Analytics Business Maturity Model

As I stated earlier, Big Data Analytics is a journey and can be implemented using a number of iterative phases, each advancing the capability via well-defined yet small steps to reduce risk. The Information Agenda team has observed a large number of analytics programs worldwide and has developed a set of benchmarks for analytics at different levels of maturity. These benchmarks have been captured using a business maturity model that allows us to specify current and target levels of maturity and what can be achieved in each phase. The model has five levels of maturity:

- *Breakaway*—A company that is generally considered to be the best in the class in its execution of key business strategies, able to exhibit the characteristics of an agile, transformational, and optimized organization. This classification excludes "bleeding-edge" or pioneering aspects; however, these aspects may also be evident in such companies. Key predictive performance indicators are used in modeling for outcomes, and information is utilized enterprise-wide for multidimensional decision making.
- *Differentiating*—A company whose execution of key business strategies through utilization of information is viewed as generally better than most other companies, creating a degree of sustainable competitive advantage. Management has the ability to adapt to business changes to a degree, as well as measure business performance. Business leaders and users have visibility to key information and metrics for effective decision making.
- *Competitive*—A company whose capabilities generally are in line with the majority of similar companies, with a growing ability to make decisions on how to create competitive advantage. This maturity level is also the starting point to establish some consistency in key business metrics across the enterprise.
- *Foundational*—A company whose capabilities to gather key information generally lag behind the majority of its peers, which could potentially result in a competitive disadvantage. Information is not consistently available or utilized to make enterprise-wide business decisions. A degree of manual efforts to gather information is still required.

	Ad hoc	Foundational	Competitive	Differentiating	Breakaway
Capability: Monitor brand sentiment	Marketing has hired a set of interns to monitor social media data	Organizational accounts to collect sentiment data on social media sites (Facebook, Yelp, etc.)	Customer data from social media is collected and analyzed using analytical tools	Organization engages in social media conversation to influence customer sentiment	Customer sentiment is integrated with product and marketing processes
Measurements					
Brand sentiment	Baseline	Collected	Measured	Influenced to positive direction	Influenced to positive direction
Identification of advocates / ambassadors	Baseline	Low	Medium	High	High
Impact on brand / revenue	Baseline	Baseline	Small	Medium	Large

Figure 6.1: Social media maturity model

- *Ad hoc*—A company that is just starting to develop the capability to gather consistent information in key functional areas, generally falling well behind other companies in the corresponding sector. Information beyond basic reporting is not available. Time-consuming, manual efforts are generally required to gather the information needed for day-to-day business decisions.

This model is an important tool in developing an enterprise-wide analytics roadmap. It allows us to specify specific capabilities developed in each phase, compare them with others in the industry, and align metrics to each level, so that the benefits can be identified using the metrics and can be quantified using either benchmarks or company-specific information.

The business maturity model lets us rapidly quantify the benefits of an analytics program. We have been tracking actual benefits using case studies and using these benchmarks in roadmap development. The maturity models and their underlying descriptions are industry-specific, as implementations and benefits differ from one industry to the next.

Figure 6.1 shows an example of the maturity model applied to the capability Monitor Brand Sentiment. At the Foundational level, the marketing organization establishes a Facebook account, which is used by customers to express sentiments; however, the sentiment information is not used in any way. At the Competitive level, the organization establishes a mechanism for collecting,

collating, and analyzing sentiment and tracking its value with the marketing events. At this stage, the sentiment is measured but not actively managed.

Lisa Mancuso, SVP of Marketing for Fisher-Price, recently talked about the company's ambassador program in an interview with *Forbes* magazine. "We know that more than two-thirds of mothers consider blogs to be a reliable resource for parenting information, so we have created a robust program to connect with parenting bloggers around the world. We call them our Play Ambassadors."[31] Such programs, when actively integrated with social media accounts, give organizations the capability to start differentiating themselves in their ability to converse with the customers.

At the Breakaway level, sentiments from social media are linked to the product and marketing processes. Sentiments are monitored in response to a product launch, pricing, changes, or a new advertising campaign. The results are directed to product and marketing processes to modify the product and its marketing. Successful marketers would use social media as a channel to experiment with different product options and use the feedback to launch the one with the best customer response.

Chapter 7
Closing Thoughts

I started this book with a definition of Big Data using the four V's: Velocity, Volume, Variety, and Veracity. Big Data growth can be attributed to three market forces: sophisticated consumers, product and process automation, and data monetization. I discussed a number of emerging use cases, including location-based services, micro-segmentation, next best action, Product Knowledge Hub, Social Media Command Center, infrastructure and operation improvement, and risk management. The solution includes a number of architecture components. Massively parallel platforms provide capabilities for data integration, storage, and analytics. Unstructured text analytics complements traditional quantitative modeling. Big data enhances the creation of customer and product MDM. Real-time adaptive analytics provides high-velocity analytics while changing its modeling parameters based on sophisticated predictive modeling of historical data.

