
Roadmaps and Maturity Models: Pathways toward Adopting Big Data

Danny T. Moore
dtmst192@rmu.edu
Robert Morris University
Coraopolis, PA 15108

Abstract

As organizations seek to adopt new technologies like big data, the use of roadmaps and maturity models provide a welcomed approach that could aid in their successful adoption. With the proliferation of big data commercial products and services, organizations are trying to reconcile their in-house enterprise infrastructure and data with the selection of appropriate commercial products and services. Executive decision makers face a formidable challenge as the history of IT project failure runs high for a number of documented and recurring reasons. The intent of this research is to investigate the roadmap and maturity model features through the absorptive capacity lens and to demonstrate how they can work in tandem to help the organization implement and sustain the big data initiative and hopefully negate many of the common risks associate with new technology adoption leading to an effective analytics capability.

Keywords: adoption roadmap, big data, absorptive capacity, maturity model, IT project failure

1. INTRODUCTION

From recent research conducted on big data, adoption within organizations has had its well-publicized pain points. Historically, IT project failures have run high at (62-70%) depending on the sources reviewed (Galorath, 2012, Asay, 2008, Krigsman, 2008). The Standish group has tracked IT failure in their "Chaos Report" over the past 15 years under the 'failure' and 'challenged' categories with summed totals ranging between 65 to 86%. Common root causes point to weak project management, complexity, unengaged senior leadership, no mechanism for resolving issues or to make adjustments from initial plans, poor organizational communications and the bad alignment to user needs to fulfill intended business purposes and functions (Gulla, 2012). The choices organizations could face may result in catastrophic results such as altogether abandonment to less severe effects as ongoing extension of timelines, budgetary overruns and

disappointed user satisfaction (Whitney, Daniels, 2013). Other supporting literature paints a similar picture as the national rollout of the Affordable Care Act with former HHS Secretary Kathleen Sebelius attempted to explain the issues in a congressional hearing on the problems with the enrollment and related IT support systems. Given this backdrop of failure, new factors and trends have only compounded the opportunity for failure given the proliferation of data growth (volume), the range of data sources including social media (access and privacy) as the types of data that include structure, semi-structured and unstructured data (data variety) are now occurring at high feed rates of speed (velocity). These three 'V's' are commonly used to define and characterize the world of big data. Secondly, this complexity is only compounded with the rapid proliferation of big data products and services evident in the table below of three big data landscapes in consecutive years from 2012 to 2014 (Feinleib, Turck and Zilis, Turck and Dong). With so many choices in products and services, how is an

executive decision maker to make sense of these product and service choices that lead to an effective analytic capability for the organization? Not to mention the already in place IT infrastructure in the organization of hardware, platforms, retention of staff, and maintenance and melding the 'new' with the 'existing'. It looks like the greater the complexity, the greater the likelihood for potential project failure.

Year	Title	Categories	Products Services
2012	Big Data Landscape	11	86
2013	Big Data Landscape 2.0	39	190
2014	Big Data Landscape 3.0	55	344

Table 1. Big Data Landscapes

2. INFORMATION SYSTEM THEORIES

To cope with this perplexing phenomenon in technology adoption, several relevant information system theories may explain and provide helpful insight into this adoption issue. Out of the 87 IS Theories, one of the early information system theories, Diffusion of Innovation Theory, explains the behaviors of early adopters, majority adopters and late adopters in their response to innovation adoption (Rogers, 1963). No doubt, these behaviors are evident today and are either compelled by competitive forces in their markets or postured to delay to learn more prior to entry to avoid unnecessary adoption risks. Another candidate information system theory published by Viswanath Venkatesh's in a 2003 theory on unified acceptance theory for adopting technology (UATAT) explains adoption is based upon utility and user expectations. But, one of the more plausible and suitable theories is Cohen and Levinthal's Absorptive Capacity theory (1990) and later expanded by Zahra and George (2003) that focuses on the ability or capacity for an organization to adopt a technology like big data. Adoption has been separated into real or potential abilities in this more recent model. Kimberly Zahller describes the power of absorptive capacity in organizational learning in this way, she writes, "learning is an iterative cycle and, as the knowledge base expands and the cognitive structures within organization memory are

