



## Measuring Your Big Data Capability: A Framework for Assessing Your Big Data Readiness

Big data technologies allow organizations to capitalize on data-driven decisions by utilizing conventional and previously unconventional data sources to *discover*, *present*, and *analyze* important business insights. Information and intelligence generated from these types of data analyses can lead to better business decision-making based on more facts. Better business decisions create greater operational efficiencies, cost reductions, and ultimately improve overall corporate performance. Using Advanced and Predictive Analytics (A&PA), organizations that are leading the big data revolution are able to capture large amounts of data (measured in petabytes) and analyze it to find solutions that enable both operational efficiencies and cost reductions, and even trigger new product development.

To increase the rate of adoption and determine an organization's big data readiness, it is important to measure its overall data maturity level. Organizations should be able to define a data strategy, identify critical data sources and an overall data infrastructure and architecture, assess core and needed skill sets, define the underlying process infrastructure that supports data analysis, and measure the technologies and applications that support the organization's data needs.

### Objective of this Discussion

This paper defines the competencies organizations should seek to develop in order to leverage the exponential growth of data. It identifies distinct levels of data maturity, and guides organizations through the major steps to progress through a "Lagging to Leading" (L2L) assessment framework. This framework describes the process to identify big data needs, determine whether current capabilities fulfill those needs, and build organizational business intelligence and data capabilities to achieve A&PA goals and objectives. We also explore essential strategic concerns and considerations every organization must address in order to define and implement a data strategy.

### Measuring Readiness: Lagging to Leading Framework

Big data has tremendous volume and is typically defined to be wide, heterogeneous, distributed, and unstructured. As a result of these characteristics, traditional data processing platforms and applications are unable to support these large amounts of unstructured data. We have developed the L2L framework (Figure 1) to assess an organization's ability and

readiness to leverage these new types of data. It will also help to identify an organization's data maturity level, and the "people, process, and technology" capabilities to support new and current A&PA initiatives.

The framework defines the capabilities in each competency against the data maturity level from lagging to leading organizations:

**Lagging.** Characterizes an enterprise that is relatively behind other organizations with regard to data maturity levels and data management practices. A company at this level can be described as strategically reactive with a rudimentary awareness of the impacts of big data.

**Basic.** Characterizes an organization beginning to identify data needs and a data collection strategy. Data analytics is the responsibility of the department requiring this insight and information – there is no overarching data governance model in place.

**Intermediate.** Characterizes an organization with a defined data strategy – beginning to align people, processes, and technology capabilities with a new vision of A&PA.

**Advanced.** Characterizes an organization with a strategy guided by accurate and timely A&PA, big data management changes have been operationalized.

**Leading.** Characterizes an organization at the highest level of data maturity. A&PA is codified throughout the enterprise and the defined strategy roadmap has allowed for the rapid growth of new data sources and the subsequent growth in related skill sets and technologies.

The framework measures an organization's big data capabilities across each of the five competency areas below:

**Strategy Development and Execution.** A plan must be in place to define what to do with the data and to support accessing, sharing and managing that content. Content types include structured, unstructured, semi-structured and metadata. The plan will also encompass the use of tools and methods to enable and automate collection, publishing and managing content.

**Data Accessibility and Infrastructure.** Data should be organized and architected in a way to allow sharing between systems and across multiple users. To do this, companies must define packaging, formatting and metadata requirements. As new sources of data and amounts of data increase, companies must define these methods to allow for data availability.

**People.** Organizations must hire or outsource the right talent. "The skills most often mentioned in connection with big data jobs... include math, statistics, data analysis, business analytics and even natural language processing."<sup>i</sup> New technologies and processes most often necessitate a change in workforce skill sets. An organization must assess the current skills against future needed skills based on the defined roadmap. As data technologies change, skills like Map Reduce programming and Hadoop setup and management could be needed.

**Process Definition.** An organization must define the process to capture validated, stored, and processed data, and to manage its accessibility, reliability, accuracy, and timeliness for all data consumers within the enterprise.

**Technology.** The technological component of data management should align to an organization's current and future required capabilities. As the data available grows, the need for integrated systems and platforms to facilitate analysis should also grow to match the enterprise's needs.

The framework objectively tracks the development of data analytics capabilities across maturity levels, and offers guidance for organizations at the beginning of the continuum by demonstrating leading practices for more mature organizations.

