DATA-DRIVEN TRANSFORMATION

2018 PERSPECTIVES AND CASE STUDIES

BIG DATA | AI | BLOCKCHAIN

Harvard Business Review

MIT Sloan Management Review

Forbes

WALL STREET JOURNAL

NVP
NewVantage Partners

Boston | New York | San Francisco | Austin | Charlotte
www.newvantage.com

CONTENTS

- **Feeding a Data-Hungry Organization at Charles Schwab** | **Forbes** | June 6, 2018
- **Getting to Trusted Data via AI, Machine Learning, and Blockchain** | **Forbes** | June 17, 2018
- **How FinTech Initiative Are Driving Financial Services Innovation** | **Forbes** | July 10, 2018
- **Time to Value: The Currency of Data Operations** | **Forbes** | July 24, 2018
- **Farmers Accelerates its Time to Impact with AI** | **Forbes** | August 1, 2018
- **Will Blockchain Transform Healthcare?** | **Forbes** | August 5, 2018
- **How Fannie Mae is Creating a Modern Data Environment** | **Forbes** | August 30, 2018
- **The State of Machine Learning in Business Today** | **Forbes** | September 17, 2018
- **Every Company is a Data Company** | **Forbes** | September 26, 2018
- **A Rising Crescendo Demands Data Ethics and Data Responsibility** | **Forbes** | October 29, 2018
- **Munich Re: How Data and AI Reduce Risk from Global Calamities** | **Forbes** | November 4, 2018
- **Allstate's Data-Driven Business Transformation Initiative** | **Forbes** | November 11, 2018
- **UBS Asset Management Taps Alternative Data to Increase Alpha** | **Forbes** | November 18, 2018
- **A 2019 Forecast for Data-Driven Business: From AI to Ethics** | **Forbes** | December 17, 2018
- **How Data Became Mainstream: The Path to Chief Data Officer 4.0** | **Forbes** | December 26, 2018

Big Companies Are Embracing Analytics, But Most Still Don't Have a Data-Driven Culture

February 16, 2018

By Randy Bean and Thomas H. Davenport

For six consecutive years NewVantage Partners has conducted an annual survey on how executives in large corporations view data. Each year the response rate increases, and the reported urgency of making effective use of data increases as well. This year the results are both more encouraging and more worrisome than in the past.

Six years ago, the primary focus of questions and answers in the survey was big data, which was relatively new on the business scene. In the 2018 survey, the primary attention has moved to artificial intelligence. AI is now a well-established focus at these large, sophisticated firms. There is both a stronger feeling that big data and AI projects deliver value and a greater concern that established firms will be disrupted by startups.

The survey includes senior executives from 57 large corporations. The industry group with the most firms represented in the survey is one of the most data-intensive: financial services. Companies from the life sciences, manufacturing, telecom, and online industries also participated. The actual respondents are changing somewhat from the first surveys: It has always involved a high proportion of C-level executives responsible for data, but this year chief data officers are 56% of the respondents, up from 32% last year. Only 12% of firms in the 2012 survey had even appointed a chief data officer.

While AI gets the headlines here and elsewhere in the world, the survey addresses both big data and AI. Terminology comes and goes, but the constant is a data explosion and the need to make sense of it. Big data and AI projects have become virtually indistinguishable, particularly given that machine learning is one of the most popular techniques for dealing with large volumes of fast-moving data. It's also the case that statistical approaches to AI — deep learning, for example — are increasingly popular. Therefore, we view traditional data analytics, big data, and AI as being on a continuum. Virtually all of the respondents (97%) say they are investing in these types of projects.

Perhaps the best news in this survey is that companies continue to believe they are getting value from their big data and AI projects. 73% of respondents said they have already received measurable value from these initiatives. That number is half again higher than in the 2017 survey, which suggests that more value is being achieved as companies grow familiar with the technologies.

The types of value received are perhaps consistent with other previous types of technology. Consistent with our view that big data and AI are extensions of analytical capabilities, the most common objectives — and those most likely to achieve success — are "advanced analytics/better decisions." Thirty-six percent had that as their top priority, and 69% of those had already achieved success with the objective.

Better customer service and expense reduction are also common objectives. Just over one-quarter of firms (27%) are pursuing some combination of innovation and disruption, speed to market, or data monetization initiatives. Data monetization programs had the lowest priority and the lowest percentage of success (27%).

One of the greatest issues for concern in the survey for large enterprises is the risk of disruption from new entrants. Almost four in five respondents said they feared disruption or displacement from firms like those in the fintech sector or firms specializing in big data. The technology judged most disruptive is AI — by far. Seventy-two percent chose it as the disruptive technology with the most impact — far more than cloud computing (13%) or blockchain (7%).

Another important and continuing issue is the slow speed with which these established firms make the shift to a data-driven culture. Virtually all respondents (99%) say their firms are trying to move in that direction, but only about one-third have succeeded at this objective. This gap appears every year in the surveys, and the level of success hasn't improved much over time. Clearly firms need more-concerted programs to achieve data-related cultural change. Many startups have created data-driven cultures from their beginning, which is a key reason why large, established firms fear disruption from them.

One of the approaches that firms have established to deal with data-driven disruption and change is to establish new management roles. However, there is still a lack of clarity about how different data-oriented roles (chief information officer, chief data officer, chief digital officer, chief analytics officer, etc.) relate to each other.

With respect to the chief data officer role, there is substantial disagreement about the major responsibilities of the role and what types of backgrounds are appropriate for CDO jobs. Thirty-nine percent say their CDO has primary responsibility for data strategy and results, but 37% assign that responsibility to other C-level executives, and 24% say there is no single point of accountability for it. In terms of backgrounds, 34% of respondents believe the CDO should be a change agent from outside the company, while 32% believe the person should be a company veteran from inside the firm. Role clarity in senior data-related roles is critical for both leading AI/big data projects and accomplishing cultural change. And while all respondents believed it important, the majority of firms still lack an enterprise data strategy.

This continuing rise in the importance and challenges of big data is one of the most important features of contemporary economy and society. The survey results over time provide interesting and useful documentation of this revolution. The rise of AI is only exacerbating this trend. The keys to success are to determine how your firm should respond, assign clear responsibilities for data strategy and results, and then move ahead to execute the needed changes in a systematic and effective fashion.

Thomas H. Davenport is the president's distinguished professor in management and information technology at Babson College, and cofounder of the International Institute for Analytics. He also contributes to the MIT Initiative on the Digital Economy as a fellow, and as a senior advisor to Deloitte Analytics. Author of over a dozen management books, his latest is *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines.*

Randy Bean Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal, and is Founder and Executive Director of the Big Data for Social Justice Foundation. You can follow him at @RandyBeanNVP.

MITSloan Management Review

How Big Data and AI are Driving Business Innovation in 2018

At the intersection of analytics and smart technology, companies are starting to see the long-awaited benefits of AI.

February 5, 2018

By Randy Bean

After years of hope and promise, 2018 may be the year when artificial intelligence (AI) gains meaningful traction within Fortune 1000 corporations. This is a key finding of NewVantage Partners' annual executive survey, first published in 2012. The 2018 survey, published on January 8, represented nearly 60 Fortune 1000 or industry-leading companies, with 93.1% of survey respondents identifying themselves as C-level executive decision-makers. Among the 2018 survey participants were corporate bellwether companies, including American Express, Capital One, Ford Motors, Goldman Sachs, MetLife, Morgan Stanley, and Verizon.

The main finding of the 2018 survey is that an overwhelming 97.2% of executives report that their companies are investing in building or launching big data and AI initiatives. Among surveyed executives, a growing consensus is emerging that AI and big data initiatives are becoming closely intertwined, with 76.5% of executives indicating that the proliferation and greater availability of data is empowering AI and cognitive initiatives within their organizations.

The survey results make clear that executives now see a direct correlation between big data capabilities and AI initiatives. For the first time, large corporations report that they have direct access to meaningful volumes and sources of data that can feed AI algorithms to detect patterns and understand behaviors. No longer dependent on subsets of data to conduct analyses, these companies combine big data, AI algorithms, and computing power to produce a range of business benefits from real-time consumer credit approval to new product offers. Companies such as American Express and Morgan Stanley have publicly shared stories of their successes within the past year.

Staving Off Disruption

Survey participants comprised executives representing data-intensive industries, notably financial services companies, which constituted 77.2% of the survey respondents. Financial services companies have long been at the forefront of industry due to the large volumes of transactional and customer data that they maintain, and they have developed robust data management and data governance processes over a period of decades. These organizations have been at the forefront in the use of analytics to manage risk, assess customer profitability, and identify target market segments. Industries such as life sciences, while newer to data management, possess vast repositories of scientific and patient data that have gone largely untapped relative to the potential for insight.

Now, many of these mainstream companies are facing threats from data-driven competitors that have no legacy processes and have built highly agile data cultures. Companies like Amazon, Google, Facebook, and Apple are among the most prominent disruptive threats to these traditional industry leaders. As mainstream companies increase their investment in big data and AI initiatives, they face a range of issues and challenges as they seek to organize to compete against data-driven competitors. This concern is highlighted in the 2018 survey results.

A clear majority (79.4%) of executives report that they fear the threat of disruption and potential displacement from these advancing competitors. In response to the threat of disruption, companies are increasing their investment in big data and AI initiatives. In the 2018 survey, 71.8% of executives indicate that investments in AI will have the greatest impact on their ability to stave off disruption (in the next decade). Although overall investments in AI and big data initiatives continue to be relatively modest for most large corporations, 12.7% of executives report that they have invested half a billion dollars in these initiatives to date. If the fear of disruption is any indication, this number can be expected to increase.

Driving Innovation through AI

Executives indicate that investments in big data and AI are beginning to yield meaningful results. Nearly three-fourths of executives surveyed (73.2%) report that their organizations are now achieving measurable results from their big data and AI investments. In particular, executives report notable successes in initiatives to improve decision-making through advanced analytics — with a 69% success rate — and through expense reduction, with a 60.9% success rate. Businesses are also using big data and AI investments to accelerate time-to-market for new products and services (54.1% success rate) and to improve customer service (53.4% success rate). Yet, just over one-fourth (27.3%) of executives report success thus far in monetizing their big data and AI investments. This remains an elusive goal for most organizations.

