

Making Sense of Big Data

A Day in the Life of an Enterprise Architect

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper
Prepared for IBM

January 2017



IT & DATA MANAGEMENT RESEARCH,
INDUSTRY ANALYSIS & CONSULTING

Making Sense of Big Data: A Day in the Life of an Enterprise Architect

Table of Contents

- Introduction 1
- New and Exciting Data Sources..... 1
- Growing List of Projects 2
- Diverse Platforms..... 2
- How to Ease the Trials and Tribulations of the Enterprise Architect 3
 - Customizable Applications to Speed Implementations..... 4
 - Training Internal Staff Rather Than Using Expensive Contractors..... 4
 - Communicating/Collaborating with Business Stakeholders..... 4
 - Empowering Self-service Architectures with a Hybrid Data Ecosystem 5



Making Sense of Big Data: A Day in the Life of an Enterprise Architect

Introduction

The life of an enterprise architect is becoming busy and difficult. Before the era of big data, the enterprise architect “only” had to worry about the data and systems within their own data center. However, over the past decade there were revolutionary changes to the way information is used by businesses and how data management platforms support the information available from modern data sources.

Cloud broke down the boundaries of enterprise data centers, with applications housed and data created outside the “four walls” of an organization. This introduced a host of complexities for enterprise architects focused on security, privacy, and control. Mobile influences continued to push data outside the data center. Maintaining data flows to each of those data access points, often as tablets or mobile phones, introduced additional troubles. Incoming data from mobile devices brought new data formats and a flood of information to the enterprise architect.

These alterations in the formats and locations of systems and data created massive change for data-driven organizations who want to develop competitive advantage. That advantage may come in the form of new data sources such as device sensor logs, social media streams, and mobile device geolocation information; create new projects to take advantage of these new data sources; and establish environments with diverse data management platforms to support these efforts.

Enterprise architects use a combination of experience, expertise, and vision to provide guidance and facilities to their organization’s technical infrastructure. Enterprise architects take into account business requirements and technical capabilities to provide the vision for an organization’s future. These future strategies need to mesh with existing technical infrastructure for future development to be attainable.

New and Exciting Data Sources

The transformations over the past decade mandated and created a range of additional data sources, both inside and outside an organization, for enterprise architects to consider.

External data sources from third-party content, often via cloud-based providers, can change their data structure without notification to downstream organizations using that information. Event log and device sensor information also have variability based on their individual configuration and frequently come in the form of multi-structured formats such as XML or JSON. Social data sources created by customers and the general public are based on textual formats and audio/video content. Both text and audio/visual are difficult to store and utilize due to the nature of the information.

Traditional relational data sources are also included in the wave of big data change, but they have their own challenges. Increasingly, the information is coming from outside the data center from third parties or cloud-based implementations of corporate data, which requires enterprise architects to seek out and learn how to utilize that information.



Making Sense of Big Data: A Day in the Life of an Enterprise Architect

Growing List of Projects

The expansion of information associated with big data led to a corresponding increase in the number of projects that enterprise architects need to design, implement, and manage. As data-driven organizations move to integrate big data sources, they find new available applications for the data. In the recent EMA/9sight end-user big data research study, respondents indicated they experienced tremendous growth from just a few years ago.

In 2013, nearly one in three organizations indicated they were working on a single big data project. One project means a single set of requirements and a single mandate for implementation and maintenance. Today, that number is just under one out of five organizations.