## Developing a Strategy

A good data strategy should move the organization to a predictive state rather than a reactive one. It is important that "lagging" or "basic" organizations take steps to first define their strategy using priorities and that the leadership acknowledges that data analytics are important and incorporate these types of projects or initiatives with a timeline and roadmap in their strategic planning. As an organization's data maturity grows, strategy becomes a critical starting point to begin each year's corporate data goals and objectives definition

process. As mentioned earlier, data strategy provides a plan for the enterprise that will guide allocated budget and resources for data management and analytics initiatives. At the leading edge of the framework, data management is a strategic imperative and a competitive necessity. An organization at this level has planned and prepared itself for or is currently undergoing big data process and technological transformations.

A great data strategy will also define the vision for a robust and comprehensive data governance model. With the ever-increasing collection of data, like non-transactional social data, often created and accessed outside an enterprise, the organization needs to update and maintain policies and guidelines surrounding the use of its data. Governance of big data policies and processes should manage compliance, privacy, and security.

## **Business Intelligence Requirements and Information**

Data is not the same as information. Information or intelligence is data that has been transformed, via analytics, into a meaningful and useful format for users and other systems. Information is content rich and provides visibility into an organization's operations, finances, competition, and revenue models for example. Online chatter, blog posts, e-mail blasts, tweets and LinkedIn "likes" are all examples of data previously not captured in a useful format for analysis. However, with the advent of new big data technologies, this type of data can be stored, and eventually analyzed. As enterprises look to monitor online and social media actions and events, a well-defined A&PA process could serve as an early-warning system for disruptive changes in the competitive landscape, such as a rival's new product or pricing strategy or the entrance of an unexpected new competitor.

Information requirements are not as simple as identifying what data should be analyzed or where to access this data. Conventional analysis documentation such as narratives, bulleted lists and outlines are a good start, but are not as effective in engaging business decision makers and subject matter experts in analysis. Therefore, it is important to develop model driven analysis techniques to identify, analyze and define business intelligence requirements. Model driven requirements are more appealing and reduce the vagueness of the requirements and visually represent requirements clearly, concisely and unambiguously.

*Key activities and considerations in defining and prioritizing requirements are:*

- Document current technology environment and business applications in use that may impact potential new technology requirements.
- Address current data issues and problems.
- Determine owners of and the accessibility to data.
- Identify big data resource requirements.
- Ensure that requirements align to the defined data strategy.

Another important aspect of defining these requirements is to understand and document data availability, accessibility and accuracy. For example, certain data must be available in real time or meet minimum acceptable levels of accessibility. Business decision makers need the capability to access information in a usable format, instead of querying data scientists and long lead times for responses.

It is important to have clearly defined user groups and systems. Not all users are power-users, nor do all business users need access to all data sources all the time. An organization should document a process by which the correct users groups have the appropriate level of access and receive adequate training. The systems and tools should align governance requirements to the overall data strategy and IT's governance and security policies.

## **Capabilities: Current and Future Data Analytics and Management**

After determining maturity levels in the lagging to leading framework, defining a strategy, and gathering requirements, an organization can begin to plan their future state. The methodology and approach to measure their capability is to perform a fit gap analysis. The organization should understand and document current data analytics and management capabilities, and the requirements of the data strategy (future state). After these current capabilities are identified, a comparison can be drawn between the current and desired state. The gap assessment that follows should document functional and technological gaps and the activities needed to bridge these gaps. At the conclusion of the capabilities assessment, organizations should also have a clear understanding of the range and depth of skills available to them internally, externally and in combination, and

should be able to identify the training and skill acquisition required to progress. The deliverables for this process include:

- Requirements developed in the previous section.
- Solution Fit Gap Analysis.
- Mitigation Approach.

Building the desired future state must also encompass building employee skills to match future needs. As displayed on the L2L framework, leading technologies or processes alone will not move the organization to a fully leading state. People are an integral part of a successful transformation. The organization should define and develop a roles and responsibilities matrix and/or RACI model (Responsible, Accountable, Consulted, and Informed) to document data management and analytics participation and accountability across functional and departmental teams, and then Identify and inventory the skill sets of current data workers supporting data collection, analytics and applicable technologies.

