



ISSN 1943-7544

Pacific Asia Journal of the Association for Information Systems

Research Paper

doi: 10.17705/1pais.13101

Volume 13, Issue 1 (2021)

Big Data Guided Unconventional Digital Reservoir Energy Ecosystem and its Knowledge Management

Shastri L. Nimmagadda^{1,*}, Neel Mani², Torsten Reiners³, Lincoln C. Wood⁴

¹Curtin University, Perth, Australia, shastri.nimmagadda@curtin.edu.au

²Amity Institute of IT, Amity University, Noida, UP, India, nmani@amity.edu

³Curtin University, Perth, Australia, T.Reiners@curtin.edu.au

⁴Otago University, Dunedin, New Zealand, lincoln.wood@otago.ac.nz

Abstract

Background: *In shale basins, petroleum systems are complex; they hold data sources in Big Data scales. The motivation of research lies with the facts of exploring effective inherent connectivity between unconventional petroleum systems. The connectivity between energy reservoir systems is ambiguous within a distinctive petroleum ecosystem. Heterogeneity and multidimensionality of unstructured data sources are additional challenges, precluding systematic modelling of diverse petroleum systems and their data integration process, including growing demand for storage systems. The research aims to establish the knowledge-based connectivity between petroleum systems through Information System (IS) articulations, visual analytics and data management.*

Method: *We investigate the knowledge-based IS guided exploration and production systems to explore the connectivity between diverse unconventional petroleum systems and forecast the reservoir energy. We articulate Design Science Information System (DSIS), bring various IS artefacts, unify multiple domains of petroleum provinces and analyze the associativity between petroleum systems. In addition, use, reuse, effectiveness and interoperability are utility properties of IS artefacts that we evaluate. We implement IS solutions in the oil and gas industries to facilitate database management and reservoir energy exploration.*

Results: *We simulate DSIS as an Unconventional Digital Petroleum Ecosystem (UDPE) as it allows us to investigate and ascertain the interplay between petroleum systems' elements and processes. Metadata cubes are computed for data views to visualize, interpret, and implement IS articulations in energy systems. We compute the structure and reservoir attribute views for interpreting energy-driven petroleum systems, prospect evaluation and business-knowledge management with a viable DSIS solution.*

Conclusions: *The DSIS emerges as a knowledge-based digital ecosystem innovation, demonstrating how it can effectively interconnect geographically controlled petroleum systems. Its development, in the exploration of unconventional shale basins, is a knowledge-based reservoir-energy management solution. This research is beneficial to IS practitioners who wish to pursue energy research in reservoir ecosystem contexts.*

Keywords: Big Data, Unconventional Reservoirs, Design Science, Digital Ecosystem, Data Mining.

Citation: Nimmagadda, S. L., Mani, N., Reiners, T. & Wood, L. C. (2021). Big Data Guided Unconventional Digital Reservoir Energy Ecosystem and its Knowledge Management.

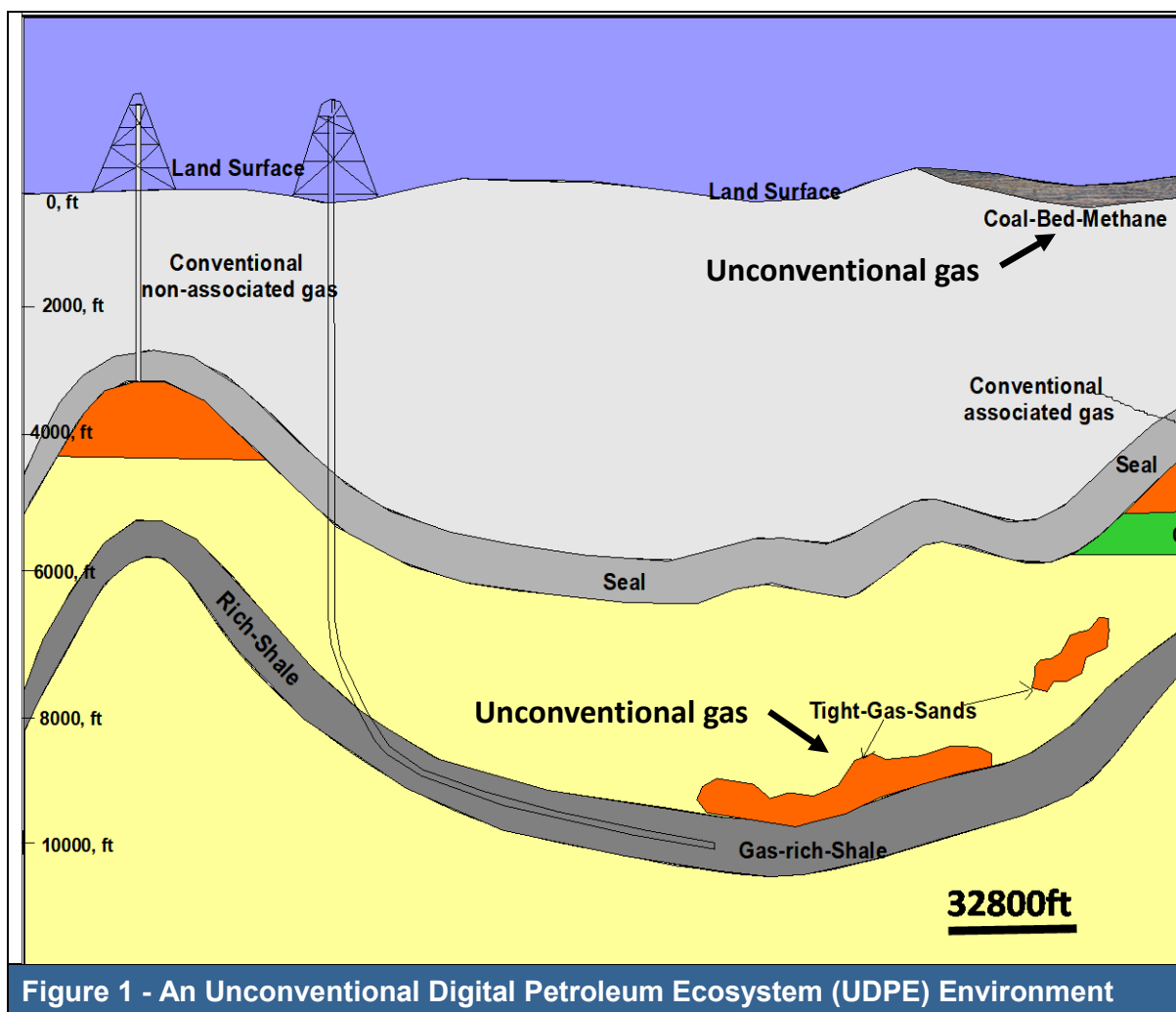
Pacific Asia Journal of the Association for Information Systems, 13(1), 1-35.

<https://doi.org/10.17705/1pais.13101>

Copyright © Association for Information Systems.

Introduction

Data modelling and system integration are critical tasks of upstream business research. It becomes particularly problematic when an unconventional digital reservoir energy ecosystem holds diverse data sources in multiple systems with varying magnitudes of elements and processes (Bento, 2018). In the current research, an ecosystem interprets a sedimentary basin with a bowl-shaped geosyncline structure (Figure 1) and several petroleum systems (Gilbert et al., 2004). Each oil and gas field may possess many oil and gas producing wells; each drilled-well has different frackable reservoir pay-zones. Each pay-zone is interpreted with various fluids - either oil or gas or both. Different attribute dimensions appear to reveal chains of events within networked reservoir zones in hierarchical data organizations or communities. Existence of gas/oil shales, coal bed methane, and heavy oil and gas hydrate resources is widespread in shale basins, as unconventional reservoirs (Li, 2011). Each of ecological basins can generate a large amount of data, about oil and gas fields and their associated multi-stacked petroleum energy-reservoirs (Cleary et al., 2012; Durham, 2013).



An Unconventional Digital Petroleum Ecosystem (UDPE) thus conceptualized within such complicated ecological settings (Dovers et al., 2001) is envisaged as connectable with numerous petroleum systems. Each system holds large volumes and varieties of data in spatial-temporal dimensions in Big Data scale (Castañeda et al., 2012; Ramiz & Yadigar, 2017). For exploring the connectivity between systems, we explore concepts of petroleum ecologies, envisaged as petroleum ontologies (Li et al., 2012; Chang & West, 2006).

Digital data coexist in both conventional and unconventional reservoir systems within a single petroleum ecosystem (Figure 1). The notion provides an opportunity to assess the entire reservoir energy ecosystem and its potentiality in prospect evaluation. Reservoir ecosystem is information system guided exploration research in this article. Digital reservoir attribute dimensions characterize differently in diverse spatial-temporal dimensions (Castaneda et al., 2012). The contextualization concept emerges in between oil, and gas fields in different ecosystems, more precisely in Big geographic Data scales. The entire narrative is a description of a digital petroleum ecosystem, construed with different multidimensional artefacts. An effective solution is required to manage digital data of reservoir ecosystems (in the terabyte scale) within a single petroleum ecosystem (Figure 1). The potentiality of ecosystems depends on how effectively the framework and related artefacts are articulated in repository systems. From a data-warehouse development perspective, the UDPE articulated in an ecosystem scale, can accommodate several artefacts through composite schemas with supertype and subtype dimensions. They are conceptualized and contextualized attribute dimensions, especially in regions where there are no geological boundaries. The novelty relies on the IS methodology and how effectively its artefacts can interconnect the geological boundaries of petroleum systems to unify into metadata for meta-knowledge.

The purpose of the research is to articulate an integrated ecosystem framework with different artefacts, emerged from multiple energy systems. It is derived from a generic Design Science Research (DSR) approach, as discussed in Peffers et al., (2008), from which IS articulations developed are implementable in multiple domain applications. The IS articulations can be adaptable and representable in rapidly changing data science conditions. In a digital oil field situation, the dimensions are logically structured with ontological descriptions. The data relationships interpretable in different artefacts or schemas are more connectable, flexible and extendable in large-scale ecosystems such as super basins (Nimmagadda et al., 2018). The necessary research episodes can uncover the connectivity between conventional and unconventional petroleum systems, coexisting within single ecosystem settings (Figure 1). The connectivity issues are resolvable by interpreting data views within UDPE contexts. In addition, the semantic, schematic and syntactic data challenges are other challenges, managed during the design and development of IS artefacts (Vaishnavi & Kuechler, 2007). In that analogy, the authors interpret the sedimentary basin (Li, 2011) as a digital petroleum ecosystem, in which multidimensional schemas are brought together in a repository system.

This paper describes the motivation of reservoir-energy research with DSIS method, how it can be used to satisfy the existence of UDPE and therefore enhance its exploration activities. The rest of the paper is structured as follows. We first outline the background and literature on such digital ecosystems and the challenges of an UDPE. We then outline the DSIS methodology used and the development of required artefacts and connections. As an illustration, we articulate the UDPE, compute the data views and evaluate its effectiveness in a real-world application.

Literature Review and Research Gaps

We examine the issues associated with existing data organizations. The alternative IS solutions motivated to manage the digital petroleum ecosystem communities are analyzed. We explore the research gaps in the current data models and domain applications. The digital ecosystem in the current energy scenarios describes an organization with symbiotic and interrogative data systems. Like any biological systems, they present self-organizing, scalable and sustainable activities and functions in various applications (Li et al., 2012). Yu et al. (2008) provide domain applications that use digital ecosystems and technologies. Sabry and Krause (2012) explore diverse systems and their development, providing energy production models over cloud computing analytical tools to manage energy supplies and optimizations. Henningson and Hedman (2014) develop a digital ecosystem with a technology-based

transformation framework, integrating business and technology ecosystem theories. They illustrate the use of framework through a case study, transforming digital payment ecosystem. Porteous and Morawczynski (2017) demonstrate the digital ecologies as super-platforms, exhibiting financial, and customer-ecosystem solution scenarios. Hasan and Kazlauskas (2009) develop a theoretical framework to explore the ecosystem approach and investigate ICT issues in climate change challenges. The cloud vendors focus on digital ecosystem values, articulating a paradigm to address high-performance computing, resources management and security demands (Briscoe & Marinos, 2009; Sabry & Krause, 2012). Similar concepts of ecologies with the representation of entities and attribute dimensions are interpreted in Dovers et al., (2001).

Several exploration tools and technologies (in particular, seismic attribute analysis and their interpretation) are more relevant for maximizing the reservoir energy production and distribution, as described in Alzate and Devegowda (2013). Arogundade and Sohrabi (2012) provide necessary tools for the development of reservoir energy resources. Seismic data analysis and interpretation have significant roles in shale-gas energy exploration and integrated reservoir characterization (Bellman, 2018), besides exploring the reservoir energy areas in the study area. The geological and geophysical challenges, mainly their linked tight reservoirs and associated unconventional petroleum systems demand new IS articulations with specialized data modelling approaches. Designing and developing large-size ontologies for complex systems is challenging (Flahive et al., 2004). Their study proposes a distributed ontology framework to tailor large size ontologies to describe complex systems. The framework has five major categories, ontology processing, ontology-location, ontology-connection, users' connection and algorithmic location. Integrating ecosystem innovations by new ontologies is the motivation that can improve ecosystem-service values (Furr & Shipilov, 2018). Strategic alignment of knowledge management focuses on the case study in the petroleum company, PETROBRAS, developing parametric, non-parametric statistics and multivariate analysis for oil and gas companies. As discussed in Costa and Rezende (2018), the value creation is the strategy in the oil and gas companies, including quality well completions as per target selections (Hemingway et al., 2012). Reservoir heterogeneity is analyzed to determine the reservoir stress anisotropy. The combination of bits, drilling fluids, drilling parameters and seismic field parameters are optimized and incorporated in the modelling process.

The unconventional reservoir energy system has not been the emphasis in the existing literature due to ecosystem complexity with unresolved diverse domains and data structures. We examine the existing literature focusing on ecologies, theoretical frameworks, design and implementation in business, digital energy and environmental service sectors. Tapping the unconventional reservoir energy from Gippsland and Otway basins in Australia has been the priority, including the safety and security of technology adoption during energy exploitation. Energy research in terms of production sharing contracts, data availabilities, mineral rights, pipeline programs, including distribution and marketing, is analyzed, as in Sandalow et al., (2014). Chinese investment opportunities with new technology motivations facilitate reservoir-energy resource development as in Xiaoming et al. (2020). Shah (2015) further reviews the IoT sensor tools with high-performance computing facilities through interconnected devices, including machine learning and Big Data analytics implementations. Subramaniam (2020) deduces an interdependence between data connectivity and the articulations of the digital ecosystems. Sensor and IoT technologies have galvanized the shape of digital ecosystems. Production and consumption ecosystems are part of the digital ecologies, and the value chains are construed from the interdependencies and the data connectivity. The author discusses three implications from the studies, scope of value creation, competition and digital transformation with monopoly power. Sunjay et al. (2014) provide comprehensive worldwide scenarios of the unconventional energy resources.

