which several orientations of dip and strike attributes are construed, offering new prospective areas with further new upstream business opportunities. The reservoirs are described in terms of porosities, interpreted at respective spatial coordinates in the investigating areas. As shown in Figure 5b, several reservoirs characterized in bubble plot views are representative of IS articulated metadata views, interpreted within a single Digital Petroleum Ecosystem (DPE).

7 Research Contribution and Limitations

The research has a two-fold contribution, IS artefact-design and implementation, developing the application in the upstream business research. New artefacts are created which do not exist in the upstream business research domain. The design science approach facilitates articulating the framework with crucial constructs, models and methods, motivating the business IS. IS articulations must match the upstream business contexts for successful exploration with drillable campaigns. The opportunity lies with the facts of evaluating utility properties such as use, reuse, assessing the effectiveness, usability, usefulness, interoperability, performance, ease of use of IS articulations in the upstream business research for implementation and exploration. The potential users of the IS articulations are oil and gas explorers, data analysts and managers involved in the management of day-to-day activities and functions in geographically distributed upstream businesses. The success of IS architecture depends on evaluable artefacts and their careful utility property analysis in upstream business modelling. Data qualities may constrain their implementation in various business contexts.

8 Conclusions

The integrated architecture can evaluate multiple petroleum systems of any basin for sustainable petroleum production. Data modelling approaches that involve ontology-based IS artefacts can resolve the syntactic, semantics and schematic inconsistencies among diverse data dimensions and their associated data structures. They collaborate with conceptualized and contextualized attributes that emerged from multiple domains of upstream businesses and petroleum systems. The architecture is effective in integrating and connecting Big Data characteristics and their attributes associated with structure and reservoirs, and other elements and processes of the petroleum systems. The IS artefacts are assessable with evaluable utility properties. The Big Data volumes and varieties of the petroleum ecosystems and their integration in the holistic architecture can extract useful information, adding values to upstream business domain knowledge of oil and gas production. The fine-grained data structuring, multidimensional data warehousing, visualization and interpretation artefacts, deduced for petroleum system metadata demonstrate their successful implementations in the petroleum provinces. However, the metadata generated from G & G data sources infers the limits of large petroleum exploration areas that can make petroleum production more sustainable with systematic design, development and implementation of IS artefacts and their linked E&P scenarios. Several metadata models are designed in Australian contexts with new knowledge on the connectivity of petroleum systems and their fields through IS guided architecture. IS artefacts and their applications with Big Data tools and data analytics are rapidly evolving, with new insights of innovative designs and their understanding in exploring new knowledge in digital petroleum ecosystems. The integrated technologies that connect the systems' elements and processes in multiple domain applications ensure structure and reservoir extents with sustainable petroleum production in the investigating areas.

References

Abou-Sayed, A. 2012. “Data Mining Applications in the Oil and Gas Industry,” *JPT*, SPE, USA.

Agarwal, R., and Dhar, V. 2014. "Editorial—Big Data, Data Science, and Analytics: The Opportunity and Challenge for Is Research," *Information Systems Research* (25:3), pp. 443-448.

Yu, B. 2019. Research on information retrieval model based on ontology. *J Wireless Com Network* 2019, 30 (2019). https://doi.org/10.1186/s13638-019-1354-z.

Dan, Y. 2011. "The utility evaluation analysis of information technology based on it's usage frequency," *2011 IEEE 3rd International Conference on Communication Software and Networks*, Xi'an, 2011, pp. 278-282, doi: 10.1109/ICCSN.2011.6013592.

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., and Rossi, M. 2018. "Design Science Research Contributions: Finding a Balance between Artifact and Theory." *Journal of the Association for Information Systems*, 19(5): 358-376. https://doi.org/10.17705/1jais.00495

Baskerville, R., Lyytinen, K., Sambamurthy, V. and Straub, D. 2011. A response to the design-oriented information systems research memorandum, European Journal of Information Systems, 20:1, 11-15, DOI: 10.1057/ejis.2010.56

Beheshti, A., Benatallah, B., Nouri, R., and Tabebordbar, A. 2018. "CoreKG: a Knowledge Lake service," *Proceedings of the VLDB Endowment*, *11*(12), 1942-1945. https://doi.org/10.14778/3229863.3236230.

