BIG DATA MATURITY MODEL – A PRELIMINARY EVALUATION

Sulfeeza Mohd Drus[1], and Noor Hafizah Hassan[2]

[1]*Universiti Tenaga Nasional, Malaysia,* sulfeeza@uniten.edu.my
[2]*Universiti Tenaga Nasional, Malaysia,* hafizah@uniten.edu.my

ABSTRACT. Due to the advancement of technology, data has been generated and exploited extensively for profit and strategic gains. A tool that helps organizations to assess their competency and capability in big data domain is detrimental for organizational sustainability and enhancement. Despite the numerous big data maturity models (BDMM) available, organizations still have difficulties in selecting and choosing the suitable BDMM to assess their maturity in big data domain. An evaluation of existing BDMMs in terms of the purpose and scope, development methodology and intended usage is needed to enlighten the BDMM adopters. This paper attempts to evaluate seven (7) existing BDMMs from three (3) main perspectives, namely generic attributes, design attributes and application attributes. Even though most of these BDMM are having similar generic attributes, they differ in the intended purpose and their scope of big data maturity assessment.

Keywords: big data, maturity model, big data maturity model

INTRODUCTION

Nowadays, data is ubiquitous; being generated and transferred from many different sources and usually come in varying structures and formats. This phenomenon has resulted in the emergent of 'big data'; whereby almost every organization is seen to concentrate on exploiting data for competitive advantage (Provost and Fawcett, 2013) as well as solving specific business problems (Salleh, Janczewski and Beltran, 2015).

According to Forrester (2015), it is predicted that big data technology market will grow at 12.8% over the next five (5) years, which is also manifested in a study by IDC that projects worldwide revenues for big data and business analytics (BDA) will grow from $130.1 billion in 2016 to more than $203 billion in 2020. Due to the significance of big data in determining organizational well-being, it is very important for organizations to assess their current big data exploitation.

Maturity model can be used for this purpose as it assists organizations to review and assess organizational methods and processes against accepted standards (Moore, 2014). Maturity model can also assess organizational competency, capability and sophistication in big data domain against a set of predefined criteria (de Bruin, Roseman, Freeze and Kulkarni, 2005).

BIG DATA

Big data is voluminous (Lnenicka, 2015; Khare, 2014), collected at high-speed and can be of variety of structures (Lnenicka, 2015) but is usually beyond the technology's capability to store, manage and process efficiently (Kaisler, Armour, Espinosa and Money, 2013); which is used to identify and solve business problems (Khare, 2014). Originally, big data is defined as having 3 V's – *volume*, *velocity* and *variety* (Laney, 2001), but later expanded to more V's – *veracity* and *value* (IBM, n.d). *Volume* is the amount or scale of data available to organizations, either from internal or external sources; *velocity* - the speed of data in terms of creating, storing, retrieving and streaming, or the rate of big data changes; *variety* in terms of diversity of data structures and formats; *veracity* – the validity and accuracy of data and *value* – usefulness of data in providing insights for decision making. Kaisler et al. (2013) add in another characteristic of big data, *complexity* which looks at the degree of interconnectedness and interdependence in big data structures whereby any changes in the big data structure could yield large changes that may affect its behavior.

A 2012 study among companies in UK and Ireland by IBM and Said Business School indicated that 63% of the respondents recognize the competitive advantage associated with big data, whereby 38% of them use big data for customer-centric outcomes and 26% for operations optimization. In a 2014 survey by Accenture, 89% of the respondents agree that "big data will revolutionize the way they do business to a degree similar to the advent of the Internet in the 1990s" and 79% concur that "companies that do not embrace big data will lose their competitive position and may even face extinction". 69% of the respondents indicated that within the next one (1) year they are planning to spend more than $10 billion in big data related investment. Thus, it can be concluded that big data is among the primary agenda of majority organizations in the current borderless and globally competitive marketplace.

