Multidisciplinary study of the problems of big data technologies in the oil and gas industry

Rasim M. Alguliyev, Ramiz M. Aliguliyev* and Fargana J. Abdullayeva

Institute of Information Technology,
Azerbaijan National Academy of Sciences,
9A, B. Vahabzade Street,
AZ1141 Baku, Azerbaijan
Email: r.alguliev@gmail.com
Email: r.aliguliyev@gmail.com
Email: a_farqana@mail.ru
*Corresponding author

Abstract: This paper is devoted to the analysis of the multidisciplinary problems of the big data technology in the oil and gas industry. Application capabilities of big data technologies in issues such as reducing the exploitation risks, crude oil price forecasting, optimal management of the oil wells, health and safety ensuring in an organisation, overcoming environmental problems and so on are investigated. [Received: April 18, 2017; Accepted: February 2, 2018]

Keywords: oil and gas; crude oil; big data; exploration and production; health, safety, environment.

Reference to this paper should be made as follows: Alguliyev, R.M., Aliguliyev, R.M. and Abdullayeva, F.J. (2020) 'Multidisciplinary study of the problems of big data technologies in the oil and gas industry', *Int. J. Oil, Gas and Coal Technology*, Vol. 23, No. 1, pp.92–105.

Biographical notes: Rasim M. Alguliyev is a Director of the Institute of Information Technology of Azerbaijan National Academy of Sciences (ANAS) and Academician-Secretary of ANAS. He is Full member of ANAS and Full Professor. He received his BSc and MSc in Electronic Computing Machines from the Azerbaijan Technical University in 1979. He received his PhD and Doctor of Science (higher degree after PhD) in Computer Science in 1995 and 2003, respectively. His research interests include: information security, e-government, data mining, big data, online social network analysis, cloud computing, evolutionary and swarm computation, and scientometrics. He is author more than 585 papers, four monographs, four patents, several books.

Ramiz M. Aliguliyev is a Head of Department at the Institute of Information Technology of ANAS. He received his BSc and MSc in Applied Mathematics from the Baku State University, Azerbaijan in 1983. He received his PhD (2002) in Mathematics and Doctor of Science (higher degree after PhD) in Computer Science (2011). His research interests include text mining; clustering; evolutionary and swarm optimisation; web mining; online social network analysis; big data analytics and scientometrics. He is author 165 papers and four books.

Fargana J. Abdullayeva graduated from the Computer Techniques and Technologies Faculty of Sumgait State University in 2002. In 2003, she was accepted for employment at the Institute of Information Technology of ANAS, where she addressed issues of information security provision. In 2004, as a respondent, she commenced his work on the subject of 'Development of methods and algorithms for providing information security of population and migration system'. In 2010, she successfully defended her PhD thesis. At present, she conducts researches on cloud computing security, anomaly detection, load balancing, task scheduling and deep learning.

1 Introduction

Recently, the exponential growth of the large amounts of real-time data has made it harder for the many oil operators to get valuable information from these data (Feblowitz, 2013; Marz and Warren, 2015; Basanta-Val et al., 2016). In most cases, data analysis, upgrading of the facilities designed for oil wells are carried out on the basis of human intervention. Here, because of the more time spent on usual analysis, experts are facing delays in the required decision-making process for the optimisation of wells. Moreover, despite considerable effort and a broad range of new approaches to safety management over the years, the oil and gas industry cannot minimise the high rate of injuries and fatalities among personnel. For discovering reasons of such accidents vast amounts of data have been collected in oil and gas sector. But the highly customised collection of this data within the sector, the existence of written responses and fragmented nature has created a problem for discovering patterns of these accidents.

Currently to overcome these problems the automatic processing of data is carried out by means of big data technology (Aliguliyev et al., 2016; Basanta-Val et al., 2015; Lv, 2017; Congosto et al., 2017).

Big data is quite a complex term, covers application issues of new means and methods to large-scale data analysis, which analysis of these data is considered to be impossible to carry out via traditional methods. Generally, big data is the data set, which has fairly large and complex structure, and their processing needs advanced data storage, management, analysis, and visualisation tools.

As in many industrial sectors, in the oil and gas sector, there is a great revival to promote initiatives in this regard. By most accounts, the oil and gas industry's data is already 'big'. Modern oil and gas seismic data centres can easily contain as much as 20 petabytes of information. Current estimates suggest that the total amount of digital data in the world including things like books, images, e-mails, music, and video is doubling every 2–3 years (Lohr, 2012).

