The R & D Tax Aspects of Big Data

By Charles R. Goulding, Charles G. Goulding and Jacob Goldman

Charles R. Goulding, Charles G. Goulding and Jacob Goldman describe the nature of U.S. R & D efforts and the R & D tax credit opportunity for firms committed to harvesting Big Data through the use of advanced analytics.

Firms have always analyzed data to support the development of new products, refine existing ones and better understand their markets. However, the explosion of data brought on by the digital age has largely outpaced the ability to make cogent use of this wealth of information.

A number of major U.S. companies have now placed huge bets on the importance of "Big Data" to our country's economic progress. "Big Data" is defined as large data sets that are too large and complex to be analyzed by conventional means.

A recent report by McKinsey heralded Big Data as a major driver of both innovation and job creation. An additional research firm, the International Data Corporation, predicts that revenue from products and services specifically targeting Big Data will soar to $17 billion in 2015, from just $3 billion in 2010.[1] Meanwhile, IBM continues its transformation from manufacturer to business solutions provider, with Big Data at the core of its offering. In particular, IBM has made it clear that it is betting on Big Data for healthcare.[2]

Charles R. Goulding Attorney/CPA, is the President of R & D Tax Savers, an interdisciplinary tax and engineering firm that specializes in R & D tax credits.

Charles G. Goulding is a Senior Analyst with R & D Tax Savers.

Jacob Goldman, LEED A.P., is an Engineer and Tax Consultant with R & D Tax Savers.

A major source of U.S. research and development (R & D) is now focused on harvesting Big Data through the use of advanced analytics. This article describes the nature of these efforts and the R & D tax credit opportunity for firms committed to such activity.

The Research and Development Tax Credit

Federally enacted in 1981, the Research and Development Tax Credit allows a credit of up to 13 percent of eligible spending for new and improved products and processes. Qualified research must meet the following four criteria:

- new or improved products, processes or software
- technological in nature
- elimination of uncertainty
- process of experimentation

Eligible costs include employee wages, cost of supplies, cost of testing, contract research expenses and costs associated with developing a patent.

Big Data = Big R & D

Manoj Saxena, business manager for IBM's Watson, says that 90 percent of the world's information was created within the last two years based on the volume of information created by computers and

electronic sensors. He goes on to say, "80% of that 90% is unstructured or semi-structured information, like doctor's notes or product reviews on Amazon."[2]

Even at the small and mid-sized level, many firms have departments dedicated to Big Data analysis, often headed by a Director of Analytics. These professionals apply advanced quantitative concepts to their data, including multivariate analysis and sophisticated algorithms, and are frequently highly educated in technical fields like mathematics and software engineering.

Business intelligence is thus now being supported to a much more sophisticated degree by scientific discovery, thanks to a new class of staff who are the scientists of the business world. When a firm engages in Big Data analysis, it brings an unprecedented scientific rigor to its decision making. The basic principles of the scientific method, which include hypothesizing and experimentation, are fundamental to the R & D tax credit.

McKinsey estimates that there is already a need for between 140,000 and 170,000 jobs for individuals skilled in deep data analysis. Likewise, McKinsey believes Big Data will result in 1.5 million jobs related to data management. Big Data harvesting is therefore creating value, which both creates job opportunities within firms and qualifies those firms for R & D tax credit incentives.

> One added advantage of Big Data R & D is that its digital nature generates a great deal of contemporaneous documentation which can be employed in support of the R & D credit.

Big Data in Action–Current Examples

Leaders in Big Data R & D include the following four sectors.

1. Medical Records/Disease Data

After Francis Collins and Craig Venter's incredible accomplishment of mapping the human genome, it became apparent that understanding gene traits would only provide value if it could be cross-referenced to large volumes of human disease data.[3] Until recently, the only material collections of disease data were in Iceland and the UK. The analysis of today's far richer spread data should improve decision making worldwide. WellPoint, the large Indiana-based insurer, has begun to use an IBM Watson-based expert knowledge system to handle required second opinion approvals. Presumably, this will lead to better decisions at much less administrative cost.

Another good example is eClinicalWorks, headquartered in Westborough, MA. Their electronic health records platform is used by 70,000 physicians and over 410,000 medical professionals, who see one million patients a day. According to a May 3, 2012, Business Week article, about one-fifth of America's medical record (EMR) health data is stored using eClinical software.[4]

These records create a real-time, digital file of patient symptoms, diagnosis, recommended treatment and treatment results. In a clear example of new product development *via* Big Data, the company has created Community Analytics which provides important data to health authorities and organizations across the country. To protect security, eCincialWorks queries the health records where they are stored either on servers or hard drives and removes identifying information.