A well-defined data analytics strategy often drives core operational changes, re-aligning data collection, tools and technology changes, and data management updates. To manage these changes it is necessary to implement a centralized data governance model, empowered to drive common data standards, definitions and usage is key. Data governance encompasses the people, processes, and technology required to create consistent and proper handling of data across the enterprise. Data governance should serve as quality control for assessing, managing, using, improving, and protecting organizational information. Leading practices in data governance are to establish data definitions and maintain consistency and accuracy across all systems. The goal is to improve data integrity and remediate inconsistencies in data usage. The organization needs to promote data literacy through education and data driven process improvement. Thus the organization can ensure data security, availability, and accessibility according to regulations and policies. Referring back to the RACI model the organization should define decision rights, accountabilities, controls, and standardized processes for maintaining data quality and consistency.

Big data insights and benefits cannot be realized without leveraging technology. In the L2L framework, the technology attributes of lagging organizations include fragmented business applications and disparate and unmanaged data sources. An organization maturing its

data analytics capabilities builds a flexible infrastructure with the ability to efficiently integrate new internal or external sources of data. The data collection tools, reporting technology, and data warehouse should align to business objectives, the broader technology strategy and ongoing support models. Actions to determine technology capabilities and manage gaps should move the organization along the framework. The organization should begin by documenting the current technology infrastructure, tools and business applications and determine how an enhanced strategy will integrate. Next is to analyze what data needs to be retained, managed and made accessible, and what data can be discarded. Then the organization should assess current security, disaster recovery, backup and recovery, performance management, infrastructure monitoring processes. Completing the assessment and gap mitigation ensures the proper technology selection and fit, as well as readiness for strategy implementation.

## Implementing the Strategy

Implementing the strategy may require new technology, human capital development, organizational re-alignment and data management change. Using the information obtained from the capabilities assessment and the subsequent gap analysis will help to create a comprehensive plan to evolve your organization from its current maturity level to the next one. The implementation of a successful strategy is an iterative approach and begins with successful change management to prepare for big data transformations. These change management activities should be conducted before any implementation.

- Identify potential areas of resistance, opportunities, and project risks.
- Identify and leverage areas ready for change
- Establish a baseline for future change. readiness assessments.
- Prioritization of change management actions and initiatives.
- Promotion of change adoption and acceptance.
- Establishment of quantitative mechanism for ongoing monitoring and measurement.

## Ownership

For each initiative developed in support of the organization's strategy, you should identify the key

stakeholders and should be able to define the critical project success criteria. Stakeholders should be categorized by role and prioritized by their A&PA value and then included in a sequential and progressive road map. A Stakeholder Analysis is the foundation for all change management activities and plays a vital role in a successful implementation. Identifying stakeholders and engaging them appropriately throughout the change process will enable a big data strategy to:

- Outline the individuals and groups that will need to be engaged, directly or indirectly, throughout the lifecycle of the project.
- Help the right people receive the right information, at the right time, in the right way.
- Proactively manage the pace and amount of change that each stakeholder must undergo to avoid “change overload.”
- Build organizational buy-in, commitment and capability for change.

Results of the Stakeholder analyses will be used as the foundation for the following deliverables: Change Impact Assessment & Mitigation Plan, Communications Strategy & Plan, and the Training Strategy & Plan. The Change Impact Assessment will identify key changes happening as a result of the new data strategy and will look at ways to mitigate potential impacts. As the organization deploys their big data initiative stakeholders will want to ensure their readiness and that the solution addresses their original business questions.

Communications will be managed through a comprehensive Communications Plan to develop awareness, understanding, acceptance, and commitment to big data and workflow changes. The Training Strategy and Plan will provide an overview of the direction, goals, and objectives for end-user learning. It will also incorporate a detailed plan that will lay out the approach in order to execute on the strategy. The organization must have a scalable end-user training strategy and solution that will increase user adoption of the underlying business processes built into the platform. The strategy should gauge the readiness ensuring data-literate users are adhering to business objectives and policies.