In some sources, big data is considered as a component of the social ecosystem (Shin and Choi, 2015). Socio-technical problems of big data cover issues about the creation of heuristic guidelines and strategies for the development of society. In many countries, big data initiatives are adopted. Thus, in Korea MSIP (Ministry of Science, ICT, and Future Planning) established a new big data centre in 2014, and further plans exist to rebuild a real-time collection and analytics system that can collect and analyse text-based social network data, stream of audio and video data and data from the Internet of Things (Shin, 2014). These initiatives have been designed to expand connections among government, industry, and social agency organisations. For example, the centre connects its online platform with a government data portal service that stores national statistics and other public information.

Software and technologies are as significant as infrastructure. Most efforts so far have been concentrated on the development of crowdsourcing, data fusion and integration, genetic algorithms, machine learning, natural language processing, signal processing, simulation, time series analysis, visualisation, massive parallel-processing databases, search-based applications, data-mining grids, distributed file systems, distributed databases, cloud-based infrastructure (applications, storage, and computing resources), and the internet.

There are numerous multidisciplinary problems of big data technology in the oil and gas industry (Nicholson, 2012) and has been widely used in solving problems such as optimal management of wells and reservoirs (Crockett and Kurrey, 2014), to ensure the health and safety in the organisation (Pettinger, 2014), to carry out geological exploration (Binotto et al., 2014) and others.

In this paper, multidisiplinary problems of big data technology in oil and gas industry are analysed. Application capabilities of big data technologies in issues such as reducing the exploitation risks, crude oil price forecasting, optimal management of the oil wells, health and safety ensuring in the organisation, overcoming environmental problems and so on are investigated.

2 Big data

The discussions were held with regard to big data in mass information media shown that there was no exact definition of big data. Clearly, size is the first characteristic that comes to mind considering the question 'what is big data?' (Gandomi and Haider, 2015). Here the structured data is just a small subset of a big data; the largest component of big data is unstructured data in the audio, image, video, and text form. The volume of data, rate (velocity) at which data are generated and so on are essential characteristics of big data. Recently the three V's (volume, variety, velocity) have emerged as a common framework to describe big data. For example, Gartner Corporation defines big data in similar terms as:

> "Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation." (IT Glossary, Gartner)

Similarly, TechAmerica Foundation defines big data term as:

> "Big data is large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information." (TechAmerica Foundation Big Data Commission)

Here volume refers to the magnitude of data, variety refers to the structural heterogeneity in a dataset, and velocity refers to the rate at which data are generated and the speed at which it should be analysed.

3 The distinguishing factors of big data from traditional technologies

Energy sector and some other organisations have been aspiring to for a long time to make better decisions. What is changing, however, these decisions are made via big data type questions are always concerns the people. First, big data differs from traditional approaches on account of the three V's: volume, velocity, and variety.

Traditionally, data collection was implemented on sampling and trying to make the collected data from that sample as accurate as possible. By contrast, the move towards big data has led to a much greater tolerance for messiness and imprecision. This more relaxed approach to vagueness has been compensated, however, by much larger volumes of data.

This technology enables to make more accurate decisions in real-time on the basis of more data. Here this data is coming from a wider variety of sources and in an increasingly broad array of formats. The data type can appear as digital and analogue form. The data fusion capability of big data technology makes it more and more possible to bring together disparate sources of data to glean fresh insights that nobody predicted (Big Data, 2014).

4 Data collection in big data

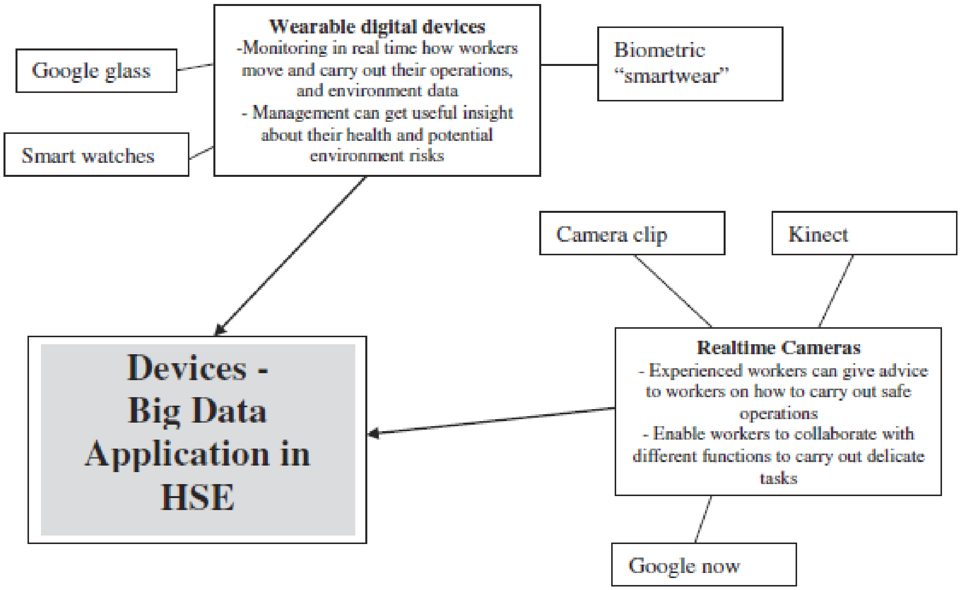
Traditional data sets have historically needed to be fairly structured, orderly, and static. Digital information in the era of big data is collecting frequently noisy, messy, raw, unstructured and dynamic.

