On Digital Opencast Mining Ecosystems (DOME) and Knowledge Management – a Big Data Perspective

Shastri L Nimmagadda*
School of Information Systems
Curtin Business School
Perth, WA, Australia

Veemelia V Murupindy Veenaikar
Department of Spatial Sciences
Curtin University
Perth, WA, Australia

Torsten Reiners
School of Information Systems
Curtin Business School
Perth, WA, Australia

**presenting author asterisked*

SUMMARY

Many opencast mines inhabit thousands of square km area, which are productive and commercial Australia wide. Hundreds of volumes and varieties of data dimensions and facts exist in the opencast mining areas. The data sources linked with various opencast mines are often heterogeneous and multidimensional. Data modelling is challenging in a Big Data scale, at times precluding the data integration process. The mineralization connected to opencast mines occurs in shafts, pit slopes, ramps and benches with varying geometries and configurations in large-scale geographic and periodic dimensions. The limits of the mineralization at places are either unknown and or ambiguously interpreted. The Big Data, in the context of the Australian mining industry, are due to the explosive growth of data sources and their uncontrolled management in many national and multinational companies. New knowledge is required for interpreting new opencast mining areas and their mineralization. For sustainable production, the knowledge of the connectivity between mineralization and its associated opencast mines is constrained. We propose an empirical modelling, analysing hundreds of attribute dimensions and fact instances of geological and geophysical vintages in the mining areas. Different data constructs and models are built for logical metadata, accommodating it in a multidimensional warehouse repository, as a DOME solution. It is an innovative solution to the mining industry's Big Data problem including the opencast mine planning and design, adding values to the existing domain knowledge with new interpretations. Various geological events attributed to the interpretation and distribution of mineralization are useful for the opencast mine managers.

Key words: Big Data, Digital Ecosystems, Opencast Mining, Sustainability, Knowledge Management

INTRODUCTION

Lack of knowledge on the opencast mines, skilful interpretation of their boundaries and the connectivity between opencast benches and mine mineralization (Aitken and Reid, 2000) including the absence of mining boom, the exploration and production of many mines have been held up in various regions in Australia. Besides, large volumes and varieties of data associated with opencast systems are underutilized and the way explorers interpret, perceive and extract knowledge from various events of mining geology and geophysics (Catchpole and Robins, 2013) has not been tactical. Innovative database approaches, such as data warehousing and mining (Anahory and Murray, 1997) are proposed to address the data integration, interoperability and the systems' connectivity with a quest to improve the mine mineralization. The existing data management practices (Berson and Smith, 2004 and Nimmagadda et al. 2014) and other application scenarios are not compatible in the prevailing in opencast mine contexts. We articulate various new data schemas and integrated frameworks as a measure of systems' development process (Damiani, 2008 and Dhar et al. 2014) for mineral exploration and reenergizing the mining industries (Catchpole and Robins, 2013). As illustrated in Fig. 1, there may be many bench systems associated with opencast and geomorphic systems, in which a sustainable management is needed through increased understanding of the systems' connectivity in geographically spread unstructured data sources.

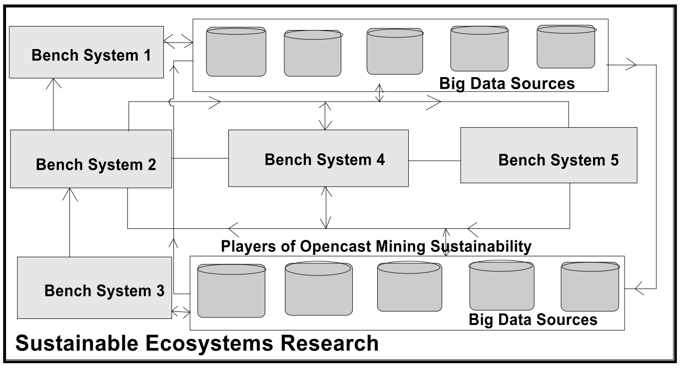

Fig. 1: Sustainable ecosystems research – Opencast mining – building a case for DOME