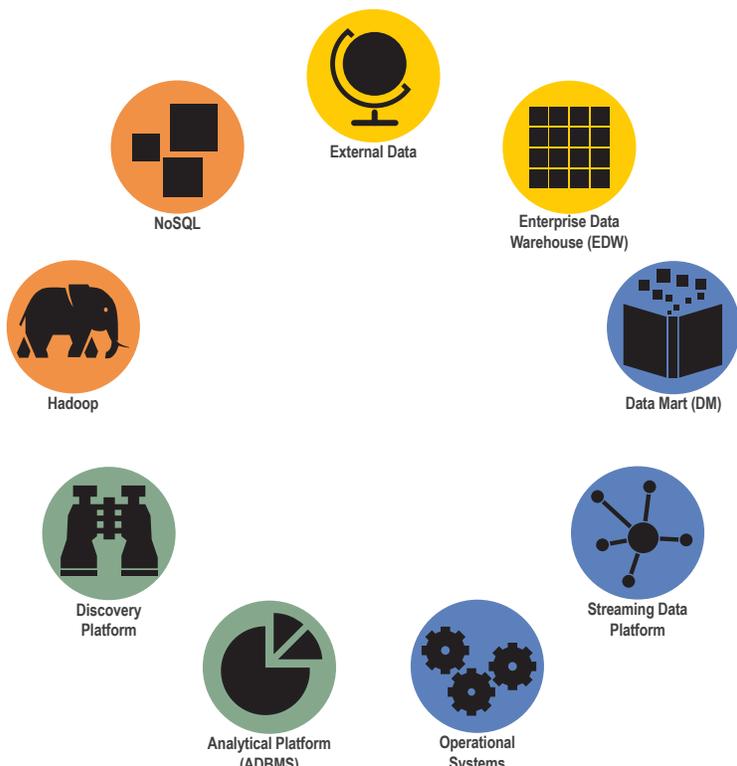
Correspondingly, the number of multiple big data projects doubled. In 2013, 32% of respondents indicated they were implementing three or more projects. That number grew to nearly 61%. This increase in projects per organization equates to at least triple the amount of requirements to collect and implement, and three times the number of data consumers to support this growth.

2x

Twice as many organizations with **three or more big data projects** in 2016 as there were in 2013.

Diverse Platforms

With the increased number of data sources and projects, enterprise architects are experiencing a similar surge in the number of data management platforms to support these big data initiatives. Over the past five years, panel respondents from the EMA/9sight study have consistently shown that they are using a variety of platforms to support big data. These platforms are not limited to just Hadoop or NoSQL environments, but include traditional relational databases such as the enterprise data warehouse and data marts, and operational support platforms such as enterprise applications like ERP and SCM. Examples of NoSQL platforms include Cassandra and MongoDB. Streaming analytical platforms to analyze data as it streams from applications and IoT devices is also considered part of this growing ecosystem. In between the traditional and non-traditional data management paradigms are data discovery platforms that allow users to explore relational and multi-structured data sources.



Making Sense of Big Data: A Day in the Life of an Enterprise Architect

With this disparate number of data management platforms, enterprise architects are mixing and matching to meet their hybrid business requirements. Event-level data stored in Hadoop and/or NoSQL is mixed with reference data within operational systems or from the enterprise data warehouse. Curated customer data from third parties for demographics or firmographics augments the existing information with the EDW, and provides additional context for Hadoop and NoSQL data. 66% of respondents in EMA'sight big data end-user research indicated they were using 2-6 of these platforms in concert to support their big data initiatives. EMA calls this concept the Hybrid Data Ecosystem (HDE).

How to Ease the Trials and Tribulations of the Enterprise Architect

When faced with the transformative adjustments associated with the big data era, organizations and their enterprise architects can choose one of two paths. The first path is to be reactive in association with big data, meaning organizations will wait until there is no choice but to change. The architectural adjustments to their technical environments and implementation strategies will be manual and slow. Also, reactive organizations focus almost entirely on technical aspects, as opposed to collaborating with business stakeholders. This limits the value of new big data sources and existing business and technical components.

The second path allows proactive organizations to use the transition as another way to build a practical architecture within their organization, which includes ways to quickly implement their big data projects and remain flexible for future change. It will include cultural components to build from within instead of going outside the organization for quick fixes for implementation staff, as well as strengthening a relationship with business stakeholders.

66%

of organizations are utilizing 2-6 different platforms in the big data environments.