Recently several new technologies have also developed, that can collect large volumes of data to manage industrial risks. In recent years, there has been a proliferation of new wearable devices such as watches, rings, glasses and heads-up displays for consumers. These wearable devices will enable experts to help less experienced industrial workers in oil and gas operational settings all around the world. For example, a relatively inexperienced worker at an oil and gas operating facility can wear special safety glasses equipped with a camera, microphone, speaker and wireless antenna to send via live data feed information about their surroundings and the system around them to some kind of central command centre staffed by seasoned veterans. In this case, the more experienced personnel can advise their less experienced colleagues on the specifics of their situation, thereby imparting the wisdom and experience of the more senior staff member without actually having them physically present at the site.

Application of wearable devices and big data methods to the HSE (health, safety, and environment) sector of oil and gas industry gives high advantages to the operation of this sector (Figure 1) (Tan et al., 2016; Frenzel, 2003):

1. monitoring in real-time how workers move and carry out their operations and environmental data
2. management can get useful insight into their health and potential environmental risks
3. experienced workers can give advice to workers on how to carry out safe operations
4. enable workers to collaborate with different functions to carry out delicate tasks.

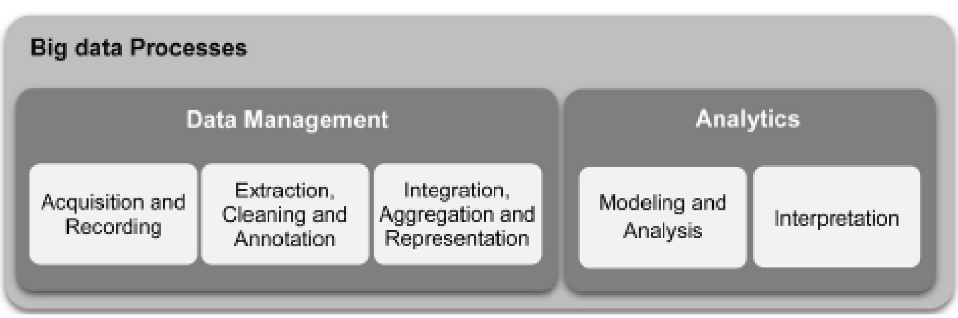
Figure 1 Big data devices for HSE



5 Big data analytics

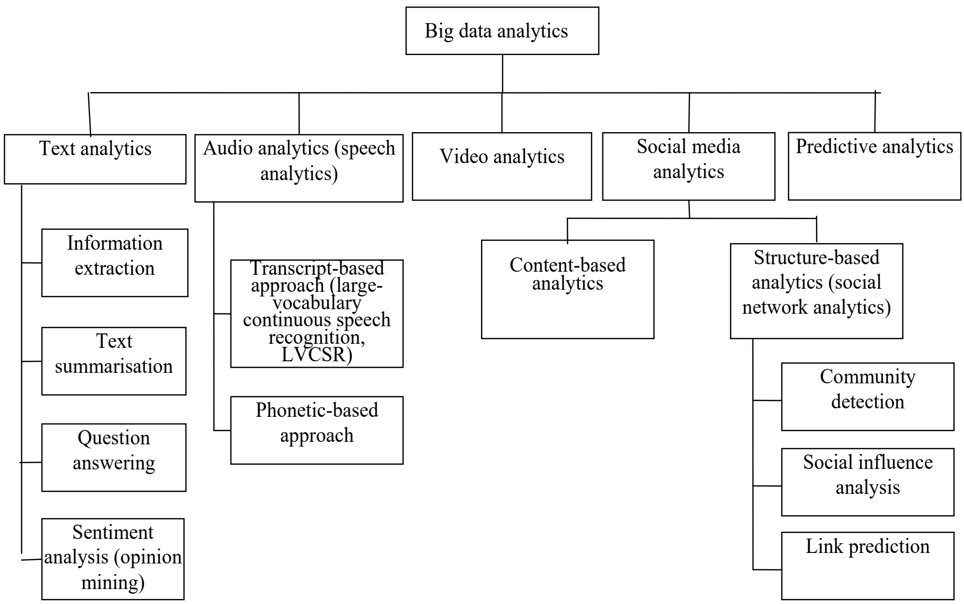
Big data is unlocked only when leveraged to drive decision making. To enable such decision making, entities need efficient means to turn high volumes of fast-moving and diverse data into meaningful insights. The overall process of extracting insights from big data can be broken down into five stages (Figure 2). These five stages form the two main sub-processes: data management and analytics. Data management involves processes and supporting technologies to acquire and store data and to prepare and retrieve it for analysis. Analytics, on the other hand, refers to techniques used to analyse and acquire intelligence from big data. For this reason, big data analytics can be viewed as a sub process in the overall process of 'insight extraction' from big data.