elaborated, decision making will become faster and more able to process greater amounts of complexity and ambiguity, granting a competitive advantage to organizations in highly uncertain or rapidly changing environments" (Zahller, 2012). The absorptive capacity theory is certainly evident from the 2011 Big Data McKinsey report that profiles big data readiness across various industry sectors. Some industries are well ahead of the game, while others appear to languish (McKinsey, 2011). A second evidence is apparent from recent surveys on big data readiness among a cross-section of industries and their executives. The NewVantage published report reveals what leaders have in place in strategic planning and funding apportioned for their big data projects (2013). To make the obvious correlation, one would expect an organization with a high absorptive capacity to be high in readiness, engaging leadership and have a great 'track record' of success on technology adoption like big data. The big data roadmap is fundamentally an implementation instrument while the maturity model is a sustainment instrument. Both of these instruments are critical and inseparable, working in tandem to ensure a complete and sustainable adoption of big data within an organization to increase absorptive capacity.

3. LITERATURE ON THE ROADMAP

Questions emerge over how to characterize and differentiate a roadmap from other planning instruments. Isn't a roadmap just another name for a project? Projects are more specific with activities and specific durations of time while the roadmap functions at a higher level and a more general, strategic compass. Are there different types of roadmaps, what are the characteristics and benefits as evident from the literature? Research on the topic revealed that some of the more informative content focused on *technology roadmaps* rather than the *big data* roadmap. Some of the richer content came from only a few proven literature sources. In 1997 at Sandia National Labs, Garcia and Bray, published, "Fundamentals of Technology Roadmapping" (1997) and from the published book by McKeen and Smith, *IT Strategy: Issues and Practices*, in their chapter 8, "Creating and Evolving a Technology Roadmap" are informative sources in the use and advantages of a roadmap (2006). Both offer their reasons for using a roadmap: "as an effective tool for technology planning and coordination and leading to better investment decisions by identifying critical technologies and

technology gaps and identifying ways to leverage R&D investments” (Garcia, Bray, 1997). Similarly, McKeen and Smith state, “it is through articulation of a technology roadmap that you learn what you did well, where you failed, and how to improve the process. McKeen and Smith add this benefit, “roadmaps limit the range of technology options to lower the decision making effort and lowering the organization’s cognitive workload while providing direction for the organization”. By using the roadmap as a high level compass, it enables the absorptive capacity of an organization. Tiffany Pham’s, *From Business Strategy to Information Technology Roadmap*, defines the IT roadmap as “an action plan that matches the organization’s business goals with specific technology solutions in order to help meet those goals” (Pham, 2013). Pham is a strong advocate of performing the due diligence to link alignment between missions, strategy and goals as reflected in the roadmap. She adds, there are different types of roadmaps such as the Lean IT roadmap that provides “an IT action plan and strategy that leverages lean values and principles”. Although she does not specifically address big data, her definitions apply: as an IT action plan with strategy that leverages big data technologies and principles. Formulation of her high level roadmap linking includes four key actions: 1. To identify the current business and IT situation, 2. To identify future business strategy and IT needs, 3. To identify business and IT gaps and 4. To identify the IT roadmap. Her chapter 7 is an excellent, detailed guide to the formation of the IT Roadmap process. From Mark Van Rijemenam’s *Think Bigger: Developing a Successful Big Data Strategy for Your Business* (2014) and website, www.bigdata-startups.com, from Dave Loshin’s *Big Data Analytics*, from the blog “How to Create a Big Data Implementation Roadmap”, authors of *Big Data for Dummies*, and the prior sources, all of these share very similar phases and steps that have been part of a converged model with four fundamental phases and with a breakout of key activities. While there were only a few exceptions, most of the commercial literature in the form of whitepapers and websites were primary advocates for the need for a big data roadmap and not contributing substantive content in contrast to the prior covered references. Here is a high level summary of what has been gathered from the literature in Table 2. Roadmaps are often graphically depicted with milestones rather than making use of firm timelines. From a search on Google images, of

‘big data roadmap’ versus ‘technology roadmap’, it revealed a much higher number of ‘hits’ with more mature and richer content than for big data. By making the roadmap a graphic illustration, it provides a key communication tool for status and progress for the organization. Evident from search, and support by “Technology Roadmapping” (Phall, Farruk and Probert, 2001), there are no standards to designing a graphic roadmap as some are quite artistic while others adhere to a tabular form like a spreadsheet.