McKinsey estimates that effective use of Big Data has a $300 billion value attached to the health care industry.

2. Security and Terrorism

In the new popular TV show, "Person of Interest," the "Machine" monitors by voice and video recording all critical activity in the major city. Although the show is fictional, the core underlying security technology, at some level, is being employed in many cities. New York City has a new $100 million dollar Domain Awareness System that covers large sections of Manhattan and portions of the outer boroughs.

The Domain Awareness System is a massive, networked array of public and private surveillance cameras, license plates, radiation detectors, arrest records, incident reports and facial recognition software. The NYPD developed the system with Microsoft; it is being marketed to law enforcement agencies throughout the world.

Examples from other fields include fraud detection platforms, which flag abnormalities in credit card usage and account transfers based on patterns gleaned from analysis of huge volumes of user data.

Federal Government Big Data

The public sector's commitment to Big Data is also evident. In March of 2012, President Obama announced the Big Data Research and Development Initiative, which allocated $200 million to Big Data initiatives within a host of pertinent departments. The Department of Defense, for example, plans to use Big Data to "[h]arness and utilize massive data in new ways and bring together sensing, perception and decision support to make truly autonomous systems that can maneuver and make decisions on their own.".[5]

3. Building Energy Management

Six major jurisdictions including California; New York City; Washington, D.C.; Philadelphia; Austin, Texas; and Seattle, Washington now have mandatory building energy benchmarking laws that, for the most part, require buildings 50,000 square foot or greater to disclose by major building category building energy use and in some cases water usage.[6]

Twenty-six other cities are considering implementing similar provisions. All of the jurisdictions require usage of user-friendly Energy Star software developed by the EPA. This means that building owners, utilities and tenants will now have a wealth of easily comparable information about one of a building's largest operating costs. A third-party service, Honest Buildings, has currently amassed a tremendous amount of information about buildings that is easily accessible at *http://www.honestbuildings.com/*.

4. Mobile Phones

The explosive growth of increasingly smarter phones and accompanying applications is generating a massive amount of data that is increasing geometrically by the minute. Every use of the phone, including traditional voice communication, text messaging, internet search, GPS navigation, video, photography, purchasing, marketing, advertising, coupon redemption and airline boarding, creates data which can be researched, analyzed, developed and put to further use.

One example of mobile phone Big Data innovation involves disease prevention. Malaria spreads so problematically because symptoms take time to present; infected persons therefore travel and socialize unaware of the significant health risks they pose to others. Researchers at the Harvard School of Public Health recently recommended to Kenyan officials that they systematically text individuals passing through areas known to be high-risk and encourage them to wear head-nets.

Conclusion

Thorough R & D tax credit study requires excellent documentation. One added advantage of Big Data R & D is that its digital nature generates a great deal of contemporaneous documentation which can be employed in support of the R & D credit. Firms undergoing initiatives to support their product and process development through Big Data should consider whether they are eligible for an R & D tax credit.

Endnotes

1 Brian Deagon, *Teradata's Newest Products Dive Deeper Into Big Data*, Investor's Bus. Daily, Oct. 18, 2012, at A4.

2 Jon Gertner, *Calling Dr. Watson,* Fast Co., Nov. 2012, at 124.

3 Kristen Philipkoski, *Celara Wins Genome Race*, Wired.com, Apr. 6, 2000, accessed at *http://www.wired.com/science/discoveries/news/2000/04/35479*.

4 John Tozzi, *Girish Navani: Examining a City's Health*, Bloomberg Businessweek, May 3, 2012, accessed at *http://www.businessweek.com/articles/2012-05-03/girish-navani-examining-a-citys-health*.

5 Mike Lennon, *Obama Administration Places $200 Million Bet On Big Data*, Security Week, Mar. 29, 2012, accessed at *http://www.securityweek.com/obama-administration-places-200-million-bet-big-data*.

6 Charles R. Goulding, Andrea Albanese and Charles G. Goulding, *Tax Aspects of New York City's Building Energy Benchmarking Project*, Corp. Bus. Tax'n Monthly, Apr. 2013, at 15.