Figure 2 Big data processes



There are big data analytical techniques for analysing structured and unstructured data. In Gandomi and Haider (2015) these analytical techniques are divided into five categories (Figure 3).

Figure 3 Big data analytics techniques



6 Multidisciplinary problems of big data technologies in the oil and gas industry

Big data is widely used in production operations, logistics, supply chain, oil and gas industry, risk-reducing areas (Schlegel and Trent, 2014; Jeske et al., 2013; Saenz and Revilla, 2014; Griffith, 2014). Thus, in Li et al. (2015), with the use of big data generated in the online environment on the basis of GSVI (Google search volume index) the crude oil price forecasting approach by considering the interest of investors is proposed. Here multiple data sources from the Google Trends, Energy Information Administration (EIA) website and Commitments of Traders (COT) reports by CFTC, are used to examine the relationship among investor attention, trading positions, and crude oil prices for 2004–2014 years. All data series covers the period from January 2004 through June 2014 at a weekly frequency. The value of GSVI at the week t is defined as follows:

$$GSVI_{K_i,t} = \frac{S_{K_i,t}}{\sum_{K_1,t}^{K_m,t} S_{K_i,t}}$$

where $S_{K_i,t}$ represents the search volume of the specific term, K_i at week *t*. $\sum_{K_1,t}^{K_m,t} S_{K_i,t}$ represents the search volumes of all terms during time *t* in Google. After extracting the GSVI series based on different search terms, the selection of GSVI series that most represents investor attention on crude oil price dynamics are considered. For this purpose, here descriptive, correlation and stability analysis is used.

Big data has a significant impact on reducing risks and cost optimisation associated with operations in HSE sector (Baaziz and Quoniam, 2013):

- prevention of accidents in drilling process
- prediction of time for maintenance of wells, optimisation of drilling parameters and failure prevention
- by using weather and labour resource planning data prevention arising from dangerous working conditions for workers and the reduction of environmental risks.

In Tan et al. (2016) application opportunities of big data and forecasting analytics in reducing the exploitation risks in the oil and gas industry are investigated and a research agenda toward improving its health and safety performance is offered:

- overcoming technical, economic or organisational barriers for big data technologies to be applied to the oil and gas industry's HSE-related challenges
- reduce the use of cloud computing in the application process of big data in oil and gas industry
- to make the application of big data analytics tools possible reduce the use of cloud computing which allows bringing together the myriad HSE-related data sets in the oil and gas sector
- overcome legal, political or organisational barriers for big data technologies to be applied to the oil and gas industry's personnel records en route to discovering HSE-related trends.

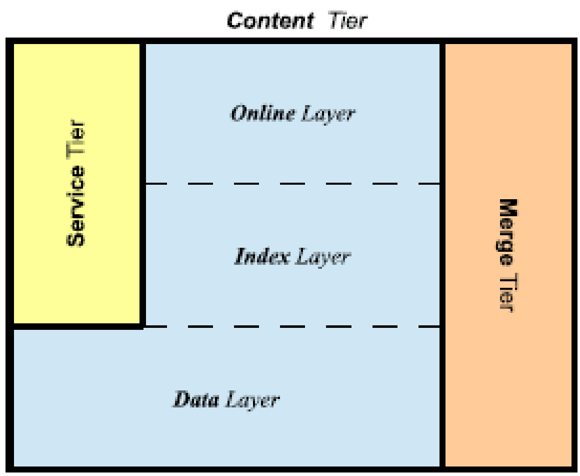
Any type architecture designed for the management of big data must take into account all characteristics (e.g., volume, velocity, etc.) of big data (Begoli and Horey, 2012). Architectures must consider following requirements:

1 architecture must store and expose big datasets containing historical data
2 architecture must gather and expose run-time data.