Briscoe, G. 2010. Complex adaptive digital EcoSystems. In <i>Proceedings of the International Conference on Management of Emergent Digital EcoSystems</i> (<i>MEDES '10</i>). Association for Computing Machinery, New York, NY, USA, 39–46. DOI:https://doi.org/10.1145/1936254.1936262

Bishop, B., Kiryakov, A., Ognyanoff, D., Peikov, I., Tashev, Z., Velkov, R. 2011. Owlim: A family of scalable semantic repositories. Semantic Web 2(1), 33–42 (2011).

Catchpole, M. and Robins, W. 2013. Oil and Gas And Energy Resource Exploration, Productivity Commission Issues Paper, the Australasian Institute of Petroleum and Metallurgy, March 2013.

Chen, H., Chiang, R. H., and Storey, V. C. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS quarterly* (36:4), pp. 1165-1188.

Cooksey, R. W. 2020. A Practical Introduction to Mixed Methods for Business & Management. *Journal of Mixed Methods Research*. https://doi.org/10.1177/1558689820940166

Dhar, V., Jarke, M. and Laartz, J. 2014. Big Data, WIRTSCHAFTSINFORMATIK, doi:10.1007/s11576-014-0428-0, Springer Fachmedien Wiesbaden 2014.

Downes, P. Hanslow, K. and Tulip, P. 2014. The Effect of the Petroleum Boom on the Australian Economy, Research Discussion Paper, The Reserve Bank of Australia, 2014-8.

Hevner, A.R., March, S.T., Park, J. and Ram, S. 2004. "Design science in information systems research," *MIS Quarterly*, Vol. 28 (1), pp. 75-105, Society for Information Management and the Management Information Systems Research Center, MN, USA, ISSN: 0276 - 7783.

Jing, N., Fan, H., Zhai, Y. and Tianyu Liu, T. 2012. "Data Warehouse Design and Optimization for Drilling Engineering," *The Open Petroleum Engineering Journal*, 2012, 5, 124-129.

Lee, A.S., Thomas, M. and Baskerville, R. 2015. Going back to basics in design science: from the information technology artifact to the information systems artifact. *Information Systems Journal*, 25(1), 5– 21.

Marthandan, G. and Tang, C. M. 2010. "Information Systems evaluation: an ongoing measure," *International Journal of Business Information Systems (IJBIS)*, Vol 6, No. 3, Inderscience, Switzerland.

Mallory, J. and Wright, R. 2017. Building Big Data Storage Solutions (Data Lakes) for Maximum Flexibility, 2017, Amazon Web Services, Inc. https://d1.awsstatic.com/whitepapers/Storage/data-lake-on-aws.pdf.

Meena, S. D. and Meena, S. V. 2016. "Data Lakes – A New Data Repository for Big Data Analytics Workloads," *International Journal of Advanced Research in Computer Science*, Volume 7, No. 5, September-October 2016.

Nimmagadda, S., and Dreher, H. 2012. On new emerging concepts of petroleum digital ecosystem. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2.

Nimmagadda, S. L., Reiners, T. and Rudra, A. 2017. An Upstream Business Data Science in a Big Data Perspective, Procedia Computer Science, Volume 112, 2017, Pages 1881-1890, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.08.236.

Oil and gas and Energy Resource Exploration (OGERE). 2012. Productivity Commission Paper, Dec (2012). www.pc.gov.au (under 'projects')

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. 2007. "A design science research methodology for information systems research," *Journal of management information systems*, 24(3), 45-77.

Syed, A. Grafton, Q. and Kalirajan, K. 2013. Productivity in the Australian Petroleum Sector, Bureau of Resources and Energy Economics, www.bree.gov.au, March 2013.

Topp, V., Soames, L., Parham, D. and Bloch, H. 2008. *Productivity in the Petroleum Industry: Measurement and Interpretation,* Productivity Commission Staff Working Paper, December.

Venable, J., Pries-Heje, J. and Baskerville, R. 2016. "FEDS: a Framework for Evaluation in Design Science Research," *Eur J Inf Syst* 25, 77–89 (2016). https://doi.org/10.1057/ejis.2014.36