Nearly one-fourth (23.9%) of respondents report that their investments in big data and AI are highly transformational and innovative for their organization, and potentially disruptive for their industry. But 43.8% of executives report that innovation and disruption initiatives involving big data and AI yield successful results for their organizations.

As mainstream companies look to the future, there is a growing consensus that AI holds the key. With 93% of executives identifying artificial intelligence as the disruptive technology their company is investing in for the future, there appears to be common agreement that companies must leverage cognitive technologies to compete in an increasingly disruptive period. Investment in AI can be expected to increase as organizations position themselves to compete in the future. Those companies that prove themselves to be adept at developing and executing initiatives using big data and AI capabilities will likely be the companies that are best positioned to deflect the threats of agile, data-driven competitors in the decade ahead.

About the Author

Randy Bean is an industry thought leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm that he founded in 2001. He is a contributor to MIT Sloan Management Review, Forbes, Harvard Business Review, and The Wall Street Journal. He tweets @RandyBeanNVP.

MITSloan
Management Review

The Problem with AI Pilots

AI technology is not just an experiment.

June 26, 2018

By Thomas H. Davenport and Randy Bean

Over the past year or so we've been engaged in an effort to tell the story of how large organizations are deploying artificial intelligence in their businesses. We were encouraged by the response to the 2018 NewVantage Partners executive survey, in which 93% of respondents said their organizations were investing in AI initiatives. Plenty of companies to write about, we thought. These were very large organizations spending goodly sums on AI and with a history of early adoption of other technologies. But when we approached many of these companies to discuss writing some case studies about their work, most of them demurred.

Most said the reason wasn't that they wanted to keep their AI activities secret, but that they weren't actually very far along and hence their projects were not worth discussing yet. They were doing lots of pilots, proofs of concept, and prototypes, but they had few production deployments. When they did have AI systems in production, most were machine learning-based systems that had been in place for many years. This is particularly true in financial services, where large-scale "scoring" has been used to evaluate customers for credit and potential fraud for well over a decade. Some said to us that they didn't really consider these projects to be examples of AI — consistent with the common view of AI that it describes technology that is never really here yet. Others say that they have robotic process automation (RPA) implementations in place, but most are relatively small, and there is also debate about whether RPA is really AI or not.

Why AI Implementation Is Challenging

But there are good reasons why production implementations of AI technology are relatively scarce. One is the maturity — or lack thereof — of the technology. Chatbots and intelligent agents, for example, are getting better all the time, but many companies still hesitate to turn their customers over to them. Instead, they ask their employees to use them for applications in HR and IT. Some make them available to their call center reps to use in the background to help answer customer questions. Eventually, they hope, they will support customer interactions directly.

If the AI initiative actually changes the relevant business process and the skills necessary to perform it, that raises another barrier to full implementation — the old bugaboo "change management." Most AI systems still involve some interaction with human workers, and educating those workers on new tasks and new skills can be time-consuming and expensive. UPS, for example, developed a complex machine learning algorithm for daily routing of its package delivery trucks, and it's still rolling it out 10 years after the algorithm was developed. Getting tens of thousands of drivers to change their behavior isn't easy. Similarly, a new claims process in an insurance company we worked with involves using deep learning models to analyze photos of car accidents. The technology works pretty well, but it doesn't work with all

types of collision damage. The interface of the AI system with existing claims adjusters, who are still needed for most damage assessments, and their existing work processes has been challenging.

Full production implementation also involves interfacing AI with production information systems and architectures. A 2017 Deloitte survey found that the number one obstacle to successful AI deployments was that it was "difficult to integrate cognitive projects with existing processes and systems." New machine learning models may have to be written as APIs or as program code modules within existing systems. Even RPA systems, which are quite easy to implement in small volumes, can become an architectural challenge when adopted in large numbers. Because they act as users of production systems, they are typically impacted by changes in those systems and may have to be reprogrammed.

Making AI Implementation More Likely

There are several ways that companies can increase the likelihood and speed of production AI implementations. Here are some of them:

- Set a time and criteria for deciding whether to go into production before the pilot starts. This will add rigor to the decision process and put pilot project advocates on notice that implementation is an important consideration from the beginning (this should help to address all the challenges mentioned above).
- Adopt technologies that can scale and that can be used by the intended audience. If, for example, it is unlikely that a chatbot will be made available to customers as a primary channel, don't adopt it with vague hopes that it will improve its performance quickly (helps with the technological maturity challenge).
- Adopt AI capabilities that are already embedded within transaction systems. Major transaction system vendors, including Salesforce, SAP, Oracle, and Workday, are adding AI capabilities to their offerings. That typically means that the AI offerings will be somewhat integrated with transaction systems from the beginning, and that they can make use of the data within those systems (helps with the integration challenge).
- If the AI system will be stand-alone, make sure it can create a relatively easy interface with your existing systems — such as an API or generated program code that works with your architecture (also assists in integration).
- Start an AI-related education and skills program now. Even though you may not be sure of the specific needs of your workers for retraining and re-skilling, you can make available education offerings now about how to understand and work with smart machines. Such programs put employees on notice that change in their jobs from AI will happen and that they should begin preparing for it (and helps with the change management challenge).

Implementation of AI projects is the only way that organizations will realize tangible business value from their AI investments. Companies are spending considerable sums on AI technology, and it should not be viewed exclusively as an experiment. It is only when companies step up to production status with AI that it will deliver ROI and productivity for their organizations.

Forbes

Bloomberg's Data Initiative: Big Data for Social Good in 2018

January 2, 2018

By Randy Bean

Big Data has become ubiquitous in recent years. Whether it is data-driven marketing, sports analytics, political campaigns, or national security threats, data has become central to any type of informed analysis and plan of action. Consequently, the arrival of Big Data has also spawned a data industry and the emergence of data professions – data analysts, data architects, data scientists, and chief data officers. Against this backdrop, governmental and social service organizations are following suit, and initiating efforts to apply sound data practices to a range of societal challenges. These can range from matching scarce resources to acute needs, detecting disparities in social justice administration, or the establishment of polices for ethical data usage. Here are some of the initiatives that are being undertaken to advance data for social good as we look ahead for 2018:

Bloomberg's Data for Good Exchange

The Bloomberg Data for Good Exchange was launched in 2015 to encourage and promote the use of data science and human capital to solve problems at the core of society. Each year, the program focuses on themes pertaining to how data science can play a role in helping drive change in the delivery of public services, city operations, public health, climate resilience and the environment, criminal justice and other areas of public concern. Over 1,000 data scientists, thought leaders, and public policy makers gather at Bloomberg's Global Headquarters in New York City for a day of discussion. The program committee for the 2017 program considered over 170 proposals for papers, panels, and presentations.

Big Data in Public Health

Bloomberg Philanthropies has been an active partner in organizing and supporting the Data for Good Exchange, in addition to sponsoring a range of initiatives including Bloomberg Philanthropies' Public Health programs. In 2015, Dr. Kelly Henning, who leads the Public Health program, delivered a keynote on the topic of Data for Health, an initiative that is enabling countries to improve public health data collection with the goal of addressing public health problems. Working with partners, Data for Health aims to help more than one billion people in 20 countries across Africa, Asia, and Latin America. With this information and training in data analysis, participating countries are able to turn insights from data into public policy, and direct resources to issues affecting public health. To date, 20 countries have partnered with the Data for Health, reaching more than 1 billion people.

Big Data in Criminal Justice

Big Data is making a difference in addressing disparities in criminal justice sentencing and in tackling challenges of poverty and crime. According to the data-driven justice initiative, more than 11 million people move through America's 3,100 local jails each year. Many are low-level, non-violent offenders, costing local governments approximately $22 billion a year. Data shows that, 64 percent of those

incarcerated in local jails suffer from mental illness, 68 percent have a substance abuse disorder, and 44 percent suffer from chronic health problems.

I hosted a panel at the 2017 Data for Good Exchange, bringing together experts in this field. Mary McKernan McKay, dean of the Brown School of Social Work at Washington University in St. Louis, and former professor of poverty studies and director of poverty policy at New York University, joined co-panelists Kelly Jin, director of the data-driven justice initiative for the The Laura and John Arnold Foundation, and former policy advisor for the Obama White House data-driven justice initiative, and Rebecca Ackerman, a data scientist with New York Defender Services, to bring their perspectives on links between poverty, mental health, and racial discrimination. With an annual prison bill of $70B in the United States, and an incarceration rate that is 5x the average rate in other developed nations, universities and community organizations are undertaking bi-partisan initiatives to address root causes that can lead to systemic change.

Algorithmic Equality

Cathy O'Neil is an outspoken advocate for greater transparency in the social uses of Big Data, and is on a mission to ensure data and algorithmic "equality". Her 2016 book, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, was intended as a "wake up call". O'Neil, who also serves as an advisor to the Data for Good Exchange, believes that many people are intimidated by math, and as a result algorithms may be employed to support biases without critical and objective consideration. She calls this "hiding behind mathematics", and laments a lack of diverse perspectives that are missing from many of the algorithms that she discusses in her book. O'Neil warns of the dangers of algorithms which have become "accepted truths", and cautions about the damage that can result when algorithms become "widespread, mysterious, and destructive". She notes that "algorithms are often opinions embedded in code, which reflect subjective biases and decisions", and believes that algorithms are having an "outsized impact of algorithms" in areas ranging from teacher evaluations to academic admissions. She calls this the "weaponization of math", and advocates that data and algorithms do no harm.

Ethical Data Sharing

Bloomberg has recently been spearheading the development of a code of ethics for data scientists. The "Community Principles on Ethical Data Sharing (CPEDS)" initiative, which was announced at the Data for Good Exchange in September 2017, will provide a set of guidelines about responsible data sharing and collaboration. Characterized as a 'Hippocratic Oath' for the industry, Bloomberg believes that data scientists should be thoughtful, responsible, and ethical agents for change. This partnership will collect input from the global data science community through social media, conversations and working groups to define the values and priorities for ethical behavior by data scientists.

