Using Big Data to manage safety-related risk in the upstream oil & gas industry: a research agenda

Sustainable Production and Consumption

Kim Hua Tan [a], Víctor G. Ortiz-Gallardo [b], Robert K. Perrons [c, d, 1]

http://dx.doi.org/10.1016/j.spc.2016.07.003

[a] Nottingham Business School, Jubilee Campus, Nottingham, NG8 1BB, United Kingdom

[b] Mexican Petroleum Institute, Eje Central Lazaro Cardenas Norte 152,
Col. San Bartolo Atepehuacan, C.P. 07730, Mexico City D.F., Mexico

[c] Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia
phone +61-(0)7-3138-2648, e-mail perrons@alum.mit.edu

[d] Centre for Strategy and Performance, University of Cambridge, United Kingdom

Abstract Despite considerable effort and a broad range of new approaches over the years, the upstream oil & gas industry has been frustrated by the sector's stubbornly high rate of injuries and fatalities. This short communication points out, however, that the industry may be in a position to make considerable progress by applying "Big Data" analytical tools to the large volumes of safety-related data that have been collected by these organizations. Towards making this case, we examine existing safety-related information management practices in the upstream oil & gas industry, and specifically note that data in this sector often tends to be highly customized, difficult to analyze using conventional quantitative tools, and frequently ignored. We then contend that the application of new Big Data kinds of analytical techniques could potentially reveal patterns and trends that have been hidden or unknown thus far, and argue that these tools could help the upstream oil & gas sector to improve its injury and fatality statistics. Finally, we offer a research agenda towards accelerating the rate at which Big Data and new analytical capabilities could play a material role in helping the industry to improve its health and safety performance.

Key words: oil & gas, safety, Big Data, HSE

[1] Corresponding author.

1. INTRODUCTION

Despite a broad range of ideas and approaches over the years to solve the problem (e.g., Cramwinckel and Thummarukudy, 2000; Flin *et al.*, 1996), the upstream oil & gas industry has found it difficult to reduce stubbornly high fatality and injury rates among personnel (Curlee *et al.*, 2005). In 2012, there were a total of 138 fatalities within the sector in the U.S. alone (King, 2013) and, because of the uncommonly global nature of the industry (Goldstein, 2009; Hatakenaka *et al.*, 2006; Yergin, 1991), this unfortunate trend extends throughout the sector's operations around the world. By comparison, the industry's on-the-job fatality rate in the U.S. is approximately 7.6 times higher than the national average (King, 2013).

A major part of the oil & gas industry's strategy towards mitigating these kinds of health, safety, and environment (HSE) incidents is to measure them in impressive detail. Vast amounts of data have been collected about these accidents (Veley, 2002) in the hope that this will improve the sector's ability to spot trends and discover patterns that can shed light on potential solutions. But early attempts at finding these high-level trends have been thwarted by three factors. First, while some of the industry's HSE data has been coded and categorized into numerical data such as lost-time incident statistics, much of it has been captured as written responses that were later converted to text (Campbell *et al.*, 2012). Second, a significant fraction of this historical data resides within highly customized applications and bespoke files whose structures are relatively unique. As DeVol (2004) suggests, many of the people collecting this data had to:

> … rely on either custom-built data systems or use a system of e-mails and spreadsheets or simple standalone database applications for HSE incident reporting and information management. Incorporating injury, environmental, property/equipment damage, and vehicle incident reporting with medical case management, industrial hygiene monitoring, auditing and inspections, investigation results with corrective actions and insurance claims management into a single integrated system was not possible or worth the investment to build (p. 1).

Third, the highly customized and fragmented nature of HSE data within the sector has created a problem that one operator refers to as "dark data"—that is, information collected during the course of business that remains in archives that frequently do not garner much attention, or that are not generally accessible or structured sufficiently for analysis (Akoum and Mahjoub, 2013).

This paper explains how new tools and approaches unfolding within the Big Data revolution could be applied to data within the upstream oil & gas industry, and puts forward five specific research questions that, if answered, could materially improve the sector's HSE performance.