Plan	Requirements and identification of Key Business Drivers
	Develop Strategy and Vision
	Research use cases that correspond to the organization that includes data, architecture, platforms, funding and staffing skills
	Develop a Proof of Concept with applications, scope, expected outcomes, architecture, resources and risks
Communication	Form a communication plan across the organization for leadership and staff to convey status and progress
	Start the development team
Project	Develop the project plan
	Evaluate and select product, services, systems, platforms as aligned from requirements
	Conduct reviews of progress and status
	Conduct testing on implementation from milestones
Follow-up	Conduct Leadership Reviews
	Adjust as Organization Learns Lessons
	Establish Maintenance
	Maturity Model Plans

Table 2. A Literature Summary Roadmap

4. BENEFITS OF THE ROADMAP

Documented benefits for the roadmap were most detailed in *IT Strategy* (McKean & Smith, 2006) offering both internal and external benefits. The roster of external benefits from a roadmap include: a check to achieve business goals and potential gaps, reduces complexity, enhances interoperability across business functional lines, increases flexibility, and speeds implementation, preserves investment and

existing systems, provides a response mechanism to market changes, focuses investment dollars, responds to new legislation and reduces the difficulties associated with deployment of new technologies. On internal benefits, they include providing common design points, build a consistent and cohesive technology base, provide the ability to move forward in planned phases, consolidate global solutions, and lower the cost of development and maintenance (McKeen and Smith, 2006).

5. THE BIG DATA MATURITY MODEL

In 1986, SEI at Carnegie Mellon initially developed the maturity framework later to become Capability Maturity Model Integration (CMMI). Now, with over 5000 implementations using this model in over 70 countries, it is a *mature* maturity model. Maturity models provide a way to review and assess an organization's methods and processes against an accepted standard. The CMMI standard makes use of a five level model to bring about improvements to organizational practices and consisting of:

Level 1 is *Initial* where processes are not controlled and unpredictable

Level 2 is *Managed* where processes exist but are often reactive

Level 3 is *Defined* where processes are standardized and typically documented

Level 4 is *Quantitatively Managed* where processes are measured and controlled

Level 5 is *Optimized*, where processes have a focus on continuous improvement

While most maturity model early adopters were in the manufacturing sector, this standard has since been widely applied across a number of industries and fields with the common thread that maturity is about the improving interaction and alignment and coordination of resources and organizational behaviors. But CMMI is not the only maturity model available. One of the first developers of a technology specific maturity model was published on "Data Warehousing Stages of Growth" (Watson, Ariyachandra, Matyska, 2001). What has ensued is a proliferation of other technology specific maturity models across a number of sectors that include big data, enterprise architecture, social media, digital asset management, telecommunications, web analytics, data center infrastructure, business process modeling and many others. From conducting a simple Google web search, almost all useable maturity models

share a common set of core features that include a tabled graphic with three key components: Levels, Domains and Attributes. Maturity levels can range from 4 to 7 but, 4 or 5 are the most common. Domains represent a complete set of areas to be reviewed. Most will focus on areas involved with people, process, and technology but, other domain areas are also included that are of key importance. The attributes are very critical descriptions that must accurately depict the domain at a specific maturity level. Multiple items within each cell should provide a very accurate and complete characterization. Frankly, of the review on many of the technology maturity levels, it is here where many maturity models appear to 'fall apart'. As there is no room for ambiguous or vague language in a maturity model which could result in confused stakeholders and lead to a disastrous implementation.

Of the three found big data maturity models in this research, two are recommended candidates for use as they fit the described features criteria. IDC's Big Data and Analytics Maturity Model (2013) by Vesset, Girard, Burghard, Osswald, Versace, O'Brien, Febowitz and Ellis use different domains comprised of Intent, Data, Technology, People, Process. The IDC Big Data Maturity model is Figure 1 on the last page of the article. The Halper and Krishnan TDWI model (2013) includes organization, infrastructure, data management, analytics and governance as distinct attributes but, the authors have only published a description of their maturity model and it is not currently a graphic model. Halper and Krishnan have also provided a 50 question benchmark from the TDWI website for firms to conduct an assessment of an organization's their maturity level.