Across the public and private sector, organizations are confronting the responsible use of data, and exploring ways in which data can be applied to a wide range of social issues, needs, and challenges. Data ethics and data for good initiatives promise to be an increasing area of focus in 2018 in the ongoing advance and application of Big Data to a broad range of business as well as societal challenges.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

Biting the Data Management Bullet at GlaxoSmithKline

January 17, 2018

By Thomas H. Davenport and Randy Bean

There comes a time in the life of almost every large organization when it has to admit that it doesn't have the data environment it needs to succeed. For GlaxoSmithKline's (GSK) Research and Development (R&D) organization, that time was in early 2015. President of the unit Patrick Vallance and his senior colleagues deliberated on whether their data environment was of sufficient quality and integration to develop new drugs in the desired fashion. They examined not only GSK's situation, but compared it to other companies who were increasingly competing on the basis of their analytical capabilities.

Their conclusion was that the data at GSK R&D needed a major transformation. To lead it they brought in Mark Ramsey as the first head (and Senior Vice President) of R&D Data. He was charged with overseeing a transformation in how data and analytics were used across the organization. Vallance and his team had a vision for data within GSK R&D, which was to make it easier to access and use for exploratory analysis and decision-making about new medicines. GSK had been relatively good at making decisions with data, but the executives felt—and Ramsey quickly agreed—that the data within R&D was too siloed and fragmented to be used effectively for exploratory purposes. In particular, R&D data was kept within silos created for particular scientists, experiments, or clinical trials. Secondary analyses of it were almost impossible.

To determine the extent of the problem and confirm his initial impressions, Ramsey used a survey instrument developed by the MIT International Society of Chief Data Officers (isCDO). As an isCDO founding board member, he had seen the value of the survey across organizations, but it had not been applied inside an organization. It included questions like how easy it was to share data across the organization, whether scientists could get data from other departments, and how possible it was to perform analytics on data across the organization. He sent it to all of the 10,000 scientists within R&D, and 30%—an unusually high number—responded. The survey responses were virtually unanimous that it was very difficult or impossible to work with data outside your personal or departmental silo.

So integrating diverse data was clearly job one for Ramsey and his team. To guide and prioritize their activities, they identified over 20 use cases for what questions the scientists wanted to answer with R&D data, and eventually selected 10 as the focus. They were judged as having the greatest value, importance to key decisions, and role in addressing important scientific questions. More broadly, the goal of the work was to provide analytics-ready data of all kinds across R&D in a timely manner.

The R&D data team also looked at what other pharma firms were up to with data in order to guide and validate their own approach. Most were focusing on "real world evidence" data from insurance claims and electronic health records. Another group was focused on clinical trial data. Yet another concentrated on DNA sequencing data. GSK was interested in all of these, but the goal was to work both within and across these data domains, rather than having each as distinct effort.

A traditional master data management approach—which Ramsey characterizes as "map and move"—would have taken too much time and effort to implement. There were millions of data elements to rationalize. Ramsey knew that companies were beginning to apply big data and analytics tools to data itself. One company with tools for that purpose, Tamr, stood out for its machine learning technology and focus on the pharmaceutical industry, among other industries. Tamr's co-founder and CEO, Andy Palmer, was once Global Head of Software and Data Engineering for the Novartis R&D organization. As a result, Tamr was very familiar with pharma industry data standards like CDISC (Clinical Data Interchange Standards Consortium). (Disclosure: both authors are advisors to Tamr and hold very small equity positions).

GSK decided to employ the "probabilistic matching" approach used by Tamr to combine data across the organization into a single Hadoop-based data lake with three different domains. First would be "assays," or data from experiments. Second would be clinical trial data. And third would be genetic data. The goal was to get 100% of the data into the lake within three months—an unheard-of objective using traditional data management approaches. But GSK was able to use the tools to understand the level of duplication and pull the data together in the desired timeframe. To work across the three domains, the R&D data team created an "integrated layer" on top of them with standardized ontologies; this was the only way to solve the use cases.

In the clinical trials domain, for example, Ramsey and his colleagues believed there was a massive amount of insight possible outside of the original goals for a particular trial. But combining trial data was difficult because there is a lot of variance in how they are conducted and their results recorded. But using industry standard formats, the data (originally in GSK internal formats) was ingested and mapped to the industry standard, and machine learning models learned the process. The team would feed in the source trial data, and what the target format should look like—and then let the machine go to work. Outcomes initially had 50/60% accuracy levels, and now in some domains they are at 100% accuracy. After the models were developed and refined, they could be applied to other data with relatively little human intervention—just some occasional judgments from an expert team.

There are other technologies in play in this implementation in addition to Hadoop and Tamr. GSK uses Streamsets to move the data into the lake—a million pipelines for thousands of data sources. There is also a crawler that goes into every data source and extracts metadata and schemas from each one, and then creates loaders and pipelines to move data to the lake. There is a repository of every source data system and how often it changes, which is the mechanism for ingestion. GSK uses Apache Hive and Hbase to manage the data lake (on one of the largest Hadoop platforms in the world), and Waterline Data for cataloging. As a result of these technologies, GSK knows where all the data is, where it's come from, and how often it's updated.

GSK uses a best-in-class approach to deliver on the overall R&D data strategy, integrating several other technologies to deliver on the use cases. Ramsey has the vision to simplify future large-scale implementations with progress in how the technologies work together. GSK hosts partner summits with the key technology companies to assure that collaboration is a key component of their development roadmaps.

Now that the data management bullet has been bitten (perhaps in a faster and less painful way than anticipated), GSK is beginning to see some of the benefits. Scientists are beginning to see what an asset they have now, and the number of use cases has expanded from 10 to 250. Many projects that use the new data environment are underway. There are significant reductions in times to get an answer to an *ad hoc* question. As GSK has rationalized clinical trial data, a team is focused on "clinical trial diversity" to

make sure the company's trials match the demographics of patients. Real world evidence from more than 30 sources is now rationalized to the industry standard—instead of being a catch-all category, as it is in many pharma firms. GSK is also using combined clinical trials data to reuse placebo patients where appropriate. They can simulate the control arm in some cases rather than having to give new patients placebos.

In the genetic data domain, GSK has established a relationship with UK BioBank, which is doing full genetic sequencing on their 500,000 patients. GSK will have data not only on their genomes, but also their health records, and will be able to undertake many studies on them in the identification of new drug targets.

Ramsey feels that the data foundation has been laid, but actually building the house—i.e., using the data for better science—will also require help from AI. He notes:

We are doing a step change on machine learning. We're looking for "blue unicorns"—people who are life scientists and also machine learning experts, we simply have to have more machine learning skills to deal with all the available data now. We're training current scientists and also recruiting. We find that our data assets make it much easier to attract the right people.

GSK R&D's data environment is something that one often hears about in startups, but is rarely found in large enterprises whose roots go back over 300 years. And its great news for all of us humans who will benefit from the scientific advances it is likely to engender.

Thomas H. Davenport is the President's Distinguished Professor in Management and Information Technology at Babson College, a research fellow at the MIT Initiative on the Digital Economy, and a senior adviser at Deloitte Analytics. Author of over a dozen management books, his latest is Only Humans Need Apply: Winners and Losers in the Age of Smart Machines. You can follow him at @tdav.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

The Chief Data Officer Dilemma

January 30, 2018

By Randy Bean

One of the most encouraging signs that leading corporations are embracing the importance of data as a critical enterprise asset has been the establishment of the Chief Data Officer role. This recognition of the CDO role has been evidenced and confirmed by the results of an annual survey of Fortune 1000 c-executives launched by advisory firm NewVantage Partners in 2012.

While only 12% of executives reported that their firm had appointed a Chief Data Officer when the survey was first conducted in 2012, there has been a sharp and steady increase in adoption of this new c-executive role over the course of the past several years. In the 2018 survey, which was recently released, nearly two thirds of executives – 63.4% -- now report their firm having a CDO. Clearly, the Chief Data Officer has become an established role within a majority of leading corporations.

The annual survey provides a finger on the pulse of those blue chip corporations that are widely viewed as leaders in investing in data initiatives -- ranging from legacy data management to Big Data and data-driven AI initiatives. Financial services firms are heavily represented. These firms have historically maintained a rich history of customer transactions and activity for what can be very high value customer relationships. Leading firms in sectors such as health care, life sciences, media, and manufacturing are also represented. 2018 survey participants included bellwether firms American Express, Bank of America, Capital One, Charles Schwab, CitiGroup, Fidelity Investments, Ford Motors, Goldman Sachs, JP Morgan, IBM, Wells Fargo, and VISA, among nearly 60 industry leaders.

In spite of the common recognition of the need for a Chief Data Officer, there appears to be a profound lack of consensus on the nature of the role and responsibilities, mandate, and background that qualifies an executive to operate as a successful CDO. Further, because few organizations -- 13.5% -- have assigned revenue responsibility to their Chief Data Officers, for most firms the CDO role functions primarily as an influencer, not a revenue generator.

This divergence of opinion on CDO responsibilities, mandate, and importance of the role underscores why the Chief Data Officer may be the toughest job in the executive c-suite within many organizations, and why the position has become a hot seat with high turnover in a number of firms.

Lacking Consensus on Data Leadership

To understand the lack of consensus regarding the Chief Data Officer position, let's delve deeper into the findings. While 39.4% of survey respondents identify the Chief Data Officer as the executive with primary responsibility for data strategy and results within their firm, a majority of survey respondents – 60.6% -- identify *other* C-Executives as the point person, or claim *no single point* of accountability.

This is remarkable and highly significant, for it highlights the challenges that CDO's face within many organizations. It is noteworthy that 15.5% of firms are still committed to the position that the Chief Information Officer is the primary executive responsible for data strategy and results.

Of greatest concern is the finding that nearly a quarter of executives – 23.9% -- report that there is no single point of accountability for data and analytics within their organization. This should not be surprising. Most corporations that have been in existence for decades or longer were not organized around data as an organizing principle, or as an enterprise asset. As firms make the adjustment to become data-driven organizations, and work to forge data cultures, there is bound to be lingering resistance to change, or a profound lack of common agreement on where data responsibility lies, or what it looks like.