For exploring the systems' connectivity, the domain ontologies (Aitken and Reid, 2000 and Nimmagadda and Dreher, 2012) are built with likely development of an ontology-based data warehouse repository, an integrated approach. Various surveys-mines-permits' data sources are considered in the multidimensional modelling process. Hundreds of dimensions and attribute instances are identified and at places conceptualized and contextualized because of the existence of unknown mine attributes and inherent data quality challenges. Besides, geological and geophysical (G & G) survey domains comprise of millions of geographic point, line and areal contour dimensions and fact data instances that vary with space and time, including data instances pertained to many associated opencast mining data entities. The goal of the research is to build constructs and models for simulating a digital opencast mining ecosystem (DOME) articulations representing computer simulated opencast mining areas and generate attribute cubes for visualization and interpretation. The motivation is the existence volumes and varieties of hugely spread and geologically complex mining areas, interpreted in Big Data size and scale. To implement the DOME solutions in various application scenarios of mining industry novel and holistic data modelling methodologies are needed including strategies of data integration from multiple domains of opencast mines.

METHOD AND RESULTS

We introduce a design science information system (DSIS) a proven information system (IS), an analytical approach by staging purposeful artefacts, addressing the ecosystems' issues and evaluating them through a variety of utility data properties (Venable et al. 2016). We relate to the ontological descriptions and their role in the digital ecosystems' modelling. We aim at evolving artefacts in the DSIS driven Multidimensional Warehouse Repository for data mining, visualization and data interpretation for new knowledge discovery. Data mining and interpretation of cognitive patterns hidden in trillions of exploration data that combined with geospatial Big Data are critical goals of the geo-informatics. Extracting useful knowledge on favorable geological-structures that trap the mine mineralization and commercial ores of the opencast mining ecosystems has implications on the application of Big Data technology in the mining business.

Further, we reiterate the DSIS is a simulation of the digital opencast mining ecosystem (DOME) for which various domain ontologies and their descriptions are made for opencast mining multidimensional data sources in different geographic dimensions. There are many productive mines worldwide, each with thousands of sq. km. of areal extents with complex topographies and geographies. Many minerals and mines are under active exploration for many decades in Western Australia, and more than 50 percent contribute to the Australian economy. Extracting new knowledge and information from volumes and varieties of data sources is crucial for sustainable opencast mining management and investigating the connectivity among opencast benches and mineralization. With the advent of new concepts and technologies, innovative ideas emerge in documenting and managing mines' databases, especially in areas, where thousands of mining pits, G & G vintages coexist, and their logical integrated interpretation becomes necessitated. Integration of Big Data tools (Chen et al. 2012) with the proposed framework, from which we explore usefulness in connecting various systems in the DOME contexts. The holistic DSIS approach and their artefacts that guide the DOME ensure exploring connections among various mining pits, benches and making the opencast systems more sustainable and productive (Chatpole and Robins, 2013). A star schema modelled in Fig. 2, demonstrates the connectivity between open pit data dimensions and other associated attributes of the mining areas. In database perspective, there are several such schemas based on the size of opencast mining provinces and connectable dimension attributes with fact tables with one-to-many data relationships. Other possible data relationships can also be explored based on business rules and constraints set in the modelling process.