6. USING THE BIG DATA MATURITY MODEL

In using the big data maturity model, both van Rijmenam and Halper and Krishnan, describe there is an initial orientation phase followed by a fluctuating phase and eventually leading to a stabilizing phase on three core domains: persons, objects and social systems to guide owners and managers to become more effective and efficient. Both sets of authors point out that it is within the mid-level of an organization's effort that they can incur the greatest risk as more resources and time are required to put maturity into place. Often, organizations can realize initial success addressing the low hanging fruit of adoption, but can incur resistance and

spending of resources in the mid-range normalizing. If organizations can persevere through the mid-phase efforts, the higher maturity levels can be realized where the greatest benefit and ROI is the reward (CrossTalk, 2012). Maturity models are not without their controversy as organizations can incur bias and over-represent their maturity in how they code (Kohlegger, Maier, Thalmann, 2009). The choice each company will face is whether to use an existing model, customize the attributes to closer fit for the organization or to completely develop their own (Decision Factor, 2013).

7. ROI: THE MATURITY MODEL BENEFIT

The motivation for a firm to use a big data maturity model is clearly from the financial return on investment (ROI). From IBM’s commercial whitepaper, “The Power of Analytics Maturity: achieve better business outcomes by raising your analytics quotient”, one of the underlying incentives for adopting a maturity model is that organization inherently can grow “step by step” in their analytics maturity result in a 36% growth in revenue and 15% greater ROI with twice the rate of growth in EBITDA from the companies they reviewed (IBM, 2011). Similarly, in the January/February 2012 of *Crosstalk*, a journal of defense software engineering includes six documented studies arguing the case to pursue the highest levels of maturity to realize the greatest payback and benefit. From a Raytheon’s case study, tangible benefits also include strategic focus, continuous evaluation occur at highest maturity level and enjoy a much higher ROI (24:1). Both CMMI and Gartner have published similar improvement benefits to substantiate these in Table 3.

Criteria	Level 1=>2	Level 2=>3	Level 3=>4
Reduce defects	12%	40%	85%
Reduce cycle time	10%	38%	63%
Reduce cost	8%	35%	75%
Schedule variance	145%	24%	15%

Table 3. IBM Benefits Summary

8. CONCLUSION

For successful adoption of big data technologies as an analytic capability, organizations need both a big data roadmap and maturity model. The roadmap provides an essential compass for initial implementation for the organization while the maturity model is an instrument for sustaining the capability. For a big data maturity model, it is about building an ecosystem of the domains of the technology infrastructure, data management, analytics, governance and organizational components of measurement, monitoring and governance (Halper, Krishnan, 2013).

Organizations that use both the big data roadmap and big data maturity model in concert will vastly improve their absorptive capacity and reduce the likelihood of project failure which is so pervasive in this high stakes, complex IT environment.

9. REFERENCES

Asay, M. (2008). 62% of IT Projects Fail. Why? Retrieved from <http://www.cnet.com/news/62-percent-of-it-projects-fail-why/>

Big Data Start-ups, Retrieved from <http://www.bigdata-startups.com/>

Campo, M., (2012). “Why CMMI Maturity Level 5?”, *Crosstalk*. Retrieved from <http://www.crosstalkonline.org/storage/issue-archives/2012/201201/201201-0-Issue.pdf>

Cohen, W. & Levinthal, D. (1990). “Absorptive Capacity: A New Perspective on Learning and Innovation”, *ASQ*, 35, pp 128-152.