Establishing the CDO Mandate

Lack of clarity and common agreement on the nature of the role of the Chief Data Officer extends to disagreement on the CDO mandate. Nearly half of executives -- 44.4% -- indicate that the primary responsibility of a CDO is to *develop* the overall data and analytics strategy for the firm. This implies a strategic role. More than a quarter of executives – 26.7% -- see the primary responsibility of the CDO as being to *coordinate* data initiatives across the firm. A lower percentage of executives – 20% -- view the role as *leading* these initiatives across the firm. There is a strong consensus – 91.1% -- that the CDO should be playing a leadership role when it comes to charting the course of a firm's data and analytics strategy.

The point of disagreement appears to come down to the *scope* of the Chief Data Officer mandate, and whether the primary responsibility of the CDO should be to plan, coordinate, or implement data initiatives. It can be expected that this fine tuning will continue to evolve and will play out as firm's progress on their data journey. Executives need to remember that developing a data-driven organization and culture is indeed a journey. It may be cause for optimism that only 6.7% of executives see the primary responsibility of the CDO as being to ensure regulatory compliance, suggesting that the CDO is steadily moving from defense to offense within most organizations. On the other hand, few executives -- only 2.2% -- see the primary CDO role as building new sources of revenue. It is simply premature and too early for most firms to undertake this in a determined fashion.

Struggling to Define the CDO Role

There was one glaring area of noteworthy and divergent perception in the survey findings. This pertains to the background and qualifications that make for a successful Chief Data Officer, with perspectives varying sharply. While 34% of executives believe the ideal CDO should be an *external* change agent (outsider) who brings fresh perspectives, an almost equivalent 32.1% of executives believe the ideal CDO should be an *internal* company veteran (insider) who understands the culture and history of the firm, and knows how to get things done within that organization. There are notable examples of both flavors of CDO within industry.

Also of note is the significant 22.6% of executives who indicated that the CDO must be either a data scientist or a technologist who is highly conversant with data. An additional 11.3% responded that a successful CDO must be a line-of-business executive who has been accountable for financial results.

Clearly there is a sharp and highly divergent matter of opinion on this topic. When asked about the long term role of the Chief Data Officer, exactly half – 50% -- reported that the CDO should sit on the executive committee, with exactly half – 50% -- disagreeing. Of those disagreeing, 37.1% believed that the CDO should report up to the executive committee, without a seat on the committee.

Perhaps most troubling for incumbent Chief Data Officers are the 12.9% of executives who responded that either the CDO was an interim role which will likely be phased out over time, or that the CDO role is unnecessary and responsibility for data should reside elsewhere.

Looking to the Future

This is the Chief Data Officer dilemma. While the role has emerged as a de facto industry standard and gained broad acceptance as evidenced by the sharp rise in CDO appointments, agreement on responsibilities, mandate, profile, and seniority continue to vary dramatically. The lack of consensus on the Chief Data Officer role aptly mirrors the diversity of opinion on the value and importance of data as an enterprise asset and how it should be managed. The Chief Data Officer can be expect to be near the center of action as firms strive to become data-driven, but should also be prepared to sit in the hot seat during these times of rapid evolution.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

How ADP Gives Data Value Back to Its Customers

March 12, 2018

By Thomas H. Davenport and Randy Bean

In the highly digitized economy in which we live today, relationships between customers and suppliers are increasingly digital. Take, for example, ADP, where the output to customers was once paper paychecks delivered by van. Today, of course, the majority of employees get their compensation directly deposited in their bank accounts, and ADP's business is primarily around digital flows of information.

Most companies that orchestrate digital information flows haven't turned them into anything of value for customers. Think about the companies you do business with, for example. Do they tell you how your spending compares to other customers at your income or service level? Do they recommend ways for you to save money on future purchases? That's highly unlikely. But ADP is an exception to this sorry situation: it provides analytics and benchmarking to its clients about the employment and payroll services it provides to them.

As you might imagine, ADP has a lot of data on which to provide some value. It pays one in every six people in the United States, and serves over 700,000 clients with millions of employees in 113 countries. It knows all sorts of things about employees—not only how much they are paid, but also things like their job tenure, attrition rates, how much they invest in retirement accounts, and at what age they retire. It's taken all that data and made it available in the cloud through a product called ADP DataCloud. If you subscribe to one of its traditional payroll or time-and-attendance services, you can get access to analytics and benchmarking capabilities. And unlike salary surveys—which are almost always self-reported data—ADP's data is from actual payroll records.

Unfortunately, all that data requires substantial massaging before it can be made available across ADP clients. It needs to be anonymized, of course, so that individual clients or employees can't be identified. A more challenging data transformation is to aggregate data across different job titles for the same job. One company might call a job "warehouse associate II;" another might call the same job "inventory selector" or "material handler." ADP addresses the unification of data from similar jobs with "probabilistic matching" machine learning tools (we've described this same technology at companies like GlaxoSmithKline and GE). So regardless of what title you give warehouse employees, you can get data on how much they are paid on average in your industry and geography.

Most of the analytics in the ADP DataCloud and benchmarking offerings are descriptive—generating reports with percentages and averages, often in visual formats. But there are also some predictive analytics available. There is a turnover probability tool, for example, that uses not just a company's own data but industry-wide data and predictive models based on it. Using factors like employees who have been in their role for too long, or who are underpaid for their job type, the model identifies employees at high risk for attrition. If the employer wants to hold on to those employees, it can investigate further or make an intervention.

The analytics and benchmarking offerings are timely ones. Human Resources functions are the most likely to be ADP's client within organizations, and HR is becoming more analytical. It started later than many other functions, but it is coming on strong—the HR expert Josh Bersin found in a recent survey that 69% of companies are working with people analytics now. There are increasing numbers of both HR people and managers who are interested in and capable of using analytics to make people decisions.

Wheelabrator Technologies is one of the companies that is increasingly interested in HR analytics and using ADP DataCloud to address that desire. Wheelabrator Technologies is the second largest U.S. energy-from-waste business, and is an industry leader in the conversion of everyday residential and business waste into clean energy. Wheelabrator Technologies employs approximately 1,200 people in the US and the UK and uses ADP for payroll and HR management software.

The company has only been using the analytics functions for about a year, but Luke Staskal, HR Director at Wheelabrator, says they are quite useful. He commented,

We are a lean organization and don't have a large overhead of resources in HR. Using ADP we have gained the ability to use our internal employee data together with external data to compare descriptive statistics on our employees with people related items – i.e. how different business performance measures like revenue, kilowatts generated, or tons of waste processed compare across facilities, and relate them to the numbers and types of people in different facilities. This analysis is beginning to help us make better decisions about staffing levels, and we expect that such data and analytics will allow us to continue to move from more tactical work in HR to supporting larger strategic issues facing the organization.

At the moment, analytics on ADP data are only possible through the company's proprietary tools on ADP DataCloud. But Marc Rind, ADP's Chief Data Scientist, is aware that most companies already have a set of analytical tools. He notes:

It's possible that we would eventually allow clients to access ADP data through their own analytics tools. That might make it easier to compare people data to business metrics, sales in a particular geography, etc. We may also at some point send alerts and predictions—of attrition, for example—directly to managers. We believe the data is of great potential value to our clients and we are just beginning to see the nature of the value.

It would be great if more digital organizations would emulate ADP and return some value from customer data to those who originally provide that data—the customers themselves. It's also very helpful for organizations to learn how they compare to similar organizations elsewhere. ADP is an early adopter of this concept of allowing customers to hold up a data mirror to themselves. Using it, they can see how they are performing and how they compare to other organizations.

Thomas H. Davenport is the President's Distinguished Professor in Management and Information Technology at Babson College, a research fellow at the MIT Initiative on the Digital Economy, and a senior adviser at Deloitte Analytics. Author of over a dozen management books, his latest is Only Humans Need Apply: Winners and Losers in the Age of Smart Machines. You can follow him at @tdav.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

How American Express Excels as a Data-Driven Culture

March 15, 2018

By Randy Bean

Ash Gupta retires this month after 41 years at American Express, most recently as President for Global Credit Risk and Information Management. He has seen a lot during his tenure, and the characteristically modest Gupta has done more than his share to help shepherd the transformation of American Express into a data-driven company. For many in the industry, Gupta is seen as being the father of data-driven risk analysis, and his efforts have contributed to the increased usage of data and analytics in financial services. I recently sat down with Gupta to reflect back on his four decades with American Express and to understand how the firm transformed into a data-driven culture and a leader in the application of data and analytics within the financial services industry.

Adapting to a Changing Industry

Gupta noted how the financial services and payments industry has changed in important ways over the last four decades, notably:

- Electronic payments, which comprise credit/debit cards and online payment, have steadily gained market share over cash on a global basis, driven in part by co-brand relationships, e-Commerce, convenience, rewards, and the security offered by Electronic Payments;
- E-Commerce and cross border commerce continues to rapidly gain share over offline commerce with the rapid proliferation of internet and smartphones;
- Credit/debit card issuing has become concentrated among a few banks. In addition, multiple companies have joined the payments/e-Commerce ecosystem.

While interactions with the customer are changing from a monologue (direct mail) to dialogue (online application), customers' online and offline identity is being synthesized for a 360-degree view. Industry-wide collaboration among data providers and marketers is driving innovation in product design and marketing, all with rising emphasis on customer privacy and responsible lending. Gupta notes that the rate of change continues to accelerate. As examples:

- Customer interactions, across the entire lifecycle, are gradually becoming app and mobile-based;
- QR codes are big enablers, especially markets like India or China that were under-served and where point-of-sale terminals are relatively expensive;
- Payments and e-Commerce have begun to merge, even as the share of merged entities currently remains small
- New and non-traditional data sources emerge, such as phone number, cookies, and digital data

- Marketplace lenders and Fin Tech companies are increasingly promoting innovations across the customer lifecycle, from targeting, to underwriting, to ongoing servicing and through credit and collections.

Gupta observes that we are nearing the end of one of the longest economic expansions since the World War II. Credit loss has begun to rise and card issuers are focusing on forward-looking economic indicators, stress testing portfolios and investing in decision science to address the potential for a recession.

Forging a Data-Driven Culture

In this environment, data and analytics (machine learning based decision science) have become central to effective marketing, servicing and risk management. Data sophistication is now essential to competitive excellence.