In most proposed new architectural models mainly trying to address the above-mentioned problems. In Martínez-Prieto et al. (2015), the architecture titled as service-online-index-data architecture (SOLID) allowing to collect, store and use of big data in real-time is proposed (Figure 4).

Here both of the requirements above claimed for big data architectures are taken into account. In order to improve the quality of database integration and management of the data under general model is carried out. For the effective integration of data, a graph-oriented resource description framework (RDF) model is used. SOLID is a high-performance architecture for real-time systems managing big semantic data, i.e., big RDF datasets. Figure 4 illustrates the SOLID configuration. The content tier is the central data core and groups three layers. The big semantic data is preserved on a data layer which realises a data store for serialised RDF. The index layer is built on top of the data layer and is responsible for providing indexed fast access to the big semantic data. This ensures efficient query resolution, typically through SPARQL, the de facto standard for querying RDF. Therefore, the combination of these two layers brings a competitive solution for the first aforementioned challenge: storing and serving big semantic data containing historical data. The online layer realises the gathering of run-time data. The merge tier integrates the run-time data into the big semantic data; the service tier implements the indexation process.

Figure 4 The solid architecture (see online version for colours)



Perrons and Jensen (2015) examines existing data management practices in the upstream oil and gas industry, and compares them with big data techniques that have emerged in organisations recently. The paper then offers suggestions and recommendations for the ways to extract more value from data.

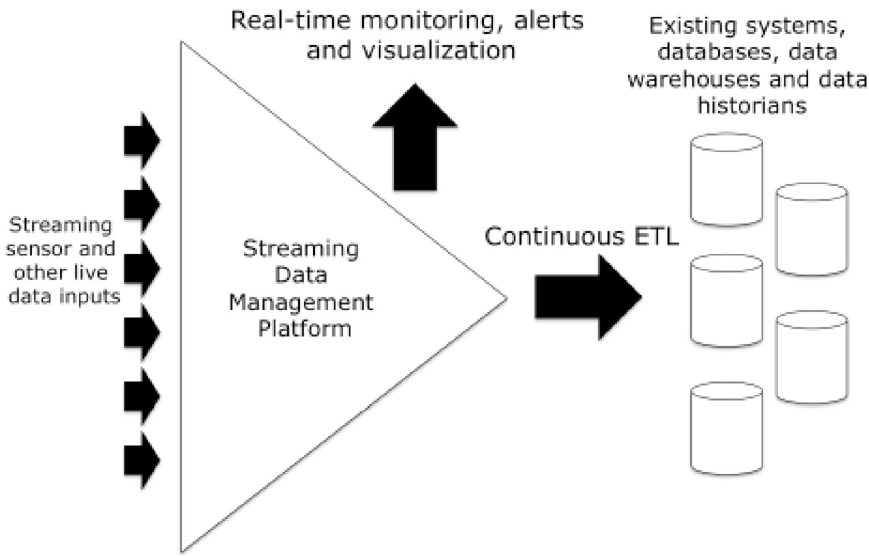
Oil and gas companies, collect and analyse the data from different resources as follows (Hamzeh, 2017):

- collecting data from different sensors during drilling stage
- traditional enterprise data from operational systems
- social media
- web searching designs
- demographic data
- historical oil and gas exploration, delivery, and pricing data (e.g., log files).

In the meeting, todays need of petroleum industry the data management (using big data platforms), integration, collaboration and performance management of data coming from various sources is considered too much important. Many studies have already taken measures in this direction. In Abadie and Beggs (2013) is offered streaming data architecture that is capable of processing and analysing big data and smart oilfield data in real-time. Further, the architecture supports the continuous, streaming integration of any and all of an organisation's sensor data, operational platforms, and data warehouses (Figure 5). The approach aggregates and analyses live, streaming data on the fly, without the need to store the data first.

Here sensors are the most common source of streaming, real-time data. But any static data source such as log files and database can be used. Incoming data from sensors after processing in stream data management platform transmitted to the existing databases.