I discussed data privacy issues and how some of the data can be masked to limit exposure. These components can be organized in a three-layer architecture, with a conversation layer that uses real-time analytics to provide low-latency decisions and an orchestration layer that synthesizes entities, controls the conversation, and offers visibility to business users via a command center. The supporting discovery is provided by unstructured and structured analysis. Last, I discussed implementation approaches, data governance, roadmap development, and maturity models.

By calling it "Big Data," our attention obviously goes first to the volume dimension. With data sizes in exabytes, the analytics requires special tools capable of scaling to such big volumes. We saw how massively parallel

platforms provided performance that naturally scales. The HDFS platform offers inexpensive data storage but requires special skills to manipulate the data. Also, as we collect more data, we increase our chances of improved identity resolution.

The velocity dimension forces us to establish an architecture where conversations can be intelligent and yet fast enough to handle the velocity requirements for the use cases. Location-based campaigns and web searches are two examples of capabilities that require low-latency response. Real-time adaptive analytics provides a robust architecture to deal with low-latency analytics while at the same time adjusting the models to accommodate changes based on historical and predictive modeling. The orchestration layer allows us to converse intelligently based on historical data, sophisticated models, and both unstructured and structured discovery.

The variety dimension focuses on unstructured data. A number of qualitative reasoning techniques can be used in conjunction with quantitative predictive modeling to incorporate findings from the unstructured data in the predictive models. In addition, qualitative reasoning in the context of time-based correlations allows us to find a specific collection of events.

The veracity dimension focuses on data quality, governance, and privacy-related issues. By incorporating a proper governance framework, we can identify faulty data and discount it before creating predictive models. The result is a thorough cleanup of the data before it is used in a critical customer-facing situation.

We looked at a number of use cases. Use of Big Data has enormous potential in product selection, design, and engineering; however, this area is still in its infancy. The most successful production applications are using Big Data to improve infrastructure, monitor customer feedback through the Social Media Command Center, and advance micro-segmentation and intelligent campaigns.

We discussed the Advanced Analytics Platform (AAP) as the overall integrated architecture that combines Big Data with traditional Business Intelligence and Data Warehouse components. Most of the greenfield organizations are leapfrogging using Big Data Analytics and have taken a revolutionary approach to their analytics architecture. However, mature organizations with significant investment in BI and Data Warehousing are using more of an evolutionary approach to the overall architecture, with a hybrid architecture that combines the traditional data warehouse architecture with the newer Big Data capabilities.

We posed three questions at the beginning of this book. Let us try to answer them now using the material discussed in the book.

1. *What is Big Data and what are others doing with it?*
   Chapters 1 and 2 provided a definition of Big Data in terms of velocity, volume, variety, and veracity and discussed the popularity of Big Data due to market forces. The use cases provided examples of how businesses are using Big Data today.

2. *How do we build a strategic plan for Big Data Analytics in response to a management request?*
   Big Data Analytics is a multi-year, multi-phase journey. It is important to have a strategic vision that aligns with industry direction and responds well to the disruptive forces. It is also important to pick a target that makes a substantial impact on the organization. However, it is equally important to select short-term projects with short durations and measurable impact. Choosing areas closer to product engineering, operations, or infrastructure will provide quick and early results. Privacy is a difficult topic that should be handled with care.

3. *How does Big Data change our analytics organization and architecture?*
   The Big Data Analytics program does not work in a silo. Integration with the current environment is probably the most difficult part of the development activity. Care must be taken in establishing a strategic architecture along the lines discussed in Chapter 5 and in experimenting to see how an integrated architecture would support business processes using a combination of Big Data and conventional analytics tools.

Big Data is still an emerging topic. However, it has already resulted in major disruptions in many markets. In the world of analytics, it has changed how we view BI. Unlike in the past, where operational information was collected in the warehouse to be analyzed and researched over the long haul, current-day technologies are bringing analytics closer to the conversation. It requires the orchestration and conversation layers of the architecture in order to respond to the velocity and volume of data.

Notes

1. Sunil Soares, "A Framework That Focuses on the Data in Big Data Governance," *IBM Data Management*, June 13, 2012. *http://ibmdatamag.com/2012/06/a-framework-that-focuses-on-the-data-in-big-data-governance*.

2. "What Data Says About Us," *Fortune*, September 24, 2012, p. 163.

3. "Top 10 Largest Databases in the World," March 17, 2010. *http://www.comparebusinessproducts.com/fyi/10-largest-databases-in-the-world*.

4. "Statshot: How Mobile Data Traffic Will Grow by 2016," August 23, 2012. *http://gigaom.com/mobile/global-mobile-data-forecast*.