Since its formation 168 years ago, transformation has been a hallmark of American Express's DNA. The firm has transformed from a freight forwarding business to a global leader in payments and personalized services. It has responded to marketplace changes and has been often an industry pioneer in customer-based risk management, membership rewards, marketing, and servicing. Gupta recalls that when he started with the firm in 1976, traveler's checks were a dominant product line and that card-services were just evolving. American Express has since launched many new products, one classic example being small business lending which launched in the late 1990s. American Express is now the number one small business card issuer in the U.S.

While the company continues to evolve at an accelerated pace, its core values remain essentially the same. Gupta describes this as "keeping the interests of the customer first in everything we do and being a good citizen in the communities which we serve". Gupta believes that an effective data-driven organization invests in capabilities, technology and analytics, along with people, recognizing that it is the synthesis of technology and human intuition that provides superior results. Above all, it shows a very deep regard for "permissible purpose" when it comes to data use and customer privacy, ensuring that the data is used responsibly.

Gupta observes that good data-driven organizations have a "test-and-learn" culture and encourage experimentation that can challenge long-held beliefs. These organizations promote listening and learning, and adequately synthesize it with the institutional knowledge and experience. For American Express, this data-driven culture is a natural course of business. At every step, from acquisition to customer management, a large volume of insightful data is created, which is synthesized with partner and third party data to produce industry leading insights that are used for personalization and excellence in customer service and risk management. Gupta remarks, "At American Express, we take our responsibility to serve customers and the public seriously, always ensuring that solutions are best-in-class and valuable to our customers".

Success in creating and maintaining a data-driven organization requires commitment from the top and a willingness to invest in people and capabilities in advance of customer outcomes. In addition, the organization must be attractive to the top talent. Gupta cites several examples of how American Express stays fresh and interesting:

- Risk 2020 -- American Express conceptualizes how the economy and marketplace might evolve in the coming years and what are the most important risk capabilities to maintain to proactively

address the weakness in the economy, a steady move towards mobile computing, cloud, artificial intelligence and deep-learning.

- Cornerstone -- This is a global, big data ecosystem is where data is organized in one place with shared global capabilities, to democratize its use across functions and geographies, recognizing that the very essence of innovation must happen at the company's DNA rather than exclusively from the top.

Gupta notes, "Challenges vary across institutions. The most common challenge is a willingness to invest upfront, recognizing that the benefits will be in the medium term, rather than in the immediate future". Another challenge is recruiting. This means retaining and developing quantitative talent in an environment where the demand exceeds supply globally. According to Gupta, we are successful when we create a culture where there is "openness to listening to analysis that is at times is not in sync with one's mutual instincts."

A Legacy of Accomplishment

In the end, though, it's not just about having the best algorithm. Gupta comes back time and again to people as the critical ingredient to building a culture of success. He stresses the importance of cooperation and striving for collective excellence in collaborating with people. He cites examples from throughout his career:

- Three decades ago, this resulted in converting the credit and fraud risk management discipline from an 'apprentice based' to a 'data and decision science-based' organization not just for American Express, but for the entire industry. This has attracted some of the brightest people to this industry and has been an important training ground for the next generation of leaders within the industry. It has allowed credit to be offered to a far broader range of consumers and businesses promoting better economic growth. It has also been an important contributor to American Express' industry leading credit and fraud write off rates across all economic environments;

- Two decades ago, it was creating a Center of Excellence in India as a collaborating team across all its global businesses to provide data and decision science partnership as American Express launched new products and services. This team remains an internal source of competitive advantage, an envy of competitors and often a benchmark for how to build a "knowledge center" in an international geography;

- In this decade, American Express has created Big Data and machine learning based decision science capabilities and teams, which have fundamentally transformed all aspects of its customer experience, including product design, marketing, risk control and servicing. Most importantly, it has created a foundation for excellence for the next decade as we invest in deep learning, artificial intelligence and blockchain.

Looking back, Gupta is proudest of the creation of the risk and information management team as an industry pioneer. He notes, "It is a team that consistently delivers the very best for our people, customers and shareholders. It is the place where our regard for each other and for all our American Express colleagues trumps our individual needs and desires, where the thinking from each of us has to be 'what can I do for the company?' rather than 'what can the company do for me?', where we collaborate and learn by listening and doing, where being second (in the industry) is not a choice".

Gupta reflects, “My greatest accomplishment has been winning the hearts and minds of my colleagues that investments in data and decision science will bring lasting benefits to our brand, our customers, our people, and forever accelerate our mode to a listening and learning organization." He continues, “I hope my legacy will leave a lasting effect -- a commitment to collaboration, learning and innovation among my American Express colleagues who I’ve had the privilege of serving over the past 41 years”.

It will be the people and the culture of American Express that Gupta will miss most, and it is this culture of collaboration and excellence in data-driven customer experience which is the legacy that Gupta hopes to be remembered for.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

How Blockchain Is Impacting Healthcare and Life Sciences

April 2, 2018

By Randy Bean and Grant Stephen

Soon after its development as the underlying architecture for Bitcoin, the concept of the blockchain was recognized as having broader value beyond enabling a decentralized alternative form of currency. For example, some organizations are beginning to use blockchain to apply advanced analytics from distributed sources without compromising the privacy of individuals.

We are now seeing a rapid expansion of blockchain-based technology offerings, many of them in the Healthcare and Pharma/Biotech spaces. A November 2017 article in Forbes, This Is Why Blockchains Will Transform Healthcare, considered some of the latest developments. Though still in the early days and, despite an enormous amount of interest, most of these value propositions remain little more than prototypes. However, the rate of evolution is surprisingly fast, and the time is right to review some of the emerging use cases and consider where things may lead.

First, what is blockchain? In a nutshell, it is a log of transactions that is replicated and distributed across multiple decentralized locations: As such, it offers a secure, high integrity, "neutral" 3rd party mechanism *for knowing what data is where and precisely how it is changing over time*. Blockchain is not a magic bullet that solves all data management problems but few areas of data sharing cry out more for improvements in efficiency and security than the data domains of Healthcare and Pharma/Biotech.

Blockchain in Healthcare

There are many use cases being explored in Healthcare but let's narrow our gaze on a few of the most promising:

Health Records

Creating usable, high integrity records associated with a patient despite their moving through different healthcare domains and systems is one of the great challenges of healthcare IT. Blockchain offers the possibility of creating a reliable place to track the changes across systems in a manner that gets around many of the concerns associated with data integration between proprietary systems. In effect, blockchain becomes the unifying glue that holds together a highly fragmented healthcare record.

The variety of blockchain for Healthcare concepts can be thought of as spreading along a spectrum from "techno-utopian fantasy" to "compelling near-term use case". One approach towards the latter half of the spectrum is the creation of a blockchain-based Master Patient Index (MPI). MPI is a topic that aggravates most healthcare organizations as it is surprisingly hard for them to reliably track individuals as they consume healthcare from different sources. Patients names are spelled differently, addresses change, insurance providers change. Tracking and adjudicating such patient data is a constant struggle. When it goes wrong… well, in healthcare you really don't want it to go wrong.

MPI is an area of low hanging blockchain fruit, especially for the many existing Health Information Exchanges that have been established to share patient data between provider systems.

Revenue Cycle, Reconciliation & Fraud

One of the greatest cost burdens lurking in the processes of US healthcare is the continuous tracking of the flow of services and money: The enormous complexity and distributed nature of our health system means that billions of dollars are expended annually trying to understand which patient received what service from which service provider and by whose authority. Disputes inevitably arise, and both the insurance industry and the providers of healthcare then expend a great deal of additional time and money adjudicating these disagreements.

Because of its independent architecture, blockchain could potentially form the foundation of a high integrity tracking capability that is updated in a near instantaneous manner. This would lead to many less errors (with both financial and patient care upsides) and substantially reduce fraud. Given the scale of opportunity here, this might generate enough interest to attract the investment and intellectual commitment required to solve this complex technical challenge.

Blockchain in Life Sciences

The direction of travel of Pharma/Biotech R&D is towards cross-organizational collaboration: The sharing of resources and insights across the borders of companies, government agencies and research institutes is central to the development of therapies. But organizations are often reluctant to share data for fear that it gets stolen. Blockchain offers the possibility for trust to be hard coded into the process of collaborative R&D in a way not possible before. More trust means more collaboration and, in turn, more productivity.

Exploring the implications that are generated for individuals by blockchain, the rise of personal genomics means that patients who have their genetic code sequenced are in a position to license access to their genome (for cash or other rewards) in a way not previously possible. Therapeutic research is increasingly focused on mining the genome for information. The potential for blockchain to be used as the backbone for licensed access to an individual's anonymized genetic sequence is one of the patient-empowering edge cases that might well move center stage sooner rather than later.

Other opportunities in hospital & pharma cybersecurity, supply management, IoT data management, and precision medicine lie on the horizon.

A Blockchain Future?

Many barriers need to be overcome before blockchain can be accepted as a mainstream technology: A lack of standards, challenges in communicating between different block chains, the need for off-chain development of open system interfaces, and getting intermediaries (who currently profit from controlling the data) to interface with blockchains are all reasons to slow adoption.

Given that medical errors are estimated to be the third leading cause of death for Americans, improvements in the handling of health data is a time-critical necessity. These are early days but with effort and experience, we can expect more of the technical limitations of blockchain to be overcome and the broad range of advantages to become more compelling. There is so much inefficiency to remove and so many improvements to patient care within our grasp, that the future for blockchain in health seems promising.

Forbes

How DataOps Is Transforming Data Management Practices

April 11, 2018

By Randy Bean

NewVantage Partners 2018 Big Data Executive Survey, demonstrates that culture and organizational impediments are leading barriers to harnessing Big Data. Over half of executives surveyed reported that organizational impediments are preventing the realization of broad business adoption of Big Data initiatives. Yet, 69.4% of those same respondents indicated that building a data-driven culture is a huge priority, and 48.3% envision the long term role of the CDO as supporting innovation and building a data-driven culture. With nearly 44% of firms investing more than $50 million into big data projects in the six years we've been running the survey, it's clear that enterprises are embracing Big Data even as their employees may not be.

So why are organizations seemingly resisting the changes that are so clearly mandated? It might be a culture failure. I spoke with Andy Palmer, a serial entrepreneur who is founder and CEO of TAMR, a next-generation data curation company, and was formerly a founder of Vertica, and former CIO at Infinity Pharmaceuticals. Palmer notes, "In every deployment that we have observed with Global 2000 companies, overcoming internal politics and human behavior with regards to data is orders of magnitude more complex than the technology implementations themselves".