## Organizational Transformation

Effective change management techniques drive successful implementation and adoption. These process changes should work to develop the ability to mobilize data from across the enterprise; to probe that data intensely, to understand its value, and prioritize the data; all this with a strict governance discipline to maintain its prevalence within the organization. The chief impact of big data is its effect on how decisions are made and who makes them. When data is lacking, not available, or too expensive to purchase it organizations should fill the gap with advantageously positioned people make decisions, on the base of their experience and interpretations. For organizations that lack the in-house knowledge, experience, and/or resources for training, forming an analytics team is difficult. Without data experienced and skilled workers, it can be problematic for an organization to establish the necessary capabilities internally and the cost of setting up an analytics teams could make outsourcing an appealing option to achieve advantage over competitors.

The organization should focus on the problem they must solve for and what drives their business. To be data driven they must migrate from relying on intuitions and instinct. This necessitates the organization must stop thinking they are more data driven than they actually are. Without the proper strategy, approach, and governance it is too easy to mistake correlation for causation and find misleading patterns in the data. To create and foster a productive work culture around A&PA leaders and managers should gauge the effectiveness of current policies in depth by investigating data sources, their people and how they utilize analytics in their workflow to realize how policies impact efficiency. The organization should establish transparency to drive productivity. By increasing transparency managers can focus on merging integration across teams. Thus allowing insights to drive evidence-based decisions to integrate teams and meet revenue targets. The organizational transformation can be challenging, but the developments, both in the technology and in the business payoff, are considerable

Developing, testing and maintaining A&PA requires strong domain experience and business knowledge. The organization should optimize knowledge transfer (from IT, for example) with a center of excellence (CoE) to share solution knowledge, plan artifacts and ensure oversight for projects can help minimize mistakes. User data requirements and interactions will vary across the enterprise. Consider the user or system’s need to

present the right content to the right user at the right time. The users must have proper training and support. The ultimate goal is for the organization to breed and foster a data driven culture with high degree of collaboration across departments. An analytics team should include experienced data scientists and statisticians. A mix of technical and business skills, whether from a single person or members of a tightly aligned team, is key to produce the most successful results.

## Technology Changes

There are many vendors offering a diverse set of data collection, cleansing, storage, security, performance, analytics and reporting technologies. Determining which technology or combination of technologies will best support your organization's data needs can be a challenge. Developing or purchasing a tool, suite of tools or individual tools and interfaces should be approached in a way that will meet the needs of your business and aligns back to your goals. Once the organization has determined the needed data to fill in data gaps observed during the initial assessment the organization should determine the best sourcing strategy for analytics generation. This can be done either in house or by purchasing the data.

Once the organization is ready a Deployment Plan that describes how the big data strategy and/ or tool will be installed and transitioned to an operational system. It should contain an overview of the system, a brief description of the major tasks involved in the deployment, the overall resources needed to support the deployment effort, and any site-specific deployment requirements. The implementation will transition people, processes, and technology from the current state to the new environment. The results of the transition should be measured against the business case objectives.

## Measuring Performance

The organization must be able to measure success based on the defined acceptance criteria created at the beginning of this transformative endeavor. Success of any data analytics or transformation initiative depends on the full adoption and transition of the data users and systems. Measuring and re-measuring will be required to ensure that newly introduced processes and technologies have been fully adopted and that these

processes and technologies are consistently meeting the needs of the business.

Critical success factors should be agreed upon during the onset of the project while developing the data strategy. They are the outcomes necessary in order to reach A&PA goals and attain expected benefits. The following are typical performance measurements for an A&PA initiative:

- Recommended Tools adoption
- Verified and agreed upon requirements
- Agreed upon timeline
- Access to appropriate resources
- Agreed upon hypothesis that aligns to business goals and objectives

In an effort to become more data mature the organization should employ their data analytics strategy in parallel with their big data strategy in order to validate new finding and insights against traditional methods. The organization should use defined business goals, create hypotheses, and identify problems and opportunities; then use analytics to test and refine assumptions and create a feedback loop. In the iterative approach they should continue to refine their tools and methods perform technology configuration, customization, and/or integration based on business and technical requirements. The pertinent test is that of utility in improving management practice

## Conclusion

With the full understanding of an organization's maturity of data analytics across the Lagging or Leading maturity levels big data and A&PA readiness can be determined. More actionable insights can be captured with a well-defined data analytics strategy and avoid the common pitfall of being tool-centric without the proper resources and organizational knowledge to utilize it. It is important to focus on the problem the organization is trying to solve for and then figure out appropriate tool. Big data is capable of providing insights from the data that were not previously thought of or analyzed with traditional methods. The more an organization thinks of itself as data-driven, the more capable they are to perform on financial and operational results. Companies that want to differentiate on the value of their product can gain insights with big data, but will still need traditional data analytics methods to validate those insights. Knowing how to utilize a big data strategy is as important as

creating one. The insights gained utilizing big data does not diminish the need for human intuition. Technology alone cannot provide actionable insights.