2. BIG DATA AND ITS APPLICATIONS FOR MANAGING INDUSTRIAL RISKS

Large data sets captured by digital devices and application software have been successfully used by managers to gain valuable insights into market, product, and consumer behavior (Mayer-Schönberger and Cukier, 2013). Many firms have been able to leverage Big Data to increase operational efficiency, inform strategic direction, bring about better customer service, develop new products and services, and identify new markets (Demirkan and Delen, 2013; Fulgoni, 2013; Lohr, 2012a). These data-driven decisions have, in turn, enabled firms to create new and inventive types of competitive advantage for themselves (Davenport *et al.*, 2012).

There are also clear opportunities for firms—including many of those in the oil & gas industry—to leverage Big Data and predictive analytics to mitigate operational kinds of risk. Managers now recognize the need to understand and reduce risks to infrastructure, industrial operations, and the supply chain. Big Data and its related technologies can potentially act as a catalyst to transform the scientific use of environmental information so that it can be applied towards risk mitigation. Table 1 shows the list of Big Data applications in risk mitigation reported in the literature.

Several complementary technologies have also made impressive gains in the past few years that are already starting to make significant impacts in how large volumes of data can be collected 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 (Griffith, 2014). These wearable devices will enable experts to help less experienced industrial workers in oil & gas operational settings all around the world. For example, a relatively inexperienced worker at an oil & 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 center staffed by seasoned veterans. The more experienced personnel can then 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.

[Insert Table 1 about here.]

3. STEERING THE INDUSTRY'S BIG DATA AGENDA TOWARDS HSE

By most accounts, the oil & gas industry's data is already "big." Modern oil & gas seismic data centers can easily contain as much as 20 petabytes[2] of information, which is roughly equivalent to 926 times the size of the U.S. Library of Congress (Beckwith, 2011). And while the industry's seismic data sets have been notoriously large and cumbersome for a long time, many of the operational aspects of the oil & gas industry are also generating significantly more data than they used to (Perrons, 2010).

Like many industries, the upstream oil & gas sector has seen a flurry of initiatives and high-profile publications (Anand, 2013; Beckwith, 2011) about Big Data, which have in turn translated into significant discussion about this topic within industry conferences (e.g., Feblowitz, 2013) and among practitioners. Critics of Big Data caution that the transformational potential of these analytical capabilities may be somewhat oversold and misunderstood (Harford, 2014; Lohr, 2012b),[3] but the oil & gas sector has already been

[2] This statistic is less impressive when you consider that Walmart, the U.S. retail giant, collects more than 2.5 petabytes of data every hour from customer transactions (McAfee and Brynjolfsson, 2012).

[3] Gary King, the director of Harvard University's Institute for Quantitative Social Science, goes as far as predicting that there "is no area that is going to be untouched" by Big Data (Lohr, 2012a)

noticeably impacted by several of the technologies underpinning these changes (e.g., Perrons and Jensen, 2015).

However, these inroads have largely been focused on more technical parts of the business such as reservoir characterization and drilling optimization (e.g., Akoum and Mahjoub, 2013; Holdaway, 2014). With the notable exception of a handful of extremely high-level conference papers (Batterson and Iovino, 2014; Pettinger, 2014), there have been few significant inroads in the application of Big Data specifically to the HSE-related parts of the industry. We therefore submit:

> **Research Question 1: What barriers—technical, economic, or organizational—need to be overcome for Big Data technologies to be applied to the oil & gas industry's HSE-related challenges with the same sense of urgency seen in more technical parts of the sector?**

Also, because Big Data and cloud computing are highly interconnected technologies that are evolving in parallel (e.g., Bochman, 2013; Bughin *et al.*, 2010), the upstream oil & gas sector's tepid advances toward the cloud (Perrons, 2015) are also highly relevant to its evolutionary trajectory with regards to Big Data. Companies in the industry have made relatively modest inroads into cloud computing by frequently preferring private clouds instead of public ones (Feblowitz, 2011; Perrons and Hems, 2013) or opting for hybridized solutions that mix cloud and existing non-cloud IT resources (Mathieson and Triplett, 2011). We believe that this particular obstacle is sufficiently important to warrant:

