Image © David Castillo Dominici | 123rf.com

Defining analytics: a conceptual framework

Analytics' rapid emergence a decade ago created a great deal of corporate interest, as well as confusion regarding its meaning.

By Robert Rose

Arguably, the term "analytics," as it is widely used today, was introduced in a research report "Competing on Analytics" [1] by Tom Davenport et al. in May 2005, and its emergence into public view coincided with the introduction of Google Analytics on Nov. 14, 2005. As can be seen in the Google Trends chart shown in Figure 1, in November 2005 searches for the term "analytics" jumped almost 500 percent. Davenport's subsequent *Harvard Business Review* paper [2] and book [3] with the same title, as well as analytics-oriented marketing programs by IBM and other companies, further raised consciousness of the word and contributed to the subsequent dramatic growth in Google searches for the term analytics. The dramatic growth in the use of the term analytics has been accompanied by a proliferation in the way analytics

is used, and phrases such as "text analytics" and "healthcare analytics" have become common.

Unfortunately, in addition to the great interest and excitement surrounding analytics, there is a corresponding amount of confusion and uncertainty regarding its meaning. Perhaps the best example of this is a statement from the beginning of a March 2011 article in *Analytics* magazine [4]: "It's not likely that we'll ever arrive at a conclusive definition of analytics," a reference to surveys of INFORMS members who at the time expressed widely divergent views on the relationship between operations research and analytics. According to the survey:

- A significant number of respondents believed that operations research is part of analytics.
- A significant number of respondents believed that analytics is part of operations research.
- Some respondents believed that operations research and analytics are the same thing.
- Some respondents believed that operations research and analytics are completely different disciplines.

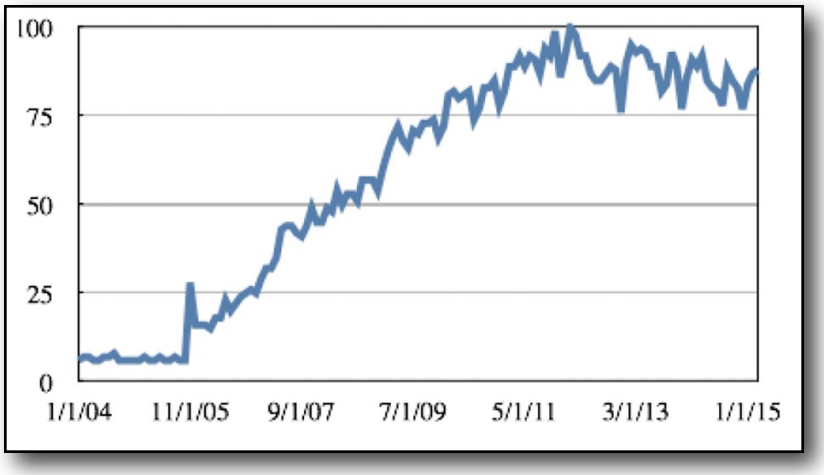

Figure 1: Google Trends chart shows the rapid rise in "analytics" searches.

This lack of consensus regarding the relationship of analytics to other disciplines is not limited to INFORMS members. In a September 2014 article in the *European Journal of Operational Research* [5], the authors state: "One possible reason for the discrepancy between the perceived opportunity analytics may offer to the OR/MS community and the amount of research in the area, as alluded to in above, may be the lack of any clear consensus about analytics' precise definition, and how it differs from related concepts."

Unanswered Questions

The uncertainty surrounding analytics has led to several vague conceptions regarding the term and many unanswered questions. Analytics is often referred to as an emerging field or an emerging discipline. If analytics is an emerging discipline, where did it come from, and how could it have emerged suddenly? Why is there no unique research associated with analytics – as distinct from research associated with statistics, computer science or operations research? Why are the examples of analytics that are cited always examples of methods from disciplines such as statistics, computer science, operations research, economics or industrial engineering?

Analytics is sometimes represented as a convergence of the quantitative decision sciences. If this is the case, it would represent a reversal of a trend in human history, lasting for thousands of years, toward specialization. What caused such a reversal, and why did it occur suddenly? More importantly, since there is no new high-level unifying theory associated with analytics, such as string theory in physics, how could individuals acquire the knowledge and master the methods of at least five or six separate disciplines?

