BIG DATA FRAMEWORK®

APMG International

BIG DATA

The Big Data Framework

An introduction to the benefits and capabilities of the Big Data Framework

Why a Big Data Framework?

Many words have been written about quick growth of data in enterprises and the value that can be obtained by analyzing Big Data. Commonly cited benefits are: Increased speed-to-market, Better cost control and Enhanced personalization of marketing campaigns.
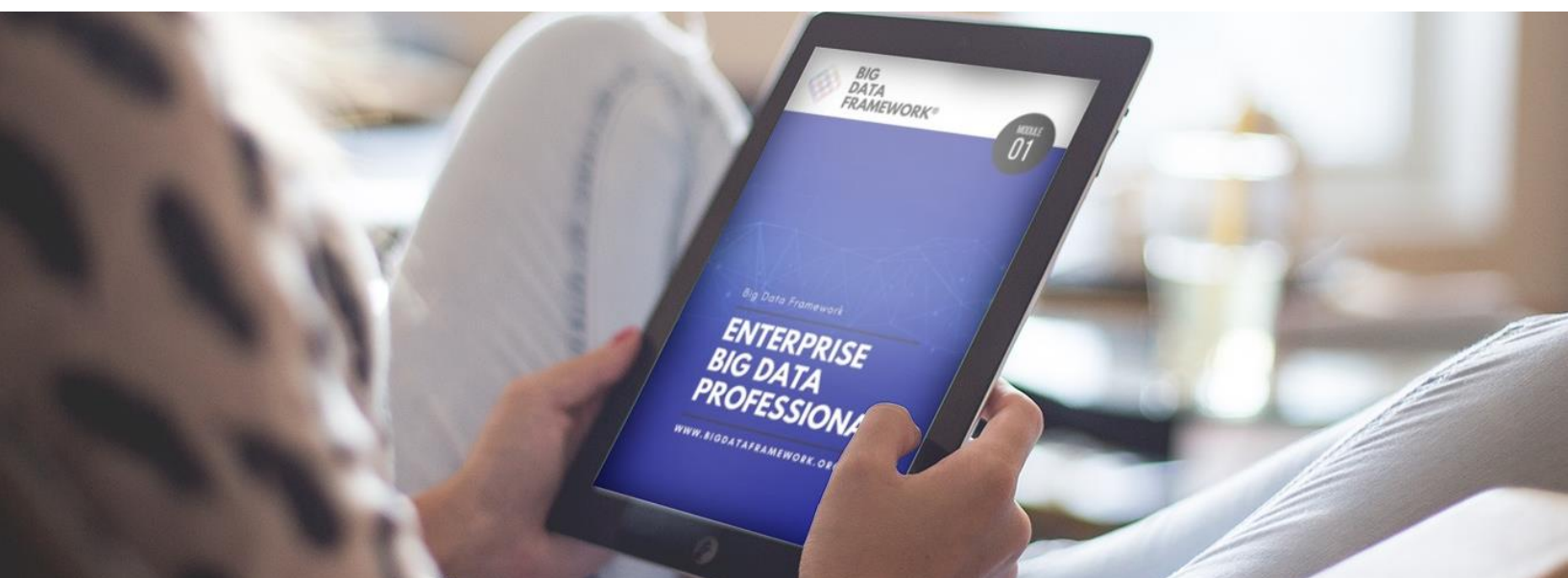
Although many enterprises see the value of Big Data – most have already started with initial projects as an extension of Business Intelligence efforts – many organizations struggle to realize a positive return on investment on Big Data initiatives.

Most enterprises have now realized that 'Big Data' is not just a project or department that can be managed by (a part of) the IT department. In order to obtain long-term value from Big Data, there is more required than hiring smart data scientists and acquiring a Big Data technology stack. Rather, data-driven thinking and data-driven decision making should become part of the organization's DNA. This is easier said than done. In a sense, transformation towards a Data Driven Enterprise can be regarded as an organizational change program. It requires a different way of thinking, in which all parts of the organization (from strategy to culture) play a vital role. Because of this requirement, the Big Data Framework was created.