> **Research Question 2: How will the oil & gas industry's relatively slow adoption of cloud computing impede its uptake of Big Data technologies, and what can be done about it?**

Also, the recurring problem of custom-built data systems and spreadsheet-based databases for HSE reporting have created a highly fragmented body of data (DeVol, 2004) that is difficult to consolidate into the massive data sets that have revealed new and insightful relationships in other parts of the marketplace. Several other sectors—most notably, the healthcare, financial, retail, and media industries—have come to realize that new and valuable insights are frequently gleaned from using new techniques to analyze massive data sets in ways that were never possible with smaller ones (Mayer-Schönberger and Cukier, 2013). Thus:

> **Research Question 3: What is the most efficient way to bring together the industry's myriad HSE-related data sets so that Big Data kinds of analytical tools and approaches can be brought to bear on this area?**

The possibility of applying new tools like those shown in Figure 1 and Big Data methods to the HSE domain presents an interesting problem, however: reported accidents typically involve people, and vital data from personnel records that might be able to shed considerable light on HSE types of phenomena are frequently more protected than many other types of data (Frenzel, 2003; Lasprogata *et al.*, 2004).

[Insert Figure 1 about here.]

Potentially life-saving new insights may also be lurking in other non-personnel kinds of data, but the high degree of human factors at play in the industry's HSE challenges (Gordon, 1998) clearly make personnel and human resources information sources a particularly fruitful place to look for data that might be linked to HSE incidents. Accordingly:

> **Research Question 4: What barriers—legal, political, or organizational—need to be overcome for Big Data technologies to be applied to the oil & gas industry's personnel records en route to discovering HSE-related trends?**

Finally, the literature has long sought to understand the link between HSE-related investments and safety performance (e.g., Zacharatos *et al.*, 2005). Do some investment strategies yield better results than others in terms of their ability to reduce accidents and injuries? As noted earlier, Big Data technologies open up the possibility of joining organizations' vast HSE data records with their financial and investment data to shed more light on the potency of different investment strategies in this domain. Thus:

> **Research Question 5: How can Big Data bring together HSE and investment data sets to reveal the relative success or failure of different HSE-related investments?**

4. CONCLUSIONS

This short communication explained how the upstream oil & gas industry has found it difficult to reduce stubbornly high fatality and injury rates among personnel, and examined existing HSE-related data management practices within the sector. In particular, we noted that data in this industry often tends to be highly customized, difficult to analyze using conventional quantitative tools, and frequently ignored. We argued that the application of

new Big Data kinds of analytical techniques could potentially reveal patterns and trends that have been hidden or unknown thus far, and suggested that these tools could help the upstream oil & gas sector to improve its injury and fatality statistics. The paper then offered a research agenda towards accelerating the rate at which Big Data and new analytical capabilities could play a material role in helping the industry to improve its health and safety performance.

It is our fervent hope that this contribution will be a catalyst for further research in the areas discussed here. The upstream oil & gas industry has been characterized as "the world's biggest and most pervasive business" (Yergin, 1991, p. 13) and, because of its uncommon size and scale, any breakthroughs that can be made with regards to the research questions put forward here would almost certainly translate to the saving of many lives and the avoiding of many serious accidents. Thus, while the research agenda suggested in this paper may potentially be interesting from a theoretical point of view, the practical implications of getting this right—that is, of successfully applying Big Data to real-world HSE problems in the upstream oil & gas sector—would be far greater still.

REFERENCES

Akoum M., and Mahjoub A., 2013. *A Unified Framework for Implementing Business Intelligence, Real-Time Operational Intelligence and Big Data Analytics for Upstream Oil Industry Operators*. Paper presented at the Society of Petroleum Engineers (SPE) Middle East Intelligent Energy Conference and Exhibition, Dubai, UAE, October 28-30. Paper number SPE-167410.