5. Kate Maddox, "Turn Ad Inspired by 'Mad Men'," *www.btobonline.com*, July 16, 2012.

6. Ben Grubb, "Can't Buy Love Online? 'Likes' for Sale," *www.stuff.co.nz*, August 24, 2012.

7. Rob Van Den Dam, *Global Telecom Consumer Survey*, IBM Institute for Business Value, 2011.

8. Ibid.

9. *http://www.iab.net/about_the_iab/recent_press_releases/press_release_archive/press_release/pr-041311*.

10. *http://www.yelp.com*.

11. Amir Efrati, "Online Ads: Where 1,240 Companies Fit In," *Wall Street Journal*, June 6, 2011.

12. Valerie Bauerlein, "Gatorade's 'Mission': Sell More Drinks," *Wall Street Journal*, September 13, 2010. Adam, Ostrow, "Inside Gatorade's Social Media Command Center," Mashable Social Media, June 15, 2010. Also see the YouTube video at *http://www.youtube.com/watch?v=InrOvEE2v38*.

13. Jeff Bertolucci, "Smart Phones, Big Data Help Fix Boston's Potholes," *Information Week*, July 25, 2012.

14. "Predictive Analytics: Police Use Analytics to Reduce Crime," *http://www.youtube.com/watch?v=_ZyU6po_E74&feature=relmfu*.

15. By "opting-in," a consumer may choose to allow use of location information, typically in exchange for a free or discounted service.

16. Don Peppers and Martha Rogers, *The One to One Future: Building Relationships One Customer at a Time*, Bantam Press, 1997.

17. Robert Lee Hotz, "The Really Smart Phone." *Wall Street Journal*, April 23, 2011.

18. Robert Andrews, “NBC: Nearly Half of Olympic Streams are from Mobile, Tablet.” August 2, 2012, Paid Content, *www.paidcontent.org*.

19. Kuang-Chih Lee, Burkay Orten, Ali Dasdan, Wentong Li, “Estimating Conversion Rate in Display Advertising from Past Performance Data.” *www.turn.com*.

20. Tom White, *Hadoop: the Definitive Guide*, O’Reilly Yahoo! Press, 2009.

21. T.R. Gruber, “A Translation Approach to Portable Ontologies,” *Knowledge Acquisition* 5, no. 2 (1993): 199–220, 1993. Also see A. Sathi, M. Fox, and M. Greenberg, “Representation of Activity Knowledge for Project Management,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 7, no. 5 (May 1985).

22. Jeff Hefflin, “OWL Web Ontology Language: Use Cases and Requirements,” *www.w3.org*.

23. John Dawes, “Close the Multi-Channel Customer Experience Gap,” *www.tealeaf.com*, January 2011.

24. Drew Fitzgerald, “Yahoo Passwords Stolen in Latest Data Breach,” *Wall Street Journal*, July 12, 2012.

25. Charles Duhigg, “How Companies Learn Your Secrets,” *New York Times*, February 16, 2012.

26. Anick Jesdanun, “FTC Finalizes Privacy Settlement with Facebook,” *Huffington Post*, August 10, 2012.

27. Garland Grammer, Shallin Joshi, William Kroeschel, Arvind Sathi, Sudir Kumar, Mahesh Viswanathan, “Obfuscating Sensitive Data While Preserving Data Usability,” USPTO Patent Number 20090132419. United States Patent and Trademark Office: *http://www.uspto.gov*.

28. William Kroeschel, Arvind Sathi, Mahesh Viswanathan, “Masking Related Sensitive Data in Groups,” USPTO Patent Number 20090132575. United States Patent and Trademark Office: *http://www.uspto.gov*.

29. Julia Angwin, “A New Type of Tracking: Akamai’s Pixel-Free Technology,” *Wall Street Journal*, November 30, 2010. *http://blogs.wsj.com/digits/2010/11/30/a-new-type-of-tracking-akamais-pixel-free-technology*.

30. “The Best Performance Is the One You Can’t See,” IBM Website, *www-05.ibm.com/innovation/fr/rolandgarros/en*.

31. Brandon Gutman, “Fischer-Price on Connecting with Moms in the Digital World,” *Forbes*, September 13, 2012.

Abbreviations

AAP	Advanced Analytics Platform
BI	Business Intelligence
BSS	Business Support System
CCI	Cognos Consumer Insight
CDR	Call Detail Record
CSP	Communications Service Provider
DMP	Data Management Platform
DSP	Demand Side Platform
DW	Data Warehouse
ETL	Extract Load Transform
HA	High Availability
HDFS	Hadoop Distributed File System
IPO	Initial Public Offering
IVR	Interactive Voice Response
KPI	Key Performance Indicator
MDM	Master Data Management
MPP	Massively Parallel Platform
NBA	Next Best Action
OLTP	On-Line Transaction Processing
OSS	Operations Support System
PII	Personally Identifiable Information
PoS	Point of Sale
SMP	Symmetric Multi-Processing
SSP	Supply Side Platform
STB	Set Top Box
STP	Straight Through Processing