What is lacking in the current literature is the implementation of IS artefacts and their linked design-science research frameworks. Conceptualization and contextualization are features of diverse data sources to manage the design-science guided ecosystems and linked multiple systems. The connectivity between reservoir energy systems is more relevant in building new business opportunities to articulate new drillable exploratory campaigns in the upstream businesses. The growing ambiguities in data analysis of petroleum systems may have added uncertainty in interpreting multidisciplinary data, which motivated us to examine the existing reservoir simulation techniques. Interpreting favorable connectivity between systems is challenging in commercial petroleum provinces. It needs factual metadata views that characterize multidisciplinary data in multiple domains.

Other challenges include poorly managed data integration methods. Inadequate spatial-temporal information and indecisive knowledge on areal extents of systems make oil and gas exploration operations challenging, for which knowledge-based-decision-support digital ecosystem solutions are required. The data characteristics as schematized in Figure 1 can support and collaborate an assembly of relational, hierarchical and network type of complex spatial-temporal data structures, including applications involving the depiction of multiple ecosystems (Khatri & Ram, 2004; Ozkarahan, 1990). Still, the heterogeneity of the data sources poses critical data modelling and visual analytics challenges.

In addition, with the increasing exploration and production activities (Durham, 2013) in petroleum-bearing provinces, the sedimentary basins emerge with large volumes and varieties of Big Data sources in many upstream petroleum companies. Massive storage devices are needed for warehousing the data instances in petabyte-scale (Nimmagadda et al., 2019). In data modelling scenarios, managing thousands of such attribute data instances, connecting them with a similar number of fact tables is a tedious process with added data challenges of exploration and production dimensions (Coronel et al., 2011; Nimmagadda et al., 2018). Integrating multidisciplinary data in spatial-temporal dimensions is a significant challenge in digital petroleum ecosystem management. For example, in a hierarchical data structure, the digital oil and gas fields are made up of field data with: each field having several surveys, each survey acquiring volumes of data instances from millions of sensors in spatial dimensions (Castaneda et al., 2012). Each survey has several survey profiles, with added volumes and varieties of data. For several drilled-well attributes, each well has many horizons (geological formations). The digital oil and gas fields can be characterized in hierarchical and relational domain ontologies (Shanks et al., 2003). The criteria are to interconnect the surveys and drilled-well data. Therefore, the exploration of digital energy and contemplation of evaluating ontology-based IS artefacts within reservoir energy ecosystems is the focus of current research.

Geerts (2011) discusses the design science research methodology (DSRM) templates with design science research (DSR), illustrating an application through retroactive analysis. The author integrates the DSRM with operational views of artefact networks (Section 4, Figure 1, and Page 149). Peffers et al. (2008) enlighten the IS articulations deduced from design science research and design science research methodology and their motivation in building process models with evaluable implications. The design process constitutes problem identification and motivation, objectives description, design and development, demonstration, evaluation, including communication entities. Similar guidelines are used to validate grammar process designs (Lee et al., 2008). Tung et al. (2020) discuss digital transformation through a makeover of business activities, processes, products and models to leverage new scopes and opportunities of digital technologies, characterized by development, growth, innovation and disruption. The concepts of cloud computing, big data and the Internet of Things are added support to digital ecosystems in oil and gas industries in the fourth industrial revolution. Other challenges include increased hydrocarbon recovery, improvised business ecosystem, and lack of operational reliability. We review the research gaps with new opportunities, as summarized in Table 1.

We adopt a DSIS framework in the current research, articulating and resolving the reservoir ecosystem challenges and their management with innovative IS artefacts.

Table 1 - Existing Research Issues and Gaps	Bridging the Gaps – New Research Scopes
<i>Data modelling and integration challenges (Martinez-Mosquera et al., 2020)</i>	<i>Schema design and development; connectivity between systems is through a warehouse schema</i>
<i>Heterogeneity and Multidimensionality Challenges (Chromiak & Stencel, 2014)</i>	<i>Domain Ontologies and Schemas' Integration</i>
<i>The complexity of unconventional petroleum systems (Bento, 2018; Jia et al., 2016)</i>	<i>Complexity Resolvable with DSIS articulations</i>
<i>Constrained IS Artefacts and Ontologies (Gregor & Jones, 2007)</i>	<i>Big Data guided DSIS articulations</i>
<i>Ambiguous interpretations (Li et al., 2008)</i>	<i>Data views are user generated, so the ambiguity is mitigated while applying interpretation artefact</i>
<i>Use, reuse, effectiveness and interoperability properties are inadequately evaluable (Kadadi et al., 2014; Venable et al., 2016)</i>	<i>DSIS solutions are easily evaluable in the current domain applications</i>
<i>Inadequate data management services (Yu et al., 2017)</i>	<i>DSIS is designed for robust research data management solutions</i>

Research Questions and Objectives

The shale gas, coal tar sands, tight gas sands and coalbed methane are unconventional resources (Durham, 2013). They are abundant worldwide; natural gas operators or explorers have not exploited reservoir ecosystem viability using the latest IS/IT tools. It is partly due to scarce and sparse geological and reservoir engineering information, regulating the natural gas policy framework and promoting reservoir energy markets. In addition, there is a shortage of IS expertise and technical know-how needed to develop the models. The Big Data tools, as given in Cleary et al., (2012), motivate us to develop new data integration methods, unifying datasets in UDPE settings into a framework. The introduction and description of the problem statement expedite us to identify research gaps, framing the following research questions:

1. How do we design knowledge-based DSIS in unconventional digital ecosystem solutions?
2. How do we articulate and manage the Big Data tools, easing the complexity of exploration data, and facilitating the upstream businesses?
3. How do we evaluate and implement the DSIS solutions in digital reservoir-based energy ecosystem?

Based on the research questions, we describe the following research objectives:

1. *Articulate an integrated ecosystem framework:* In the UDPE contexts, different domains, types and sub-types of attribute dimensions are defined. We articulate a knowledge base and flexible methodological framework (Gregor & Jones 2007; Venable et al., 2016), describing the UDPE. We then analyze how the IS articulations can effectively intervene and facilitate exploring the connectivity between reservoir energy systems in spatial attribute dimensions.

2. *Share a common understanding of information and knowledge structure: Domain ontologies are expressed in various multidimensional relational-data structures, assimilating and sharing data relationships in multiple systems.* Domain experts, data analysts, oil and gas explorers and project managers should manage, share knowledge and use/reuse data structures in several domain applications. New data grids may support map creation in geographic environments. Different plot and map views with aggregated user queries can help facilitate their interpretation in new knowledge domains and support decision support systems.
3. *Enable the use and reuse of constructs/models and domain knowledge:* Keeping in view, the complexity of the application domain, the proposed artefacts may have design challenges that can affect the overall framework. In UDPE contexts, we ensure they are flexible and evaluable by use and reuse. Models in several domains/systems require signifying aggregated data view representations in spatial-temporal dimensions. The digital ecosystem representation includes the construction of models in time-and depth-intervals with the description of associated measures and units (Khatri and Ram 2004). If a group of explorers develops petroleum ontologies in detail, other domain experts can reuse them in related domains. For example, the domain knowledge acquired in a particular field is interpreted with a specific model or vice versa. It may be used or reused in other similar/dissimilar ecosystem scenarios.
4. *Digital ecosystem explores the connectivity between multi-stacked energy reservoirs:* One of the foremost objectives is to facilitate the investors with new business opportunities in the investigating areas using the UDPE assets and their management.

As a part of Research Objective 1, for addressing the connectivity between unconventional petroleum systems, the DSIS is articulated with different artefacts, representing multiple petroleum systems. Further, the research aims to implement the Big Data paradigm in broader digital reservoir-energy ecosystem scenarios (Research Objective 2). We must resolve the exploration data challenges in multiple domains before analyzing issues of reservoir connectivity.

The Motivation of the Reservoir-Energy Research

IS research in energy systems is relatively new, and limited literature is available. The design science research and literature on design and development have been enlightened by IS researchers. Limited research on evaluation and implementation of artefacts is another motivation of the current research investigation in the reservoir-energy domain application. The emergence of Big Data is an added support to develop reservoir-energy research. Data intricacies that preclude the data integration process has further encouraged us to explore new IS tools and technologies. Multiple domains and hundreds of linked attributes in industry scenarios expedite with new modelling methodologies. For example, for each business process, function and task, many-valued chains of events emerge in supply chain systems.

To connect systems, we need a more robust, holistic DSIS approaches to resolve the challenges of managing multiple systems through digital reservoir-energy ecosystems. DSIS articulations, their evaluable use, reuse, effectiveness and interoperability are other motivations of reservoir energy research. For example, multiple domains of the UDPE (Figure 1) may have closely connectable and unifiable attribute dimensions in ecosystem contexts, but heterogeneity may have affected the integration process. Digital ecosystem process explores the connectivity between multi-stacked fractured reservoirs of the unconventional shale-gas basins.

The methodologies needed for developing digital reservoir-energy ecosystem solutions are discussed in the following sections to address and assimilate the connectivity between systems. Based on connectivity, investors in petroleum industries look for new and sustainable reserves in unconventional shale basins. They often aim at specific production targets, improving their management for prolific frack reservoirs.

Methodology

Most interactive workstations are integrated interpretative systems that manage very large data systems. They collaborate with a cloud-facilitated integrated framework with Oracle-driven software to handle very large-size database management systems (Inmon, 2005; Lee et al., 2006). The unconventional petroleum systems are needed to store in high-performance computing databases. The notion of connectivity is described in the data integration process (e.g. Castaneda et al., (2012); Gilbert et al. (2004)). For establishing system connectivity, the data relationships are described as attribute dimensions. Within the context of a petroleum ecosystem, either element or process is interpreted as a dimension. The connectable data attribute events are conceptualized and contextualized in geographic dimensions (Khatri & Ram 2004). Ontologies described as data relationships are characterized as multidimensional data structures. In other words, to analyze a petroleum ecosystem, ontologically described data relationships are interpreted, envisaging reservoir connectivity ecology. The petroleum elements, such as structures, reservoirs, seals, source; processes, such as migration and timing of deposition and energy accumulations, may belong to multiple systems (Nimmagadda et al., 2019). The elements and processes that characterize a sedimentary-basin make up digital energy ecosystem representation.

Therefore, with reference to Research Objectives 1 and 2, an approach is sought to pursue connections among multiple reservoirs and traps within a petroleum ecosystem, as chains of event attribute dimensions. The design of an integrated conceptualized framework can systematize compatible artefacts and the research outputs through the IS articulations. Further, Hadoop Distributed File System (HDFS) high-level tools are compatible with matching artefacts of DSIS for managing the unstructured data sources by Hadoop, MapReduce and Spark ML (Shvachko et al., 2010).

In Figure 2, we present various research activities “design, judge, infer and defend” and outcomes of the activities, such as, “artefacts, prototype, approach and validation (instantiation)” in the methodological framework. The articulations associated with research activities and outcomes are interconnected with artefacts linked with large-size data sources that make up Big Data guided DSIS for unconventional digital ecosystem solution development (Figure 2). As demonstrated in Figure 2, knowledge-based multidimensional data constructs and models are adaptable to the integrated framework with evolved geological and business rules (Indulska & Recker, 2008; Vaishnavi & Kuechler, 2007). The dimension models, involving 7Vs attributes (or Big Data characteristics), involve various schemas that can link multiple research activities and outcomes of proposed DSIS architecture. The products of the DSIS research framework represent knowledge-based UDPE contexts.

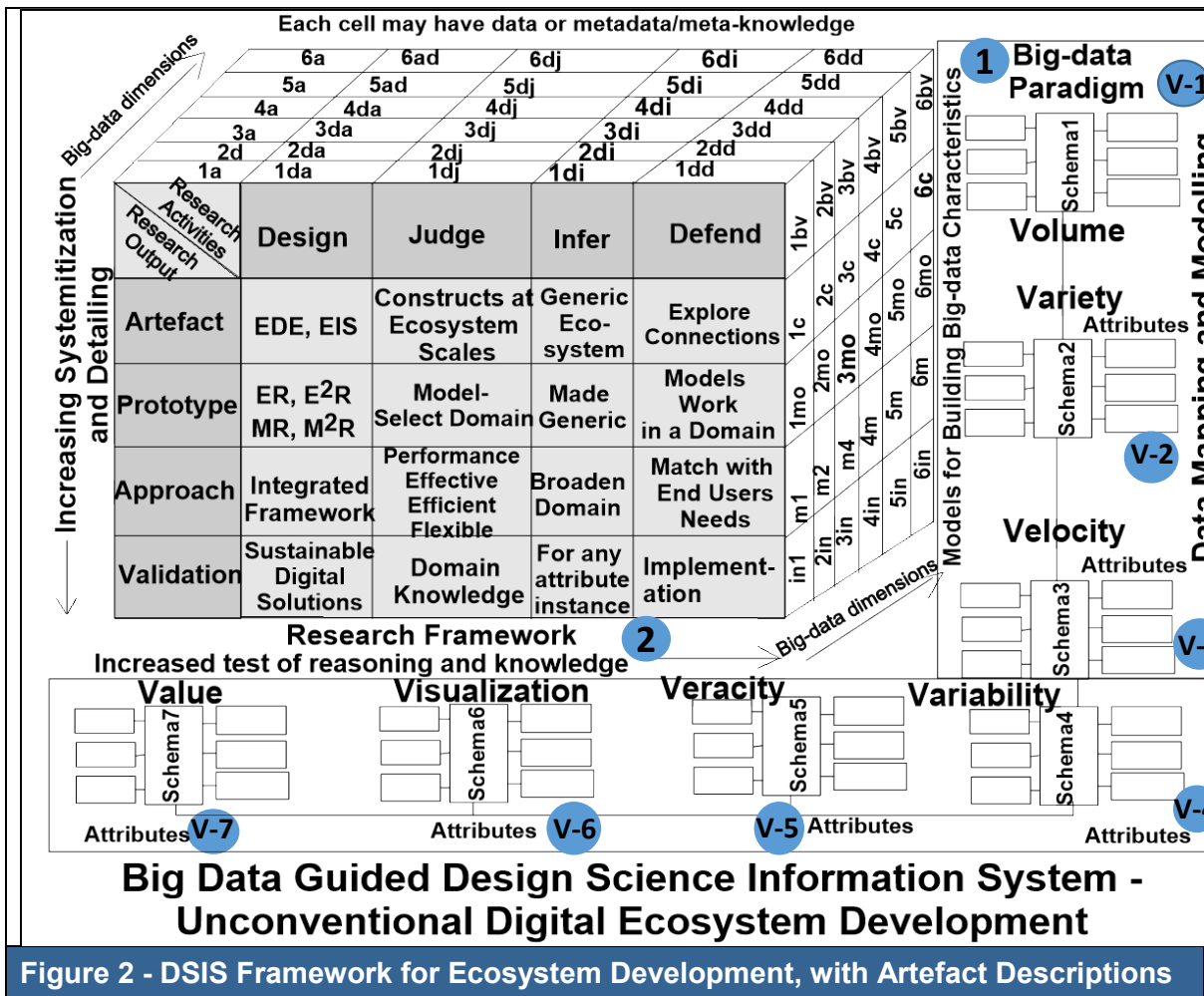


Figure 2 - DSIS Framework for Ecosystem Development, with Artefact Descriptions

The authors articulate research activities and the deliverable results, as construed in cuboid structure within a matrix form. Different attribute dimensions and their instances of Big Data represent various cells of a cuboid structure, as described in the DSIS framework. For example, the cells of the cuboid structure (Figure 2) stores data or metadata instance in Big Data scale interpreted from digital reservoir energy ecosystems that can construe meta-knowledge. Various schemas and sub-schemas facilitate the connectivity process (Figure 2), logically linking DSIS framework architecture's data structures. In the UDPE project, two major itemized stages (1 and 2) are evolved with the schemas associated with Big Data attributes, as represented in V-1 to V-7, respectively with volumes, variety, velocity, variability, veracity, visualization and value characteristics. The fine-grained multidimensional data structuring with added domain ontology descriptions can make the DSIS approach more robust in reservoir-energy management applications. The constructs and models need constant updates based on reservoir energy linked petroleum systems. Systems require adaptability in IS design considerations to accommodate flexible business rules that occur due to rapid geological changes.