BIG DATA MATURITY MODEL

According to Mettler, Rohner and Winter (2010), maturity can be considered as a progressive process exhibiting an entity's ability from some starting point to the desired completion point. In addition, Fraser, Moultrie and Gregory (2002) defined maturity as 'notion of development from some initial state to some more advanced state…. suggesting that the subject may pass through a number of intermediate states on the way to maturity'. Thus, it can be concluded that maturity is an incremental process from some infancy stage to a more sophisticated stage, whereby an entity has to undergo some transitional stages.

In the context of big data, a tool such as maturity model that could help organizations to assess their big data adoption and initiative is imperative for continuous improvement (Mettler et al., 2010) as well as for gaining or retaining competitive advantage (de Bruin et al., 2005). de Bruin et al. (2005) classify maturity model as i) *descriptive* – assesses the 'as-is' situation; ii) *prescriptive* – focuses on relationships to business performance and how maturity improvement can positively affect business value; and lastly iii) *comparative* – enables benchmarking across industries or regions. Although this classification of maturity model can be perceived distinctively, it can also be discerned progressively.

Mettler (2009) identifies three (3) factors or aspects in a maturity model, namely i) *process maturity* – to which extent a specific process is explicitly defined, managed, measured, controlled and effective; ii) *object maturity* - to which extent a particular object reaches a predefined level of sophistication; and iii) *people capability* – to which extent the workforce is able to enable knowledge creation and enhance proficiency.

A number of big data maturity models (BDMM) can be found in literature such as TDWI Big Data Maturity Model, EMC's Big Data Business Model Maturity Index, IBM's Big Data

and Analytics Maturity Model and Hortonworks Big Data Maturity Model. Most of these BDMM are using maturity grids with phases between one (1) and five (5) to assess the big data maturity in an organization.

According to de Bruin et al. (2005), despite the availability of numerous maturity models, "there is little documentation on how to develop a maturity model that is theoretically sound, rigorously tested and widely accepted". This notion is also evident in a study by Mettler et al. (2010) that develops an information systems (IS) maturity model classification system as a means to diminish the difficulties in searching, selecting and retrieving a specific maturity model. This paper attempts to systematically evaluate the existing Big Data Maturity Models (BDMM) by adopting the approaches of de Bruin et al. (2005); Mettler (2009); and Mettler et al. (2010). It is hoped that this systematic evaluation of BDMM can be used as a point-of-reference for organizations in choosing the most appropriate BDMM to be adopted. Table 1 illustrates the elements that are used in this evaluation of BDMM. The detailed description of each element is described in Appendix A.

Table 1. Big Data Maturity Models (BDMM) Evaluation Criteria.

Perspectives	Element	Source
Generic Attributes	Name[1] Acronym[1] Primary source[1] Scope [1, 2] Origin[1, 2] Audience[1, 2] Driver[2] Respondents [2]	Mettler, Rohner and Winter (2010) [1] De Bruin, Rossman, Freeze and Kulkarni (2005) [2]
Design Attributes	Focus of assessment [1, 2, 3] Method of assessment [1, 3] Reliability and validity of assessment [1, 3] Suggestions for design improvement [2]	Mettler, Rohner and Winter (2010) [1] Mettler (2009) [2] De Bruin, Rossman, Freeze and Kulkarni (2005) [3]
Application Attributes	Method of application [1, 2] Purpose of application [2] Supporting material [1] Suggestions for maturity improvement [1]	Mettler, Rohner and Winter (2010) [1] De Bruin, Rossman, Freeze and Kulkarni (2005) [2]

As depicted in Table 1, this paper evaluates the existing BDMMs based on three (3) main perspectives, namely i) *generic attributes*, which provides the high level overview of BDMM; ii) *design attributes*, which focuses on the elements that are being used in designing the BDMM, so that it could help the readers to understand the underlying concept behind the BDMM developments; and lastly iii) *application attributes*, describes on the deployment aspect of BDMM which allows readers to select the suitable maturity model to be adopted. A total of seven (7) BDMMs were evaluated in this paper, which are exhibited in details in Table 2, 3 and 4 based on the *generic attributes, design attributes* and *application attributes* perspectives respectively.