The Big Data Framework is a structured approach that considers six core capabilities that enterprises need to consider to obtain long-lasting value from Big Data. Rooted in the scientific domain, the Big Data Framework provides fundamental knowledge and insights in the core components of Big Data. The topics of the framework range from the technical components of setting up a Big Data architecture towards the soft skills required to set up a Big Data center of excellence. By taking a capabilities approach, enterprises can grow their competencies over time, firmly embedding the practice of Big Data into the enterprise.

In this white paper, an overview is provided of the benefits of a framework approach towards Big Data. Subsequently, an introductory overview is provided of the six core capabilities of the Big Data Framework.

> There is more required than hiring smart data scientists and acquiring a Big Data technology stack.

Benefits of a Framework Approach for Big Data

Frameworks provide structure. The core objective of the Big Data Framework is to provide a structure for enterprise organizations that aim to benefit from the potential of Big Data. In order to achieve long-term success, Big Data is more than just the combination of skilled people and technology - it requires structure and capabilities.

The Big Data Framework was developed because although the benefits and business cases of Big Data are apparent - many organizations struggle to embed a successful Big Data practice in their organization. The structure provided by the Big Data Framework provides an approach for organizations which takes into account all the organizational capabilities of successful Big Data practice.

The main benefits of applying a Big Data framework include:

- Provision of a structure for organizations that want to start with Big Data or aim to develop their Big Data capabilities further.

- Inclusion of all organizational aspects that should be considered in a Big Data organization.

- Neutrality. The Big Data Framework is vendor-independent, it can be applied to any organization regardless of their choice of technology, specialization or tools.

- A common reference model that can be used across departmental functions or country boundaries.

- Identification of core and measurable capabilities in each of its six domains so that the organization can develop over time.

Big Data is a people business. Even with the most advanced computers and processors in the world, organizations cannot grow successfully without appropriate knowledge and skills. The Big Data Framework aims to increase the knowledge of everyone who is interested in Big Data. The modular approach and accompanying certification scheme develops knowledge about Big Data in a similar, structured fashion.

Big Data is a people business.

The Six Core Capabilities of the Big Data Framework

The Big Data framework is a structured approach that consists of six core capabilities that organizations need to take into consideration when setting up their Big Data organization. The Big Data Framework is depicted in the figure below:
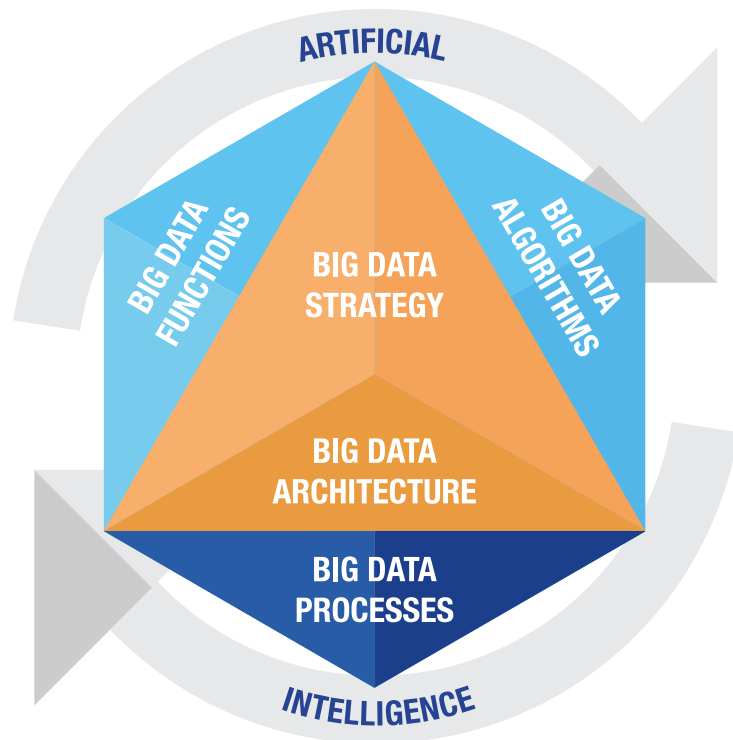
1. Big Data Strategy

Data has become a strategic asset for most organizations. The capability to analyze large data sets and discern patterns in the data can provide organizations with a competitive advantage. Netflix, for example, looks at user behavior in deciding what movies or series to produce. Alibaba, the Chinese sourcing platform, became one of the global giants by identifying which suppliers to loan money and recommend on their platform. Big Data has become an essential element of Big Business.