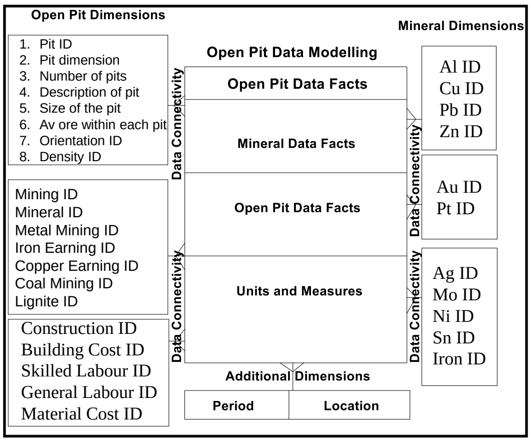

Fig. 2: A schema drawn for making connections between attribute dimensions

Big Data Role in the Mining Operations

As we identify various features of the Big Data "volume, variability, velocity, visualization, veracity and value" (Agarwal et al. 2014 and Chandrashekaran et al. 1999) from multiple data sources, hundreds of attribute dimensions including conceptualized and contextualized attributes emerge from large volumes in the star-schema dimensional modelling. For the systems dealing with the Big Data of the opencast mining exploration industries, the field of geo-informatics plays an inclusive role in the study of fundamental geological problems owing to the exponential explosion of sequence and structural information with time and geography (Nimmagadda, 2015). There are two major challenging areas in geo-informatics: data management and knowledge discovery.

Extracting new knowledge on favorable geological-structures that hold mineralization and commercial ores of the opencast mining ecosystems have implications on the feasibility and applicability of Big Data analytics in the mining business. To substantiate the reasoning of Big Data, we demonstrate the existence of multiple systems (Fig. 1) within a single ecosystem. As shown in Fig. 3, data from multiple domains and systems link with bench geometries of the opencast mining and associated geomorphic ecosystems. The economic viability and the sustainability of the mineralization in the mining areas are assessed with the systems connectivity. Sustainability in which case may depend on the production of mineralization with periodic time dimension (Downes et al. 2014), each mining district of a particular system or systems can produce minerals for longer periods without any interruption. But commercially, the sustainable exploration and subsequent drilling campaigns can help make discoveries or continue to know the limits of the existing opencast mining systems or add new mineral reserves to the existing ones (Mineral and Energy Resource Exploration, 2012).

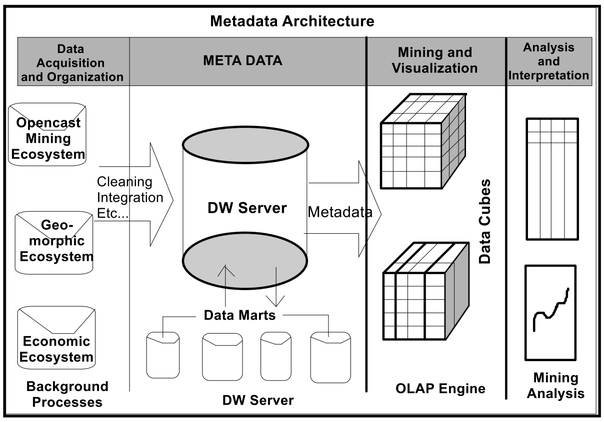

Fig. 3: Data warehouse architecture for integrating opencast ecosystems

Big Data Implementation

Once the warehoused metadata is ready, the data storage and computed data views for visualization and interpretation of new knowledge and its management are assessed. Onsite workstations use various real time high-performance computing and data processing capabilities. For implementing the integrated framework in the current contexts, we evaluate again the Big Data tools, regarding the storage, data processing, and other computing requirements. The Big Data storages (for volumes and varieties) and their requirements in mining industries are different, compared with conventional resources' business storage systems. The data structures continuously evolve in mining areas with data varieties, multiple data types, and their data relationships. Based on the types of structuring, data storage, and retrieval methods, we consider the scalability, extensibility and compatibility criteria. Onsite workstations use various computing and data processing capabilities. The grids are computed, and map views are extracted from digital ecosystems' metadata that may consist of several opencast mining data views and mineralization areas. In this context, the cloud clusters capture the required data from multiple opencast mines in large-scale geographic dimensions. The clouds, categorizing the specific clusters are migrated to our workstations to build models and integrate them into warehouse metadata structures. We uninterruptedly monitor the G & G data qualities and their veracity on workstations, for better visualization and metadata interpretation, ensuring that the new knowledge obtained in various DOME contexts is implementable in various geographic contexts. An implementation of such digital ecosystem is demonstrated through the data integration process flowchart as shown in Fig. 4. The Big Data volumes and varieties are extracted from the opencast mining ecosystem as demonstrated in Fig. 4. The G & G and mining data sources are integrated into the warehouse environment to generate metadata cubes (Pujari, 2002 and Zhong et al. 1996). These data cubes are further explored for mineralization connections through visualization, interpretation and multiple opencast mining systems in real time. Data schemas ensure representing bench geometries (BG), mineral occurrence (MO) areas and mining logistics (ML) from the Australian situations and integrate them in a MO-BG-ML data instances' warehouse (Fig. 4). Data mining (Pujari, 2002), visualization and interpretation are performed on interactive interpretation workstations in real time.