As depicted in Table 2, it can be observed that all of the BDMMs are being developed to meet the internal organizational requirements (i.e. improve business performance, cost reduc-

tions, etc.) as well as external requirements (i.e. to gain competitive advantage, advancements of technology, etc.). It is also very apparent that majority of the BDMM are generic, which means that they can be applied in any domains or industries; and are being developed by people from the industries. Based on the generic attributes of the BDMMs, it can be concluded that all seven (7) BDMMs are almost similar with each other. Thus, the generic attributes of BDMM cannot be used as the sole factor in determining which BDMM is most suitable to be adopted by organizations.

Table 2. Big Data Maturity Models Evaluation – Generic Attributes Perspective.

Name	Acronym	Primary source	Scope	Origin	Audience	Driver	Respondents
TDWI Big Data Maturity Model	n/d	Halper and Krishnan (2013)	General	Practitioner (TDWI)	Internal	Internal and external requirements	Management, Staff
Big Data Business Model Maturity Index	BDBM	Schmarzo (2015)	General	Practitioner (EMC)	Internal	Internal and external requirements	Management, Staff
Big Data Maturity Model	n/d	Radcliffe (2014)	General	Practitioner (Radcliffe Advisory Services)	Internal	Internal and external requirements	Management, Staff
Big Data and Analytics Maturity Model	n/d	Betteridge and Nott (2014)	General	Practitioner (IBM)	Internal	Internal and external requirements	Management, Staff
Big Data Maturity Framework	n/d	El-Darwiche, Koch, Meer, Shehadi, Tohme (2014)	General	Practitioner (Booz & Company)	Internal	Internal and external requirements	Management, Staff
Big Data Maturity Model for Zakat Institution	n/d	Sulaiman, Che Cob and Ali (2015)	Domain Specific - Zakat Institution in Malaysia	Academia	Internal	Internal and external requirements	Management, Staff
Hortonworks Big Data Maturity Model	n/d	Dhanuka (2016)	General	Practitioner (Hortonworks)	Internal	Internal and external requirements	Management, Staff

Table 3. Big Data Maturity Models Evaluation – Design Attributes Perspective.

Acronym	Focus of assessment	Method of assessment	Reliability and validity of assessment	Suggestions for design improvement
TDWI BDMM	Process, Object	Maturity grid – 5 levels with 5 primary dimensions (Level 1 – Nascent, Level 2 – Pre-adoption, Level 3 – Early Adoption, Level 4 – Corporate Adoption, Level 5 – Mature/Visionary; Dimensions: Organization, Infrastructure, Data Management, Analytics, Governance)	Validated (using benchmark survey)	n/d

Acronym	Focus of assessment	Method of assessment	Reliability and validity of assessment	Suggestions for design improvement
BDBD	Process	Maturity grid – 4 levels (Level 1 – Insights, Level 2 – Optimization, Level 3 – Monetization, Level 4 – Metamorphosis)	Verified	n/d
Radcliffe's BDMM	Process	Maturity grid – 6 levels (Level 0 – In the Dark, Level 1 – Catching Up, Level 2 – First Pilot, Level 3 – Tactical Value, Level 4 – Strategic Leverage, Level 5 – Optimize and Extend)	Verified	n/d
IBM's BDAMM	Process, Object	Maturity grid – 5 levels with 6 elements (Levels: Ad-hoc, Foundational, Competitive, Differentiating, Breakaway; Elements: Business Strategy, Information, Analytics, Culture and Execution, Architecture, Governance)	Verified	n/d
Booz's BDMF	Process	Maturity grid – 4 levels (Level 1 – Performance Management, Level 2 – Functional Area Excellence, Level 3 – Value Proposition Enhancement, Level 4 – Business Model Transformation)	Verified	Enablers of environment readiness; organizational internal capabilities
Sulaiman's BDMM for Zakat Institution	Process	Maturity grid – 5 levels (Level 1 – Ignorance, Level 2 – Coping, Level 3 – Understanding, Level 4 – Managing, Level 5 – Innovating)	n/d	n/d
Hortonworks BDMM	Process, Object	Maturity grid – 4 levels with 5 domains (Levels: Aware, Exploring, Optimizing, Transforming; Domains: Data Sponsorship, Data and Analytics Practices, Technology and Infrastructure, Organization and Skills, Process Management)	Verified (based upon previous consulting experiences)	n/d