We articulate the artefacts, as mentioned in the Design Science guided IS (DSIS) and construed in Figure 2. Chromiak and Stencil (2014) describe the heterogeneity challenges with missing features of the integration process that require new storage requirements, including unifying the large size data grids needed to build a framework. The authors propose new data structures and framework architectures with high-performance computing features using IS articulations. Data storage, data management, data maintenance, data integration and data interoperability are critical community development entities between systems. Shale gas and hydraulic fracturing have been the focus of such research efforts. The authors in

Middleton et al. (2017) have done data mining for 23 years of historical data covering 20000 production wells. Disruptive technology innovations have led to improved production from sick wells. It may be possible to increase shale gas recoveries by better fracturing technologies, minimizing environmental impacts due to carbon sequestrations. To design and develop the methodological framework in UDPE scenarios, we need a new research direction with emerging artefacts, as discussed in the following sections.

Petroleum ontology descriptions in Big Data assemblage

Big data analytics is an emerging trend in the upstream and downstream oil and gas businesses (Mohammadpoor & Torabi 2018). Seismic and micro-seismic data are integrated with drilled-well data and asset management attributes. As a part of the digital transformation of the oil and gas industry, we examine the data analytics features in Nguyen et al. (2020). Enhanced seismic data processing provides improved upstream business with a better understanding of data applications. Bai et al. (2018) describe the art of adopting data acquisition methodologies and storage through Hadoop systems to manage the data heterogeneity, particularly in industry 4.0 revolution. Masa et al. (2018) consider Big Data Analytics (BDA) one of the IT initiatives in the oil and gas industries. With BDA, the authors demonstrate improved performances and processes by 8-30% in oil and gas businesses. With the motivation of data-driven ecosystem, matured stages of BDA framework, delivery of quality information and processed data views can facilitate upstream business research outcomes. Systematic literature review of data modelling and management is a motivational factor of the research, described in Martinez-Mosquera (2020). Identification of data source, modelling and database construction is the focus of the study. As discussed in Zhonghua (2017), seismic attribute analysis has a role in seismic exploration technology, especially when it is tied up with Hadoop tools through which innovative data processing features can deliver quality attribute models.

In worldwide unconventional shale-gas basins multiple petroleum (information) systems exist with various elements, processes with volumes of their fact instances (Coronel et al., 2011). Relevant data come from seismic sensors and data arrays (both onshore, transition and offshore areas), where data volumes are presentable in various formats (Nimmagadda et al., 2019). The significant volume of data gathered relevant to the attributes are brought together as described in a schematic view (Figure 3), using digital portals in spatial-temporal dimensions. The process checks the quality of data at intermediate stages of data processing, ensuring IS artefact quality. We need to ascertain the accuracy, accountability and data reliability at every stage of data management and real-time processing. There may be inconsistencies affecting the adaptability of Big Data dimensions and their fact instances in IS articulations. The characteristic property “variability” emerges to mitigate the discrepancies, conflicts or ambiguities that may have arisen while integrating diverse data sources with new domain applications. When the data accumulated in large volumes are reconciled with qualities, they can be ready for modelling, data warehousing and mining. These artefacts facilitate the visualization and interpretation at later stages for new knowledge exploration and its management.

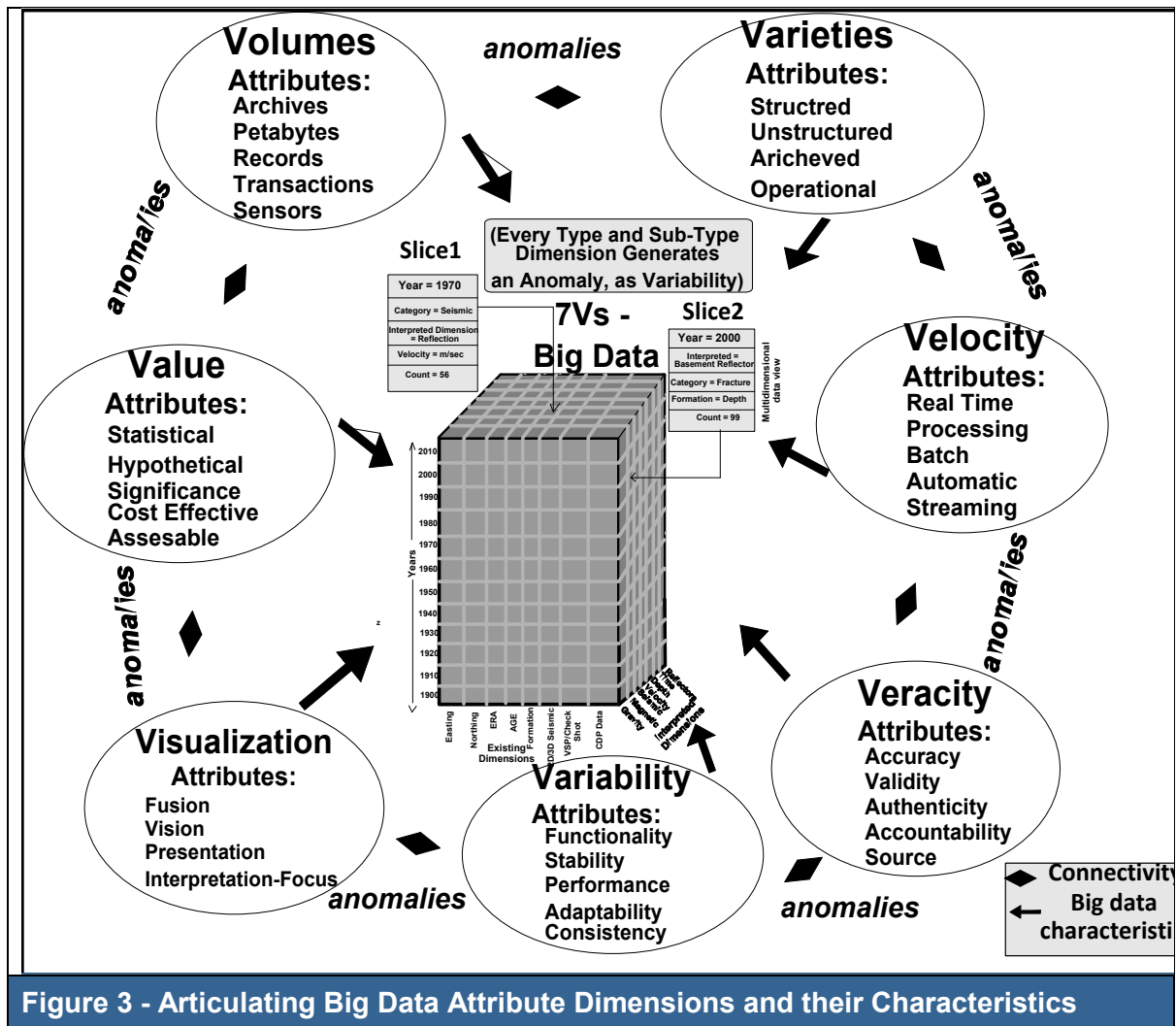


Figure 3 - Articulating Big Data Attribute Dimensions and their Characteristics

The growth of unstructured data in many applications motivates the authors to adopt the Big Data tools in digital energy ecosystem contexts. The data characteristics and their connectivity with UDPE are made possible through their dimension attributes and modelling. Each volume may have one or more varieties of data or types, each of which can take the workflow in a new direction. Accordingly, real-time processing and interpretation (including live streaming) can change. The volumes, varieties and data movements between project modules depend upon data veracity (Figure 3). The anomalies attributed to characteristics property contrasts of multiple data sources are connectable in a cuboid structure that can make the data mining and visualization effective at later stages. As illustrated in Figure 3, we interpret various Big Data dimension attributes. We consider they are relevant to the data sources of digital reservoir energy ecosystems. 7 Vs are commonly used to describe such Big Data, representing spatial-temporal controlled digital reservoir energy ecosystems. Among them, volumes and types of data are typical in the description of UDPE (Figure 3). In Big Data ventures, the petroleum bearing basins are characterized with volumes and varieties of data. The velocity and veracity attributes characterize project performance measures that can deliver quality research outcomes within project timeframes (Durham, 2013; Fosso Wamba et al., 2017). The variability is another characteristic attribute of the Big Data, contributing to interpretation in anomalies emerging between data characteristics (Figure 3). As suggested in (Fosso Wamba et al., 2017), various anomalies are interpreted as conceptualized attributes between Big Data characteristics. The anomalies may have been generated between different data characteristics in the digital ecosystems.

Table 2 - Big Data Sources in an Unconventional Digital Ecosystem							
Field	VO	Size	VA	DM	CS	C	V
Nimm1	15	90	20	90	15	1000	100
Shak1	12	72	10	55	13	1095	109
Kak1	17	54	8	25	8	990	85
Mak1	20	66	7	20	10	800	45

Field: Name of oil/gas field; VO: Number of Volumes; Size: gigabytes; VA: Varieties (Number); DM: Data Models (Number); CS: Composite Schemas (Number); C: Cells in Cubes; V: Number of Data Views.

Various data anomalies deviate from geophysical responses, obtained due to rocks hosting oil and gas deposits and their physical property contrasts (Sunjay et al., 2014). The visualization is a graphical artefact that can benefit the interpretation and add values to geological knowledge discovery. For example, instances associated with the Hydrocarbon Indicators (HCI) can attribute to seismic anomalies, providing development of porosity zones within unconventional petroleum provinces. As an example, we examine the unstructured data sources in various digital petroleum ecosystem scenarios (Table 2). The challenge is modelling and integrating real-time data events, and their knowledge management in spatial-temporal dimensions (Khatri & Ram 2004). The symbol \longleftrightarrow indicates predictable connectivity construed in between systems. Each system may have multiple links with spatial dimensions. Typical exploration data described in a unique digital energy ecosystem are shown in Figures 4a and 4b.

Oil and gas fields occupy large geographic regions and generate multidisciplinary data (Khatri & Ram, 2004; Li, 2011). Various domains and systems emerge with the increase in upstream businesses in spatial-temporal dimensions, as shown in Figure 4a.

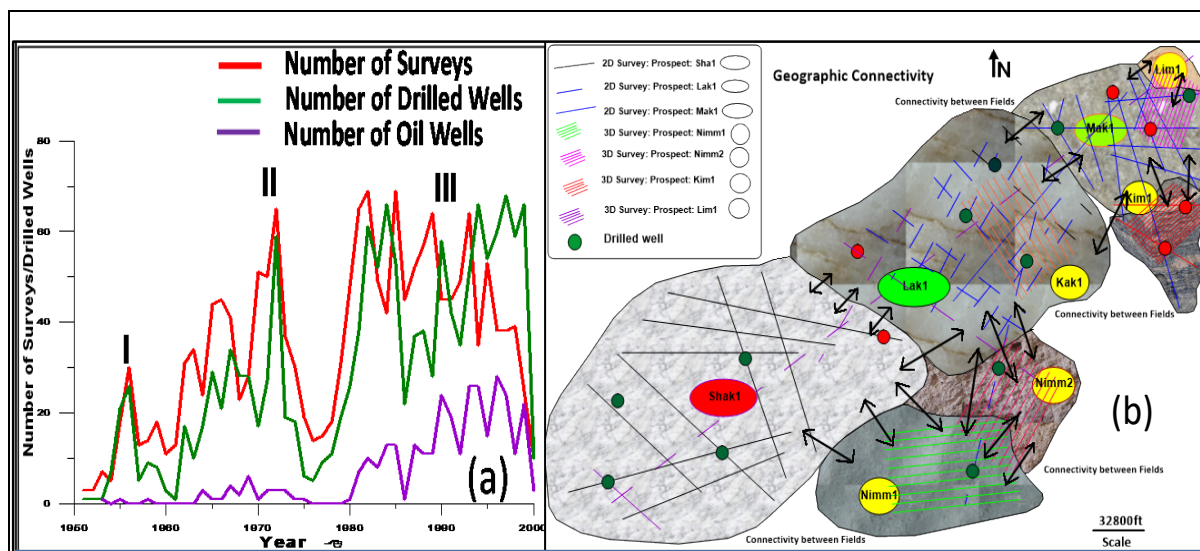


Figure 4 - (a) Typical Data Used in the Modelling Process (b) Conceptualization of an Ecosystem Representing the Unconventional Big Data Sources

The number of surveys, drilled wells, and oil wells are distinctive attribute dimensions of the current digital energy ecosystems. As indicated by I, II and III in Figure 4a, the three peaks describe attribute instance strengths at different periodic intervals, implying the contextualization of unconventional features in multiple models. In an investigating area, as shown in Figure 4b, there may be energy systems 1 – 5, Shak1, Nimm1, Nimm2, Kak1, Mak1 and Lak1. Each overlaps others with a large amount of spatial-temporal controlled seismic

data and drilled-well data. They are multiple entities and/ or dimensions; the attribute dimensions may be connectable through the modelling process, facilitating the data integration process in the investigating area. In addition, the exploration data (depicted in Table 2) are corroborated with respective production data instances.