Anand P., 2013. Guest Editorial: Big Data is a Big Deal. *Journal of Petroleum Technology* **65**(**4**), 18-21.

Batterson L.M., and Iovino C., 2014. *The Era of Big Data: What Does it Mean for a Safety Professional?* Paper presented at the American Society of Safety Engineers (ASSE) Professional Development Conference and Exposition, Orlando, Florida, USA, June 8-11. Paper number ASSE-14-612.

Beckwith R., 2011. Managing Big Data: Cloud Computing and Co-Location Centers. *Journal of Petroleum Technology* **63**(**10**), 42-45.

Bochman A., 2013. Securing Mobile, Big Data and the Cloud in the Energy Sector. Retrieved 7 January 2014 from http://www.greentechmedia.com/articles/read/securing-mobile-big-data-and-the-cloud-in-the-energy-sector

Bughin J., Chui M., and Manyika J., 2010. Clouds, Big Data, and Smart Assets: Ten Tech-Enabled Business Trends to Watch. *McKinsey Quarterly* **56**(**1**), 75-86.

Campbell H., Polo J., and Bouly G., 2012. *HSE Management System: Keep It Simple!* Paper presented at the Society of Petroleum Engineers (SPE) International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production, Perth, Australia, September 11-13. Paper number SPE-156739.

Cramwinckel J., and Thummarukudy M., 2000. *Development of an ISO 14001 Certified HSE Management System in a Large E&P Company*. Paper presented at the Society of Petroleum Engineers (SPE) International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production, Stavanger, Norway, June 26-28. Paper number SPE-60990.

Curlee C.K., Brouillard S.J., Marshall M.L., Knode T.L., and Smith S.L., 2005. *Upstream Onshore Oil and Gas Fatalities: A Review of OSHA's Database and Strategic Direction for Reducing Fatal Incidents*. Paper presented at the SPE/EPA/DOE Exploration and Production Environmental Conference, Galveston, Texas, USA, March 7-9. Paper number SPE-94416-MS.

Davenport T.H., Barth P., and Bean R., 2012. How "Big Data" is Different. *MIT Sloan Management Review* **54**(**1**), 22-24.

Demirkan H., and Delen D., 2013. Leveraging the Capabilities of Service-Oriented Decision Support Systems: Putting Analytics and Big Data in Cloud. *Decision Support Systems* **55**(**1**), 412-421.

DeVol E.C., 2004. *Deriving Value from Enterprise Wide HSE Data*. Paper presented at the Society of Petroleum Engineers (SPE) International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production, Calgary, Alberta, Canada, March 29-31. Paper number SPE-86807-MS.

Feblowitz J., 2011. Oil and Gas: Into the Cloud? *Journal of Petroleum Technology* **63**(**5**), 32-33.

Feblowitz J., 2013. *Analytics in Oil and Gas: The Big Deal about Big Data*. Paper presented at the SPE Digital Energy Conference and Exhibition 2013 at The Woodlands, Texas, USA from 5-7 March 2013. Paper number SPE 163717.

Flin R., Mearns K., Gordon R., and Fleming M., 1996. Risk perception by offshore workers on UK oil and gas platforms. *Safety Science* **22**(**1**), 131-145.

Frenzel J.C., 2003. Data security issues arising from integration of wireless access into healthcare networks. *Journal of Medical Systems* **27**(**2**), 163-175.

Fulgoni G., 2013. Big Data: Friend or Foe of Digital Advertising? Five Ways Marketers Should Use Big Data to Their Advantage. *Journal of Advertising Research* **53**(**4**), 372-376.

Goldstein A., 2009. New Multinationals from Emerging Asia: The Case of National Oil Companies. *Asian Development Review* **26**(**2**), 26-56.

Gordon R.P., 1998. The contribution of human factors to accidents in the offshore oil industry. *Reliability Engineering & System Safety* **61**(**1**), 95-108.