In order to achieve tangible results from investments in Big Data, enterprise organizations need a sound Big Data strategy which addresses the questions; How can return on investments be realized? Where do we need to focus effort in Big Data analysis and analytics?

The possibilities to analyze are literally endless and organizations can easily get lost in the zettabytes of data. A sound and structured Big Data strategy is the first step to Big Data success.

Data has become a strategic asset for most organizations.

2. Big Data Architecture

In order to work with massive data sets, organizations should have the capabilities to store and process large quantities of data. In order to achieve this, the enterprise should have the underlying IT infrastructure to facilitate Big Data. Enterprises should therefore have a comprehensive Big Data architecture to facilitate Big Data analysis. How should enterprises design and set up their architecture to facilitate Big Data? And what are the requirements from a storage and processing perspective?

The Big Data Architecture element of the Big Data Framework considers the technical capabilities of Big Data environments. It discusses the various roles that are present within a Big Data Architecture and looks at best practice for design. In line with the vendor-independent structure of the Framework, this section will consider the Big Data reference architecture of the National Institute of Standards and Technology (NIST).

3. Big Data Algorithms

A fundamental capability of working with data is to have a thorough understanding of statistics and algorithms. Big Data professionals need to have a solid background in statistics and algorithms to deduct insights from data. Algorithms are unambiguous, computerized, problem-solving rules. Algorithms can perform calculations, data processing and automated reasoning tasks. By applying algorithms to large volumes of data, valuable knowledge and insights can be obtained.

A fundamental capability of working with data is to have a thorough understanding of statistics and algorithms

The Big Data algorithms element of the framework focuses on the (technical) capabilities of everyone who aspires to work with Big Data. It builds a solid foundation which includes basic statistical operations and provides an introduction to different classes of algorithms:

- Implementation (Recursion vs. Iterative)

- Design Paradigm (Brute force, divide-and-conqueror, Dynamic Programming, Greed)

- Complexity (Constant, Linear, Polynomial, Non-Polynomial, Exponential)

Consider more than just the skills and technology.

4. Big Data Processes

In order to make Big Data successful in an enterprise organization, it is necessary to consider more than just the skills and technology. Processes can help enterprises to focus their direction. Processes bring structure, measurable steps and can be effectively managed on a day-to-day basis. Additionally, processes embed Big Data expertise within the organization by following similar procedures and steps. By embedding it as 'a practice' of the organization, analysis becomes less dependent on individuals, greatly enhancing the chances of capturing value in the long term.

5. Big Data Functions

Big Data functions are concerned with the organizational aspects of managing Big Data in enterprises. This element of the Big Data framework addresses how organizations can structure themselves to set up Big Data roles and discusses roles and responsibilities in Big Data organizations. organizational culture, organizational structures and job roles have a large impact on the success of Big Data initiatives. We will therefore review some 'best practices' in setting up a Big Data organization.

In the Big Data Functions section of the Big Data Framework, the non-technical aspects of Big Data are covered. You will learn how to set up a Big Data Center of Excellence (BDCoE). Additionally, it also addresses critical success factors for starting Big Data project in the organization.

6. Artificial Intelligence

The last element of the Big Data Framework addresses Artificial Intelligence (AI). One of the major areas of interest in the world today, AI provides a whole world of potential. In this part of the framework, we address the relation between Big Data and Artificial Intelligence and outline key characteristics of AI.

Many organizations are keen to start Artificial Intelligence projects, but most are unsure where to start their journey. The Big Data Framework takes a functional view of AI in the context of bringing business benefits to enterprise organizations. The last section of the framework therefore showcases how AI follows as a logical next step for organizations that have built up the other capabilities of the Big Data Framework. The last element of the Big Data Framework has been depicted as a lifecycle on purposes. Artificial Intelligence can start to continuously learn from the Big Data in the organization in order to provide long lasting value.

Learn more about the Big Data Framework

This white paper provides an introduction to the benefits and capabilities of the Big Data Framework. If you are interested to learn more about the framework, you can download a full copy of the Enterprise Big Data Professional Guide at www.bigdataframework.org or visit the APMG-International website at www.apmg-international.com.