The production data are part of the upstream businesses, in which the drilled- and producing-wells are connected to respective seismic survey profiles. We plot them in a schematic scalar-line plot-view (Figure 4a). Adding more dimensions to the Big Data helps minimize the ambiguity and uncertainty in exploration and production (E &P) portfolios of the oil and gas industry. As illustrated in Figure 4b, a digital ecosystem concept brings data attributes together in multiple domains and connects through different multidimensional schemas. Several data types are deduced in geology and reservoir-engineering domains to build knowledge-based IS artefacts (Gilbert et al., 2004). The volumes and varieties of data are considered to develop sub-schemas and schemas (Li, 2011). Standard schemas include star, snowflake and constellation types (Coronel et al., 2011). Based on varieties of Big Data, various schemas are categorized. An example of a schema is presented in Figure 5, describing different attribute and fact instances in a time-depth domain in well-log data. A particular schema chosen in the modelling depends on the grain size of the data structure needed for data mining. The composite schemas can significantly reduce the storage space, besides integrating models into the data-warehousing environment.

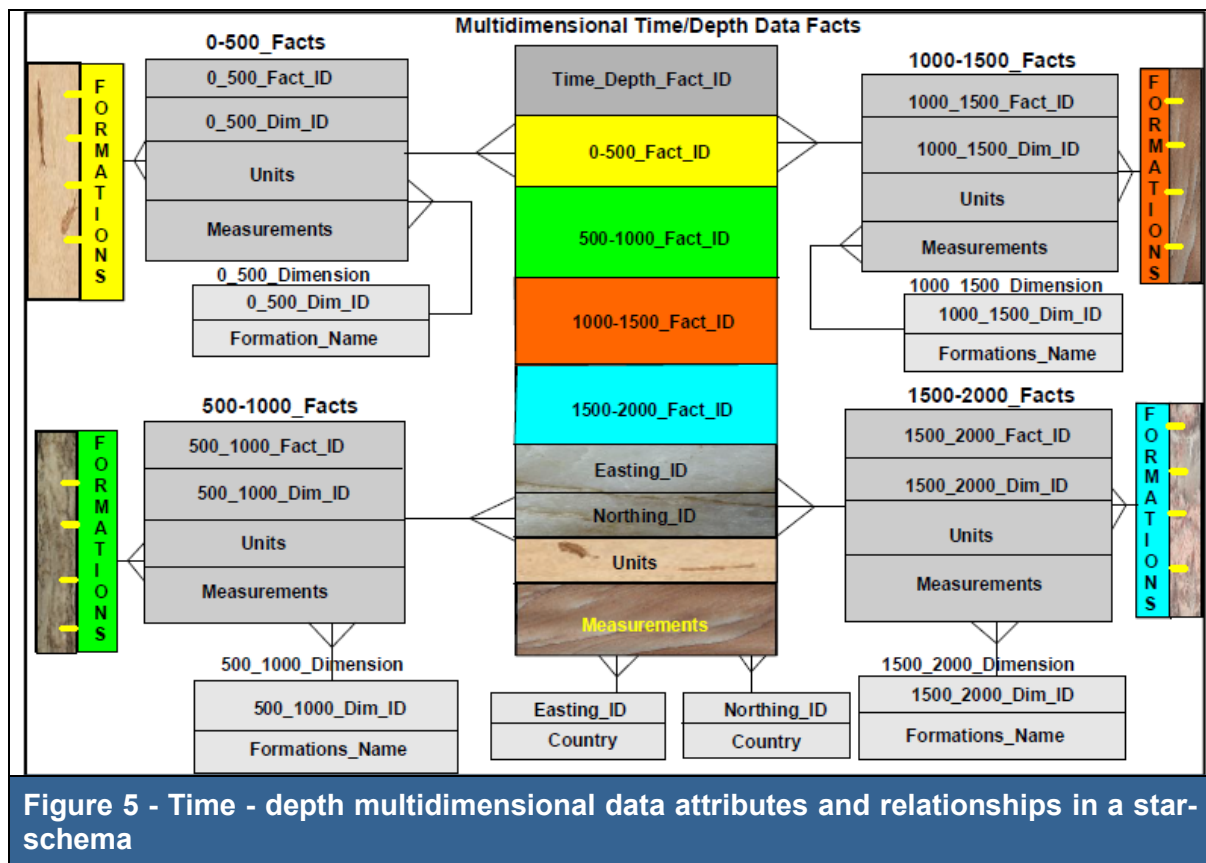


Figure 5 - Time - depth multidimensional data attributes and relationships in a star-schema

As per Research Objective 2, we analyze the data sources of ecosystems that exhibited heterogeneity characteristics. The amount of data and information considered in describing the data relationships in warehouse schemas (Table 2), simplifies large-size ontological structures and their classes. The seismic and drilled-well data systems with hundreds of attribute dimensions and volumes of fact instances are part of the Big Data, accommodated in various time-depth intervals in spatial domains (Figure 5). The factual instances at different depth intervals 0-500, 500-1000, 1000-1500, 1500-2000 represent data models in various

The seismic and borehole data attributes and their instances are typically represented in spatial-temporal dimensions for unification. Schematically, ontologies characterized in the form of schemas that interpret diverse attribute relationships in multiple domains and systems, are termed as IS artefacts. We use data relationships as dimensions in dimensional models, as shown in Figures 5, 6 and 7. We have drawn them in the form of star-schemas, but with an ontology focus, narrating how data relationships take the schema designs and drive the development and implementation stages.

We further examine the framework with a couple of schemas prepared for an ecosystem (sedimentary basin). The ecosystem attribute dimensions connect multiple petroleum ecosystems and their associated reservoir energy systems (Figures 5 and 6). Facts are instances of the attribute dimensions deduced from volumes and varieties of geology and geophysics, including oil and gas exploration and production entities (Li, 2011). It can be productive when building data relationships between similar and dissimilar dimensions to describe their connections (Figure 6). Petroleum ontologies remove the potential conflicts and constraints arising during conceptualization and interpretation of terms in various data sources to provide a structured vocabulary that can integrate all data and knowledge into unified metadata (Pujari, 2001). Petroleum ontologies additionally specify different topological relationships among structure, and reservoir/source to illustrate how mandatory semantic constraints are incorporated in schemas.

As presented in Figure 6, the snowflake schema model connects various attributes of the UDPE, including their semantic relationships. As described in Figure 6, the appraisal asset attributes are used over interconnecting dimensions through a snowflake schema with several data relationships interpreted and documented in 3D seismic fact tables, collaborated with the asset appraisals. In an unconventional reservoir system, the source does not have a necessary limitation with trap or structure but depends on fracture-reservoir (present within the source rock itself). The elements and process of energy source must topologically touch one or more instances of fractured reservoir attributes to retain the hydrocarbon content (Castañeda et al., 2012). Production of unconventional oil or gas field must surround at least one fracture-reservoir or frack-connectivity among dense fractures. Unlike conventional reservoirs, they have a mandatory relationship and connection with a trap in a producing field, described in (anonymous1) and formulated in the following sections.

The conceptual models are further standardized for internal representation of petroleum ontologies, based on Eapen (2008). The UDPE precisely describes a series of events in the contexts of petroleum exploration that can be used to interpret the reservoirs, or drillable fracks in any unconventional petroleum province. The petroleum ontologies in the UDPE contexts describe how:

1. As an unconventional reservoir, it is connectable to seismic sequences and structural information.
2. It is mitigating the structure, reservoir and source elements in the integration process.
3. It exhibits internal and external data associations within unconventional reservoir fracks.
4. The interpretation of constraints affect the petroleum trap/s with seal confirmation.

In a knowledge-based repository system, hundreds of logical multidimensional data relationships emerge while analyzing digital energy ecosystems and connecting them within the integrated framework. One of the schemas (Figure 6) demonstrates the connectivity between various attribute dimensions. E and P scenarios and their volumes of data instances emerge to instantiate the DSIS framework's artefacts and assess them in UDPE contexts. In the data schemas, there are multiple dimensions narrated and interpreted conceptually with various relationships. The associations link the fact tables in one-to-one, one-to-many (as indicated in a symbol \rightarrow) and many-to-many data relationship types (Coronel et al., 2011). As illustrated in a dimensional model in Figure 6, multiple fact tables are connected through

common attribute dimensions. For storage and flexibility of attribute dimensions, the snowflake schemas are made compatible during modelling. Though it is easier to implement the schemas, the computational performance may be adversely affected. However, we limit the number of schemas that may be optimum enough to achieve connectivity, including their use and reuse extendable in diverse domains. The process prevails with varieties of data existing in reservoir-energy data volumes (Nimmagadda et al., 2019).

Integration of unconventional petroleum ontologies

Tools and concepts used to design and develop data-warehouses in different domains are given in Inmon (2005); Lee et al. (2006). We have added a couple of more schemas, relevant to resources development for exploration, appraisal, drilling, completion and production entities. The appraisal is needed to understand the extent, including the size of oil and gas accumulation. The evaluation is relevant to spatial-temporal dimensions, with which the coal seams and shale-gas reservoirs are managed geographically. They correspond with the composite schemas presented in Figures 4, 6 and 7 that support the data integration process. As per the Research Objectives 1-3, the need has further arisen to integrate multiple petroleum systems and their sources with large-size ontologies, as described in Eapen (2008). The size refers to the number of systems collaborated within the schema.

Table 3 - Attributes of Entities of Unconventional Resources Development in the Upstream Business				
Exploration Facts	Appraisal Facts	Drilling Facts	Production Facts	Completion Facts
Basin ID	Shale Reservoir ID	Bottom Hole Assembly (BHA)	Production Analysis	Completion Design
Exploration Type	Heterogeneity	Drilling Fluid	Candidate Recognition	3D stimulation
Source Maturity	Reservoir Quality	Well Placement	Refrack	Fluid Rock Interact
Seismic_ID	Completion Quality	Mud Logging	Production Frack	Micro-seismic
Permeability_ID	Optimization	Depth	Water Management	Model Status
Total Organic Carbon (TOC)	Life Cycle	Target	Environment Management	Source effectiveness
Connecting Facts	Connecting Facts	Connecting Facts	Connecting Facts	Connecting Facts
Appraisal Fact ID, Drilling Fact ID, Production Fact ID and Completion ID	Exploration Fact ID, Drilling Fact ID, Production Fact ID and Completion ID	Appraisal Fact ID, Exploration Fact ID, Production Fact ID and Completion ID	Appraisal Fact ID, Drilling Fact ID, Exploration Fact ID and Completion ID	Appraisal Fact ID, Drilling Fact ID, Production Fact ID and Exploration ID

Table 4 - Asset appraisal Attributes in the Mid- and Down-stream Businesses		
Attribute Dimensions	Fact ID	Connectivity Attributes
Gas Content	GC_Fact_ID (F1)	GasCom_Fact_ID
Rock Eval Pyro	Rep_Fact_ID (F2)	TOC_Fact_ID
TOC	TOC_Fact_ID (F3)	Rep_Fact_ID
Gas Composition (F4)	GasComp_Fact_ID (F4)	GC_Fact_ID
Core Description (F5)	CoreDes_Fact_ID	SpecialLog_Fact_ID
Sorption Isotherm (F6)	SorpiSoth_Fact_ID	GasComp_Fact_ID
Proximal Analytics (F7)	ProxyAnaly_Fact_ID	MacerAnaly_Fact_ID
Mineral Analytics (F8)	MineralAnaly_Fact_ID	GasComp_Fact_ID
Vitrenite Reflection (F9)	VitriniRefect_Fact_ID	TOC_Fact_ID
Calorific Value (F10)	CaloriVal_Fact_ID	GasComp_Fact_ID
Maceral Analytics (F11)	MacerAnaly_Fact_ID	MineralAnaly_Fact_ID
Bulk Density (F12)	BulkDens_Fact_ID	MineralAnaly_Fact_ID
Conventional Log (F13)	ConvenLogs_Fact_ID	SpecialLogs_Fact_ID
Special Log (F14)	SpecialLogs_Fact_ID	ConvenLogs_Fact_ID
Pressure Transition (F15)	PresTrans_Fact_ID	3DSeismic_Fact_ID
3D Seismics (F16)	3DSeismic_Fact_ID	ConvenLogs_Fact_ID

The attributes involved in the integration process are shown in Figure 7a in a dimensional model. Silos and their connectivity, surrounding the integration process are relevant to entities, dimensions and fact instances, as presented in Tables 3 and 4. The dimensions are connectable through data schemas (Figures 7a and 7b). It ensures the unification of multiple systems into a single repository system and develops an effective knowledge-based metadata structure.

