International Journal of Modern Research in Engineering and Technology (IJMRET)
www.ijmret.org Volume 4 Issue 1 ǁ January 2019.

Realization of Data Science in Small Industry Using DSMM

Syed Atif Ali Shah[1], Zubia[2], Syed Asif Ali Shah[3], Zafar Khan[4], Hashim Sajid Khan[5]

Northern University 1,4, Qurtuba University 2, Abasyn University 3,5

Abstract: *This paper proposes a novel method for assessing competence ranking of Data Maturity Model of growing organizations. Our approach uses fine grained appraisal techniques to rank different small and medium organizations that fall below the minimum maturity level. Moreover, it adds the credulous features i.e., Stature, Dependence and Assurance level for a company's maturity. The benefit of the proposed model is to set diminutive objectives to achieve higher levels of the model. This technique not only reclassifies the levels but also exposes various pledge factors.*

Keywords: Data Science, Data Science Maturity Model, Medium and Small Industry

I. INTRODUCTION

One of the first demands of companies when they think of investing in data science is what return they will get. As a data science team can work and also the forms of work to be performed. The data science maturity model helps organizations make decisions to deliver real value to their interests. Industrial capability of an organization is measured against DSMM. The said model has four levels and five dimensions that can be applied at all organizational levels. A generic model of DSMM is shown in figure 1.

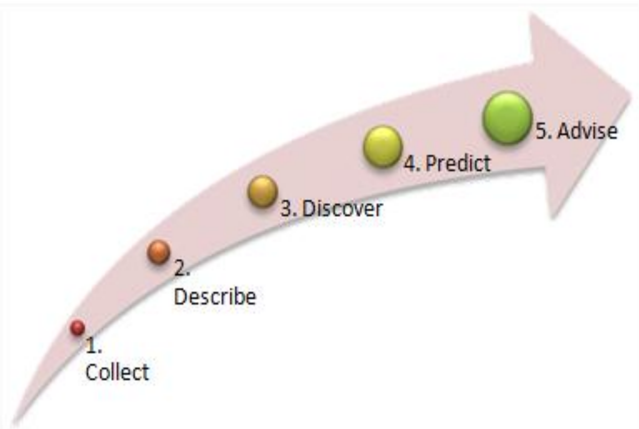


Fig 1

In the above model, we have four levels, from ad hoc exploration, where there are some practices, however, purely analytical. Level four shows best practices where all data is processed and stored according to the organization's strategic directions, producing data products with robust data engineering and data science, acting proactively.

DSMM models provide guidance for developing or improving data that meet the business goals of an organization. According to the practitioners, DSMM helps "integrate traditionally separate organizational functions, set data improvement goals and priorities, provide guidance for quality data, and provide a point of reference for appraising current data CMU claims DSMM can be used to guide data improvement across a project, division, or an entire organization.

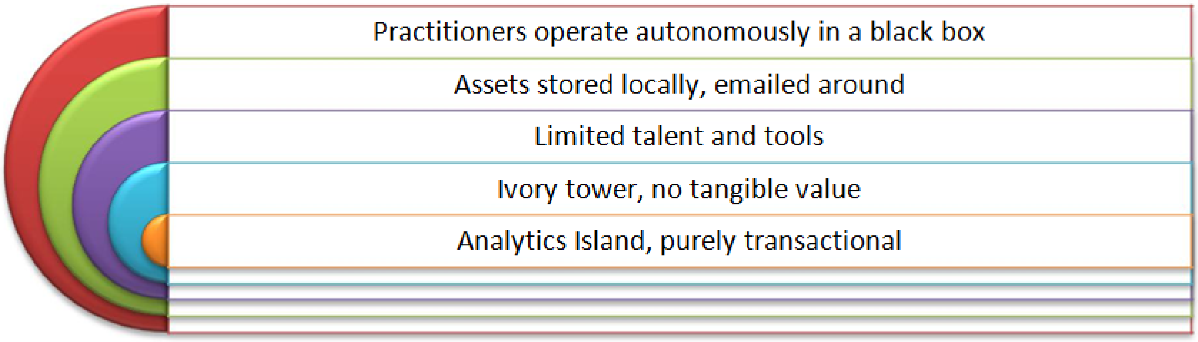
Fig 2

A DSMM model may also be used as a framework for appraising the data maturity of the organization. DSMM originated in software engineering but has been extremely widespread in the last decade to hold other areas of interest; this generalization of improvement concepts makes DSMM extremely abstract. To help data science practitioners and leaders identify their existing gaps and direct future investment, we proposed a framework called the Data Science Maturity Model (DSMM) for small industry. The DSMM assesses how reliably and sustainably a data science team can deliver value for their organization.