Data should be a shared asset, but many companies struggle to treat it as such. Data transcends traditional organizational structures and lines of business, and managers find it difficult to reconcile its governance against traditional business structures. It is not uncommon for data management projects to digress into organizational turf battles. This lack of sharing can result in many different versions of reality, where managers compete to promote their own. When data users don't trust the data or each other, it's hard to unlock value.

The Emergence of DataOps

Emerging technology providers think that they've found a path forward for building trust through a discipline called Data Operations, or "DataOps." TAMR's Palmer has been a pioneer in the field of DataOps, which he describes as "the framework of tools and culture that allow data engineering organizations to deliver rapid, comprehensive and curated data to their users". He continues, "DataOps enable users to help curate and correct data when they consume it by providing feedback from the point of consumption". By engaging users in the process and facilitating their feedback to improve data quality, organizations allow data engineers, data stewards, managers and business users to work together to drive transformational analytic outcomes.

In his 2015 article, From DevOps and DataOps, Palmer explains how the concept of DataOps borrows heavily from the DevOps movement which transformed the way many internet and SaaS companies build and release software. By focusing on the people, processes, and tools used to build, test, and release software, these companies were able to achieve a level of feature velocity that left many of their competitors in the dust. DevOps methodology engages the customers of the software early and often in

the design process, iterating quickly and increasing the utility and adoption of the eventual releases. DataOps aims to do the same when it comes to delivering analytic velocity for end users in the enterprise using agile data management practices. Palmer notes that as with DevOps, there are a series of both open source and proprietary software projects that are enabling DataOps, including AirFlow, Flink, Beam, Sqoop, Flume and Anaconda. Innovative companies in the DataOps ecosystem include Astronomer, Streamsets, Tamr, DataKitchen, Immuta and Alation.

DataOps projects start with an analytic outcome in mind, and then focus on unlocking value from across data sources and lines of business, enabling users to quickly build analytics that they need. Focusing on analytic outcomes, rather than tools, helps cut through politics and get employees on the same page. DataOps firms recognize that the process and tools they offer need to help build and reinforce a data-driven culture; and must involve consumer earlier in the process to help define the desired analytics. Finally, they also give users ownership in the process, and when a user spots something in the data that isn't quite right, they can flag it and that information is transferred back to the data models themselves.

Collaborative Benefits

Traditional approaches to data engineering don't have mechanisms to allow users to easily provide feedback on data quality from the point of data consumption -- be it spreadsheets, reports or visualization tools. Palmer notes that, "DataOps allows end users to help curate data, and this bi-directional flow allows users to feel engaged in the process and more confident in the results". If a user sees the benefit from the unified data they are consuming, they are more likely to offer up their own spreadsheets and databases to be included. Palmer concludes, "By engaging end users often, it's less likely a user will discard an entire an report or visualization when they encounter an error; instead; they'll want to (and know how) to help fix the upstream issue".

DataOps vendors believe in automating the menial jobs associated with cleaning data and harness the powers of people to tackle the more complicated chores. As GlaxoSmithKline's digital transformation showed, a DataOps mentality can make big projects happen faster and with less pain than anticipated. The first step is admitting that you have a problem.

If Big Data projects aren't yielding returns fast enough, or at all, the answer might not be the technology stack, but people's mindsets. Managers would be well served to approach their projects with an eye towards improving buy-in and adoption from users. Spending less time thinking about technology, and more time addressing people and cultural issues may help organizations deliver data that users will use, and result in meaningful returns from their Big Data investments.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

How Big Data Became 'Big Bad Data'

April 25, 2018

By Randy Bean

Three years ago, in a May 2015 Wall Street Journal article entitled Tracing Some of Big Data's Big Paradoxes, I quoted Washington University in St. Louis Law Professor Neil M. Richards in summing up the opportunities and challenges represented by Big Data, "Big Data will create winners and losers, and it is likely to benefit the institutions who wield its tools over the individuals being mined, analyzed, and sorted."

Now, nearly three years later, the dark side of Big Data is dominating the headlines, and we can't say that we weren't warned. The Monday April 23 edition of The Wall Street Journal carries a book review by Gregg Easterbrook entitled Big Data, Big Problems. The review begins: "'Big Data" is the Big Bad of our moment. Companies and governments amass enormous troves of information about our online and offline activities, so they can understand them better than we do. Recently we learned that creepy firms like Cambridge Analytica mine Big Data from websites such as Facebook. Facebook itself seems increasingly creepy, grounded in lying to the public about what happens to the data it collects". So here we are three years later.

How did Big Data go from promise of the future to "*Big Bad Data*"? Professor Richards hinted at the dark side of Big Data back in 2015 when he noted, "Big Data promises to use data to make the world transparent, buts its collection is *invisible*, and its tools and techniques are opaque, shrouded by layers of physical, legal, and technical privacy by design". Now in 2018, we have this view from Gregg Easterbrook: "In the future, will Big Data help physicians cure diseases or help health insurers deny claims? Make factories and products safer or accelerate layoffs? Ultimately spawn some kind of hostile artificial intelligence? Right now, it's fair to suppose that many people would favor putting the Big Data genie back into the bottle".

Cautionary warnings about the perils of Big Data have now become omnipresent. The latest issue of The New York Review of Books carries an essay entitled, Reining in Big Data's Robber Barons. Author Jennifer Cobbe comes out of the blocks forcefully, arguing: "Google, Facebook, Amazon, and other tech giants have constructed the most extensive and intrusive surveillance apparatus the world has ever seen. And we are the target". Cobbe describes what she calls "surveillance capitalism", a term she credits to Harvard academic Shoshana Zuboff. She concludes, "The questionable practices of surveillance corporations and their refusal to act responsibly have brought us to a turning point. This is a moment of decision: Will it be our Internet, or theirs?"

The new critique also has drawn in early champions of Big Data such as Alex "Sandy" Pentland and David Shrier of MIT's Media Lab, who were early proponents of ethical data privacy practices. In an April 11 editorial in Newsweek, Facebook's Arrogance Crisis, the authors note, "By failing to measure social and governance risk in an appropriate fashion, Facebook has endangered not only its market capitalization and shareholder return, but also its future cash flows". They continue, "Ethical guidelines could also have helped moderate Facebook's risk profile".

The proliferation of Big Data can be expected to continue to accelerate in the decades ahead. Where consumers and companies had reveled in the promise of personalized experience and convenience, the risks of untrammeled Big Data use and abuse now overshadow potential benefits. Big Data is not going away. Organizations and individuals must weigh the benefits and consequences of making data freely available for open use and access. It is incumbent upon the data industry and individuals as never before to develop sensible approaches and safeguards for managing Big Data as an asset. Data can be used to great benefit if governed wisely, and to great harm if deployed without regard to its consequences.

I used to tell a story to people who didn't understand my work in the field of data, and found the terminology and concepts to be mystifying. I explained how my wife, who was in the medical profession (with its own mystifying terminology), would say to her colleagues, "I don't understand what he does. I think he must work for the CIA!" Later, after it was revealed government agencies were maintaining data on private citizens, she updated her response to say that she was now convinced that I worked for the NSA. I retold this story to an executive the other day when he looked up at me and deadpanned, "And, it turns out that all that time you were working for Facebook!" Big Data has devoured itself. We have come full circle.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Understanding Blockchain 101: Untangling Myth from Reality

June 6, 2018

By Randy Bean

Blockchain is the latest business and technology Holy Grail du jour. It is this year's Artificial Intelligence or Big Data – both now "so yesterday" when measured by fervor. I recently wrote in this space about How Blockchain is Impacting Healthcare and Life Sciences Today. Clearly, there is great potential, and with this come great expectations.

Blockchain holds the promise of creating new levels of data trust, while providing a shared history of activity ("a ledger"). Blockchain can improve the efficiency and reliability of transactions and exchanges of data between multiple parties, and can track the flow of activity and actions taken upon that data. In a time of unregulated and ungoverned data usage and exchange, Blockchain holds the potential to introduce a historic record of activity leading to greater data veracity.

I recently hosted an executive roundtable breakfast on the topic of Blockchain in Healthcare and Life Sciences. The breakfast drew Chief Data Officers, Chief Analytics Officer, Chief Information Officers, and senior data science executives from firms including Aetna, Alexion, Astellas, Cigna, CVS, Partners HealthCare, and Takeda among others. These executives shared their excitement, and some skepticism, over opportunities to apply Blockchain to issues of traceability and lineage in areas such as clinical trials, health information exchanges, and to combat counterfeit pharmaceuticals.

Two summers ago, in July 2016, Massachusetts Institute of Technology (MIT) hosted a working group session in Cambridge, MA, in response to a call to action at that time from the then White House Commission on Cybersecurity. MIT Connection Science, organizers of the working group, published a summary whitepaper entitled Towards an Internet of Trusted Data: A New Framework for Identity and Data Sharing. The 2016 MIT discussions sought to understand Blockchain in the context of the larger economic and societal transformation, as we have moved from an industrial economy to an information-based economy over the course of the past half century. This information economy is increasingly characterized by digital interactions based on online data and transactions. As data has proliferated (Big Data), so too has the need to manage, distribute, govern, access, and utilize data, in a secure, protected, and trusted manner. It is within this context that Blockchain has emerged as a potential solution.

Yet, for all of the promise and potential represented by Blockchain, there remains a critical gap in both understanding and execution. I attended an industry program last fall in New York which featured a discussion of how Blockchain might be a factor in disrupting the financial services industry. At the conclusion of the discussion, the panel moderator commented, "Well, I still don't understand Blockchain". From my experience, this reaction is not uncommon. Understanding the role and application of Blockchain remains a challenge for many executives, both business and technical, who seek to grasp the significance and potential business impact of Blockchain to their organizations.

One obstacle to understanding Blockchain is that explanations can come across to some as sounding highly technical or "architectural" in nature, making it difficult for non-technical executives to understand how Blockchain may be relevant. Executives want to know how Blockchain will translate into business value and business benefits? Will it make my organization and our data more usable, reliable, safer, or more dependable? What will be the business impact of Blockchain? Why will Blockchain matter? For Blockchain to take hold, the rest of us non-specialists will need to develop a richer understanding and appreciation of what Blockchain is and is not, what it can do and what it cannot do.