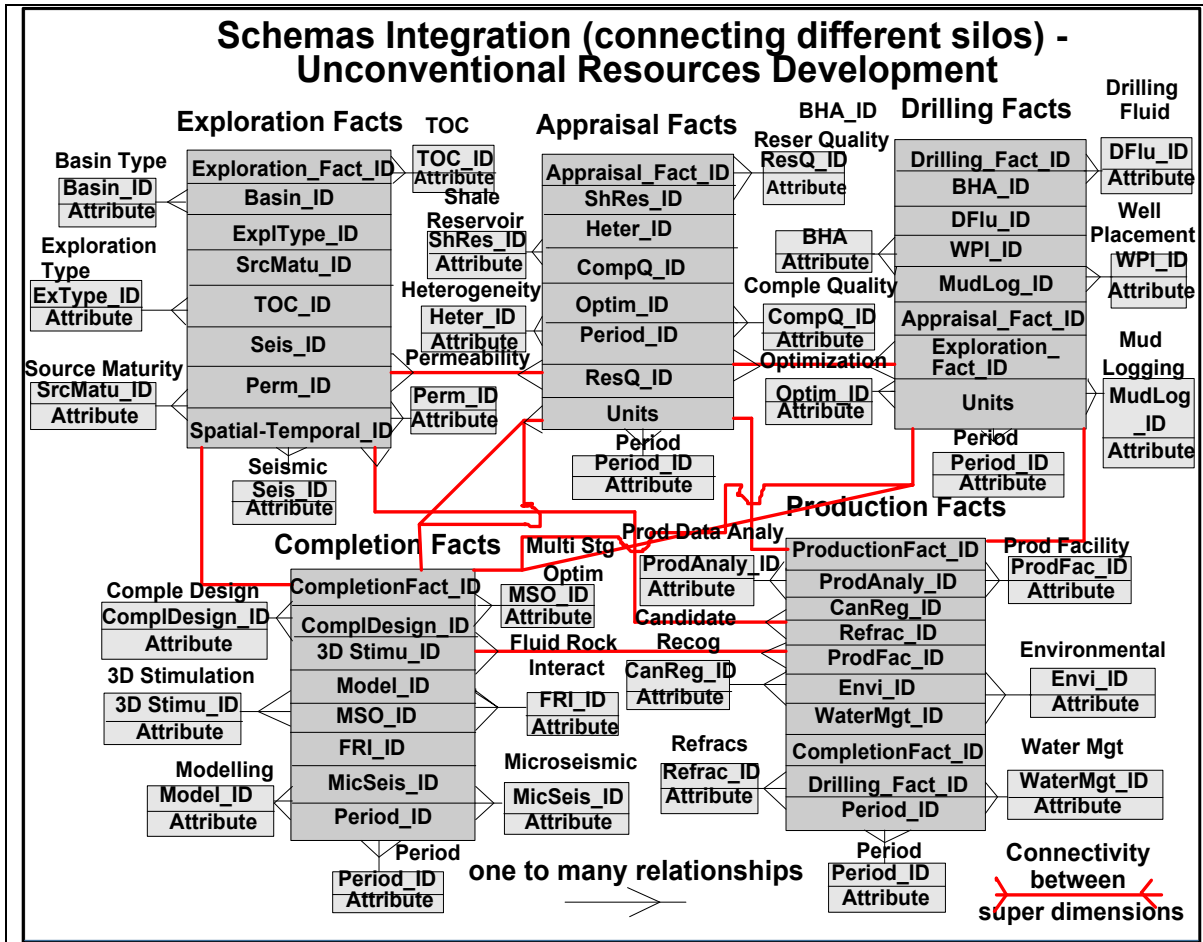
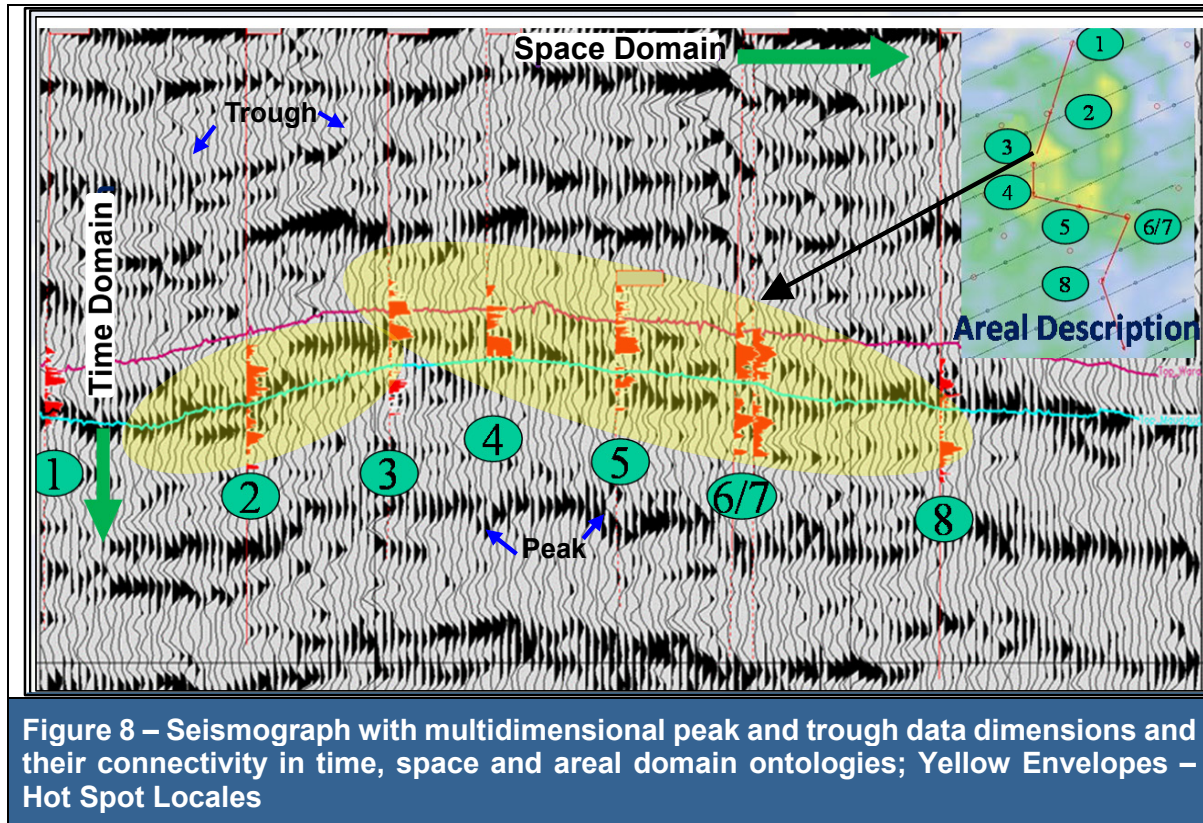


Figure 7a - Multiple schemas connected from variety of Unconventional Resources Development

For connecting different attribute dimensions at appraisal stage (Figure 7b), we design multiple fact tables to strategize the reservoir-energy assets in the appraisal phase with schemas F1-F16 with respective fact tables of attribute instances (Table 4). Each fact table is interpreted with a relevant dimension table, as demonstrated (Figure 7b). Multiple star-schemas are designed to integrate all asset appraisal attributes.



A sedimentary basin characterizes with various entities; and their attribute data instances from seismic, drilled-well, petrophysical and production entities are connectable in repositories through DSIS articulations. The DSIS framework described in Figure 9 allows us to generate the necessary logical data models to interconnect various domain ontologies and structures. The interconnection enables processed and knowledge-based information to be shared among professionals and distributed geographically varying digital clouds (Briscoe & Marinos, 2009). The silos may be from multiple ecosystems that capture multidimensional logical data structures. As shown in Figure 9, the Big Data-guided IS articulation in the UDPE setting, is formulated for implementing research outcomes to validate by various evaluation properties (Venable et al., 2016). The artefacts are building blocks of the overall UDPE setting with which the IS articulations are made adaptable.

Besides, the integrated articulations unify different petroleum systems that communicate and interact with each other in real-time within the framework. The IS artefact designers support the ecosystem concepts and define the scope, depth, comparability, and accuracy of data entering the warehouse articulations (Figure 9). The data range describes types of petroleum systems and their linked varied geological, geophysical and geochemical data in multiple periods (time-dimension), and geographic locations (spatial-dimension). The depth of data refers to the level of details needed in the modelling. For flexing data structures, data are in similar and dissimilar attribute dimensions; separate geographic locations can either use or reuse the same classifications. It is paramount to integrate data schemas in multiple geographic regions to make up multidimensional repositories, no matter how different data are collected across such sites, as illustrated in Figure 9. As per Research Objective 2, for easing complexity and security, petroleum system analysts and geo-modellers must use compatible software systems to map the grids and model data consistent with repository structures (Figure 6). Obtaining required accuracy in modelling is however challenging for types of data in any given geological situation.

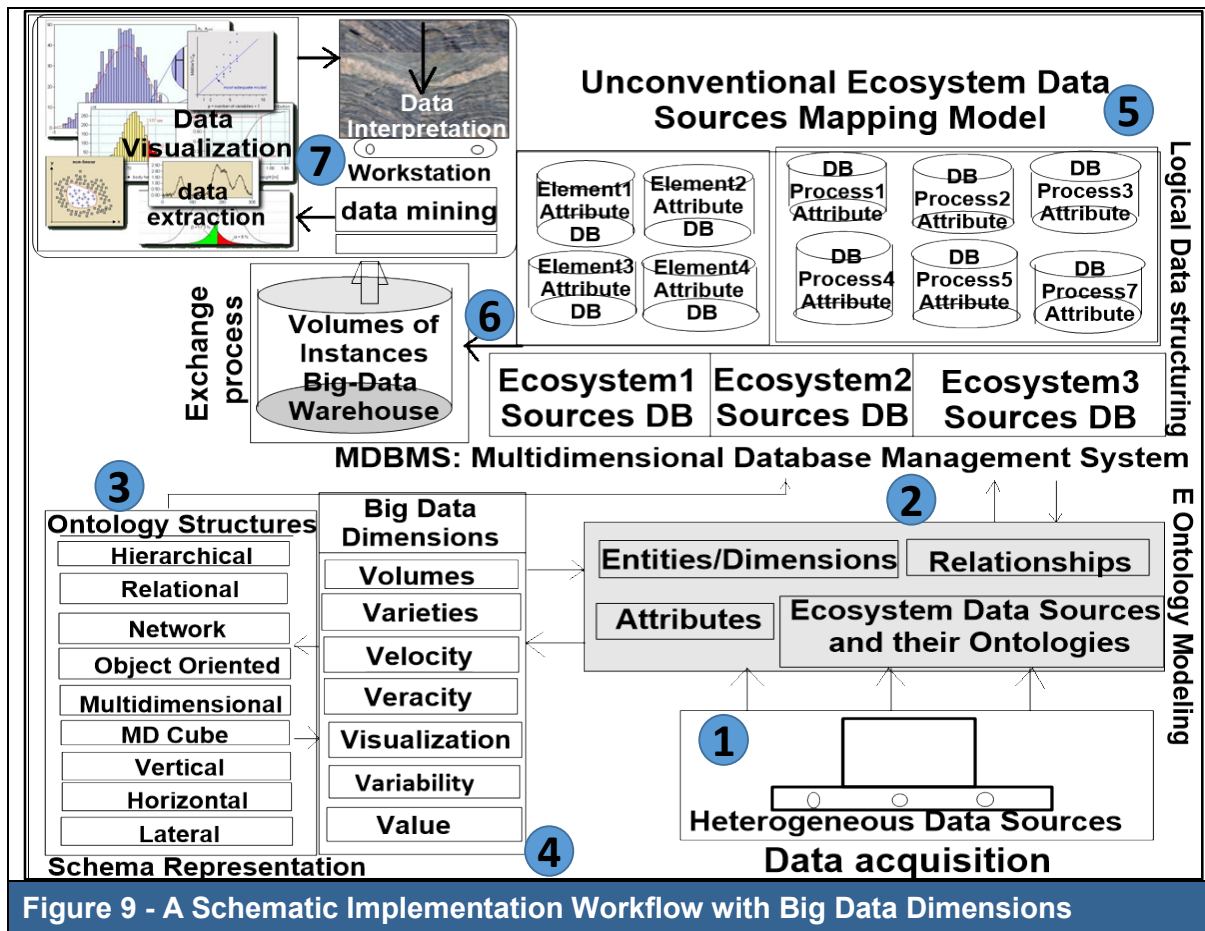


Figure 9 - A Schematic Implementation Workflow with Big Data Dimensions

As shown in Figure 9, the implementation workflow is articulated using different artefacts, such as DSIS framework architecture, connectivity process, multidimensional star and snowflake schemas, including the necessary seismograph, as described in Figures 1-8 Tables 2-5. Various phases involved in deducing the UDPE metadata are shown in the implementation workflow. The phases start from data acquisition (1) to data mining, visualization and interpretation (7). Other phases include modelling (2), types of data structure methods (3) needed with the interpretation of data dimensions (4), with mapping process, as practiced in the phase (5). In phase (6), we warehouse the large-size data instances of UDPE with subsequent data mining, visualization and interpretation artefacts, as shown in phase (7). Implementation of reservoir energy solutions and their management are discussed in the following sections.

Implementation of the DSIS Framework and Knowledge Management

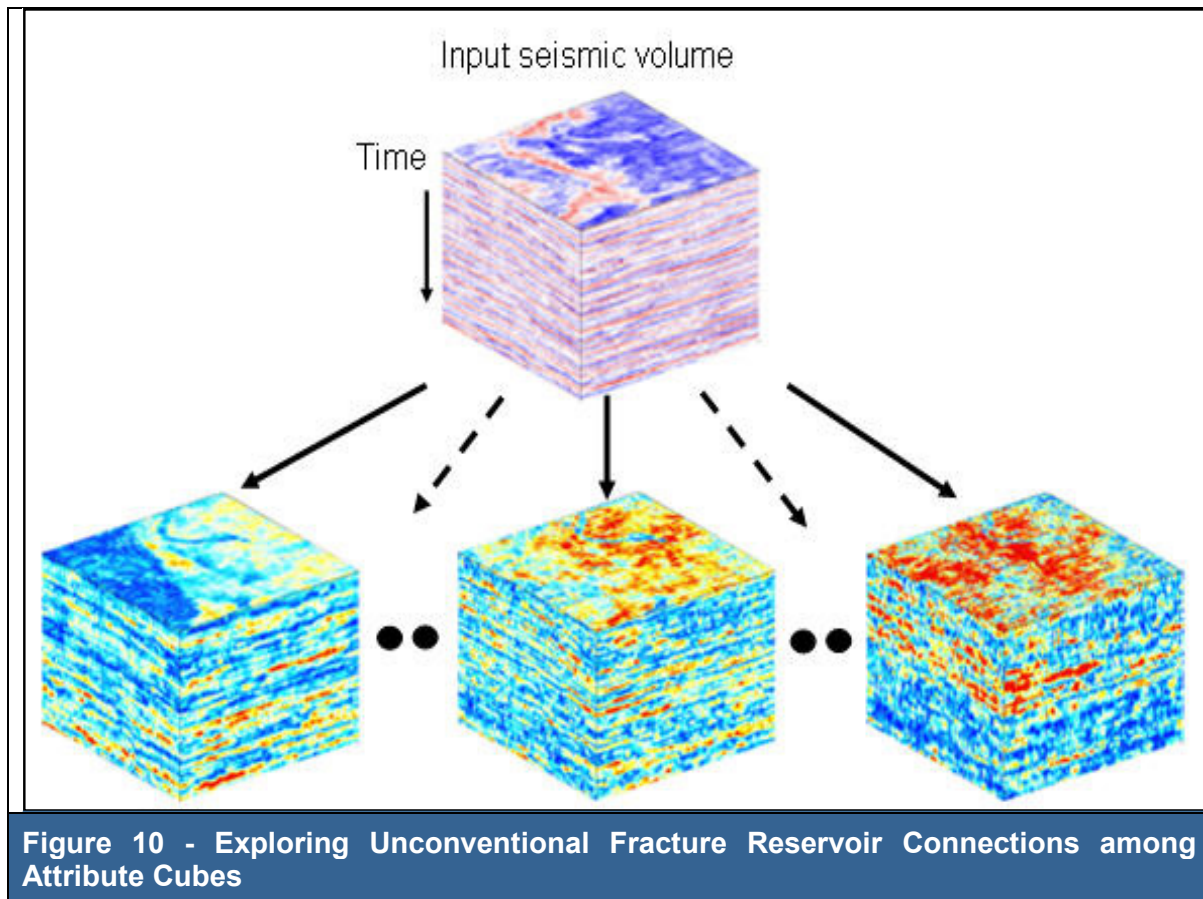
Given the context of UDPE, a DSIS framework is suitable. Offermann et al. (2009) discuss knowledge-based information systems, research domains with relevant perspectives and methods. The combined research is based on design-science. The authors develop a formalization process that combines both qualitative and quantitative research methods, including case studies. Knowledge sharing is the key strategy with the purpose of its promotion in the oil and gas industry. Grant (2013) identifies two major types of knowledge management practices in information communications technology and person-to-person knowledge management, all facilitating knowledge management in the oil and gas industries. The authors in Li-Ying et al. (2008) focus on interpretation aspects of exploration with a quest to minimize the ambiguities that may arise during the exploitation of drillable exploratory drilling campaigns. A theoretical framework is deduced to integrate with a novel typology of exploration and exploitation entities. Ramanigopal (2013) describes the importance and effectiveness of knowledge-driven entities in modern business environments. Ceptureanu et al. (2018) examine the issues and challenges affecting the oil and gas industry and its knowledge management (KM) processes. Conceptual models made from various entities, including data storage systems, have practical industrial implications. The implementation of KM processes allows the establishment of models in the oil and gas industries. Gregor and Jones (2007) describe the design theory and its research components with instantiation and intervention of IS products and applications. The article provides a sound knowledge-based with further rigor and legitimacy of IS articulations as an ongoing scope and opportunity of design theory in diverse IS applications. Valdez-de-Leon (2019) construe linear value chains that emerge with vertically integrated organizations through digital technologies and their transformations. Value chains are interpretable in digital ecosystems. The knowledge and strength of ecosystems are improved through value creation from digital giants and organizations. From these studies, a practicable framework is made with new paradigms of digital ecosystems. Bento (2018) analyzes the oil and gas industry complexity with manageable integrated operations (IO). Such operations can facilitate knowledge management in spatial-temporal dimensions, minimizing the uncertainty in exploration risks. Concepts of data organizations are interpreted in Jia et al. (2016) to accommodate new models of systems, including the Total Petroleum Systems (TPC).