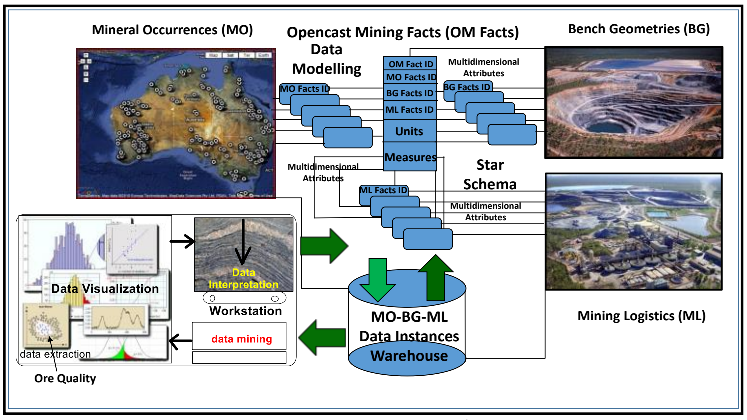

Fig. 4: DOME implementation framework

In the opencast provinces, data schemas are generated from bench geometries (BG), mineral occurrence (MO) areas and mining logistics (ML) from the Australian situations and integrate them in a MO-BG-ML data instances' warehouse. The third application domain is geomorphic information system and its management (Nimmagadda and Dreher, 2009) for interpreting the areal extents and overburden rock properties of the earth's surface. The issue of controlling large-scale opencast mining through ripping of overburdened host rocks is another challenging application. The Big Data integration, modelling semantics base conceptualized data relationships among multiple attribute dimensions, and mining of several interpretable data views of quality geological and geomorphic structures, effective for holding the mineralisation are key areas of the current DOME research applications. Several large bubbles, having Cu grades of large pits, the structural topography of mining pits, bench grades, and bench frequency attributes exhibit several clusters, as interpreted in the mining areas (Zhong et al. 1996). As shown in Fig. 5 the bubble plot views are immensely useful for mining professionals and mineral asset managers.

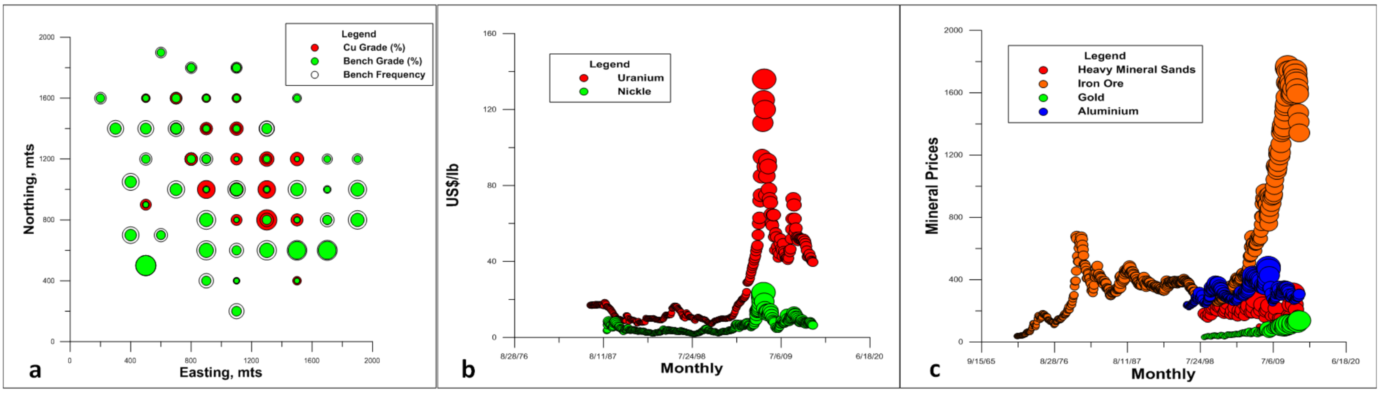

Fig. 5: The bubble plot views (extracted from DOME metadata) of mineral grades and between mineral pricing with periodic dimension attribute