Griffith E., 2014. Search: How Do I Punch This Rivet Hole? *Fortune*, Oct. 27 issue. Retrieved on 31 December 2014 from http://fortune.com/2014/10/09/wearable-technology-blue-collar-jobs/.

Harford T., 2014. Big Data: Are We Making a Big Mistake? *Financial Times*, March 28 issue. Retrieved from http://www.ft.com/intl/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz2xV4TdTmn

Hatakenaka S., Westnes P., Gjelsvik M., and Lester R.K., 2006. *From "Black Gold" to "Human Gold": A Comparative Case Study of the Transition from a Resource-Based to a Knowledge Economy in Stavanger and Aberdeen.* Report no. MIT-IPC-06-004.

Cambridge, Massachusetts, USA: Massachusetts Institute of Technology Industrial Performance Center.

Holdaway K.R., 2014. Harness Oil and Gas Big Data with Analytics. John Wiley & Sons, Hoboken, New Jersey.

Jeske M., Grüner M., and Weiß F., 2013. *Big Data in Logistics: A DHL Perspective on How to Move Beyond the Hype*. Report issued by DHL Customer Solutions & Innovation, Germany. Retrieved on 3 January 2015 from http://www.dhl.com/content/dam/downloads/g0/about_us/innovation/CSI_Studie_BIG_DATA.pdf

King P., 2013. Oil and Gas Deaths Reached Record High in 2012. *EnergyWire*, August 23 issue. Retrieved from http://www.eenews.net/stories/1059986375

Lasprogata G., King N.J., and Pillay S., 2004. Regulation of electronic employee monitoring: Identifying fundamental principles of employee privacy through a comparative study of data privacy legislation in the European Union, United States and Canada. *Stan. Tech. L. Rev.* **2004**, p. 4.

Lohr S., 2012a. The Age of Big Data. *New York Times*, February 11 issue. Printed on p. SR1 of the New York edition of the newspaper. Retrieved from http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html

Lohr S., 2012b. How Big Data Became So Big. *New York Times*, August 11 issue. Also published on p. BU3 of the August 12 issue of the New York Edition with the headline "Amid the Flood, a Catchphrase is Born.". Retrieved from http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html

Mathieson D., and Triplett C., 2011. Are There Clouds in Our Blue Sky Research Programs? *Journal of Petroleum Technology* **63**(**9**), 16-18.

Mayer-Schönberger V., and Cukier K., 2013. Big Data: A Revolution That Will Transform How We Live, Work and Think. John Murray Publishers, London, United Kingdom.

McAfee A., and Brynjolfsson E., 2012. Big Data: The Management Revolution. *Harvard Business Review* **90**(**10**), 60-68.

Perrons R.K., 2010. Perdido Ties Together Shell Digital Oilfield Technologies. *World Oil* **231**(**5**), 43-49.

Perrons R.K., 2015. How the Energy Sector Could Get It Wrong with Cloud Computing. *Energy Exploration & Exploitation* **33**(**2**), 217-226.

Perrons R.K., and Hems A., 2013. Cloud Computing in the Upstream Oil & Gas Industry: A Proposed Way Forward. *Energy Policy* **56**, 732-737.

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* **81**, 117-121.

Pettinger C.B., 2014. *Leading Indicators, Culture and Big Data: Using Your Data to Eliminate Death*. Paper presented at the American Society of Safety Engineers (ASSE) Professional Development Conference and Exposition, Orlando, Florida, USA, June 8-11. Paper number ASSE-14-689.

Sáenz M.J., and Revilla E., 2014. Creating More Resilient Supply Chains. *MIT Sloan Management Review* **55**(**4**), 22-24.

Schlegel G.L., and Trent R.J., 2014. Supply Chain Risk Management: An Emerging Discipline. CRC Press, Boca Raton, Florida, USA.