Coalbed Methane (CBM) gas is extractable because of the existence of unique attributes of coal seams and their matrixes from which gas is released through cleats or fracture systems to the wellbores. New technologies and cost-effective methods enable us to explore more production from CBM assets, largely due to IS applications. Without risking the environment, faster returns on investments through reserves can be recoverable from organically rich gas shales. Understanding the heterogeneity of fractured shale reservoir geometries has enabled us to detail the unconventional reservoir energies in green environments. As a part of Research Objective 3, several logical rules and constraints applicable for scaling UDPE models and multiple levels of information stored in ontology models are incorporated in the warehouse repository schemas. IS artefacts designed and developed as in Figures 5, 6 and 7 are the focus of energy domain application. DSR theory, developed in Figures 1 and 2, is implemented in the upstream business research to motivate the down-and mid-stream activities of oil and gas industry. The artefact is one of the components of the DSIS design and development. It is articulated from generic design science research methodology. Initial conceptualization is from artefact design, which is contextualized based on domain application. The IT artefact's contribution lies with how effectively the DSIS is implemented to interpret the data views of metadata in terms of drillable hot-spot locales in the investigating areas. For queries, information retrieval and presentation, we consolidate metadata of outcomes IS framework and validate different attribute data cubes as illustrated in the following sections.

Validation of the DSIS Approach in the UDPE Contexts

For implementing DSIS articulations, the data views retrieved from metadata cubes of UDPE are analyzed for data trends relevant to reservoir connections. Data visualization and interpretation artefacts are vital to inspect through attribute analysis with geological knowledge in ecosystem contexts. Largely, the DSIS has a decisive role in strategizing the exploration and field development in particular in digital reservoir solutions, making huge impacts in the integrated interpretation projects, especially during prospect identification and risk evaluation stages. The data interpretation and knowledge discovery are the final stages of reservoir management project, facilitating the framework's implementation (Research Objective 3). The analysis of reservoir distribution in different knowledge domains of UDPE and basin scenarios must support and sustain the DSIS articulations.

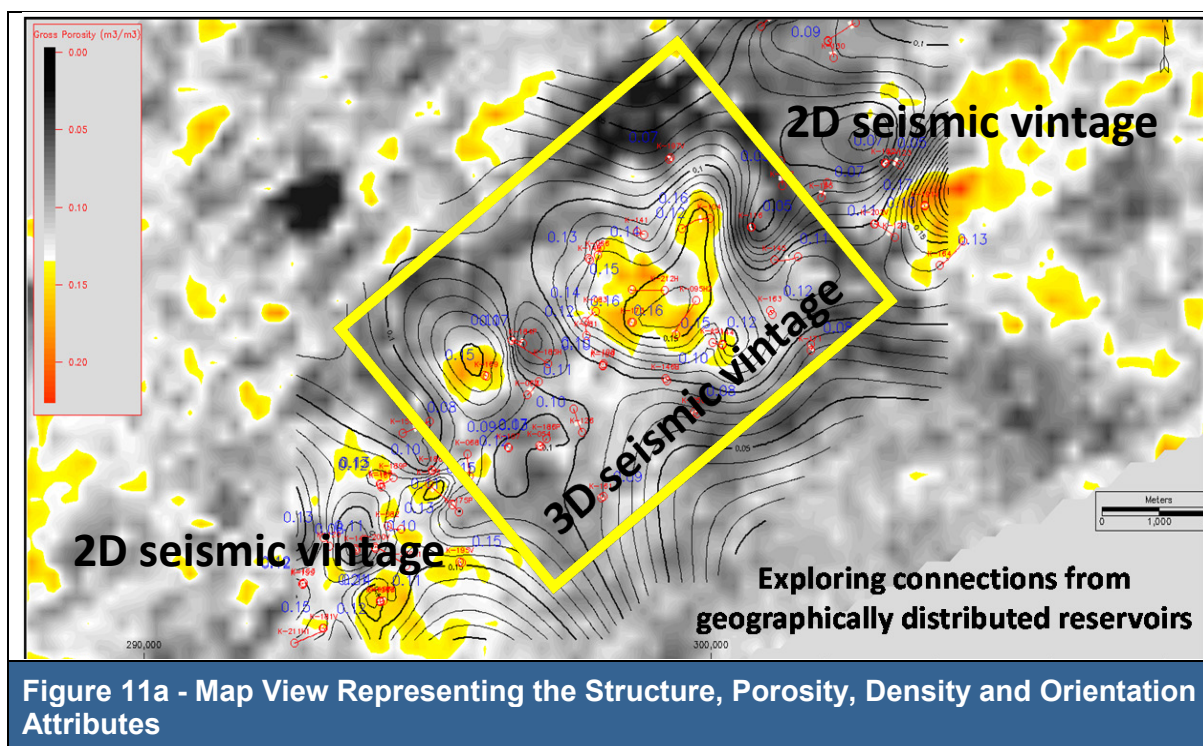
The multiple reservoir connections and their areal extents are interpreted with seismic attributes at volume and surface levels (Castañeda et al., 2012). Interpretation of different map and plot views supports the frackable-shale reservoir connections, explicitly perceivable at field and basin levels.



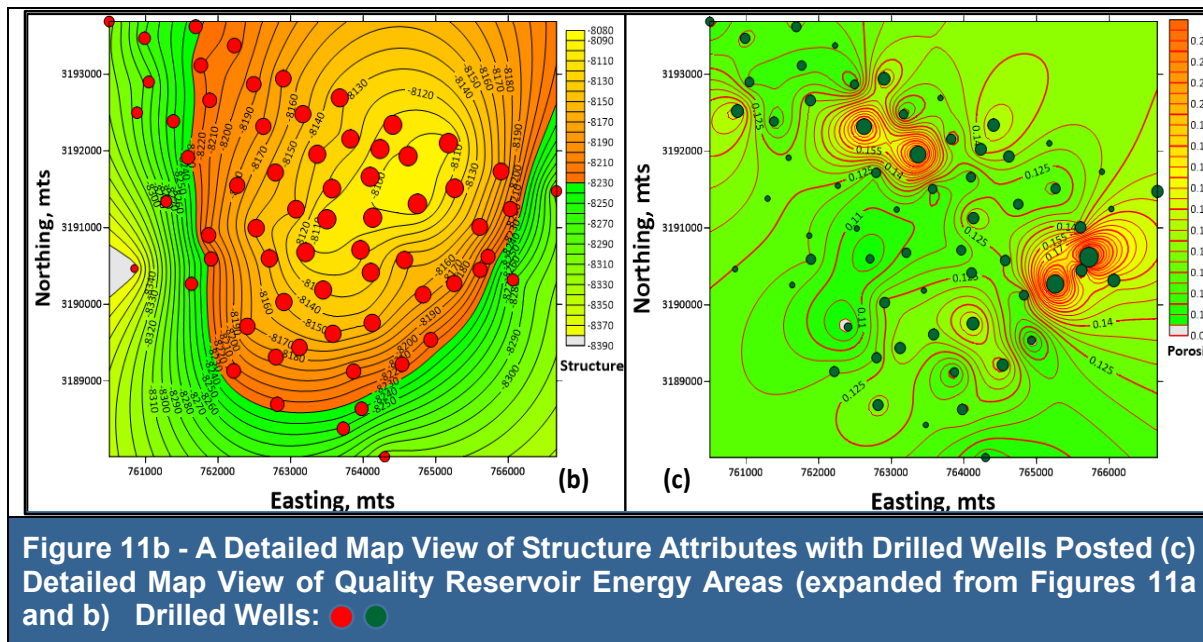
As shown in Figure 10, the computed seismic attributes and their related data views are evaluable by presenting the reservoir cubes in space and depth dimensions. Pujari (2001); Matsuzawa and Fukuda (2000) describe the volume and surface attribute cubes, used for mining slices and dices to interpret and evaluate new geological features, visually attributable to UDPE in the study area. The study ascertains the reservoir potentiality, including visualization of quality reservoirs with their big knowledge.

Ecosystem Visual Analytics and Big-Knowledge

For decision support systems in upstream businesses, knowledge management is crucial to optimize the financial resources and risk of interpreting and locating drillable prospects. Data transformation tools, such as statistical regressions and, bubble plot graphical representations, can translate the Big Data into big-knowledge. The data instances interpreted for structure, reservoir, source and seal elements (Figure 2) are documented in several classifications. Grouping of elements and processes is construed to interpret each reservoir energy system in an ecosystem category. For example, the geological structure and its attitude attribute instances have facilitated us in presenting their ontological descriptions to make reservoir connections and their locales in the study area. Similarly, the specific reservoir attribute instances documented for an interpreted horizon (composed of fracks) in a basin are structured (in data structuring sense) through reservoir ontology descriptions (Shanks et al., 2003; Nimmagadda et al., 2018). For in-depth knowledge of energy ecosystems, we emphasize fine-grain structuring for effective data mining. Slicing and dicing are further performed on metadata for tracking knowledge-based attribute dimensions, interpreted in shale basin contexts. The strength of element and process attributes of reservoir energy systems and their connectivity are assessed through high-resolution map views. Frackable reservoirs interpreted in structural compartments, establish reservoir connections with associativity between fracking attributes and their instances (Figure 11a).



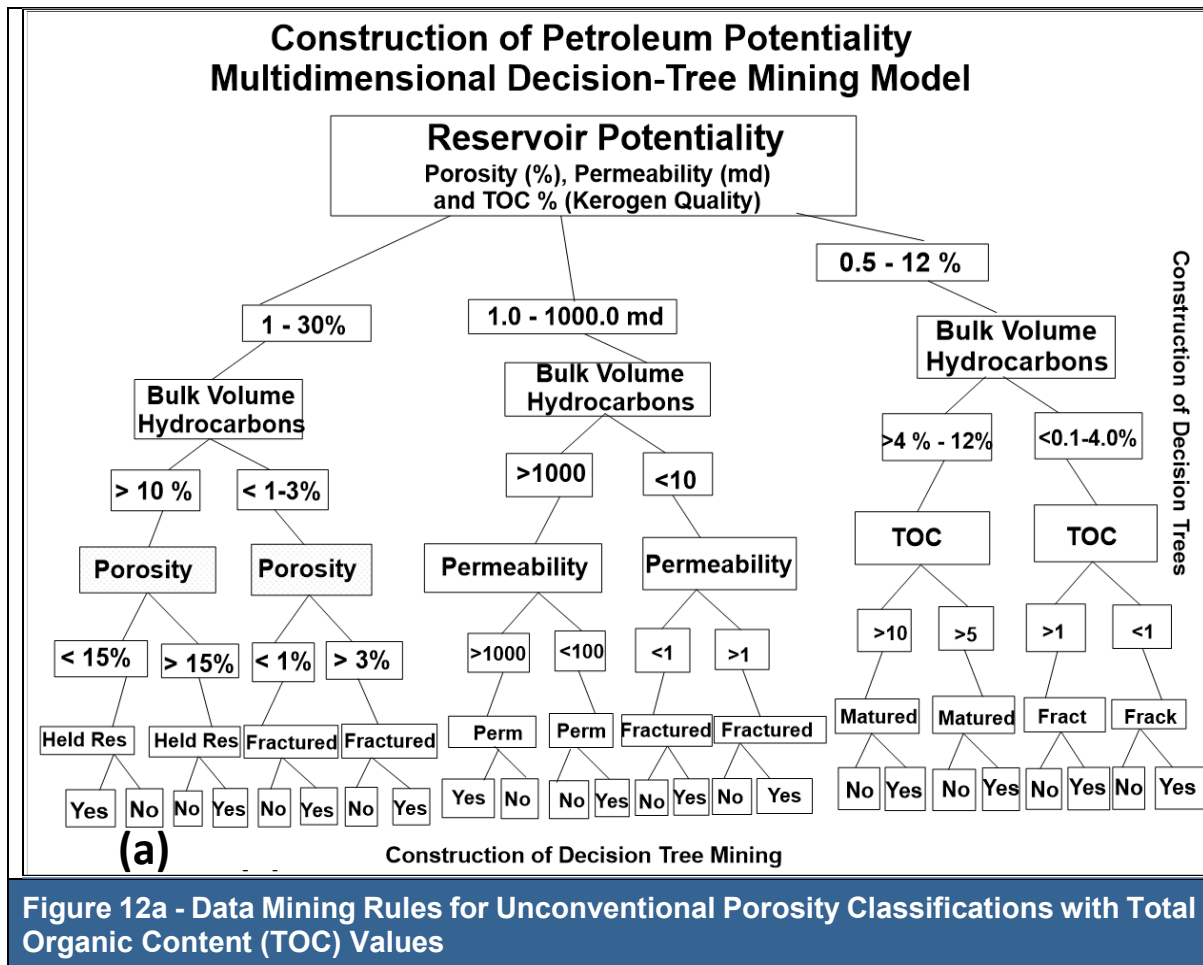
As shown in Figure 11b, the bubble plot view displays multiple variables, as attribute dimensions of reservoirs on a scattered plot. The diameter of each bubble varies in size, providing ways to present additional dimensions of data with a quest to offer new insights of reservoir information from map views. Multiple reservoir connections are interpreted that can facilitate the energy explorers to plan for resources development.