About the author

Jan-Willem Middelburg is the Regional Director for the Asian Pacific Division of Pink Elephant, based in Kuala Lumpur. He is responsible for Pink Elephant's regional strategy and operations in Asia, an appointment he holds as of January 1st, 2016. Prior this appointment, he was responsible for restarting the office of Pink Elephant in the Netherlands.

In 2017, Jan-Willem published the Service Automation Framework for the design and delivery of automated services, which is internationally considered the first leading publication on Service Automation. As author of the book, he is a frequent keynote speaker and moderator at universities and technology conferences around the world. Jan-Willem holds a bachelor (BSc) in Industrial Engineering and a Master (MSc) in Supply Chain Management from the Erasmus University Rotterdam School of Management.
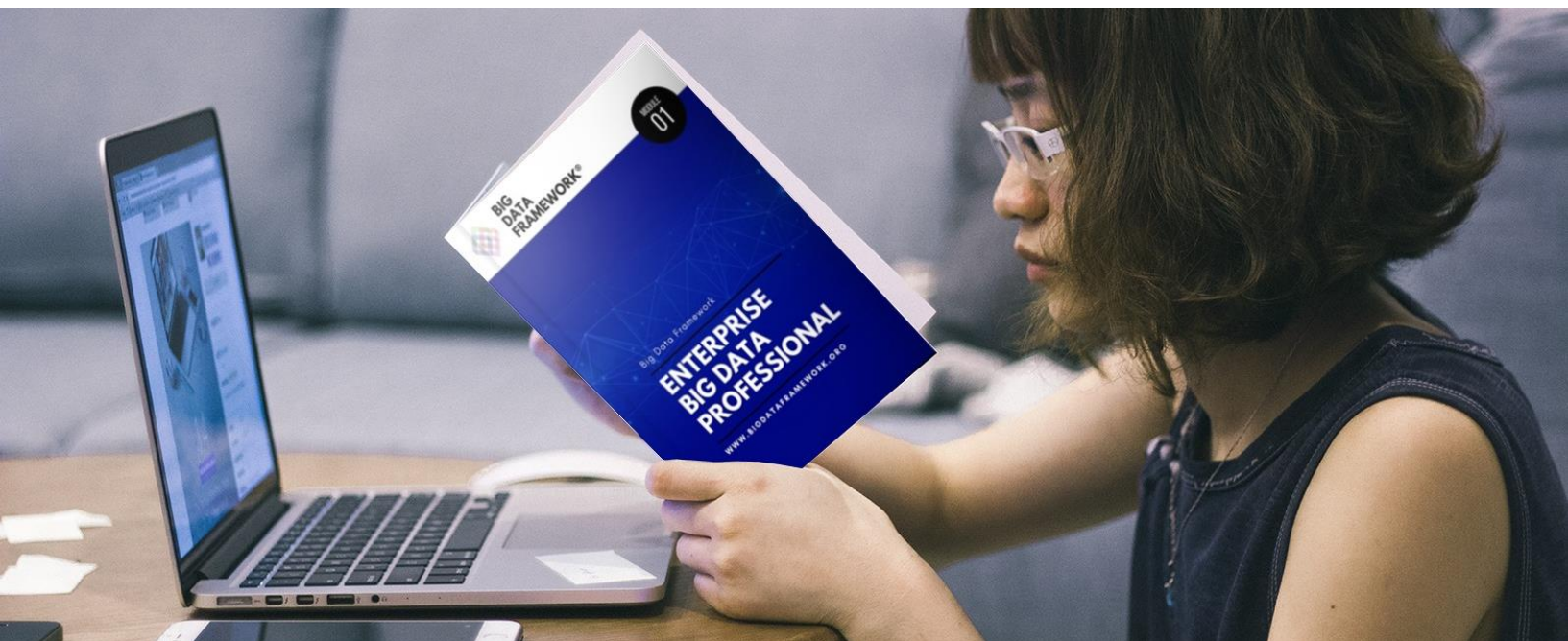
FOLLOW US ONLINE

APMG International

@APMG_Inter

www.apmg-international.com

company/apmg-international

facebook.com/APMGinternationalLTD