analyzed to produce actionable business information. This data analysis need may explain ...

Randy Johnston • Jun. 15, 2018
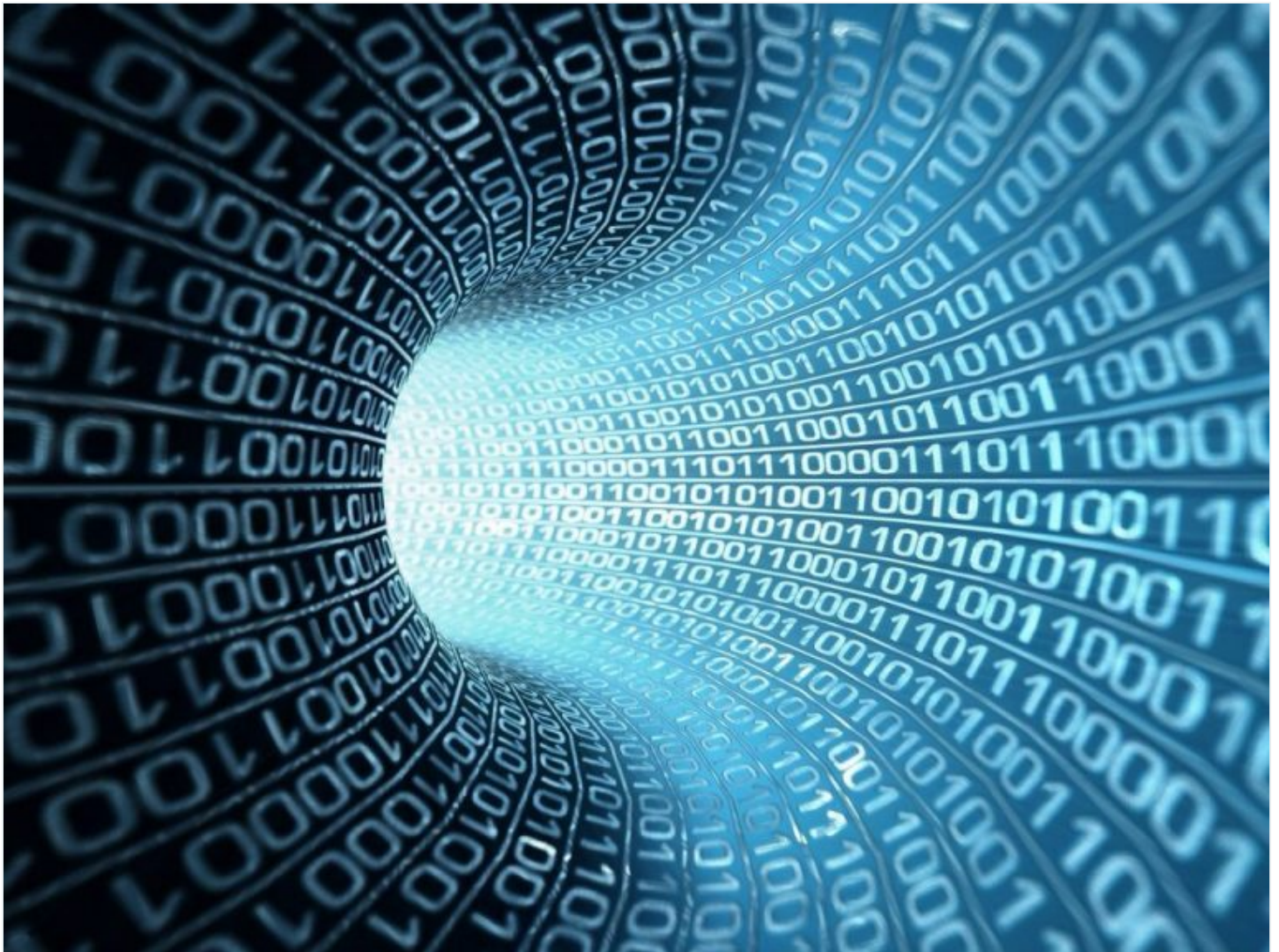
From the June 2018 Issue.

If Cognitive Computing drives much of the emerging technology computer science research, most of the data science research is focused on Big Data. Big Data can be analyzed to produce actionable business information. This data analysis need may

explain the number of data scientists that are being hired by mid-sized and large

inaccuracies increases with data volume growth.

Your data scientists may need to check that the data is relevant, connected (meaning it is related and complete), accurate (but the data can be precise/imprecise), and that there is enough data to work with. If the data is ready for processing, according to Brandon Rohrer, Senior Data Scientist of Microsoft, data science answers five questions:

1. Is this A or B?
2. Is this weird?
3. How much – or – how many?
4. How is this organized?
5. What should I do next?

There are four types of Data Analytics that can be run on Big Data including:

1. Descriptive Analytics: What's happening in my business?
2. Diagnostic Analytics: Why is it happening?
3. Predictive Analytics: What's likely to happen?
4. Prescriptive Analytics: What do I need to do?