To overcome the uncertainty surrounding analytics, we need an overarching conception of analytics that answers the preceding questions while remaining consistent with how the term is used. Moreover, since the terms "descriptive analytics," "predictive analytics" and "prescriptive analytics" are often used interchangeably with data science, and since prescriptive analytics is often associated with operations research, a framework is needed to relate all of these terms.

Analytics, Analytics, Analytics

The first step in gaining such an understanding is the recognition that the term "analytics" is used in at least three different ways, and therefore, requires three separate definitions:

1. ***Analytics is used as a synonym for statistics or metrics.*** Examples are website analytics (how many views or clicks) or scoring analytics (number of points scored per 100 possessions).
2. ***Analytics is used as a synonym for data science.*** Examples are data analytics and predictive analytics.

If analytics is **represented** as a **convergence** of the **quantitative decision sciences,** it would **represent a reversal** of a **trend toward specialization.**

Basing our understanding of analytics **on the way the term** is being used **will allow us** to think **more clearly about analytics,** but there is still **a problem.**

3. ***Analytics is used in a very general way to represent a quantitative approach to organizational decision-making.*** This is Davenport's "Competing on Analytics" usage.

Attempting to cover all three of these usages with a single definition has led to much of the confusion surrounding analytics. The first usage refers to a type of measurement or counting. The second usage refers to the processing of large amounts of data with advanced software technologies and sophisticated statistical and computer science techniques. The third usage refers to a management philosophy that emphasizes a quantitative approach to decision-making.

Basing our understanding of analytics on the way the term is being used will allow us to think more clearly about analytics, but there is still a problem: embedded in the third usage is an apparent paradox.

Among analytics thought leaders there is one area in which there is agreement – analytics is related to many different disciplines:

- Davenport, Cohn and Jackson, in the previously mentioned May 2005 research report "Competing on Analytics" [1] offer statistics, operations research, industrial engineering, econometrics and mathematical modeling as examples of analytics.
- Rahul Saxena, co-author of the December 2012 book "Business Analytics," on slide No. 5 of a SlideShare presentation [6], lists 14 disciplines as being antecedents of analytics. The list includes business intelligence, computer science, statistics, operations research, industrial engineering, and finance planning and analysis.

If these authors are correct, how can we explain the fact that disciplines such as statistics, computer science, operations research, industrial engineering and economics continue to exist?

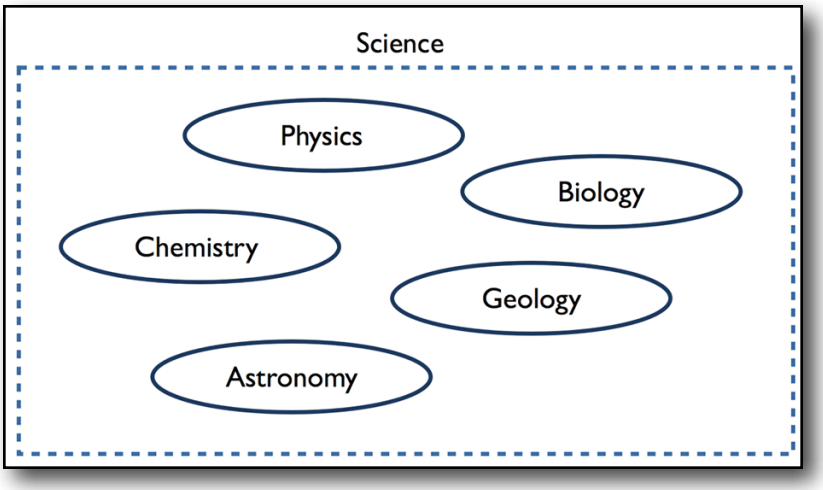

Figure 2: Visualizing separate disciplines of science that have common elements.

An Analogy

To explain this apparent paradox, an analogy will be helpful. If you see the words chemistry, biology, science, physics, geology and astronomy, are you confused about their meanings or their relationship to each other? I think not. Further, I suspect you will visualize something similar to the diagram shown in Figure 2.

