

# BUILDING A DIGITAL ANALYTICS ORGANIZATION

Create Value by Integrating Analytical Processes,  
Technology, and People into Business Operations

JUDAH PHILLIPS

## Praise for *Building a Digital Analytics Organization*

“The allure of Big Data is immense. There is SO MUCH DATA! Yet, data collection does nothing by itself. It actually does financial harm left in the wrong hands. Fix that, and you win. Let Judah show you how to build an organization where Big Data’s primary imperative is to drive Big Action.”

—**Avinash Kaushik**, author of *Web Analytics 2.0* and *Web Analytics: An Hour a Day*

“Without a digital analytics organization, you aren’t optimizing your business, your site, or your app. Reading this book can help you. I just hope you adapt before one of your competitors does.”

—**Bryan Eisenberg**, author of *Call to Action*, *Waiting for Your Cat to Bark*, and *Always Be Testing*, Keynote Speaker, Cofounder of the Digital Analytics Association, and Publisher of UseTheData.

“Since founding Gomez 15 years ago, I’ve had the privilege of working with many pioneers in the field of web site and e-commerce excellence. I am impressed with the managerial principles and analytical techniques Judah has developed and presented in this book. A must read for the 21st century analyst and executive who wants to learn how to create value and deliver excellence with digital analytics.”

—**Julio Gomez**, Founder of Gomez, Inc., General Manager at Attivio

“Before ‘Big Data’ and ‘Data Science’ became buzz words, Judah was extracting actionable insights from immense data sets and revolutionizing the field of business analytics. Unfortunately, the digital version of his nerdy brain is still too big for a download, so reading and internalizing this book is the second best option for anyone interested in improving their business via an intelligent approach to data and analytics.”

—**Yaakov Kimelfeld**, Ph.D., Chief Research Officer of Compete

“Among the most valuable things you’ll find in Judah’s book is his experience. *Building a Digital Analytics Organization* is packed with real-life guidance and wisdom from his years of work as a practitioner and manager in the analytics field. From defining measurement needs, to analyzing data, to comparing analytics tools, Judah has done just about everything. We can all put his experience to use as a guide as we build out our own analytics organizations.”

—**Justin Cutroni**, author of *Google Analytics*, Analytics Evangelist at Google.

(overhead or sunk). And some might argue that tools don't actually do analysis at all ever, only people do.

My friend Avinash Kaushik is a keen advocate of the 90/10 model for allocating an analytics budget, where 90% of the money is spent on people and 10% on tools. While that ratio may make sense as a philosophy, in practicality it is hard to match that suggestion—though if possible it would be an excellent level of resourcing. He was smart and forward-thinking in setting the bar high—and in the right frame with which to resource analytics. As I said, people are the most important. Others have identified a split of 60/40 between budgeting for people and tools. The best ratio for investment of tools to people depends on your business requirements—so in reality, no ratio is going to be accurate for all cases. A 50/50 split in cases where software licenses or cloud-based services are involved could be the appropriate level. The analytics leadership needs to determine what the best and correct ratio is. In fact, the larger the corporation and more complex the data, the more likely it is that the balance of budget will tip in favor of analytics technology because it costs a lot when collecting and processing big data volumes. Regardless of the accuracy of any suggested ratio, the larger point understood is that some investment in tools is necessary, but it should be balanced against the fact that people do analysis, not tools. So how then can digital analytics tools be understood in more clarity if all tools that an analyst ever learns about are analytics tools?

The answer is by categorizing tools in the ReDCARPS framework. Tools can be understood as doing one or more of the following: **Requirements**, **Definitions**, **Collection**, **Analysis**, **Reporting**, **Prediction**, or **Storage**.

- **Requirements:** Technologies used for capturing, sorting, prioritizing, scheduling, and tracking analytics work, projects, and programs from start to finish.
- **Definitions:** All digital data must have definitions that enable understanding by the various stakeholders who create and use the data. Definitions must include a business definition, an operational definition, and a technical definition. See Chapter 6, “Defining, Planning, Collecting, and Governing Data in Digital Analytics.”