Like all of the emerging technologies we have covered in these columns, Big Data has pros and cons.

On the positive side:

- Extremely large data sets may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions
- Analysis of data sets can find new correlations to "spot business trends, prevent diseases, combat crime and so on" according to The Economist, June 2017

On the down side:

The term Big Data has been in use since the 1990s, with some giving credit to computer scientist John Mashey, formerly of Bell Labs, for coining or at least making it popular. A notable challenge is how to make the reporting simple enough for smaller businesses or firms to be able to process data effectively. Alternatively, with small businesses, the amount of data may remain small enough that there is insufficient data for the algorithms to produce meaningful data analytics.

Why?

Big Data can provide value for making business decisions. There are five characteristics of Big Data:

- Volume: big data doesn't sample; it just observes and tracks what happens
- Variety: big data draws from text, images, audio, video; plus it completes missing pieces through data fusion
- Velocity: big data is often available in real-time
- Veracity: the data quality of captured data can vary greatly, affecting the accurate analysis
- Value: Technology and Analytical Methods for big data transformation as well as usefulness

How?

So how do Big Data approaches work?

- In 2000, Seisint Inc. (now LexisNexis Group) developed a C++-based distributed file-sharing framework for data storage and query
- In 2004, Google published a paper on a similar architecture called MapReduce that uses a parallel processing model.

– With MapReduce, queries are split and distributed across parallel nodes and

What does this mean to the practice of accounting and to accountants? We have several working examples available:

- The video, audio, and textual information made available via Big Data can provide for improved managerial accounting, financial accounting, and financial reporting practices
- In managerial accounting, Big Data will contribute to the development and evolution of effective management control systems and budgeting processes
- In financial accounting, Big Data will improve the quality and relevance of accounting information, thereby enhancing transparency and stakeholder decision making
- In reporting, Big Data can assist with the creation and refinement of accounting standards, helping to ensure that the accounting profession will continue to provide useful information as the dynamic, real-time, global economy evolves
- In the press, C-Span is using Amazon's vision system to compile a database of politicians, so they can name them quickly when they appear on screen

Continue Reading Online at: www.CPAPracticeAdvisor.com/12412520

Cloud storage of data and large-scale data sets provide the source for processing with the software below. Small accounting software designers expected that the amount of data accumulated in QuickBooks Online, Xero and other products would provide enough data, that Big Data analytics could be run on the complete data set and provide insight to the small business owner or the accountant providing guidance to the business owner.

As development continues and Big Data transitions from an emerging technology to

5. **Actian Analytics Platform**
6. **Attivio Active Intelligence Engine**
7. **Google Bigdata**
8. **Wavefront**
9. **Opera Solutions Signal Hubs**
0. **Daatmeer**
11. **FICO Big Data Analyzer**
2. **IBM Big Data**
3. **Amazon Web Service**
4. **DataTorrent**
5. **Oracle Bigdata Analytics**
6. **Palintir Bigdata**
7. **Cloudera Enterprise Bigdata**
8. **Amdocs Insight**
9. **Splunk Bigdata Analytics**
0. **Syncsort**

The best example of tools for accounting that are working today is:

- Distributed data processing using Hadoop, which is pretty much the standard for processing large data sets across distributed systems
- Processing data streams using Spark or Flink, and then graduate to Beam
- Machine learning using Google's TensorFlow
- Big Data tool chain integration using Talend Open Studio
- Data Lakes using Kylo

Here's a summary of what you need to know about Big Data:

Key Information   **TECHNOLOGY: Big Data**

Why is the new technology	Analyze extremely large data sets to

Risks	can lead to incorrect strategic conclusions
Where/when to use	To find trends in large amounts of data
How much?	Can be thousands to start, or free on open source
When expected in mainstream	Three to five years
Displaced technology or service	Data Warehouse
Other resources	Accounting Today, CPAPA

Big Data capture and processing into meaningful information is still complex and needs to be simpler.

Recommended Next Steps

Consider what would be meaningful information for your firm or your clients. Don't be too restricted by thinking about your current financial reporting or dashboard technologies. What would help you run your firm better? What information would provide insight to you so you could advice your clients better? Products will need to provide a way to satisfy this need.

Big Data processing for small and medium business still needs some breakthrough products to make the technology practical and useful to smaller firms and businesses. The tools that are working are all for larger businesses with larger data sets with

larger budgets. The Emerging Technology of Big Data has great promise, but right

(NASBA) as a sponsor of continuing professional education on the National Registry of CPE Sponsors.