CONCLUSIONS

The methodology is effective in integrating and connecting dimensions and their attributes, associated with the structure, mineralization and other elements and processes of the opencast mining systems. The DOME methodology ensures a sustainable exploration and production of mineralization in Australian contexts. The fine-grained data structuring, data mining, visualization and interpretation artefacts deduced for the DOME metadata demonstrate their implementation in the opencast mining field areas. Several constructs and models are designed with the Australian contexts with a quest obtaining new knowledge on the connectivity of opencast mines through DOME approach including knowledge of geological structuring associated with the mineralization.

REFERENCES

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

Aitken, S., and Reid, S., 2000, Evaluation of an Ontology-Based Information, Retrieval Tool, Proceedings of ECAI'00. Berlin, Germany.

Anahory, S. and Murray, D. 1997, "Data warehousing in the real world: a practical guide for building decision support systems", Pearson Education, Addison-Wesley, UK, pp. 255-360.

Berson, A. and Smith, J. S., 2004, "Data warehousing, data mining & OLAP", Mc Graw – Hill Education (India) Pty Ltd, pp. 205-219, 221-513.

Catchpole, M. and Robins, W., 2013, MINERAL AND ENERGY RESOURCE EXPLORATION, PRODUCTIVITY COMMISSION ISSUES PAPER, the Australasian Institute of Mining and Metallurgy, March 2013.

Chandrasekaran, B., Johnson, R. and Benjamins, R., 1999, Ontologies: what are they? why do we need, them?. IEEE Intelligent Systems and Their Applications, 14(1), Special Issue on Ontologies, pp. 20– 26.

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

Damiani, E., 2008, Key note address on 'Digital Ecosystems: the next Generation of Service Oriented Internet", IEEE-DEST, Phitsanulok, Thailand, Feb 2008.

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

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

Mineral and Energy Resource Exploration, Productivity Commission Paper, Dec 2012, www.pc.gov.au (under ‘projects’)

Nimmagadda, S.L, and Dreher, H., 2009, Ontology based data warehouse modelling for managing carbon emissions in safe and secure geological storages, a paper presented in the international *SEG symposium* – Imaging and Interpretation, in a forum “science and technology for sustainable development”, held in Sapparo, Japan, Oct 2009

Nimmagadda, S. L. and Dreher, H., 2012, On new emerging concepts of Petroleum Digital Ecosystem, *Journal Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2012, 2 (6): 457–475 doi: 10.1002/widm.1070.

Nimmagadda, S.L. Rudra, A. and Dreher, H.V., 2014, Big Data Information Systems for Managing Embedded Digital Ecosystems (EDE), a book chapter in a book entitled “Big Data and Learning Analytics: Current Theory and Practice in Higher Education” The Springer International, The Netherlands.

Nimmagadda, S.L., 2015, Data Warehousing for Mining of Heterogeneous and Multidimensional Data Sources, Verlag Publisher, Scholar Press, OmniScriptum GMBH & CO. KG, p. 1-657, Germany.

Pujari, A.K. 2002, “Data mining techniques”, University Press (India) Pty Limited, Hyderabad, India.

Venable, J. Pries-Heje, J. and Baskerville, R., 2016, FEDS: a Framework for Evaluation in Design Science Research, Research Essay, European Journal of Information Systems (2016) 25, 77-89.

Zhong, T. Raghu, R. and Livny, M., 1996, An efficient data clustering method for very large databases, *Proceedings of ACM SIGMOD* International conference on management of data, ACM, NY, USA**.**