Feinleib, D., (2012). “The Big Data Landscape for 2012”, *Forbes*, Retrieved from <http://www.forbes.com/sites/davefeinleib/2012/06/19/the-big-data-landscape/>

Galorath, D. (2012). “Software Project Failure Costs Billions.. Better Estimation & Planning Can Help” Retrieved from <http://www.galorath.com/wp/software-project-failure-costs-billions-better-estimation-planning-can-help.php>

Garcia, M., Bray, O., (1997), *Fundamentals of Technology Roadmapping*, Sandia

- National Laboratories, Retrieved from <http://www.sandia.gov/PHMCOE/pdf/Sandia'sFundamentalsofTech.pdf>
- Gulla, J. (2012), Seven Reasons IT Projects Fail, Retrieved from http://www.ibmsystemsmag.com/power/Systems-Management/Workload-Management/project_pitfalls/
- Halper, F., Krishnan, K. (2013). "TDWI Big Data Maturity Model Guide: Interpreting Your Assessment Score", Retrieved from <http://tdwi.org/whitepapers/2013/10/tdwi-big-data-maturity-model-guide.aspx>
- Hurwitz, J., Nugent, A., Halper, F., Kaufman, M.,(2013). *Big Data for Dummies*, NJ: John Wiley & Sons
- Krigsman, M., (2012). Retrieved from <http://www.zdnet.com/blog/projectfailures/whos-accountable-for-it-failure-part-one/15451>
- Kohlegger, M., Maier, R., Thalmann, S., (2009). "Understanding Maturity Models Results of a Structure Content Analysis", Semantics Proceedings '09, September 2-4, 2009, Graz, Austria
- IBM, (2013). "Improve decision making with content analytics", Retrieved from http://www.ibm.com/smarterplanet/us/en/business_analytics/article/it_business_intelligence.html
- IBM Whitepaper, (2011). "The Power of Analytics Maturity: Achieve Better Business Outcomes by Raising Your Analytics Quotient". Retrieved from <http://www-01.ibm.com/software/analytics/rte/analytics/>
- IDC, Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East, (2012). IDC study sponsored by EMC, Retrieved from <http://www.emc.com/about/news/press/2012/20121211-01.htm>
- Loshin, D. (2013). *Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph*, Morgan Kaufman, NY
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung-Byers, A., (2011). Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey, May 2011, Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
- McKeen, J. and H. Smith, (2006). "Creating and Evolving a Technology Roadmap", *IT Strategy: Issues and Practices*, 2nd ed. Prentice Hall, NY
- NewVantage Partners, (2013). "Big Data Executive Survey 2013: The State of Big Data in a Large Corporate World", Retrieved from <http://newvantage.com/wp-content/uploads/2013/09/Big-Data-Survey-2013-Executive-Summary.pdf>
- Pham, T., D. Pham, A. Pham, (2013). *From Business Strategy to Information Technology Roadmap*, CRC Press, Boca Raton, FL
- Phaal, R., Farrukh, C., Probert, D., (2004). "Technology Roadmapping—A planning framework for evolution and revolution, *Technological Forecasting and Social Change*, 71, pp 5-26
- Rogers, E. (1962). *Diffusion of Innovations*, Free Press, New York
- Turck, M., Zilis, S., (2012), The Big Data Landscape (2.0). Retrieved from <http://mattturck.com/2012/10/15/a-chart-of-the-big-data-ecosystem-take-2/>
- Van Rijmenam, M. (2014). *Think Bigger: Developing a Successful Big Data Strategy for Your Business*, NY: AMACOM
- The Big Data Roadmap To A Winning Big Data Strategy (2013). Retrieved from <http://www.bigdata-startups.com/the-big-data-roadmap/>
- Venkatesh, V. (2000). "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies", *Management Science* , vol 46, no. 2, pp 186-204
- Watson, H, Ariyachandra, T., Matyska, R. , "Data Warehousing Stages of Growth", *Information Systems Management*, vol 18 (3), 2001, pp 42-50. Retrieved from

<http://ism-journal.com/ITToday/datawarehouse.pdf>

Whitney, K. and Daniels, C. (2013). "The Root Cause of Failure in Complex IT Projects: Complexity Itself", *Procedia Computer Science* 20 , pp 325-330

Zahra, S. and G. George, (2000). "Absorptive Capacity: A Review, Reconceptualization,

and Extension:", *Academy of Management Review*, vol 27, no 2, pp 185-203

Zahller, K. (2013). "Scientia potential est: Organizational Learning, Absorptive Capacity and the Power of Knowledge", *Information Systems Theory: Explaining and Predicting Our Digital Society*, vol 2, Springer, NY