For a non-technical person capability is the measure of expertise. The expertise or skills are directly proportional capability level and vice versa. In developing countries a general statistics show that only few organizations working at level 4. The number of organizations increases as we go down the DSMM levels. A huge portion is occupied by level 0 and level 1 organization.

At Level 1 i.e. the initial level practices are out of scope. To improve from level 1 to level 2 it takes a lot of time, resources and efforts. To achieve level 2 there are seven formal data areas to be practiced accordingly, these several areas are further divided into many endorsed activities. Below are the key data areas.

II. DSMM LEVEL 1 OR ADHOC EXPLORATION



The above level is strictly defined and is distinct data areas of DSMM level 1. For stepping to level 2 from level 1 an organization has to work through five different data areas.

Level	Structured Processes	Discoverability & Compounding	Analytical Speed & Agility	Breadth & Depth of Impact	Organizational Cohesion
1 Ad Hoc Exploration	Practitioners operate autonomously in a black box	Assets stored locally, emailed around	Limited talent and tools	Ivory tower, no tangible value	Analytics island, purely transactional
2 Repeating, but Limited	Recurring workflows discussed, no enforcement	Assets stored centrally, but lack metadata / permissions	Some tools and talent investment	Static reports in a few business areas	Some collaboration with line managers
3 Defined and Controlled	Formalized process, manually enforced	Assets stored and tagged centrally with metadata and permissions	Rapidly test ideas with novel methods / tools	Results translated into multiple operational workflows	Analytics are key stakeholders in strategic decisions
4 Optimized and Automated	Best practices codified into infrastructure, transparency for all	All asset versions stored/tagged, searchable, reproducible	Cutting-edge tools, comfortable at the analytical frontier	Data products drive org with robust safeguards	Analytics enmeshed in business and proactively anticipates needs

Table 1

An organization that has worked on various data areas (but not on all), still considered on level 1, though practically it is more capable than the one that doesn't worked even on single data area. Organizations with enormous capability difference are still considered on same level, hence not desired anyways. The said issue is resolved by 'Realization of Data Science in Small Industry Using DSMM'. The matrix on the table 1 shows how we map the DSMM.

III. ROPOSED MODEL

Many organizations have been underwhelmed by the return on their investment in data science. This is due to a narrow focus on tools, rather than a broader consideration of how data science teams work and how they fit within the larger organization. For gigantic size projects customer needs higher DSMM levels, whereas for medium size projects DSMM level 2 is also a rational choice. Purpose of this paper is to introduce the difference of capabilities of small scale industry, working at lower levels of DSMM. It answers the following question. How to rank different organizations that fall below certain level of DSMM model?

Furthermore, it adds the credulous factors, i.e., Stature, Dependence and Assurance level for an organization's capability. The assessment techniques are used to mark the potency and limitations of current data, also expose the advancement risks, and establish capability and maturity ranks. Usually they are used as a part of a data improvement program or for rating potential suppliers. These techniques identify the assessment data as consisting of grounding;

- On-site behavior;
- Foundation clarification, conclusion, and ratings;
- Final reporting; and
- Ensuing activities.

The set of credentials related with a meticulous version of the DSMM incorporates a requirements specification called the Appraisal Requirements for DSMM (ARC).

I. Stature

After the formal assessment of DSMM Level, stature for each data area is calculated. The stature is the number of goals achieved against the total number of goals (both specific goals and generic goals).The Data Area Stature or simply stature is represented in percentage. As shown in fig 3, each stature ranging from 0 to 100.

Fig 3

II. Competence

Nowhere we going to define the internal capability of DSMM Level X and call it Competence Rank. Here X represents certain level of DSMM i.e. from level 1 to level 5. To find the Competence ranking for DSMM level X we need to know the scoring of each data area. Then we calculate the geometric mean of all the statures, which is 'Competence Rank' for DSMM Level X. Competence rank shows the capability of an organization working below DSMM level X. The introduction of Competence Ranking technique has opened a door to distinguish the higher capability against the lower one.