As depicted in Table 3, it can be observed that all BDMMs are using the maturity grid with levels of between 1 and 6; and most of them focus only on the process aspect on big data in their assessments. Most of the BDMMs are developed based on the vast experiences of the authors in the industry, and thus can be considered to be verified in terms of their reliability, rigor and generalizability.

As depicted in Table 4, it can be observed that most of the BDMMs are used in self-assessing the organization's current big data environment; but some BDMM also look at how big data adoption and initiative can be elevated to improve business performance and values; as well as benchmarking with other entities for competitive advantage purpose.

Table 4. Big Data Maturity Models Evaluation – Application Attributes Perspective.

Acronym	Method of application	Purpose of application	Supporting material	Suggestions for maturity improvement

Acronym	Method of application	Purpose of application	Supporting material	Suggestions for maturity improvement
TDWI BDMM	Third party	Comparative	Textual description	Implicit improvement
BDBD	Self-assessment	Prescriptive	No supporting material	n/d
Radcliffe's BDMM	Self-assessment	Prescriptive	No supporting material	n/d
IBM's BDAMM	Self-assessment	Descriptive	No supporting material	n/d
Booz's BDMF	Self-assessment	Prescriptive	No supporting material	n/d
Sulaiman's BDMM for Zakat Institution	Self-assessment	Prescriptive	No supporting material	n/d
Hortonworks BDMM	Third party	Comparative	Textual description	Implicit improvement

CONCLUSION

The evaluation of the existing big data maturity models (BDMM) from the generic, design and application attributes perspectives is hoped to help organizations to competently and effectively choose and select the appropriate BDMM, thus would be able to reap the full potential and leverage on the enormous amount of big data readily available for gaining competitive advantage and insights to the future endeavor.

ACKNOWLEDGMENTS

REFERENCES

Accenture. (2014). Accenture Big Success with Big Data Survey. Retrieved from *https://www.accenture.com/ie-en/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Industries_14/Accenture-Big-Data-POV.pdf*

Betteridge, N. & Nott, C. 2014. Big data and analytics maturity model. Retrieved from *http://www.ibmbigdatahub.com/blog/big-data-analytics-maturity-model*

de Bruin, T., Rosemann, M., Freeze, R., & Kulkarni, U. (2005). Understanding the main phases of developing a maturity assessment model. In *ACIS 2005 Proceedings - 16th Australasian Conference on Information Systems.*

Dhanuka, V. (2016). Hortonworks big data maturity model. Retrieved from *http://hortonworks.com/wp-content/.../04/Hortonworks-Big-Data-Maturity-Assessment.pdf*

El-Darwiche, B., Koch, V., Meer, D. & Tohme, W. (2014) . Big data maturity: An action plan for policymakers and executives. Retrieved from http://www.strategyand.pwc.com/reports/big-data-maturity

Forrester. (2016). Forecasts Big Data Tech Market Will Grow ~3x Faster Than Overall Tech Market. Retrieved from https://www.forrester.com/Forrester+Forecasts+Big+Data+Tech+Market+Will+Grow+3x+Faster+Than+Overall+Tech+Market/-/E-PRE9484

Fraser, P., Moultrie, J., & Gregory, M. (2002). The use of maturity models/grids as a tool in assessing product development capability. *In Engineering Management Conference (IEMC'02)*. IEEE International (Vol. 1, pp. 244-249). IEEE.