		Maturity				
		Lagging	Basic	Intermediate	Advanced	Leading
Capability	Strategy	<ul style="list-style-type: none"> <li>Organizational strategy is reactive and changes frequently</li> <li>Leadership is aware of big data concepts, methods and tools, but has not committed resources to explore further</li> </ul>	<ul style="list-style-type: none"> <li>Strategy is defined using pre-set value-based priorities</li> <li>Leadership acknowledges big data analytics is important but strategic implications unclear</li> <li>Planning incorporates some big data projects or initiatives with a timeline and roadmap</li> </ul>	<ul style="list-style-type: none"> <li>Strategy is proactive, directionally-aligned and incorporates some big data insights</li> <li>The organization is defining a data strategy and is assessing current capabilities against the strategy</li> <li>Leadership is committing resources to implement the data strategy</li> </ul>	<ul style="list-style-type: none"> <li>Big data is used to set strategy based on predictive analytics with computational modeling and specialized algorithms</li> <li>The organization has implemented change in its people, process and technology capabilities to support the big data strategy</li> <li>Leadership agrees that innovation in data and data analysis is a core value for the organization</li> </ul>	<ul style="list-style-type: none"> <li>Strategy is self learned, driven on big data insight and leverages prescriptive analytics</li> <li>Leadership is committed to organizational-wide big data analytics and views it as strategic imperative as well as a competitive necessity</li> <li>Big Data is a yearly budget item and is considered operational; not a project, program or initiative</li> </ul>
	Data	<ul style="list-style-type: none"> <li>The type, quantity and source of data necessary for analytics is not fully understood</li> <li>The organization is not able to solve for bad data</li> <li>Data analytics returns inconsistent or contradictory versions of the truth</li> </ul>	<ul style="list-style-type: none"> <li>Data quality, consistency and accuracy is beginning to be addressed</li> <li>Historical data is available in a single repository for use in validating hypotheses</li> <li>Early stages of identifying, collecting and storing big data</li> </ul>	<ul style="list-style-type: none"> <li>Data is beginning to be centrally stored and organized, with a supporting data dictionary and meta data</li> <li>Multiple business applications and data sources are being sourced for cross-organizational data</li> <li>Data collection is returning data that is content rich, high in quality, accurate and usable</li> </ul>	<ul style="list-style-type: none"> <li>Single access point is established to collect, cleanse and extract data, including organization-wide data definitions</li> <li>Data is being aggregated and analyzed into actionable information</li> <li>Data usage processes and framework established to define required data inputs</li> </ul>	<ul style="list-style-type: none"> <li>Data needed for analytics is collected and available right-time, and all data collected is used</li> <li>Big Data has high integrity, is secure, complete, from disparate sources including real time</li> <li>Big Data is transformed into a meaningful and useful format with mission-critical output</li> </ul>
	People	<ul style="list-style-type: none"> <li>Communication between departments to collect, store or normalize data is inconsistent at best</li> <li>Analytics resources are not centrally organized or managed</li> <li>Organization has not determined what skills are required for big data analytics</li> </ul>	<ul style="list-style-type: none"> <li>Some communication across departments regarding standardizing data sets and sharing data for cross-department analytics</li> <li>Organization beginning to identify skill set needs to support big data analytics</li> <li>Early stages of data stewardship and ownership for source data content</li> </ul>	<ul style="list-style-type: none"> <li>Cross-department communication regarding shared data is the norm</li> <li>Organization understands necessary data analytics skills and experience</li> <li>Data stewardship is defined and accepted for source data content ownership</li> </ul>	<ul style="list-style-type: none"> <li>Departments are re-organized to support big data-driven operations</li> <li>Resources skilled in big data analytics are hired or trained and organized centrally or matrixed</li> <li>Data stewardship for source data content ownership is the norm</li> </ul>	<ul style="list-style-type: none"> <li>The organization permeates a data driven culture with high degree of collaboration across departments</li> <li>The analytics team includes experienced data scientists and statisticians</li> <li>Analytics are using big data to identify patterns and insights with as well as posing a problem for Big Data to solve</li> </ul>
	Process	<ul style="list-style-type: none"> <li>Standards on data collection, storage and interpretation are siloed at the department level</li> <li>Organization is in the early stages of project-based data governance</li> <li>No formal data management strategy and direction on what to do with the data</li> </ul>	<ul style="list-style-type: none"> <li>Processes supporting data analytics are being evaluated with a focus on re-alignment towards big data analytics</li> <li>Data governance is programmatic</li> <li>Centralized data management strategy is in development</li> </ul>	<ul style="list-style-type: none"> <li>Operational process changes to drive big data analytics are identified</li> <li>Data governance includes data from across the organization</li> <li>Data management standards and protocols are implemented</li> </ul>	<ul style="list-style-type: none"> <li>Operational changes to business processes are implemented</li> <li>Well defined and adopted data management and governance plan is in place</li> <li>Big data analytics is resulting in measureable, actionable insights that continuously drive change</li> </ul>	<ul style="list-style-type: none"> <li>Organization-wide data governance program, reports directly to Leadership</li> <li>Continuous and creative analytics development and improvement</li> <li>Operational processes are re-configured around Big Data</li> </ul>
	Technology	<ul style="list-style-type: none"> <li>Business applications are fragmented, with disparate and siloed data sources</li> <li>Multiple technology environments, including databases, platforms, and applications, must be supported</li> <li>Security is administered at the application level</li> <li>External reporting and analytics data tools are randomly selected</li> </ul>	<ul style="list-style-type: none"> <li>Core transaction source system data integrated into Enterprise Data Warehouse (EDW), providing cross-organizational data</li> <li>Non-standardized reporting and analytics tools</li> <li>Data management, security and provisioning administered through EDW</li> <li>Single EDW environment to support and maintain data</li> </ul>	<ul style="list-style-type: none"> <li>Standardized tools for reporting and data exist but are not integrated or uniformly utilized</li> <li>Platform/multi-platform enhancements handle larger, complex unstructured data sets with increased processing power</li> <li>The search for more robust big data analytics technology has begun</li> </ul>	<ul style="list-style-type: none"> <li>Analytics data visualization tools with in-memory processing implemented, resulting in for more dynamic and predictive analysis</li> <li>Massively Parallel Processing (MPP) database environments within data center established for executing coordinated, complex computations in parallel</li> <li>Security, disaster recovery, backup and recovery, performance management, and proactive infrastructure monitoring in place</li> </ul>	<ul style="list-style-type: none"> <li>The infrastructure is evolving, with ability to integrate new sources of internal or external data for analytics</li> <li>Complex analysis is codified through algorithms and not manually performed</li> <li>Big Data Tools include advanced analytic platforms, data virtualization, in-memory analytics, text mining, stream mining and geospatial analytics</li> </ul>