- **Collection:** Technologies that collect digital data from methods such as packet sniffing, database writes, log files, JavaScript, application programming interfaces (APIs), server-to-server connections, data feeds and other methods, including extracting, transforming, and loading (ETL) data.
- **Analysis:** The application of analytical methods to data derived from the science (and art) of analysis. Analysts and analysis tools enable data science to be applied to digital data—methods such as structured equation models, an ANOVA (Analysis of Variance), determining measures of dispersion, creating regression analysis, or applying advanced machine-learning, data-mining algorithms, and statistical methods. You know when you see analysis or are using an analysis tool because the output of the tool is something you are going to consider as an input to answering a business question. To render analysis in a human-understandable way, an analysis tool enables the analyst to work with data from one or more systems to create analysis. The analyst makes sense of what can sometimes be complex data, metrics, and visualizations from an analysis tool, and then synthesizes and simplifies the key data to present to stakeholders.
- **Reporting:** Tools that take data at various granularities and present it in a human-readable format across some device—and often allow for some level of drilling down and exploration into the data contained within the report itself. The dimensions in the data may be crossed or filters, metrics could be added, and the data can even sometimes be lightly segmented. Reporting tools include business intelligence (BI) tools that report data into pre-arranged and pre-scheduled reports, such as Cognos or Business Objects, to desktop spreadsheeting and data visualization tools (such as Prezi, PowerPoint, and Tableau).
- **Prediction:** Tools that enable data collected about past and current events—regardless the analytical method or types of data (see Chapter 5, “Methods and Techniques for Digital Analysis”)—to be used to predict one or more future events. Predictive analysis and predictive analytics tools tell you “what may happen next.” You know you are using a predictive analytics tool when the output informs you about a possible future state—and isn’t focused predominantly on explaining what can be understood about the current or past data or its trends, movements, and patterns—and often only so much so as to

investigate past data to create the variables from data collected in the past for input into the predictive model.

- **Storage:** Tools that take data and put it on some type of storage medium, such as a storage area network (SAN) or local disk or storage in the cloud in a SaaS. Storage is the disk space needed for recording and persisting your big data. Typically a minimum range for storing digital data is 13 months in order to allow for year-over-year comparative analysis.

The line between *analysis-only* tools and *predictive-only* tools is not a clean line in a similar way to how *reporting* differs from *analysis*. (See Chapter 7, “Reporting Data and Using Key Performance Indicators,” for a detailed discussion on the differences between reporting and analysis.) Prediction involves analysis and analysis involves reporting. Thus, prediction, of course, involves reporting the results of the predictive model that applies analytical methods. These different concepts should then be understood in the context of how each helps to create value as part of the Analytical Value Chain. You can use analytics tools to predict, but you would not necessarily use a predictive tool to make sense of data you’ve already collected or to explain a change in a time-series. But current and past data always serve as inputs to predictive models.

The referenced tools tend to do more than one thing, and vendors often sell complementary tools that enable you to do any of the ReDCARPS, as shown in Table 4-1.

**Table 4-1** The ReDCARPS Framework Applied to Common Digital Analytics Tools

	IBM	Adobe	Webtrends	Google
Requirements	X			
Definition	X	X	X	X
Collection	X	X	X	X
Analysis	X	X	X	X
Reporting	X	X	X	X
Prediction	X	X		
Storage	X	X	X	X

Every analytics practitioner and team has a number of tools they prefer. For example, many digital analysts are familiar with Google Analytics or Omniture; however, other types of digital analytics (such as email, customer, and advanced) require knowledge of larger and different toolsets (SQL, SAS, R, and so on). The more familiarity you have with the different historic, current, and emerging analytics tools, the more career options you will have.

Although an analyst may prefer a type of tool, many tools perform similar functions. Thus, don't get too tied to and involved with any one tool. An overemphasis on one tool can bias how you conceive and understand digital analytics. What may be difficult and require "guru" skills in one tool may be as simple as a few clicks in training from another tool.

Familiarity with tools is helpful and explains why people have preferences for their specific tools, but different tools are necessary to create different types of analysis. You would not commission a sculptor of bronze to paint a watercolor, just as you would not hire a SAS programmer to socialize business analysis to an executive audience. Instead, of course, you would hire the right horse for the right course; thus, it is likely that to successfully execute on an analytical vision, you need more than one tool to do so, just like an analytics team needs people with different, but complementary skills to execute the analytical value chain.