IDC (2016). Double-digit growth forecast for the worldwide big data and business analytics market through 2020 led by banking and manufacturing investments. Retrieved from *http://www.idc.com/getdoc.jsp? containerId=prUS41826116*

Halper, F., & Krishnan, K. (2014). TDWI Big Data Maturity Model Guide.

IBM. (2012) IBM study finds steep rise in UK companies recognizing the value of dig data. Retrieved from *http://nthn.firetrench.com/2012/10/ibm-study-finds-steep-rise-in-uk-companies-recognizing-the-value-of-big-data/*

IBM. (n.d.). Extracting business value from the 4 V's of big data. Retrieved from *http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data*

Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013). Big data: issues and challenges moving forward. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on (995-1004).*

Khare, A. (2014). Big data: magnification beyond the relational database and data mining exigency of cloud computing. *In 2014 Conference on IT in Business, Industry and Government (CSIBIG) Proceedings*

Laney, D. (2001). 3D data management: controlling data volume, velocity and variety. Retrieved from *https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf*

Lnenicka, M. (2014). Reframing architecture frameworks using cloud computing and open (big) data. International Journal of Information Technologies, Engineering and Management Science, 1 (2), 1-5

Mettler, T. (2009). A design science research perspective on maturity models in Information Systems.

Mettler, T., Rohner, P., & Winter, R. (2010). Towards a classification of maturity models in information systems. In Management of the Interconnected World (333-340). Physica-Verlag HD.

Moore, D. (2014). Roadmaps and maturity models: pathways toward adopting big data. *In Proceedings of the 2014 Conference for Information Systems Applied Research ISSN (2167-1508)*

Provost, F. & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big Data, 1 (1), 51 -59

Radcliffe, J. (2014). Leverage a Big Data Maturity Model to Build Your Big Data Roadmap. Retrieved from *http://www.radcliffeadvisory.com/research/download.php?file=RAS_BD_MatMod.pdf*

Salleh, K. A., Janczewski, L., & Beltran. F (2015). SEC-TOE Framework: Exploring Security Determinants in Big Data Solutions Adoption. *In Pacific Asia Conference on Information Systems (PACIS) 2015 Proceedings.*

Schmarzo, B. (2014). "Big ideas" big data business model maturity index. Retrieved from *https://infocus.emc.com/william_schmarzo/big-ideas-big-data-business-model-maturity-index/*

Sulaiman, H., Che Cob, Z. & Ali, N. (2015). Big data maturity model for Malaysian zakat institutions to embark on big data initiatives. In 4th International Conference on Software Engineering and Computer Systems (ICSECS) Proceedings

APPENDIX A

Perspectives	Element	Description
Generic Attributes	Name	Full name of the maturity model
	Acronym	Abbreviation of the full name
	Primary source	Author(s) of the maturity model
	Scope	The specific domain where the maturity model will be applied
	Origin	Stakeholder(s) that is/are involved in the maturity model development
	Audience	For whom the maturity model is developed for
	Driver	Why the maturity model is developed
	Respondents	From whom the data/input during the maturity assessments are collected
Design Attributes	Focus of assessment	The aspects that are assessed by the maturity model – process, object or people
	Method of assessment	Instruments and measurements that are used to assess the big data maturity in organization
	Reliability and validity of assessment	To test the reliability, rigor and generalizability of the maturity model *Verified* – the maturity model has sufficiently represents the developer's conceptual description and specification *Validated* – the maturity model has accurately represent the real world situation
	Suggestions for design improvement	Suggestions on how the maturity model can be improved by the developers
Application Attributes	Method of application	How the maturity assessment is conducted: self-assessment, third party assessment or by certified practitioner
	Purpose of application	Reasons for conducting the maturity assessment *Descriptive* – assessing the 'as-is' situation *Prescriptive* – maturity is related to business performance and positively affect business value *Comparative* – enables benchmarking across industries or regions
	Supporting material	Any supplementary documentations/tools that come with the maturity model that can be used by organization in assessing the maturity level
	Suggestions for maturity improvement	Suggestions on how the results of maturity assessment can help organizations to improve their level of maturity