Figure 5 Stream data processing architecture



In the oil and gas industry data coming from many different sources forms the data lake. Without data management, cleansing, integration and governance, a data lake can become data swamp. Data swamps cannot be guaranteed to be reliable in quality, lineage, and security. To overcome this problem in (Brulé, 2015) conceptual model titled as data reservoir is proposed. The main goal of this model is to provide management, tracking, cleaning and integration of data. In data reservoir model data processing can be performed in three analytical ways:

1 structured analytics

2 real-time analytics

3 deep insight analytics.

Oil and gas companies anticipate IT security breaches by using predictive analytics and bolstering security with data from the global protection systems including video monitoring, access control and anti-intrusion (Feblowitz, 2013). Also, deploying complex event processing (CEP) technology to monitor security concerns in the oil and gas industry in real-time is considered as urgent issues (Singh et al., 2015). In this regard, mainly the following operations are carried out:

- combining data from multiple sources to infer events or patterns that indicate a current or imminent threat
- making faster decisions, supported by faster delivery of decision support information, to identify possible threats
- prediction and prevention of cyber-terror acts.

Data quality is often cited as being the major challenge in computerisation of processes in the oil and gas industry. In Cameron (2014), the data quality problem is divided into three areas:

1 Boredom and bloat: routine, uninteresting, static and unchanging information.

2 Noise: background data that obscures and hides information. Unclear or non-standard definitions that introduce and reinforce semantic noise.

3 Rubbish: incorrect, corrupted data and missing data.

Summarising above mentioned problems list of multidisciplinary issues of big data technologies in the oil and gas industry can be described as follows (Table 1).

Table 1 Multidisciplinary problems of big data in the oil and gas industry

Subject area	*Application problem*	*Literature*
Exploitation risk mitigation	Big data and predictive analytics were used in mitigation of exploration risks in oil and gas industry.	(Griffith, 2014; Tan et al., 2016)
Accurate forecasting of crude oil prices	Determination of relationship among investor attention, trading positions, and crude oil prices.	(Li et al., 2015)
Ecological problems of big data	Big data's social and technical problems cover developing issues of heuristic guidelines and strategies. Analysis of big data phenomenon in terms of socio-technical systems theory is provided.	(Shin and Choi, 2015)
Development of real-time big data architectures	Any type of architecture designed for the big data must take into account big data database which stores and exposes big datasets containing historical data; furthermore it must gather and expose run-time data. Here the architecture titled as SOLID allowing to collect, store and use of big data in real-time is proposed.	(Martínez-Prieto et al., 2015; Abadie and Beggs, 2013)
Continued improvement of the processing rate of computational devices	For the development of ever more powerful microprocessors the importance of continued progress in miniaturising their components are noted.	(Ball, 2000)
Acquisition of structured data from unstructured data.	By representing big data sets as networks of geometrical nodes and edges so that the data can be rationalised using a suite of mathematical tools known as topological data analysis (TDA). Simply put, TDA is a way of getting structured information out of the unstructured data so that machine learning algorithms can be applied to it.	(Carlsson, 2009; Ouellette, 2013)

Table 1 Multidisciplinary problems of big data in the oil and gas industry (continued)

Subject area	*Application problem*	*Literature*
Development of big data application platforms.	The development of big data software platforms setup on the basis of Google's MapReduce or its open-source rival Apache™ Hadoop systems. These tools make it possible to break large data sets into smaller chunks and distribute them to several computing devices. The results of the calculations arising from each of the smaller chunks can then be reintegrated at the end of the process. This approach frequently uses cloud computing infrastructure as a platform for transferring these data chunks to different computing devices and then bringing back the results.	(Schönberger and Cukier, 2014)
Reducing costs to data storage	For reliable storage of personal files, consumers are increasingly turning towards the Internet. Services like Dropbox, Google Drive, and Amazon Cloud Storage are making it easier for people to not only backup documents and photos but also share across multiple devices. These services also provide redundant storage.	(Komorowski, 2014)
Optimal management of oil wells	The ways to the implementation of the data analysis via automated methods is proposed.	(Crocket and Kurrey, 2014)
Ensure the health and safety of the organisation	For this reason, the injury metrics and safety metrics are defined.	(Pettinger, 2014)
Big data and cloud computing convergence	Recently integration big data with the cloud computing has gained great popularity. Here the possibility reliable access to the data repository allows users to store and share data in the cloud. This reduces the need of the user to the power and cost of the computing device and simplifies the complexity of the architecture.	(Intel IT Center, 2015)
Logistics	Global courier company DHL for early detection of potential risks related to weather conditions, influenza outbreaks in its supply chains used big data technology.	(Jeske et al., 2013)
Data quality	Approach to improve data quality is proposed.	(Cameron, 2014)
Data management	For the efficient data management conceptual data reservoir model is proposed.	(Brulé, 2015)
Security	The approach deploying complex event processing (CEP) technology to monitor security concerns in the oil and gas industry in real-time is proposed.	(Singh et al., 2015; Cloud Security Alliance, 2013)

7 Conclusions

Oil and gas companies to overcome technical difficulties in the problems faced in their everyday activities, in upgrading performance indices to explore the potential of big data technology should create collaboration between multidisciplinary research directions.