Data Operations

noun

da·ta op·er·a·tions \, 'dā-tə ä-pə-'rā-shəns

Similar to systems administration or network operations, data operations are concerned with management in terms of flow, storage, and retention of data within an organization. However, data operations are not necessarily concerned about the strategic use of data; merely the maintenance and recovery of the data in question. This can be analogous to a management strategy of following a set process (doing things correctly) but not necessarily leading with vision (doing the right things).

Making Sense of Big Data: A Day in the Life of an Enterprise Architect

#1

Top project implementation strategy: customizable applications

Customizable Applications to Speed Implementations

When proactive organizations and enterprise architects put together their big data initiatives, they look for solutions that speed time to implementation *and* prepare for future adjustments to their environment(s). In the past two surveys, EMA/9sight panel respondents indicated that their top preference is to utilize third-party applications with configurations that provide “instant” implementation rather than utilizing custom application code to implement future adjustments.

With this implementation choice, enterprise architects gain the flexibility to quicken big data projects and build on corporate lessons learned as the organization moves forward. For example, a configurable advanced analytics big data solution allows for the initial installation to be implemented and utilized in short order. As the organization learns more about the analytical model(s) and how they apply to the business, change is inevitable. A solution configurable without the requirement of custom coding allows flexibility and speed for internal resources to make “tweaks” to the configuration.

Training Internal Staff Rather Than Using Expensive Contractors

With the ability to easily configure the software associated with their big data projects, organizations are choosing to develop their internal staff (18.2% of survey panelists in 2016) rather than pull in outside resources. Organizations can attract external expertise in Hadoop, NoSQL, and other big data technologies to fill staffing needs. However, developing big data talents internally allows the people with intimate knowledge of corporate goals to apply both their existing expertise and their new skills. This will give enterprise architects the best combination of staff to promote big data initiatives and feel invested in those activities, instead of having isolated sets of business and technical knowledge.

18%

Nearly one in five organizations are using organic staff development to cultivate staff for big data initiatives

3

Business stakeholders are the #1 consumer of big data projects years in a row

Communicating/Collaborating with Business Stakeholders

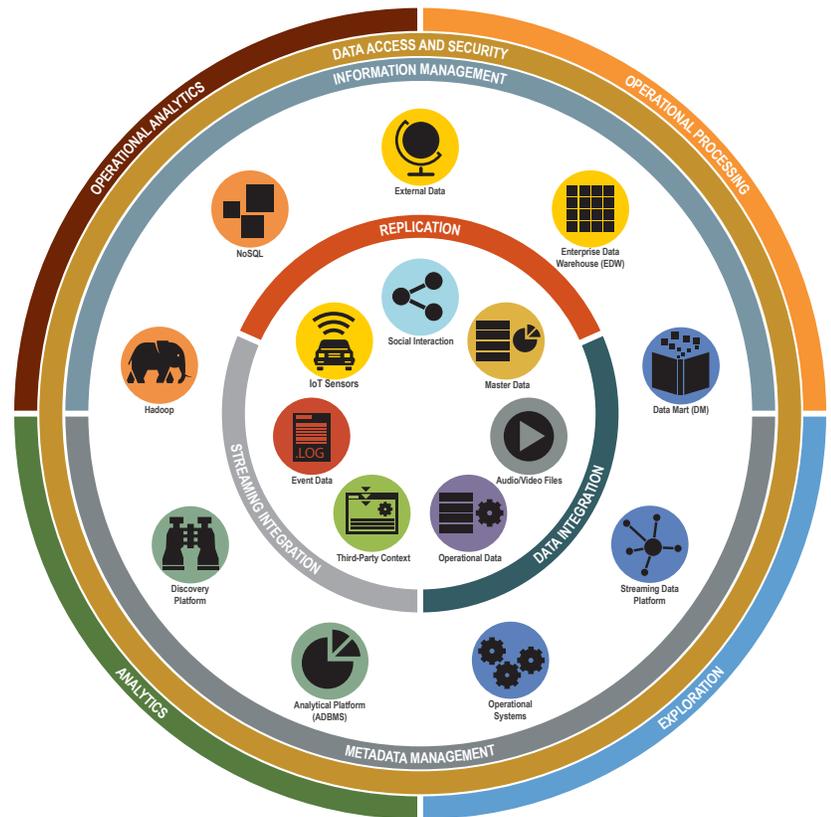
Working with internal resources on big data projects extends beyond developing staff. In three consecutive surveys, the EMA/9sight panel indicated that business stakeholders (not technologists) were the main consumers of the results and information from big data projects. From this perspective, proactive organizations arm their enterprise architects with tools to collaborate with business stakeholders. This provides better alignment with those stakeholders such as operations, human resources, and finance teams. Collaboration also