The word "science" conceptually groups together the natural sciences. Although physics, chemistry, astronomy, geology and biology are all separate disciplines that use different methodologies and require the mastery of large amounts of discipline-specific knowledge, they do have elements in common. Each of them uses the scientific method, mathematical modeling and peer-reviewed

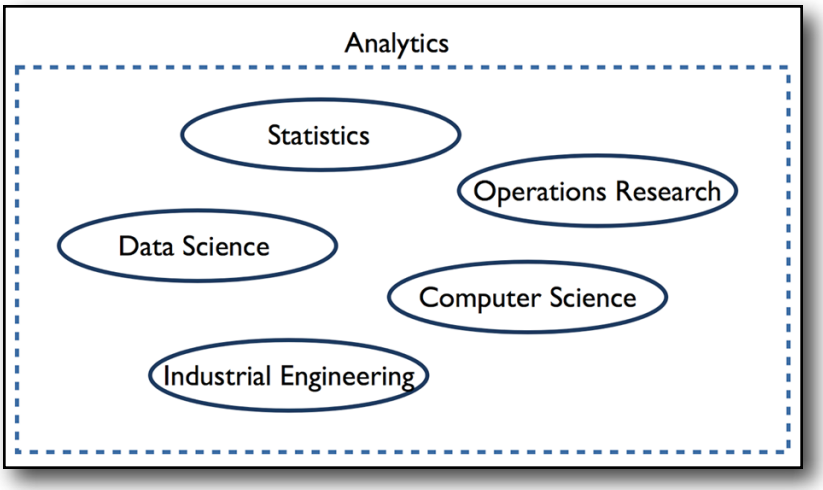

Figure 3: Visualizing separate disciplines of analytics that have common elements.

research. It is therefore appropriate to group them together; it is also quite useful to be able to refer to the natural sciences collectively. We can talk about the state of science education or discuss whether or not we should increase our investment in scientific research.

Analytics includes many disciplines, while at the same time, those disciplines continue to exist.
Image © Sergey Nivens | 123rf.com

Paradox Explained

For the same reason that it is useful to have a term that collectively represents the natural sciences, it is useful to have a term that collectively represents the quantitative decision sciences. And this is exactly the role played by the term analytics in its broadest usage: It conceptually groups together the quantitative decision sciences (see Figure 3).

Statistics, data science, industrial engineering, operations research and computer science are separate disciplines that use different methodologies and require the mastery of large amounts of discipline-specific knowledge. As in the case of the natural sciences, they do have elements in common such as the scientific method, mathematical modeling and peer-reviewed research. When viewed as a conceptual grouping of the quantitative decision sciences, the term analytics, in its broadest usage, allows us to make statements such as: "We will compete on analytics." We can meaningfully refer collectively to different quantitative decision sciences, possibly in different departments or different geographical locations, in the same way that we might collectively refer to separate scientific research projects.

Viewing analytics in this way explains the paradox of how analytics can somehow include many disciplines, while at the same time, those disciplines continue to exist. This understanding allows us to define analytics and explain its relationship to the quantitative decision sciences. Also, it is now possible to explain the sudden emergence of analytics: In the age of the Internet, new concepts can suddenly emerge and go viral.

We now have an overarching conception of what analytics is, but we do not yet have a framework that can relate the broadest usage of analytics to the various "flavors" of analytics (descriptive, predictive, prescriptive), data science and operations research. To create such a framework, we need one more important concept.

Problem Centricity

In a Dec. 17, 2014, INFORMS podcast [7], Glenn Wegryn observes that analytics is divided into two distinct camps. He notes that they tend to come from different organizational backgrounds, and he describes them in the following way:

- ***data centric*** – use data to find interesting insights and information to predict or anticipate what might happen; and
- ***decision centric*** – understand the business problem, then determine the specific methodologies and information needed to solve the specific problem.

As is clear from its description, the decision-centric category could also be named problem centric, and to make it clear that it encompasses systems and processes, that is how I will refer to it. Since analytics, in its broadest usage, conceptually groups together the quantitative decision sciences, the data-centric

Analytics conceptually **groups together** the **quantitative decision sciences;** it represents **disciplines,** but is not itself **a discipline.**

and problem-centric classification can also be applied to the quantitative decision sciences. Speaking about data science, Anthony Goldbloom (founder of Kaggle) said, "You want to extract all the signal that's possible out of a dataset" [8]. In a recent interview in *OR/MS Today* [9], Professor Edward Kaplan (president of INFORMS) said, "Operations researchers think in terms of problems" and "operations research is the scientific study of operations." These descriptions suggest that data science (data centric) and operations research (problem centric) fit nicely into the preceding classification.