To help you further understand the landscape of analytics tools beyond the ReDCARPS framework, analytics tools can be divided into the following types:

- **Internal tools:** These are always deployed in a software environment—perhaps in a data center or in a cloud—that is controlled, managed, and configured for all intents and purposes by people employed by the company for which you work.
- **Site analytics tools:** Site analytics tools collect data from a digital site via a number of methods, store and aggregate it, and then report it. Some tools enable data exploration, such as crossing of or drilling down or filtering of dimensions in the data and the usage of analytical methods.
- **Business intelligence (BI) tools:** BI tools tend to exist under the control of IT and technology and include the databases and systems that support data storage and processing. Analytics

teams make use of BI data and may combine BI data with other internal or external data. BI tools also have capabilities for ETL, storage, reporting, visualization, and analysis—and in some cases, the analytical capabilities that can be applied directly to the data can be very advanced.

- **Advanced applied analysis tools:** Applied analysis tools operate off of existing data that may require data to be specially transferred or prepared to be modeled and analyzed within the tool. Advanced analysis tools allow the user to apply statistically rigorous models and algorithms to digital and other types of data. A common example is SAS.
- **Desktop analytics, reporting, and testing tools:** Many analysts use tools available in common desktop publishing programs or programs created by smaller software companies to perform a specific task. For example, analysts can use Excel to manipulate, format, and visualize data and may use a browser-based tool such as Fiddler or Charles to test data collection. As the power of computing increases, desktop tools can also contain an impressive set of statistical methods, such as regression, on data.
- **Homegrown, internally created tools:** Sometimes, off-the-shelf tools and BI tools just don't do what may be required to analyze data; thus, companies customize off-the-shelf technology or build software solutions from scratch to deliver analysis. These solutions can be fully homegrown or cobbled together from multiple vendor systems or even highly customized versions of off-the-shelf enterprise software products.
- **Online advertising and marketing tools:** This category of internal tools includes technologies for analyzing online display, (re)targeting, optimization, bidding, planning, executing, fulfilling, emailing, Customer Relationship Management (CRM) tools, and more. These tools could be deployed and managed by your company—in other words, an ad-server you host, not one that is provided by another company.

External tools include the following:

- **Competitive intelligence:** Many tools exist that help a company understand the competitive landscape, who are the competitors, where they are located, and, of course, key metrics and data related to market share, share of wallet, corporate performance, and a host of other metrics specific to the business and industry.

- **Data enhancement:** Tools that provide data not available to the company. This new data is joined with internal data to create more meaningful and valuable data. Companies include Choicepoint, Experian, Rapleaf, and so on. For example, a company may have a customer list of other businesses and may purchase mailing data from a data enhancement company to improve customer information quality.
- **Audience measurement:** Tools for competitive audience measurement include a number of companies who collect data directly from digital experiences, by observing a statistically significant audience and estimating digital behavior and metrics (panel data collection) and by a combination of both census and panel data. These companies include comScore, Nielsen, Compete, Quantcast, Google, and so on.
- **Cloud-based and SaaS online advertising and marketing tools:** SaaS technologies for analyzing online display, (re)targeting, optimization, bidding, planning, emailing, CRM, and more. These tools are hosted in the cloud by the vendor.

## To Build or to Buy?

A core question about analytics technology and available tools is whether it makes sense to build or buy? In other words, do you spend capital to purchase analytics technology for collection, reporting, analysis, prediction, and storage usually by investing in a software license or a SaaS technology? Or do you and your engineering, IT, development, quality assurance (QA), and business teams create from scratch the analytical tool ecosystem you require using custom software development and engineering?

To build or to buy seems like a simple question, but it is not. Software that everybody can buy if they have the money means that everyone may have the same measures and data with which to analyze an experience. Take for example Google Analytics, IBM Coremetrics, and Webtrends: These tools all report visitors, visits, and time-based metrics and enable custom metrics to be collected and reported. Such commonality among tools is useful for particular industries and activities in which a common vocabulary is helpful for enabling commerce. However, the counts of data from these three vendors will all differ,