This gap in understanding Blockchain is matched by a current gap in execution as well. A recently published study from Gartner Group confirms that only 1% of the 3,138 Chief Information Officers (CIO's) who were surveyed had already "invested and deployed" Blockchain initiatives. Further, no more than 22% of CIO's reported any type of "planning or experimentation" with Blockchain. An overwhelming majority of the CIO's who were surveyed (77%) reported "no interest in the technology and/or no action planned to invest or develop it". A further breakdown indicates that 43% of CIO's report that Blockchain is "on the radar, but no action planned", while 34% of CIO's report "no interest" at all in deploying Blockchain. Clearly, the execution gap is real today. Nonetheless, Gartner strongly cautioned that companies that ignore Blockchain do so at their own peril.

Since the MIT working session of 2016, fears relating to data security, data leaks, and rampant data misuse have moved from theoretical threats to daily realities, as can be attested by the most cursory review of the news. The timing for Blockchain might never be more urgent than now, but organizations must chart a course to experiment and innovate around the edges, establishing Centers of Excellence, Innovation Labs, or consortiums to validate the promise of Blockchain. Ultimately these efforts will lead to wider adoption, as organizations address both the technology and cultural challenges that represent an obstacle to any new technology and business approach.

For now, Blockchain holds great promise. In the ensuing years, we will understand whether Blockchain can fulfill its ambitious, and perhaps revolutionary, potential.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Feeding a Data-Hungry Organization at Charles Schwab

June 6, 2018

By Thomas H. Davenport and Randy Bean

Consumer financial services has long been one of the most data-intensive industries, and Charles Schwab has long been a leader in it. It is one of the largest broker-dealer firms in the United States, with $3.36 trillion under management. It has also long been a leader in the use of technology as a way to improve customer service, operational efficiency, and investing outcomes.

But this is no time for Schwab to rest on its laurels, and it certainly isn't doing that. Based in San Francisco, the firm is surrounded by digitally native firms at every turn. Its executives, employees, investment advisors, and clients are acutely conscious of and hungry for data, analytics, and insights—perhaps more so than in other financial firms based outside of the San Francisco Bay Area. As a result, Schwab has made substantial investments in its digital, data, and analytics capabilities over the last several years in order to satisfy these data-hungry stakeholders.

Schwab has been investing in several different areas. Several years ago its executives recognized they needed more advanced analytical capabilities, an area headed by John Carter, SVP Analytics & Business Insight. The group laid out a path to build out new data science capabilities, and hired a number of data scientists.

At about the same time, Schwab felt that it needed to modernize its data environment and Andrew Salesky was named head of Global Data. The Global Data Technology organization was also formed, reporting to the CIO. This led to incorporating a Hadoop-based big data lake and other new technologies. Salesky notes that Schwab, like any other organization wanting to do more with advanced analytics and AI, needed to make some "long tailed" investments in its data. In terms of data integration and master data management, the primary focus at Schwab has been on client data and developing consistent enterprise definitions of key data assets.

One more major organizational change took place in 2017, the creation of a Chief Digital Officer position (Neesha Hathi was named to the role). Under her leadership a Digital Services organization was created to develop new digital capabilities.

Not all data, analytics, and artificial intelligence projects are included in these three groups—there is a separate organization, for example, for compliance and fraud prevention analytics—but the great majority of them are. The three groups described above have primary responsibility for customer and advisor applications, and for making sense of the data they use and generate.

Artificial intelligence at Schwab is found primarily in customer service. Although the firm's Intelligent Portfolios "robo-advisor" offering employs some analytics to select and rebalance portfolios, Schwab's focus has never been stock-picking, either with AI-based or human expertise. Instead the company uses machine learning and natural language processing (NLP) to provide some online services and advice. According to Carter: "These techniques are used across the enterprise to provide clients with a

better customer experience and help us improve our operational efficiency. For example, when you reach our call center, our algorithms will use a wide range of data -- in real-time -- to deliver the most appropriate information to our service representatives based on a client's needs."

Schwab also uses another form of NLP to understand customer concerns and satisfaction levels. The company's data scientists analyze unstructured data in the form of text, speech, social media, client verbatim forms, and other sources in order to identify customer issues and opportunities. Schwab has long been a user of the Net Promoter Score, and it relates unstructured data to the NPS scores and other satisfaction measures that customers provide.

Schwab's Digital Services organization is also examining other ways to use AI in customer interaction. It already offers an Amazon Alexa skill that provides market updates, quotes, and the values of personalized stock lists.

Digitally-driven innovation at Schwab has taken place for many years, but the company is pushing into new areas. Schwab announced the creation of two digital accelerator hubs in Austin, TX and San Francisco in 2018, which will be treated as startups and will have the goal of developing better digital experiences for retail investors and advisors. With its headquarters near the heart of Silicon Valley in the Bay Area, it seems likely that Schwab will continue to find innovative ways to make investing easier and more successful with the help of data, analytics, and artificial intelligence.

Thomas H. Davenport is the President's Distinguished Professor in Management and Information Technology at Babson College, a research fellow at the MIT Initiative on the Digital Economy, and a senior adviser at Deloitte Analytics. Author of over a dozen management books, his latest is Only Humans Need Apply: Winners and Losers in the Age of Smart Machines. You can follow him at @tdav.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

Getting to Trusted Data Via AI, Machine Learning & Blockchain

June 17, 2018

By Randy Bean

Establishing trust in data is an essential requirement for businesses and entities for whom credible, reliable information is the lifeblood. As enterprises seek to manage data as an asset, it becomes increasingly vital that data sources are trusted and verifiable.

I wrote a few weeks ago about the MIT initiative to establish a framework for trusted data, and the resulting position paper, "Towards an Internet of Trusted Data: A New Framework for Identity and Data Sharing". The authors highlight the criticality and need for "trustworthy, auditable data provenance" where "systems must automatically track every change that is made to data, so it is auditable and completely trustworthy". One of the key recommendations of the study was to improve the process and quality of data sharing. One suggestion was to move the algorithm to the data, explaining "The concept here is to perform the algorithm (i.e. query) execution at the location of data (referred to as the data-repository). This implies that raw-data should never leave its repository, and access to it is controlled by the repository/data owner".

Tom Dunlap has been at the center of issues of data trust, standardization, and normalization for well over a decade. Dunlap most recently served as a managing director at Goldman Sachs, where he was global head of enterprise data strategy and reference data operations during his seventeen-year tenure with the firm. Among other responsibilities, Dunlap served on Goldman Sachs operations data digitization council and financial reform steering group. He also serves as a member of the Financial Research Advisory Committee at the US Treasury Department's Office of Financial Research.

From his catbird seat at the heart of the action in financial services, Dunlap developed some informed perspectives on issues of data trust and data reliability. He sees the financial services industry progressing on a path to enriched data quality and reliability. Dunlap notes, "From the top on down, financial services firms are viewing data as a corporate asset, where data is seen as being foundational to achieving not only compulsory needs with regulatory reporting, but also as improving the client experience and enabling commercial initiatives". Dunlap sites as an example the introduction of Legal Entity Identifier (LEI), which is being employed by financial services firms to manage systemic risk. In addition, financial services firms are tracking data lineage and definitions of data, with the result that data can be traced from production through consumption, to accurately understand the points at which data is being used and how that data is being transformed during its lifecycle. The result, notes Dunlap, is that "data can now be trusted, and verified, from the source, with fewer data quality problems being experienced". The benefit is that higher levels of data quality translate into faster time-to-market for activities including product profiling and pricing, and faster trade executions. The net result is that client experience has improved.

As data has proliferated, so have the variety of new data types under review, including what are known as "unstructured" data sources. Examples would include documents, pictures, texts, and other free-form images. It is in addressing the challenges of managing unstructured data that Artificial Intelligence (AI) and machine learning are enabling breakthroughs. Dunlap cites the example of "derivative contracts",

where formats may differ across financial institutions. AI and machine learning capabilities can be used to look within documents to automatically detect key data elements, such as legal entity names and economic terms. Firms are applying AI and machine learning to search for these data points, perform language translations as needed, match Legal Entity Identifiers, and load the resulting output into categories that have been assigned predictive levels of completeness and accuracy, which are usually quite high. Over time, AI and machine learning algorithms become very good at knowing what key data attributes to look for, where to bucket these attributes across workflows, and delivering recommendations on data enrichment. The result is that data capture and matching processes which had taken a full day to complete have now been reduced to a matter of minutes, even seconds in some instances.

Blockchain offers an alternative model to access data and a different way to imbue trust in data quality. David Shrier has been a trailblazer in the movement to establish trusted data. In addition to serving on the MIT commission which produced the policy paper on trusted data, Shrier is a lecturer and futurist with MIT Media Lab, an advisory member to the Financial Industry Regulatory Authority (FINRA), and an associate fellow at Oxford University, where he is engaged in the delivery of global online Fintech and Blockchain initiatives through Oxford Fintech and Oxford Blockchain Strategy. Shrier observes, "Blockchain is a completely different kind of database, one with the potential for greater transparency into the data for multi-stakeholder environments, and greater cyber-resilience if certain types of Blockchain and other technology are combined". He continues, "The old-school concepts of data lake, data warehouse, and data mart still rely on the concept of having a centralized database which provides for a single point of failure and an attractive attack surface for hackers".

Shrier goes on to note, "We are just beginning to explore the potential of Blockchain to help transform society. Blockchain has given birth to a new model of funding, of distributed capital formation, for businesses called ICOs (initial coin offerings). This is particularly important in Europe, for example, where today 70% of the funding for businesses relies on banks. In the US, most innovation funding is concentrated in Silicon Valley, and ICO's hold the potential to democratize innovation funding if the regulators don't shut it down". He continues, "Consumers can have better digital identity, lower cost financial services, new employment and community models, better control over their assets, and more, through Blockchain systems". Shrier concludes, "It's still very early in the development of applications for consumers. In 1994 internet tech, we had no conception of Airbnb or Uber, and I think we're in a similar stage with Blockchain technology".