IDC's Big Data and Analytics Maturity Model: Overview of Measures by Stage					
	Ad Hoc	Opportunistic	Repeatable	Managed	Optimized
Stage Description	-Ad hoc -Siloed proof of concept or pilot projects -Undefined processes -Lack of resources -Individual effort	-Defined requirements and processes -Unbudgeted funding -Project management and resource allocation inefficiency	-Recurring projects -Budgeted and funded program management -Documented strategy and processes -Stakeholder buy-in	-Project, process, and program performance measurement influences investment decisions -Standards emerge	-Continuous and coordinated BDA process improvement value realization
Intent (strategy, sponsorship, jurisdiction)	-No strategy exists -Unbudgeted project based on individual effort -Business case undefined	-Department-level, siloed strategy -Project-level budget with localized midlevel management support -No enterprise support for measurement tools or methods	-Business-unit –level strategy exists but not widely accepted -Business-unit-level budget with LOB support -The cost-benefit analysis without common measurement tools or methods	--Cross-business unit level strategy -Enterprisewide budget with upper management support -Enterprise wide measurement tools and methods	-Enterprisewide documented, accepted strategy -Executive support, budgeted and ad hoc funding -Widely accepted tools and processes for business case development -ROI measurement
Data (relevance, quality, availability)	-Easily available data is utilized, but it is incomplete -Data requires substantial manual effort to prepare for consumption	-Multisourced structured data or unstructured content exists -Data lacks timeliness and veracity	-Data collection, monitoring and integration processes are in place -Consistent data governance data and security practices have not been established	-Metrics to manage data quality, timeliness and veracity exist -Metrics to govern data collection, monitoring, and management processes	-Enterprisewide access to on-time, trusted , and comprehensive multistructured and multisourced data sets
Technology (adoption, performance, functionality)	-On-premise technology requires substantial effort to maintain and tune to derive desired performance -Functionality (cloud or on-premise) is limited or too generic to provide appropriate performance -Adoption is project specific	-New project-specific technology is acquired and deployed -This technology is fit for purpose but is not integrated with other deployed technology -Adoption is localized	-Multiple fit-for-purpose technologies have been deployed and are integrated -Adoption is selective	-A wide range of fit-for-purpose technologies have been deployed -Performance of these technologies is monitored and tuned as needed -Relevant technology broadly adopted	-A wide range of fit-for-purpose technologies have been deployed and pervasively adopted -Software and hardware have been optimized -A high level of automation exists in systems management for existing workloads and for dynamic scalability
People (skills, culture, organizational structure)	-A few individuals with some but not all the necessary skills are scattered throughout the organization -Lack of management and support	-Teams with some but not all the necessary skills -Lack of intra-organization coordination -Departmental management support for projects	-Skills acquisition, training, and management are governed by a stated strategy -Internal skills are augmented with external service providers -Staff is primarily decentralized	-Executive management support exists for a centralized BDA group -A broad range of internal technology skills exist and are augmented with external vendors -Analytics skills are mostly decentralized	-All the necessary experts BDA human resources exist -Executive priority is placed on BDA -Management encourages and promotes BDA use -A centralized group exists with primary responsibility for BDA and for coordination of any decentralized resources and vendors
Process (tracking, analysis, decisioning)	-Focused on creating information repositories and accessing the siloed information therein -Lack of support for predictive analytics or scenario evaluation -IT and business pursue their own projects with a lack of coordination	-Data analysis is emphasized at the expense of data tracking, preparation, and decision support processes -IT and business begin to collaborate on defining requirements and funding projects	-Strategy lays out the need for cross-functional collaboration between IT and business and among different business groups -BDA processes are extended to handling multistructured data -First attempts to monitor and document decision processes and decision outcomes	-Metrics for evaluating processed quality and success have been established -BDA leadership team is responsible for process coordination -Collaboration, workflow, and rules management technology augment core information management and analysis processes	-Processes are categorized into performance management and experimentation -Appropriate support, staffing, technology, and funding exists for each -Decision management techniques enable continuous process improvement and integration of analytics into business processes

©2013 IDC

Figure 1. IDC Big Data Maturity Model