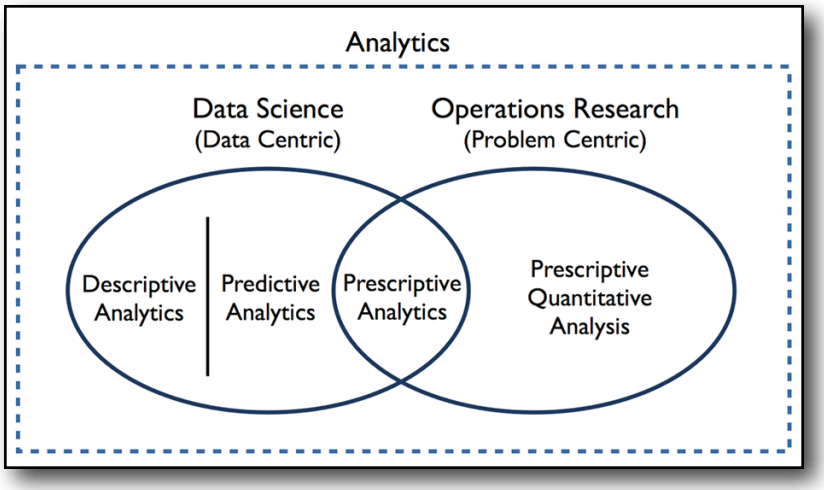

Figure 4: Conceptual framework of analytics.

A Conceptual Framework

Keeping the above in mind, and remembering that one of the usages of the term analytics is as a synonym for data science, a conceptual framework can now be constructed that relates analytics, descriptive analytics, predictive analytics, prescriptive analytics, data science and operations research (See Figure 4).

Several aspects of the diagram shown in Figure 4 should be noted:

- Since the term analytics is used in multiple ways, there is no conflict caused by its use above the diagram (broadly referring to quantitative decision-making) and within the diagram (a synonym for data science).
- Data science, and its two-word analytics synonyms, refer to the processing of large amounts of data with advanced software technologies and sophisticated statistical and computer science techniques.
- Prescriptive analytics is the domain where data-centric and problem-centric paradigms intersect, i.e., problems that require scientific study and mathematical modeling, and the processing of large amounts of data with advanced computer science and statistical techniques.
- Many problems that require scientific study and modeling analysis do not require the processing of large amounts of data and are represented by the category prescriptive quantitative analysis.

Summary

The uncertainty surrounding analytics can be eliminated by keeping the following points in mind:

- Analytics is used in three different ways and therefore requires three definitions.
- In its broadest usage, analytics conceptually groups together the quantitative decision sciences; it represents disciplines, but is not itself a discipline. Therefore, there is no research that is unique to analytics, and there are no methods that are unique to analytics.
- Analytics emerged suddenly since it is a concept that went viral.
- Operations research is a problem-centric discipline; data science is a data-centric discipline. Prescriptive analytics is where data-centric and problem-centric paradigms intersect. ORMS

Robert Rose *(robertl.rose@verizon.net) has been successfully using operations research to solve business problems for many years. A longtime member of INFORMS, he has formed a company, Optimal Decisions LLC, where he is currently developing O.R.-based decision support tools in the areas of employee scheduling and forward buying optimization.*

NOTES & REFERENCES

1. Davenport, T.H., Cohen, D., Jacobson, A., "Competing on Analytics," Babson College Research Report, May 2005.
2. Davenport, T.H., "Competing on Analytics," Harvard Business Review, January 2006, pp. 99-107.
3. Davenport, T.H., Harris, J.G., 2007, "Competing on Analytics: The New Science of Winning," Harvard Business Review Press.
4. Boyd, E.A., "What is analytics?," 2011, *Analytics* magazine, March/April 2011, pp. 7-8.
5. Mortenson, M.J., Doherty, N.F., Robinson, S, 2015, "Operational research from Taylorism to Terabytes: A research agenda for the analytics age," *European Journal of Operational Research,* Vol. 241, Issue No. 3, March 2015, pp. 583-595.
6. Rahul Saxena, "Building an Analytics CoE (Center of Excellence)," SlideShare presentation, July 20, 2013.
7. Glenn Wegryn, "Top 5 Analytics Trends for 2015," INFORMS Podcast, Dec. 17, 2014.
8. Anthony Goldbloom, "Lessons from 2MM machine learning models," YouTube video, Dec. 11, 2015.
9. Horner, P., 2015, "Meet the 'member in chief,' " *OR/MS Today,* Vol. 42, No. 6, pp. 36-41, December 2015.