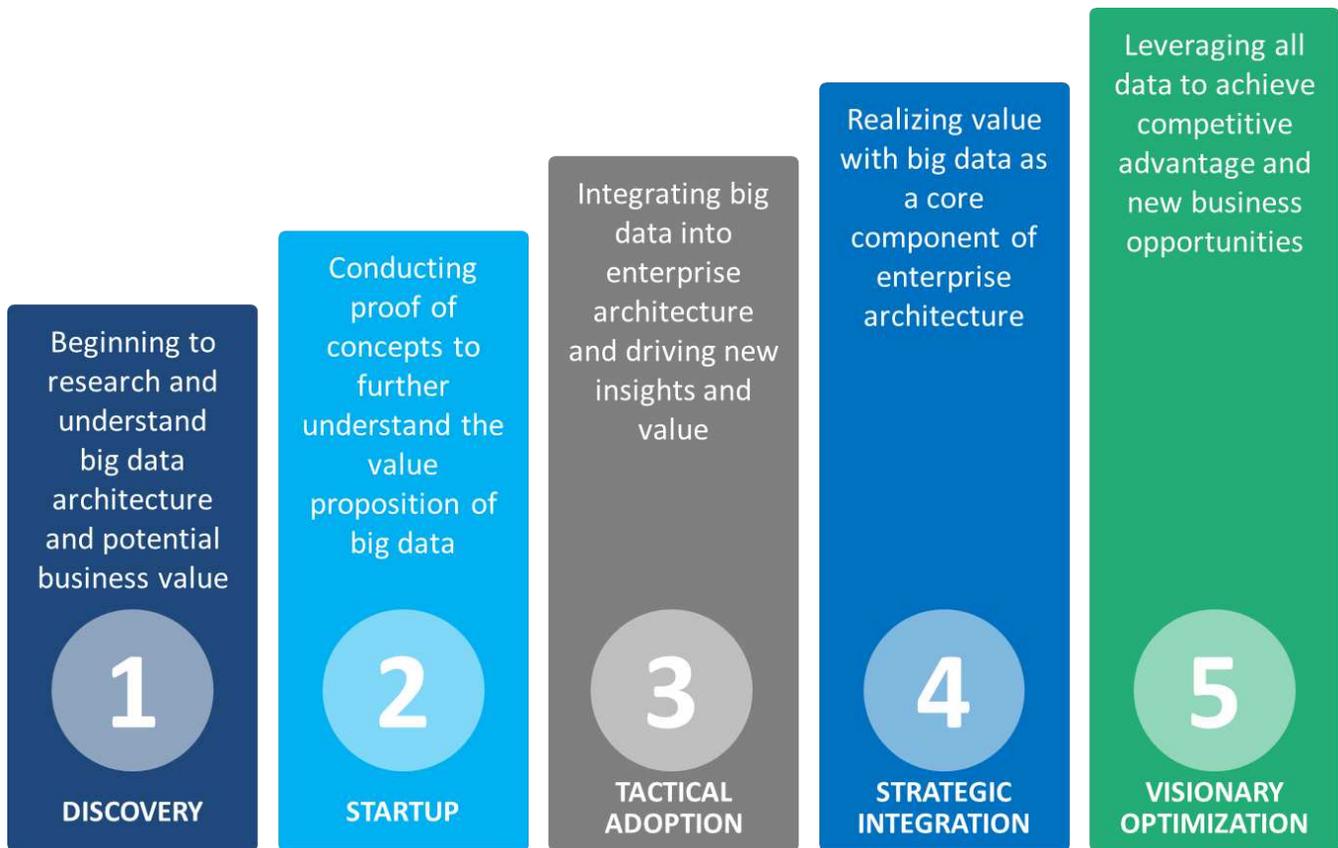


BIG DATA MATURITY MODEL

Benchmark Your Organization with Your Industry Peers



CBIG's Big Data Maturity model is a 5-stage, evolutionary progression to get your organization thinking about possibilities—it can be used as both a guide in determining where your company stands in relation to its current big data initiatives, and as a roadmap for future growth.



STAGE 1 – DISCOVERY

Description

Organizations at Stage 1 are beginning to research, review, and understand what big data is and its potential to impact the business. They evaluate what big data space is and how can big data be leveraged, and used for the corporation's benefit. At the enterprise level, they are beginning to realize their big data storage requirements are growing and will not handle the growth with current technology. They may not realize use cases that can applied. They see big data as data storage or Hadoop only.

People

A few people may be trained or knowledgeable on big data or in relational databases. Data best practices not defined and not being created.

Process

No process at corporate level, and no routine best practices for data and databases.