Table 1: Lagging to Leading Assessment Framework.

## About the Author

Holly Macke is a Management Consultant at Kenny & Company with over eight years of experience delivering consulting services and technology solutions in both the public and private sections. She has led projects relating to IT strategy, IT portfolio management, governance, process improvement and optimization, business strategy, business transformation, and testing strategy development and implementation. Holly holds a Bachelor of Science cum laude in International Affairs from the Georgia Institute of Technology and a Master of Arts in Economics from Georgia State University.

## About Kenny & Company

Kenny & Company is an independent management consulting firm providing Strategy, Operations and Technology consulting services to our clients. Our management consulting practice, experience and insight enable us to provide early stage venture capital investments to select startup companies, and through our philanthropic endeavors to give back to our communities.

---

*This article was first published on [www.michaelskenny.com](http://www.michaelskenny.com) on November 24, 2014.*

*The views and opinions expressed in this article are provided by Kenny & Company to provide general business information on a particular topic and do not constitute professional advice with respect to your business.*



Kenny & Company has licensed this work under a Creative Commons Attribution-NonCommercial 3.0 United States License.

## References

---

<sup>1</sup> Harbert, Tam "Big data, big jobs? A slew of new jobs is expected to open up in big data, but not everyone in IT will qualify. Here's what employers will be looking for" <http://www.computerworld.com/article/2492676/big-data/big-data--big-jobs-.html?page=2>.