Bubble plot views exhibit the strength of two or more attribute variables in scalar plots. As shown in Figures 11b and 11c, each bubble or extents of bubble clusters can vary in diameter, providing a way to represent additional dimensions in the data representations. For example, bubbles indicate the porosity attribute areas with varying “easting” and “northing” attribute dimensions, representing each bubble size and signifying density and magnitude of structure and porosity attributes using rule models.

Rule-Based Decision-tree Mining Model

A decision tree mining is a classification scheme (Pujari, 2001) that generates a tree structure with a set of mining rules or business constraints, representing models in different classes in a given dataset. We implement the constructs and models, with an evaluable Microsoft decision tree mining models and, Grapher visualization tools to explore and analyze the information from homogeneous branches of a tree. We find mining rules, with conditional control and constraint features. Several numerical and categorical attributes are deduced to evaluate the fracks (fractured reservoirs) based on porosities and kerogen content in unconventional reservoir plays. Hydrocarbon plays and non-plays, each has common categorical attributes in the data cubes. Several leaf nodes are characterized in the decision-trees; each leaf node of the tree represents a mining rule, as described in the following sections.



The decision tree rules deduced from a model in Figure 12a are:

1. Rule1: If the shale has more than 5% porosity, it is a *play*.
2. Rule 2: If the TOC is less than 5%, the shale has a play.
3. Rule 3: If the TOC and porosity are each less than 5% and 5%, respectively, and the permeability is less than 0.1%, the shale-play has a relatively poor fractured reservoir
4. The Kerogen type III is more than 3%, with less than 5% porosity, permeability less than 0.1% and TOC less than 5%, all appear more favorable.
5. Rule: 4: Based on Rule 1, Rule 2 and Rule 4, if the shales are good fractured reservoirs, Rule 4 holds good.
6. Rule: 5: If the attributes are not favorable as narrated in Rule 3, then the shale play does not hold good.

The exploration and production data available in the public domain are acceptable to test the classifier accuracy (Li, 2011). Accuracies of rules are calculated from training test datasets. Rule 4 emerges 90% accurate, compared with the accuracy of the other rules. TOC and permeability are other weighing attributes supporting Rule 4. Different multidimensional shale reservoirs (reservoir plays from various fields) and their properties have further been analyzed. Each attribute has unique shale reservoir properties.

Despite differences in attribute strengths, as demonstrated in Figures 11b and 11c, the durability of individual characteristics matches well with specific reservoir-play dimensions. It is a schematic view drawn for a frack model, aiding the decision support system, which type of fracking and kerogen content best classify and suit making a valuable financial deal to tap energy from unconventional reservoir energy systems. As described in Figure 12b, we arrive

at a polynomial regression model to ascertain a non-relationship between dependent and independent variable attributes, such as associativity between the depth (ft.) and production index. Despite outliers, the regression trend matches well with the corroboration between the depth and production index.

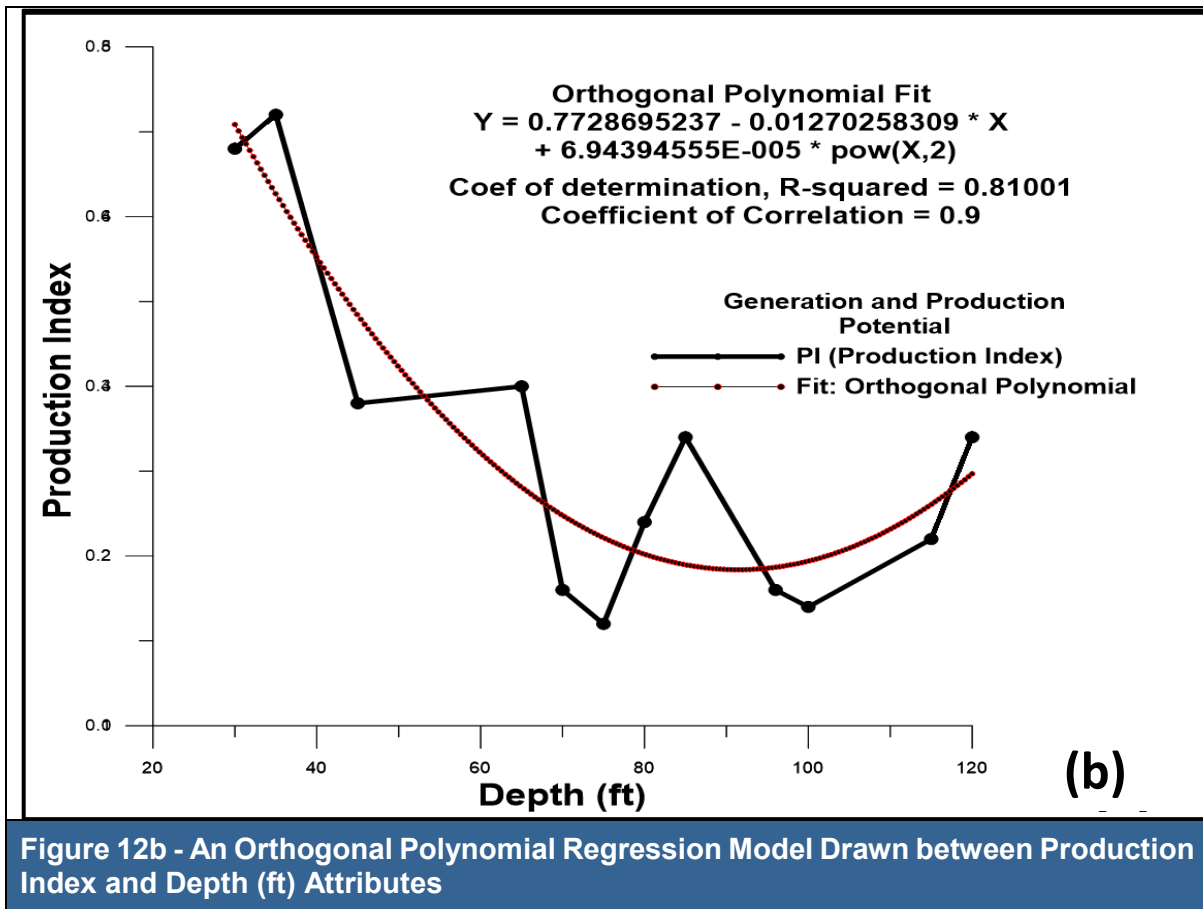


Figure 12b - An Orthogonal Polynomial Regression Model Drawn between Production Index and Depth (ft) Attributes

The Visualization and Value attained from Big Data Analytics

The metadata models are made more knowledge-based with meaningful geological information that can add value to the reservoir energy project by presenting exploration products for interpretation and new knowledge discovery. The potentiality of digital energy ecosystems is appraised assimilating the value of big knowledge of reservoir prospectivity. As described in Table 5, the hydrocarbon potential is assessed in shale bearing digital petroleum ecosystems. Production Index (PI) assesses the unconventional ecosystem's maturity that can produce commercial and viable oil and gas deposits.

Table 5 - Potentiality of Unconventional Digital Petroleum Ecosystem

Field	T	K	PO	PE	SW	PS	T	PI
Nimm	1	150	1.5	150	70	500	45	0.18
Shak1	2	300	2	245	60	400	35	0.20
Kak1	3	550	0.5	320	75	600	60	0.35
Mak1	5	600	1.25	455	80	700	55	0.46

T: TOC (%); K: Kerogen (mg/g); PO: Porosity (%); PE: Permeability (mD); SW: Water Saturation: percentage; PS: Pressure (PSI); T: Reservoir Thickness (ft.); PI: Production Index.

The data view instances extracted from cuboid metadata are tabulated in Table 5. Quality of structures is evaluable for commercial traps with frackable reservoirs and producible hydrocarbons.

Table 6 - The Value Extracted from the UDPE

Field	Number of Structures	Number of Fracks	Number of Wells	Volumes (MB)
Nimm	5	158	10	100
Shak1	4	245	11	150
Kak1	6	357	9	154
Mak1	3	634	7	200

MB: Estimated million barrels of oil equivalent.

The knowledge-based domain ontologies are described with several multidimensional data structures. Their strengths are evaluable in the UDPE business contexts in a sedimentary basin. We have plotted porosity attribute instances from multi-stack unconventional reservoirs, with visual analytics focus in an investigating area using bubble plot visualizations (Figures 11a and 11b). As described for the unconventional reservoirs in the investigating sedimentary basin, each bubble's size has significance in terms of characteristic porosity attributes. Interestingly, in the current study area, the unconventional reservoirs follow certain porosity data attribute trends, allowing the explorers to explore new drillable prospective areas. The targets are successful because of smart IS framework articulations and their implementation in Big Data scale. Large-size geographically controlled frackable reservoirs are effective candidates for testing and production. One of the characteristics of the digital ecosystem, PI is analyzed in shale-gas potentiality. The PI is a composite attribute, made up of free hydrocarbon percentage and residual petroleum potential of the source element (Table 5). PI of geological formations ranges from 0.16 to 0.46, with an average of 0.25, indicating that the matured source is sufficient for generating hydrocarbons from shale-gas basins. The value of the UDPE is presented in Table 6.

Contribution and Benefits of Digital Reservoir Energy Ecosystem

Different data structures accommodated in multidimensional repositories make digital ecosystems unified on a sedimentary basin scale. Systems can be connectable through digital clouds and their computing nodes. The constructs, models, and methods are the final deliverables of the current research application. The articulations of Big Data-guided DSIS framework simulated in UDPE settings are pathways, guiding the petroleum explorers as a knowledge-based digital reservoir energy solution. Interoperability is evaluable with different composite schemas at sedimentary basin scales. It can demonstrate the capacity to hold two or more petroleum systems and their linked data structures at different geographic locations. Data attributes interpreted among several reservoir energy scenarios provide significant associations, trends and porosity relationships. Among them are the structure, reservoir, production attribute dimensions considered in visual analytics and the interpretation of high-dense fracks, packed with accretive reservoir energies. Locales of packed frack-reservoirs are outcomes of digital energy solutions for investment.

The Conclusions, Limitations and Recommendations

The UDPE emerges as an evaluable reservoir-energy ecosystem solution. The DSIS articulations are evaluable in the contexts of energy-reservoir development in shale gas basins. The manifestation of ecosystems in the UDPE contexts cannot be appraised without the support of documentation, organization and integration of multiple data sources and their characteristics in making reservoir connections. Unique features of digital energy ecosystem representation include the description of domain ontologies, intelligent storage and integration in a warehouse environment. Different domain ontologies integrated into the framework architecture are from data sources related to shale gas, shale-gas processes, structure, reservoir capacities and geologic characterization attributes. Data modelling, schema selection, data warehousing and mining, visualization and interpretation artefacts articulated within the DSIS framework bring together and unify ecosystems with new insights of reservoir information. As a deliverable research outcome of DSIS, the connectivity attributed to reservoir energy ecosystems can help explorers to invest in the investigating areas. Feasibility and applicability of exploration and field development are assessed for each sedimentary basin with potential hydrocarbon-bearing geological structures. Advantages of the use and reuse of data structures are emphasized in the research. The data mining and interpretation of data views drawn from data warehouse rely on an effective mapping of multiple attribute dimensions and their logical modelling. Attribute dimensions such as structure (including fracks causative to faulted structures and their compartments), reservoir, and seal appear critical in assessing the quality and potentiality of the existence of unconventional digital ecosystems and associated shale-gas plays. However, the data qualities are challenging while appropriating fact instances in the data models. Warehouse repositories can meet the challenges of large size shale gas sedimentary basins, which comprise multiple petroleum systems, including oil and gas fields that share common elements and processes. We recommend implementing the methodology in such basins where geologically favorable shale gas prospects exist. A common agreeable conceptual representation of unconventional reservoir ontology is needed to resolve the heterogeneity of petroleum data sources of shale gas basins, make meaningful information solution and exchange it among distantly located data warehouse repositories.