Making Sense of Big Data: A Day in the Life of an Enterprise Architect

permits the shared responsibility of several elements of a big data initiative. By linking with the development of internal staff, organizations can develop technical and business resources to define, configure, and maintain big data project environments.

Empowering Self-service Architectures with a Hybrid Data Ecosystem

With access to business stakeholders experienced with big data sources and talented technologists trained for big data implementations, there is a great opportunity for enterprise architects. They can leverage the collaboration between these two groups to create something unique. This partnership provides the foundation for self-service projects associated with big data initiatives. In these areas, enterprise architects can deliver a flexible big data framework complete with metadata management, processing and data access tools, and data visualization and storytelling facilities.



Self-service environments enable business users to access and work with corporate data even though they do not have a background in data management, statistical analysis, business intelligence, or data mining. Allowing data consumers to make decisions based on their own requirements and analysis allows organizations' information technology teams to focus on providing strategic environments that empower data consumers, rather than place IT emphasis on tactical tasks, manual implementation methodologies, and the ability to react quickly to capitalize on business opportunities.

Making Sense of Big Data: A Day in the Life of an Enterprise Architect

Business-oriented data consumers can manage their own pursuits, which can take the form of data discovery or exploratory activities to find new uses for big data sources. They can be analytical projects to align costs via cost-management activities, or advanced analytical projects to find the next set of attributes for high-revenue customers.

This is not a one-off implementation like a “spreadmart” or an unmanaged shadow IT project. Instead, it is a supervised and administered environment strategically provided to business stakeholders so they can meet their own needs while utilizing corporate assets. This may be in the form of new big data sources and platforms, and capturing valuable metadata that can be utilized across the organization.

Enterprise architects create self-service capabilities for big data. By designing and implementing application environments with configurable software components, empowering technologists through skills development, and actively sanctioning interaction between technology implementation teams and the various lines of business, enterprise architects provide environments where business stakeholders can have requirements met at the “speed of business” rather than the speed of an implementation backlog. The business can focus on how to best use new big data resources without waiting on tactical IT workflows while maintaining the implementation components for distribution and continued development across the organization.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help EMA's clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals, and IT vendors at www.enterprisemanagement.com or blogs.enterprisemanagement.com. You can also follow EMA on [Twitter](#), [Facebook](#), or [LinkedIn](#).

This report in whole or in part may not be duplicated, reproduced, stored in a retrieval system or retransmitted without prior written permission of Enterprise Management Associates, Inc. All opinions and estimates herein constitute our judgement as of this date and are subject to change without notice. Product names mentioned herein may be trademarks and/or registered trademarks of their respective companies. “EMA” and “Enterprise Management Associates” are trademarks of Enterprise Management Associates, Inc. in the United States and other countries.

©2017 Enterprise Management Associates, Inc. All Rights Reserved. EMA™, ENTERPRISE MANAGEMENT ASSOCIATES®, and the mobius symbol are registered trademarks or common-law trademarks of Enterprise Management Associates, Inc.

Corporate Headquarters:

1995 North 57th Court, Suite 120
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com
3515.012617