III. Assurance Level

In the next step we are going to find the assurance level of an organization. Minimum stature of the data area among all data areas (of DSMM Level X) is called the assurance level. It assures that all data areas are working higher than said level and thus increases the assurance of customer.

IV. Dependence

Now we calculate the Standard Deviation (SD or σ "sigma") of the statures of all data areas. This SD is then subtracted from the maximum standard deviation and call as dependence of the DSMM Level X. If dependence is maximum, it means that organization is working to improve all the data areas equally. It has gained the same capability in all data areas. While on the other hand, if it has minimum dependence then it means for some data areas it has more capability than other Data areas. Need of this significant term varies project to project.

IV. CONCLUSION

This paper introduces the interesting capability factors i.e. Stature, Competence Rank, Assurance Level and the Dependence. These factors are used to distinguish between higher and lower capabilities of organizations, practicing below DSMM Level X. Thus provides more details about an organization than DSMM model. This way customer is more flexible and satisfied for selecting development organization, without being endured the depth of mechanics. Another important aspect of this approach is to design number of different classification models for different purposes. For a common person we can call Expertise or skills classification model of different organizations or Expertise Level of an organization. To help data science practitioners and leaders identify their existing gaps and direct future investment, we proposed a framework called the Realization of Data Science in Small Industry Using DSMM. The DSMM assesses how reliably and sustainably a data science team can deliver value for their organization. The model consists of four levels of maturity and is split along five dimensions that apply to all analytical organizations. By design, the model is not specific to any given industry.

REFERENCES

[1.] Jörg Becker,Ralf Knackstedt, Jens Pöppelbuß Developing Maturity Models for IT Management 21 May 2009 Springer.

[2.] Javier Garzás, Mark C. Paulk An exploratory study of why organizations do not adopt DSMM.

[3.] E. Chandra, D. Francis Xavier Christopher, S.D. Vijaykumar Study of DSMM based data framework for quality models, IEEE Explore.

[4.] Nonaka, I., Takeuchi, H.: The Kowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press, Inc., New York (1995)

[5.] Reengineering the Industrial CMMI, S Atif A S. Volume 4, Issue 3, Summer 2018, Page 135-208. Journal of Advances in Computer Engineering and Technology (JACET). http://jacet.srbiau.ac.ir

[6.] Industrial Electrical/Electronic Engineering

Organization's Competence Ranking of CMMI Level 2 Using SCAMPI. S Atif Ali Shah. 3rd Abasyn International Conference on Technology and Business Management (AiCTBM 18).http:// peshawar.abasyn.edu.pk/conference.php

[7.] Tuomi, I.: Corporate Knowledge: Theory and Practice of Intelligent Organizations. Metaxis, Helsinki (1999)

[8.] Formulating DNA Chains Using Effective Calculability, S Atif A S. 1 (2018) > Shah. INTERNATIONAL JOURNAL OF COMPUTER (IJC). http://ijcjournal.org

[9.] Ahern, Dennis; Jim Armstrong; Aaron Clouse; Jack Ferguson; Will Hayes; Kenneth Nidiffer (2005-03-02). DSMM FORMAL Distilled: Appraisals for Data Improvement. Addison-Wesley Professional. p. 240. ISBN 0-321-22876-6.

[10.] Decoding 3-Dimensional Euclidean Space using 3DTM. S Atif Ali Shah. 3rd Abasyn International Conference on Technology and Business Management (AiCTBM 18).http:// peshawar.abasyn.edu.pk/conference.php

[11.] Freeze R, Kulkarni U (2005) Knowledge management capability assessment: validating a knowledge assets measurement instrument. In: 38th Hawaii international conference on system sciences.

[12.] Lee J, Lee D, Sungwon K (2007) An overview of the Business Data Maturity Model (BPMM). In: International workshop on data aware information systems (PAIS 2007). Huang Shan (Yellow Mountain), China, 384–395

[13.] Renken J (2004) Developing an IS/ICT management capability maturity framework. In: Research conference of the South African Institute for Computer Scientists and Information Technologists (SAICSIT). Stellenbosch, 53–62

[14.] Rosemann M, de Bruin T, Power B (2006) A model to measure business data management maturity and improve performance. In: Jeston J, Nelis J (eds) Business data management.