

DATA SCIENCE MATURITY

Tapping the Collective Wisdom of Your Data

by
Steven Mills
Chief Scientist
Booz Allen Hamilton

Most organizations do not take full advantage of the power of their data. They use it in only basic ways, such as creating reports and analyzing past trends. But what if they could do much more? What if they could find hidden patterns and connections across all of their data? What if they could use those insights to anticipate future trends? What if they could take it yet a step further, creating thousands of possible scenarios and then identifying the ones most likely to achieve the organization's specific goals?

This is data science maturity, the step-by-step process through which an organization gradually gains more knowledge about its data to create new opportunities and solve complex problems. It is comparable to how an organization traditionally gains insight: by continually testing and refining its institutional knowledge with new information and experience. When senior leaders make a major decision, they draw on this collective wisdom. Data science maturity involves a similar iterative building process, and what emerges is a kind of collective wisdom about an organization's data.

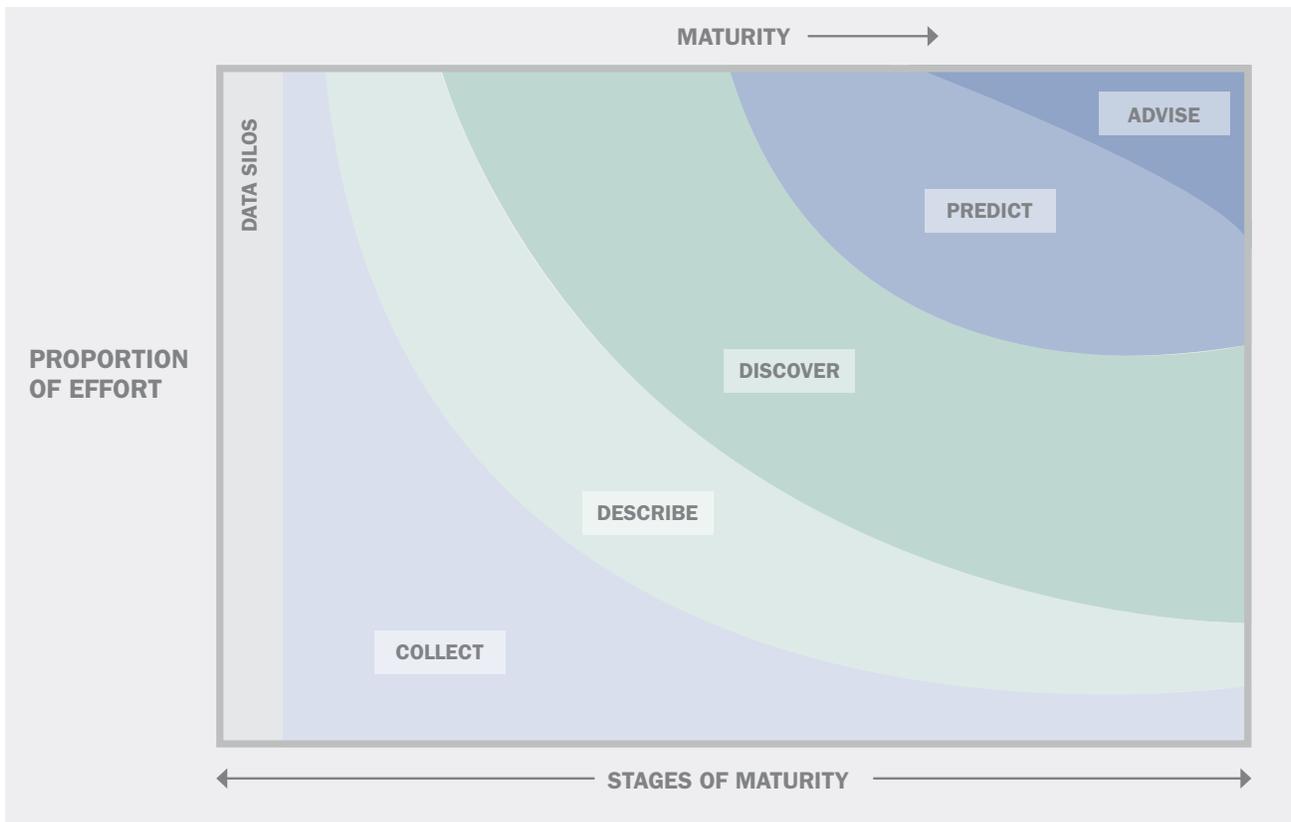
HOW DATA SCIENCE MATURITY WORKS

Data science maturity is characterized by five stages, each stage a necessary building block for the next. The first two stages are Collect and Describe, in which organizations gather data and organize it in a useful way. Most organizations are at this stage, where data analysis tends to focus on what happened in the past.