Veley C.D., 2002. *Applying a New HSE Measurement System*. Paper presented at the Society of Petroleum Engineers (SPE) International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production, Kuala Lumpur, Malaysia, March 20-22. Paper number SPE-74050.

Yergin D., 1991. The Prize: The Epic Quest for Oil, Money, and Power. Touchstone, New York.

Zacharatos A., Barling J., and Iverson R.D., 2005. High-Performance Work Systems and Occupational Safety. *Journal of Applied Psychology* **90**(**1**), 77.

Areas	Risk Mitigation Examples	References
Manufacturing Operations	Raythem uses built-in sensors on assets to have a real-time monitoring of manufacturing assembly operations. For example, only qualified operators are allowed to carry out certain operations. To ensure quality and reduce operations risks, the sensor will also limit the number of screw turns for certain assembly operations.	Schlegel and Trent (2014)
Logistics	DHL uses Big Data for early detection of potential risks (weather conditions, influenza outbreaks) in supply chains. DHL provides customers with an overview of potential disruptions of their individual supply chains.	Jeske, Grüner, and Weiß (2013)
Supply chain	Cisco uses Big Data to integrate risk awareness into the product and the value chain. The approach enables Cisco to anticipate emergencies by protecting important segments of the supply chain with built-in resiliency and levers to pull when a disruption occurs.	Sáenz & Revilla (2014)
Oil and Gas	A hands-free checklist, which workers can follow while they are assembling equipment, can save time and reduce risks arising due to mistakes.	Griffith (2014)

Table 1: Examples of Big Data Risk Mitigation

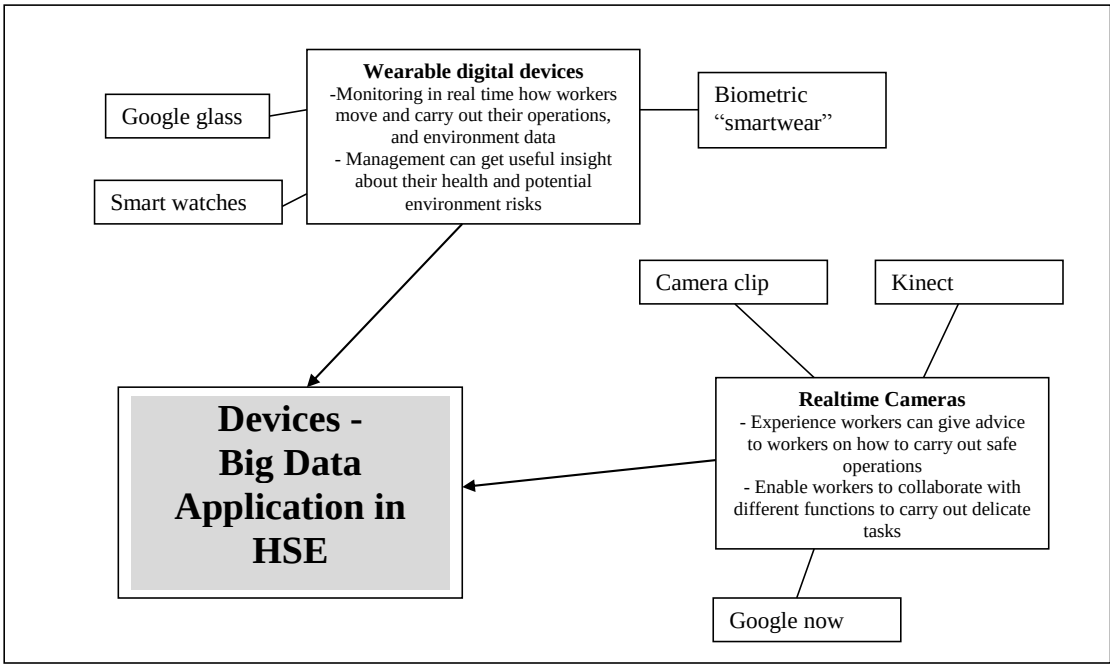

Figure 1: Big Data Devices for HSE