Technology

Relational databases, sub 1 TB level, data growth is not planned or handled with current assets and technology. Applications will not handle higher level data growth beyond current data sizes.



STAGE 2 – STARTUP

Description

Organizations at Stage 2 continue to understand what big data is but are now taking steps to design, plan, and install an initial cluster (Hadoop, MPP databases) and evaluate and select the Hadoop distribution and toolset. They are reviewing analytics and use cases and obtaining data to support the use cases. At the enterprise level, they are taking steps to plan and architect the big data storage requirements. They are beginning to realize use cases that can applied, or analytics, and they are beyond seeing big data as data storage or Hadoop only.

People

A few people may be trained or knowledgeable on big data and the initial concepts. Beginning to research into and define data best practices as related to big data. Several people are getting trained in Hadoop or MPP databases and beginning to come up to speed on technology.

Process

Beginning to define a corporate level process for big data architecture and systems, training, certification.

Technology

Reviewing the architecture, design, and installation of Hadoop and MPP databases in the technology space. Reviewing how to integrate EDW and other enterprise level data repositories. Obtaining storage that will scale beyond the 1 TB data storage space. Applications are being designed that can use the scale and performance of a MPP system.



STAGE 3 – TACTICAL ADOPTION

Description

Stage 3 organizations understand the value proposition of big data, what it is, and continue evaluation but they are still learning at the tactical level. No strategic level planning or implementation at this level. At the enterprise level, they begin to realize their big data storage requirements are growing and they look at how to leverage that data as an asset. They select use cases to implement as an initial basis. Integrated to EDW and other data repositories.

People

Additional people are being trained and learning big data technologies. Big data best practices are being defined but not yet implemented.

Process

Beginning to implement processes for big data implementation at corporate level and establishing best practices. Maintenance for databases and data management.

Technology

Installed and implementing, at the tactical level, Hadoop and MPP databases. Data size beginning to grow beyond 1-10 TB level. Data is currently integrated from EDW and being retrieved from databases into Hadoop. Data growth is being planned with new storage assets and technology. Applications are being designed that can use the larger data amounts through parallelization (Hadoop or MPP databases).



STAGE 4 – STRATEGIC INTEGRATION

Description

Stage 4 organizations understand and have implemented big data; they see returns on the value proposition, and have use cases developed and implemented for the majority of those use cases. Use cases that can be applied have been identified and implemented with ROI, and showing results. At the enterprise level, they fully realize their big data storage requirements and have planned for growth. In addition to big data capabilities, analytics as well as predictive analytics have been developed and are being used for fully predictive capabilities in the identified use cases. Hadoop and/or MPP database clusters have been created, implemented and connected to data sources such as structured, semi-structured, or unstructured data sources such as EDWs, databases, log files, machine data files.

People

Many people have been trained and have hands on experience working with big data (Hadoop and or MPP databases). People are trained on and following the data best practices.

Process

Processes and best practices are fully defined at the corporate level, and are routinely followed. Some best practices may not be fully followed but most are.

Technology

Big data technologies such as Hadoop, MPP databases, and analytics, are fully implemented and running on several projects and use cases. Additional use cases are being created and implemented. Data storage technology has been installed and being used. Applications are using MPP capabilities to take advantage of the big data technology.



STAGE 5 – VISIONARY OPTIMIZATION

Description

Organizations at Stage 5 have a visionary and optimized understanding of big data, and use for their benefit big data capabilities based on use cases and in a fully predictive way on use cases. At the enterprise level, they have fully realized their big data storage requirements and planned for this, and implemented hardware software to handle large amounts of data (500 TB to >PB scale). Use cases have been created, implemented, and seeing full big data and predictive analytics capabilities. At Stage 5, the business is now fully transformed into a “predictive enterprise” where the results are generated that are allowing business to predict as opposed to react to customer events.

People

Many people are trained or knowledgeable on big data or in relational databases. Data best practices are fully defined and adhered to. Business is proactive, in people resources being trained and visionary in looking at technology training needs.

Process

Best practices and big data process is fully defined at corporate level, and routine for data management for databases and data management.

Technology

At the enterprise level, Hadoop, MPP databases, EDW, other technologies are fully connected and integrated into the business use cases and fully support these use cases for full ROI realization. Relational databases, are replaced with Hadoop and MPP databases and “data marts” and “data lakes” and are now are at 100s TB to petabytes levels. Data growth is fully planned and realized with current assets and technology. Applications can run against higher level data amounts and are parallelized for growth.

Author: Steve Thompson
Vice President of Big Data Solutions and Practice Lead
CBIG Consulting

For more information, visit our web site at www.cbigconsulting.com