The next stage is Discover. Organizations use the data they collected and described to search for valuable patterns and connections. This kind of knowledge can be game-changing—in detecting fraud, for example, or in discovering a business opportunity that competitors do not see.

The insights gained in the Discover stage prepare an organization for the next stage, Predict. Here, data scientists use computer models to create and think through any number of possible scenarios, and then determine which ones are most likely to occur.

THE DATA SCIENCE MATURITY MODEL



In the final stage, Advise, analytics are used to sift through the most likely scenarios and recommend the optimal course of action based on the organization's resources and goals. Ultimately, decision makers will still rely on the collective wisdom of their organizations—that does not change. But now, they can tap the collective wisdom of the data as well.

As an example of how data science maturity works, say that a manufacturing company has factories in several countries. In the Collect and Describe stages, the company might gather data comparing the profitability of items manufactured in the various countries. The company can use this information to decide where to expand or shift production.

In the Discover stage, the company could map out the interrelationships of dozens of factors that influence production in each country, such as economic and political conditions, labor pools, supply-chain issues, currency rates,

and perhaps even climate change. Next, in the Predict stage, the company might use all of the knowledge gained so far to create and evaluate a broad range of foreign investment scenarios. This is where data science becomes particularly forward-looking. Though the past is still important, data analysis is less focused on understanding the past and more focused on anticipating what will happen.

In the Advise stage, analytics might be used to compare all possible investment scenarios based on the company's finances, resources, strategic plans, and other relevant factors. The result of this effort might be a detailed action plan, prepared by data scientists and subject matter experts, that recommends precisely where and how the company can best invest. The company's senior management could use this plan as a tool in their decision-making process, essentially combining the collective wisdom of the organization with the collective wisdom of the data.

Data science maturity is not simply a matter of progressing from one activity to the next. With each stage, the data becomes richer, the analytics grow increasingly sophisticated, and both become more closely integrated with the needs and goals of the organization. This continuous learning process is achieved through the collaboration of data scientists and organizational subject matter experts. Together, they build the collective wisdom of the data and determine how it can best help the organization.

The nature of the analytics themselves reveal how the building process works. For example, an analytic in the Predict stage typically combines predictive algorithms with algorithms created earlier to describe and discover. In a sense, the analytic is actually a workflow that embodies all prior effort and insight.

PURSUING THE ART OF DATA SCIENCE

Data science maturity requires entirely new approaches to data and analytics. Most organizations now use conventional computing methods that typically store data in narrow, hard-to-connect silos. Data science instead employs advanced cloud technologies that break down the silos and consolidate all of the data available from both internal and external sources. This provides organizations with the ability to search through vast amounts of data all at once. It is what makes the Discover, Predict, and Advise stages of data science maturity possible.

While organizations routinely collect and describe their data with conventional technologies, the results are not nearly as robust as with the advanced cloud technologies. Without those new technologies, organizations cannot pursue the art of data science and the powerful insights it can provide. For this reason, organizations must carefully plan their

investments, particularly in data infrastructure, storage and management, so that when they are ready to move up to the next stage of data science maturity, they will have the technology and expertise to take them there. Investments made in the early stages can either limit or pave the way for the future.

Winning organizations in the information economy will be those that can harness the collective wisdom of their data. Each stage of data science maturity brings them closer to that goal.

FOR MORE INFORMATION

Steven Mills

Chief Scientist

Mills_Steven@bah.com

301-543-4709

This document is part of a collection of papers developed by Booz Allen Hamilton to introduce new concepts and ideas spanning cloud solutions, challenges, and opportunities across government and business. For media inquiries or more information on reproducing this document, please contact:

James Fisher—Senior Manager, Media Relations, 703-377-7595, fisher_james_w@bah.com

Carrie Lake—Manager, Media Relations, 703-377-7785, lake_carrie@bah.com