Although the rapid increase of big data uses in state and research area there is not approved any official standard which ensures interoperability of big data systems. The main purpose of the application of interoperability standard is to eliminate problems as the determination of the transparent source of the data, carrying out of data exchange among multiple domains, integration and analysing of data.

To ensure privacy in big data environment in leading countries around the world is adopting various documents. The main purposes of these documents are allowing collection and use of personal data without permission of the person. In this regard, for example, the Korea Communications Commission released big data privacy guidelines in 2014 that allow industries to collect and use an individual's private information and history without consent.

Acknowledgements

This work was supported by the Science Foundation of the State Oil Company of Azerbaijan Republic (SOCAR). Project's title: Investigation of Big Data Analysis Technology for Oil and Gas Industry.

References

Abadie, V.H. and Beggs, R. (2013) 'Real-time streaming data management as a platform for large-scale mission critical sensor network applications', *SPE Digital Energy Conference*, pp.1–11.

Aliguliyev, R.M., Imamverdiyev, Y.N. and Abdullayeva, F.J. (2016) 'The investigation of opportunities of big data analytics as analytics-as-a-service in cloud computing for oil and gas industry', *Problems of Information Technology*, No. 1, pp.9–22.

Baaziz, A. and Quoniam, L. (2013) 'The information for the operational risk management in uncertain environments: case of early kick detection while drilling of the oil or gas wells', *International Journal of Innovation and Applied Studies*, Vol. 4, No. 1, pp.52–67.

Ball, P. (2000) 'Chemistry meets computing', *Nature*, Vol. 406, No. 6792, pp.118–120.

Basanta-Val, P., Audsley, N.C., Wellings, A.J., Gray, I. and Fernández-García, N. (2016) 'Architecting time-critical big-data systems', *IEEE Transactions on Big Data*, Vol. 2, No. 4, pp.310–324.

Basanta-Val, P., Norberto, F.G., Andy, J.W. and Neil, C.A. (2015) 'Improving the predictability of distributed stream processors', *Future Generation Computer Systems*, Vol. 52, No. 2742, pp.22–36.

Begoli, E. and Horey, J. (2012) 'Design principles for effective knowledge discovery from big data', *Proc. of the IEEE/IFIP Conference on Software Architecture* (*WICSA*) and *European Conference on Software Architecture* (*ECSA*), pp.215–218.

Big Data (2014) *Seizing Opportunities Preserving Values*, Executive Officer of the President, The White House, Washington, p.85.

Binotto, A.D., Sultanum, N. and Cerqueira, R. (2014) 'Cloud-based remote visualization of big data to subsurface exploration', *Proc. SPE Large Scale Computing and Big Data Challenges in Reservoir Simulation Conference and Exhibition*, pp.1–6.

Brulé, M.R. (2015) 'The data reservoir: how big data technologies advance data management and analytics in E&P', *SPE Digital Energy Conference and Exhibition*, pp.1–7.

Cameron, D. (2014) 'Big data in exploration and production: silicon snake-oil, magic bullet, or useful tool?', *SPE Intelligent Energy Conference & Exhibition*, pp.1–9.

Carlsson, G. (2009) 'Topology and data', *Bulletin of the American Mathematical Society*, Vol. 46, No. 2, pp.255–308.

Cloud Security Alliance (2013) *Big Data Analytics for Security Intelligence*, 22pp [online] https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Big_Data_Analytics_for_Security_Intelligence.pdf.

Congosto, M., Basanta-Val, P. and Sanchez-Fernandez, L. (2017) 'T-hoarder: a framework to process Twitter data streams', *Journal of Network and Computer Applications*, Vol. 83, pp.28–39.

Crockett, B. and Kurrey, K. (2014) 'Smart decision making needs automated analysis: making sense out of big data in real-time', *Proc. of the SPE Intelligent Energy Conference and Exhibition*, pp.1–8.

Feblowitz, J. (2013) 'Analytics in oil and gas: the big deal about big data', *Proc. of the SPE Digital Energy Conference*, pp.1–6.