References

- Alzate, J.H. & Devegowda, D. (2013). Integration of surface seismic, microseismic, and production logs for shale gas characterization: Methodology and field application. *Interpretation*, 1(2), SB37-SB49.
- Arogundade, O., & Sohrabi, M. (2012). A review of recent developments and challenges in shale gas recovery. In *SPE Saudi Arabia Section Technical Symposium and Exhibition*. Society of Petroleum Engineers.
- Bai, Z., Cui, S., & Zhao, C. (2018). Design and Implementation of Sensory Data Collection and Storage Based on Hadoop Platform. In *International Conference in Communications, Signal Processing, and Systems* (pp. 870-874). Springer, Singapore.
- Bellman, L.M.W. (2018). Integrated shale-gas reservoir characterization: A case study incorporating multicomponent seismic data. *Interpretation*, 6(2), SE23-SE37.
- Bento, F. (2018). Complexity in the oil and gas industry: a study into exploration and exploitation in integrated operations. *Journal of Open Innovation: Technology, Market, and Complexity*, 4(11), 11-17.
- Briscoe, G. & Marinos, A. (2009). Digital ecosystems in the clouds: towards community cloud computing. In: IEEE, (corp. ed.) *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies (DEST 2009)*. Institute of Electrical and Electronics Engineers (IEEE), New York, USA, 103-108.
- Burke, G.R. (2013). Making Viability Sustainable, PhD Thesis, Department of Humanities, Curtin University, Perth, WA, Australia.
- Castañeda, G. O. J., Nimmagadda, S. L., Cardona Mora, A. P., Lobo, A., & Darke, K. (2012). On Integrated Quantitative Interpretative Workflows for interpreting structural and combinational traps for risk minimizing the exploratory and field development plans. In *Bolivarian Geophysical Symposium*.
- Ceptureanu, S.I. Ceptureanu, E.G. Olaru, M. & Popescu, D.I. (2018). An Exploratory Study on Knowledge Management Process Barriers in the Oil Industry. *Energies*, 11(8), 1977.
- Chang, E., & West, M. (2006). Digital Ecosystems A Next Generation of the Collaborative Environment. *iiWAS*, 214, 3-24.
- Chromiak M. & Stencil K. (2014). A Data Model for Heterogeneous Data Integration Architecture. In: Kozielski S., Mrozek D., Kasprowski P., Malysiak-Mrozek B., Kostrzewa D. (eds) *Beyond Databases, Architectures, and Structures*. BDAS 2014. *Communications in Computer and Information Science*, 424. Springer, Cham. https://doi.org/10.1007/978-3-319-06932-6_53.
- Cleary, L, Freed, B, & Elke, P. (2012). "Big Data Analytics Guide: 2012", Published by SAP, CA 94607, USA.
- Coronel, C., Morris, S., & Rob, P. (2011). Database Systems, Design, Implementation and Management, Course Technology, Cengage Learning, USA.
- Costa, R., & Rezende, J.F. (2018). Strategic alignment of knowledge management and value creation: implications on to an oil and gas corporation. *RAUSP Management Journal*, 53(2), 241-252.
- Creties, D. J., De Golyer, Mac, N & Boyer, C.M. (II). (2008). Coal-bed and shale gas reservoirs, SPE, 103514, Distinguished Lecture Series, USA.
- Dovers, S., Norton, T. W., & Handmer, J. W. (2001). Ignorance, uncertainty and ecology: key themes. *Ecology, uncertainty and policy: Managing ecosystems for sustainability*, 1-25.

- Durham, S. L. (2013). An unconventional idea, open to interpretation. *AAPG Explorer, March Series, USA*.
- Eapen, B. R. (2008). ONTOderm - A domain ontology for dermatology. *Dermatology Online Journal, 14* (6).
- Flahive, A., Rahayu, W., Taniar, D., & Apduhan, B. (2004, December). A distributed ontology framework for the grid. In *International Conference on Parallel and Distributed Computing: Applications and Technologies*, 68-71. Springer, Berlin, Heidelberg.
- Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. Ji-fan., Dubey, R. & Childe, S. J. (2017). Big data analytics and firm performance: effects of dynamic capabilities. *Journal of Business Research, 70*, 356-365.
- Furr, N., & Shipilov, A. (2018). Building the right ecosystem for innovation. *MIT Sloan Management Review, 59*(4), 59-64.
- Geerts, G. L. (2011). A design science research methodology and its application to accounting information systems research. *International Journal of Accounting Information Systems, 12*(2), 142-151.
- Gilbert, R., Liu, Y., Abriel, W., & Preece, R. (2004). Reservoir modeling: Integrating various data at appropriate scales. *The leading edge, 23*(8), 784-788.
- Grant, R. M. (2013). The development of knowledge management in the oil and gas industry. *Universia Business Review, (40)*, 92-125.
- Gregor, S. & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems, 8*(5), 312-335.
- Hasan, H., & Kazlauskas, A. (2009, June). Digital ecosystems: ICT's contribution to addressing climate change. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies*, 157-162.
- Hemingway, J., Lewis, R. E. & Donald, A. (2012). Scientific target selection optimizes horizontal shale completions, Unconventional Exploration Technology, E & P, Hart Energy, USA.
- Henningson, S. & Hedman, J. (2014). Transformation of Digital Ecosystems: The Case of Digital Payments, *Information and Communication Technology – EurAsia Conference*, 46-55.
- Indulska, M. & Recker, J. C. (2008). Design Science in IS Research: A Literature Analysis. In Gregor, Shirely and Ho, Susanna, Eds. *Proceedings 4th Biennial ANU Workshop on Information Systems Foundations*, Canberra, Australia.
- Inmon, W. H. (2005). *Building the Data Warehouse*, 4th Edition, Wiley, eBook, 1-576.
- Jia, C., Zheng, M., & Zhang, Y. (2016). Some key issues on the unconventional petroleum systems. *Petroleum Research, 1*(2), 113-122.
- Kadadi, A., Agrawal, R., Nyamful, C., & Atiq, R. (2014, October). Challenges of data integration and interoperability in big data. In *2014 IEEE international conference on big data (big data)*, 38-40.
- Khatri, V., Ram, S., & Snodgrass, R. T. (2004). Augmenting a conceptual model with geospatiotemporal annotations. *IEEE transactions on knowledge and data engineering, 16*(11), 1324-1338.
- Lee, J., Wyner, G.M., & Pentland, B.T. (2008). Process Grammar as a Tool for Business Process Design. *MIS Quarterly, 32*(4), 757-778.
- Lee, T. J., Pouliot, Y., Wagner, V., Gupta, P., Stringer-Calvert, D. W., Tenenbaum, J. D., & Karp, P. D. (2006). BioWarehouse: a bioinformatics database warehouse toolkit. *BMC*

Bioinformatics, 7(1), 1-14.

- Li, G. (2011). *World Atlas of Oil & Gas Basins*, Wiley-Blackwell, p. 1-496.
- Li, W., Biennier, F. & Badr, Y. (2012). Digital Ecosystems: Challenges and Prospects, *Proceedings of the International Conference of Management of Emergent Digital Ecosystems*, DOI: 10.1145/2457276.2457297.
- Li-Ying, J., Vanhaverbeke, W., & Schoenmakers, W. (2008). Exploration and exploitation in innovation: Reframing the interpretation. *Creativity and innovation management*, 17(2), 107-126.
- Lin, A. (2014). "Principles of big data algorithms and application for unconventional oil and gas resources," in Proc. *SPE Large Scale Computing and Big Data Challenges in Reservoir Simulation Conference and Exhibition*, 87–95.
- Martinez-Mosquera, D., Navarrete, R., & Lujan-Mora, S. (2020). Modeling and management big data in databases—A systematic literature review. *Sustainability*, 12(2), 634.
- Masa, E., Busch, P., Guzman, G., & Sanzogni, L. (2018). Qualitative analysis of Big Data Analytics in the Oil & Gas industry. In K. S. Soliman (Ed.), *Proceedings of the 31st International Business Information Management Association Conference, IBIMA 2018: Innovation Management and Education Excellence through Vision 2020* (pp. 287-306). Madrid: International Business Information Management Association (IBIMA).
- Matsuzawa, H., & Fukuda, T. (2000, April). Mining structured association patterns from databases. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 233-244.
- Middleton, R. S., Gupta, R., Hyman, J. D., & Viswanathan, H. S. (2017). The shale gas revolution: Barriers, sustainability, and emerging opportunities. *Applied energy*, 199, 88-95.
- Mohammadpoor, M., & Torabi, F. (2018). Big Data analytics in oil and gas industry: An emerging trend. *Petroleum*.
- Nguyen, T., Gosine, R. G., & Warriani, P. (2020). A systematic review of big data analytics for oil and gas industry 4.0. *IEEE Access*, 8, 61183-61201.
- Nimmagadda, S. L., Reiners, T., & Wood, L. C. (2018). On big data-guided upstream business research and its knowledge management. *Journal of Business Research*, 89, 143-158.
- Nimmagadda, S. L., Reiners, T., Wood, L. C., & Zhu, D. (2019). On Big Data guided Unconventional Digital Ecosystems and their Knowledge Management. In *PACIS*, 7.
- Offermann, P., Levina, O., Schönherr, M., & Bub, U. (2009, May). Outline of a design science research process. In *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, 1-11.
- O'Leary, D. E. (1999). REAL-D: A schema for data warehouses. *Journal of Information Systems*, 13(1), 49-62.
- Ozkarahan, E. (1990). *Database management: concepts, design, and practice*. Prentice-Hall, Inc.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77.
- Porteous, D., & Morawczynski, O. (2017). Inclusive digital ecosystems of the future. *FIBR Project White Paper*, (2).
- Pujari, A. K. (2001). *Data Mining Techniques*, University Press, Hyderabad, Telangana, India.
- Ramanigopal, C.S. (2013). *Knowledge Management for the Oil and Gas Industry –*

Opportunities and Challenges. *Advances in Management*, 6(3).

- Ramiz, A. & Yadigar, I. (2017). Conceptual big data architecture for the oil and gas industry. *Problems of Information Technology*, 3-13.
- Sabry, N. & Krause, P. (2012). A digital ecosystem view on cloud computing, *6th IEEE Digital Ecosystems and Technologies*, IEEEExplore, Italy.
- Sandalow, D., Wu, J., Yang, Q., Hove, A., & Lin, J. (2014). Meeting China's Shale Gas Goals. *Columbia University Center on Global Energy Policy (September 2014)*, 5-6.
- Shah, M. (2015). Big Data and the Internet of Things. *ArXiv*, abs/1503.07092.
- Shanks, G., Tansley, E., & Weber, R. (2003). Using ontology to validate conceptual models. *Communications of the ACM*, 46(10), 85-89.
- Shvachko, K. Kuang, H. & Radia, S. (2010). The Hadoop Distributed File System, *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*, Nevada, USA.
- Subramaniam, M. (2020). Digital ecosystems and their implications for competitive strategy. *Journal of Organization Design*, 9, 1-10.
- Sunjay, S., Susheel, K. & Nikhil. J. (2014). Shale Gas: An Unconventional Gas Reservoir. 10.13140/RG.2.1.1975.2805.
- Tung, T. V., Trung, T. N., Hai, N. H., & Tinh, N. T. (2020). Digital transformation in oil and gas companies-A case study of Bien Dong POC. *Petrovietnam Journal*, 10, 67-78.
- Vaishnavi, V. & Kuechler, W. Jr. (2007). *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*, NY: Auerbach Publications, Boca Raton, FL, Taylor & Francis Group.
- Valdez-De-Leon, O. (2019). How to develop a digital ecosystem: a practical framework. *Technology Innovation Management Review*, 9(8), 43-54.
- Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A framework for evaluation in design science research, *European Journal of Information Systems*, 25, 1-13.
- Xiaoming, W., Shangyou, N., & Yunshi, W. (2020). Special Report 1: A Study of China's Energy Supply Revolution. In *China's Energy Revolution in the Context of the Global Energy Transition*, 41-211. Springer, Cham.
- Yu, F., Deuble, R., & Morgan, H. (2017). Designing research data management services based on the research lifecycle—A consultative leadership approach. *Journal of the Australian Library and Information Association*, 66(3), 287-298.
- Zhonghua, M. (2017). Seismic Data Attribute Extraction based on the Hadoop platform, *Computer Science, IEEE 2nd International Conference on Cloud Computing and Big Data Analysis*, 180-184.

About the Authors

Shastri L. Nimmagadda is presently a Research Fellow (adj.) at School of Management (Business Information Systems), Curtin University, Perth, WA, Australia. Shastri worked for the Schlumberger Company in multiple geo-markets worldwide as an Expert in Geosciences. He worked for Wafra Joint Operations Petroleum Company in Kuwait (with Chevron Energy and Kuwait Oil Company) and Santos JO in Sydney Australia. Shastri previously worked for several national petroleum operating and service companies in India, Australia, Uganda, Kuwait, Abu Dhabi, Egypt, Malaysia, Colombia, Indonesia, Russia, including several assignments in the USA. In addition to industry practice, Dr Shastri has vast academic experience. He did his M Tech and PhD in Exploration Geophysical Technology from the Indian Institute of Technology, Kharagpur. He obtained Master of Information Technology and PhD in Information Systems with distinction from the Curtin University of Technology, Australia. His industry research interests are data modelling, Digital Ecosystems and Technologies (DEST), data integration, warehouse modelling, data processing, interpretation and knowledge mapping, including research in domain applications. He already published and presented more than 150 research and technical papers in various international journals and international conference proceedings in the areas Business Information Systems, Digital Ecosystems, Oil & Gas Exploration and Geoscience Applications.

Neel Mani is an Associate Professor at Amity Institute of Information Technology (AIIT), Amity University Uttar Pradesh (AUUP), Noida. He has a doctorate degree in Computer Science from Dublin City University, Dublin Ireland, Dr. Neel Mani pursued his postdoctoral research at University College Dublin, where he investigated digital, precision agriculture, and crop science through a strong multi and inter-disciplinary approach and bioinformatics with specialization in Big data genomics and its AI Applications. He is a science graduate with three post graduations in computer science. Neel's research interests including Software Reuse, Machine Learning, Intelligent Device and Web with Big Data Analytics (IoT with micro sensors), Natural Language Processing, and Software Engineering. He is currently a professional member of ACM. He has more than 16 years of experience including industry experience in application development, predictive analysis, and data migration (structural, behavior) in the heterogeneous environment, and more than 8 years of experience in the research and academics. He has taken and supervised plenty of industrial and research projects at postgraduate programs including Post Graduate Programs in Computer Science, IT, Computer Applications, Management Studies, and Systems."

Torsten Reiners is a Senior Lecturer in Logistics and Supply Chain at the School of Management and director of the Logistic Research Cluster exploring the implication of the urban expansion on logistics infrastructure and agricultural supply chains. With the background in operations research, simulation, mathematical modelling, algorithm development, data visualization, and data analytics using, among others, clustering, logic analysis of data, and sentiment analysis, his research is exploring cross-disciplinary challenges and the application of theoretical frameworks in new, academic or business, contexts. Current research involves, among other, the use of virtual reality and phenomenology in health and safety training in logistics, sentiment analysis to counteract the bull-whip effect in supply chains, event-studies on recalls and sustainable energy, sustainable information systems in oil and gas, and waste prevention in food supply chains. Recent publications (published and in revision) include work on ontologies in oil and case, sentiment analysis in the supply chain, event studies, and logic analysis of data.

Lincoln C Wood is an Associate Professor at the University of Otago (New Zealand), Associate Dean of Postgraduate Programmes (Otago Business School), and an Adjunct Research Fellow at Curtin University (Western Australia). His research investigates the design of sustainable operational systems, improving sustainable and social outcomes through operations and supply chain management practices. His work has particular applications in the construction and logistics sectors and healthcare delivery systems. A stream of his research examines the role of technology in lifting operational performance. Dr Wood is an Associate Editor at the Journal of Supply Chain Management, Editor of the International Journal of Applied Logistics, and the co-Editor of the International Journal of Socio-technology and Knowledge Development.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.