Frenzel, J.C. (2003) 'Data security issues arising from integration of wireless access into healthcare networks', *Journal of Medical Systems*, Vol. 27, No. 2, pp.163–175.

Gandomi, A. and Haider, M. (2015) 'Beyond the hype: big data concepts, methods, and analytics', *International Journal of Information Management*, Vol. 35, No. 2, pp.137–144.

Griffith, E. (2014) 'Search: how do I punch this rivet hole?', *Fortune*, 27 October [online] http://fortune.com/2014/10/09/wearable-technology-blue-collarjobs/ (accessed 31 December 2014).

Hamzeh, H. (2017) *Application of Big Data in Petroleum Industry* [online] https://www.academia.edu/27175563/Application_of_Big_Data_in_Petroleum_Industry (accessed 20 July 2016).

Intel IT Center (2015) *Big Data in the Cloud: Converging Technologies* [online] https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/big-data-cloud-technologies-brief.pdf.

Jeske, M., Grüner, M. and Weiß, F. (2013) 'Big data in logistics: a DHL perspective on how to move beyond the hype', *Trend Research* [online] https://www.supplychain247.com/paper/big_data_in_logistics_a_dhl_perspective_on_how_to_move_beyond_the_hype.

Komorowski, M. (2014) *A History of Storage Costs* [online] http://www.mkomo.com/cost-per-gigabyte (accessed 19 May 2014).

Li, X., Ma, J., Wang, S. and Zhang, X. (2015) 'How does Google search affect trader positions and crude oil prices?', *Economic Modelling*, Vol. 49, pp.162–171.

Lohr, S. (2012) 'The age of big data', *New York Times* [online] http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html.

Lv, Z., Song, H., Basanta-Val, P., Steed, A. and Jo, M. (2017) 'Next-generation big data analytics: state of the art, challenges, and future research topics', *IEEE Transactions on Industrial Informatics*, Vol. 13, No. 4, pp.1891–1899.

Martínez-Prieto, M.A., Cuesta, C.E., Arias, M. and Fernández, J.D. (2015) 'The solid architecture for real-time management of big semantic data', *Future Generation Computer Systems*, Vol. 47, pp.62–79.

Marz, N. and Warren, J. (2015) *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*, 1st ed., Manning Publications, New York, USA, 328pp.

Nicholson, R. (2012) 'Big data in the oil & gas industry', *IDC Energy Insights*, 20pp [online] https://www-01.ibm.com/events/wwe/grp/grp037.nsf/vLookupPDFs/RICK%20-%20IDC_Calgary_Big_Data_Oil_and-Gas/%24file/RICK%20-%20IDC_Calgary_Big_Data_Oil_and-Gas.pdf

Ouellette, J. (2013) 'The mathematical shape of things to come', *Quanta Magazine*, pp.1–9 [online] https://www.quantamagazine.org/the-mathematical-shape-of-big-science-data-20131004/.

Perrons, R.K. and Jensen, J.W. (2015) 'Data as an asset: what the oil and gas sector can learn from other industries about 'big data'', *Energy Policy*, Vol. 81, pp.117–121.

Pettinger, C.B. (2014) 'Leading indicators, culture and big data: using your data to eliminate death', *Proc. of the ASSE Professional Development Conference and Exposition*, pp.1–8.

Saenz, M.J. and Revilla, E. (2014) 'Creating more resilient supply chains', *MIT Sloan Management Review*, Vol. 55, No. 4, pp.22–24.

Schlegel, G.L. and Trent, R.J. (2014) *Supply Chain Risk Management: An Emerging Discipline*, 1st ed., CRC Press, Taylor & Francis Group, London; New York, 334pp.

Schönberger, V.M. and Cukier, K. (2014) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt, Boston, Massachusetts, 272pp.

Shin, D. (2014) 'A socio-technical framework for internet-of-things design', *Telematics and Informatics*, Vol. 31, No. 4, pp.519–531.

Shin, D.H. and Choi, M.J. (2015) 'Ecological views of big data: perspectives and issues', *Telematics and Informatics*, Vol. 32, No. 2, pp.311–320.

Singh, S., Pandey, S., Shankar, R. and Dumka, A. (2015) 'Application of big data analytics to optimize the operations in the upstream petroleum industry', *Proc. 2nd International Conference on Computing for Sustainable Global Development* (*INDIACom*), pp.1074–1079.

Tan, K.H., Ortiz-Gallardo, V.G. and Perrons, R.K. (2016) 'Using big data to manage safety-related risk in the upstream oil & gas industry: a research agenda', *Energy Exploration & Exploitation*, Vol. 34, No. 2, pp.282–289.