The biggest issues surrounding the use of personal data today come from not knowing where this data is stored, who is looking at it, or what is being done with this information. While the new European data protection law, the General Data Protection Regulation (GDPR), begins to address these issues, there is still a need to provide technology infrastructure that will enable trusted data sharing. Blockchain approaches, as described in the MIT Trust Data initiative, provide a path to a trusted data framework which can ensure:

- more secure personal information
- better access to data through a personal data store
- an unchangeable audit trail of who's done what with personal information.

Shrier reflects in conclusion, "Society as a whole can benefit from more reliable, distributed data and information. In this era of fake news and state actor interference in elections, creating technology-driven trust offers the potential to restore faith in our shared institutions".

Forbes

How FinTech Initiatives Are Driving Financial Services Innovation

July 10, 2018

By Randy Bean

It was on this date, July 11, in the year 1804 in Weehawken, New Jersey, that Alexander Hamilton, architect of the American financial system, met his fate at the hands of Aaron Burr. Some would argue that the financial services industry has changed very little since then. Change has some slowly to financial services, and innovation has proceeded at an evolutionary, some would say glacial, pace. That may be changing however. Recent decades have born witness to the introduction of technology solutions that have accelerated the transformation of consumer experience. Technological advancements such as the Automated Teller Machine (ATM), online banking and bill payment, and mobile banking have enhanced consumer experience in banking and payments, while alleviating costs, increasing convenience, and streamlining processes. The shock of the great recession, coupled with the emergence of new technology applications in self-service, online and mobile banking, machine learning, Big Data, and artificial intelligence (AI) may now be sowing the seeds of financial services disruption.

Fear of disruption is a growing concern for financial services firms. A 2018 executive survey found that nearly 80% of top executives feared that their firms were at risk of disruption and displacement from highly agile, data-driven competitors. Three quarters of the executive respondents represented the largest financial services firms. This rising fear of disruption and potential displacement can be attributed in part to the increasing threat of encroachment coming from the big tech giants – Amazon, Google, Facebook, and Apple. I explored this threat in a Forbes article written last fall, Financial Services Disruption: Gradually And Then Suddenly. There is a new and additional threat as well. As consumers grow to expect greater customization and personalization of their financial services experience, a new wave of innovators is opening up a more expansive vision for financial services. Much of the innovation in financial services these days is being driven by a new set of entrants -- FinTech startups.

The Emergence of FinTech

FinTech (short for Financial Technology) can be characterized as the movement to bring transformative and disruptive innovation to financial services through the application of new and emerging technologies which address consumer needs through automation. Due to factors including consolidation in the financial services industry and regulatory constraints, financial services firms may find themselves constrained from being able to focus their energies on innovation initiatives. FinTech startups have the advantage of not being encumbered by legacy systems and processes. As a result, FinTech firms are generally able to move faster and develop solutions that compete directly with traditional methods of delivering financial services. With customer acquisition costs high, and regulatory hurdles to overcome, financial services firms are faced with a choice whether to build their own capabilities or seek out FinTech partners to help drive innovation initiatives. This opening has provided an opportunity for FinTech firms to provide new applications either directly to customers, or in partnership with large financial services institutions.

Large institutions must consider how they can move quickly to address consumer needs in an industry on the cusp of change, either through partnerships, acquisition, or internal initiatives. Most firms are taking a hybrid approach.

Jean Donnelly is executive director of FinTech Sandbox, a non-profit innovation center that was established in Boston in March 2015 with the mission of bringing to market the latest technologies that will transform financial services. According to Ms. Donnelly, "FinTech Sandbox helps startups obtain critical data to build and test products. We have a network of over 37 data and infrastructure partners (and growing) who will give access to their premium products to startups we work with to develop and test their applications". Ms. Donnelly continues, "We also help get the best solutions in front of our financial institution partners, to explore opportunities for partnership, proof of concept engagements, as well as investment opportunities". Ms. Donnelly notes that sponsors of FinTech Sandox include leaders in market data, technology infrastructure, and financial services that seek to explore the innovative solutions that FinTech startups are developing. Among the sponsors of FinTech Sandbox are financial services stalwarts such as Fidelity Investments, Franklin Templeton, Thomson Reuters, and State Street Corporation. FinTech Sandbox is also collaborating with startup accelerators such as MassChallenge on a new initiative announced last week that will match later-stage FinTech startups with industry leaders. The initiative is intended to accelerate the development of products and service solutions that have the potential to transform financial services.

Innovation at the Edge

The current FinTech universe encompasses startup firms in addition to initiatives coming from within established financial services incumbents, many of whom have launched Innovation Centers or Excellence. Capital One has been an innovator in financial services for several decades based on the application of data-driven analytics. Cap One has operated a Big Data Lab for several years now as it seeks to continue to apply innovative technologies and approaches to its business. The maturity of FinTech initiatives vary. My colleague Tom Davenport and I have written about AI and machine learning solutions from incumbent firms including Morgan Stanley and Charles Schwab, both of whom have been at the forefront of data and AI-driven financial services innovation in recent years.

A continued topic of debate revolves around who is in the best position to innovate in financial services, and whether that is large incumbents or FinTech startups. FinTech firms have the advantage of not being encumbered by legacy systems and processes, and can therefore move faster to develop custom solutions. Firms like PayPal, Square, and Lending Club are examples of FinTech startups that have successfully grown into leading market competitors. Manish Gupta is founder and CEO of Corridor Platforms, an integrated credit risk management platform which has been designed to enable real-time credit underwriting and customer management. Mr. Gupta was formerly the executive vice president and global head responsible for information management and data products, Big Data, and advanced decisioning at American Express. Mr. Gupta observes, "FinTech companies are better suited to initiate disruption by innovating and improving key drivers in an established financial value chain. They can be more agile, focused, and unencumbered by legacy issues like fixed cost, old infrastructure, and technology". Gupta cites the example of marketplace lenders who have improved customer experience through the generation of multiple counter offers in real-time. He further notes that the initial disruption caused by FinTech startups can accelerate the innovation in established companies and make them rethink their own structure and partnerships. "A mix of FinTech energy and long-term experience from established companies elevate the overall value proposition and value chain" says Gupta.

The Future of FinTech

What does the future hold for FinTech? AI, machine learning, and Big Data are becoming central to FinTech solutions as firms look to new areas of financial services innovation. One area that is ripe for innovation is marketplace lending. Ash Gupta, former Chief Risk Officer and President of Global Credit Risk for American Express, notes, "Fintech and digital lenders have created new excellence in customer experience and have permanently changed customer expectations with regards to speed of credit decisioning and the choice and flexibility available with regards to range of offers and prices". Recent examples of FinTech firms that have found opportunities to innovate include:

- Quantopian – A community of over 100,000 quant users sharing a platform where they can develop and share algorithms and trade them in the market;
- Kensho – An AI engine focused on putting data and analytics at the fingertips of non-technical users to bring data science to a larger financial services audience;
- Elsen – A high performance computing engine that can take the complex models for hedge funds and traders and process results in minutes.

Financial technology provides a foundation for financial services innovation and transformation, yet firms still must overcome the final barriers and roadblocks to widespread business adoption – cultural change. Nearly two-thirds of executives state that business adoption remains a challenge, yet only 1 in 5 executives say that technology represents the barrier to successful adoption. With 4 out of 5 executives pointing to cultural challenges -- people and process -- as the principle barriers to business adoption, financial services firms will continue to face challenges as they seek to leverage FinTech solutions to stave off agile, data-driven competitors.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Time to Value: The Currency of Data Operations

July 24, 2018

By Randy Bean

Business executives at mainstream corporations are quickly grasping a central premise that executives of data-driven companies have well understood -- the speed with which the enterprise can get value from its data matters. And, it matters a lot. Accelerating that timeframe, or improving on time to value, is or should be a core focus of any Chief Data Officer. As Mark Clare, former Chief Data Officer of HSBC's and JP Morgan Chase's retail divisions expressed it, "This is about speed and cost. Both are key requirements. Speed to market, and speed to value."

Accelerating speed to market can be challenging enough, but when it is executed at the scale of data that corporate enterprises manage, the process can seem impossible. As Clare describes it, "I had oversight across numerous markets and hundreds of data sources." In his efforts to fuel his firm's insights, he was operating at a scale that was daunting. Clare continues, "At that time, both technology and legacy, manual processes wouldn't scale to the challenge." As a result, Clare had to forge his own approaches so that he could deliver timely results while he waited for the technology landscape to catch up. Solving the huge demand for analytics drawn from vast amounts of data sources while delivering insights in time to be actionable, is a core challenge for enterprises that seek to compete with agile, data-driven competitors. Chief Data Officers like Clare are turning to new approaches and technologies to tackle parts of the problem. These executives are learning that iterative, agile tactics yield the most dividends.

Next generation data engineering technologies and processes are part of the emerging Data Operations, or DataOps movement. The DataOps movement is about enabling organizations to deliver more comprehensive data to their business analysts and decision makers. This is done by applying some of the same techniques employed by the DevOps movement in software engineering to the disciplines of data modeling and data engineering. As Clare expains it, "DataOps really excites me because it creates formality around how to be a lean, agile data organization." Interestingly, a recent survey noted that 73% of companies are planning to invest in DataOps initiatives to support their Artificial lntelligence (AI) and machine learning initiatives. The survey went on to indicate that a similar percentage of firms intended to hire professionals with DataOps skills as they seek to automate all aspects of data engineering.

Clare believes that DataOps represents the future. During his time at one large financial services firm, Clare undertook what was expected to be a multi-year project and delivered it in one quarter of that time by adopting an agile approach which streamlined processes through automation and agility. This enabled the business to deliver actionable insights across lines of business, and helped the enterprise optimize costs and increase revenue opportunities. The business benefits are clear. Today, DataOps practices are being used to accelerate the time that it takes to realize business value for a range of business activities. One example of how DataOps is having a business impact is in the transformation of fraud detection with consumer credit cards. Executives understand that fraud detection is an essential capability, but implementing a fraud detection system that doesn't offend or inconvenience customers can be difficult.

Denial of a credit card payment, or shutting down a credit card, represents a massive consumer inconvenience, with the result that customers may change banks in response. Banks that invest heavily in rules-based approaches to fraud detection may only see when certain rules are violated, without having a complete picture of the consumer's pattern of spending. If a bank uses a set of rules to approve credit transactions, rather than employing dynamic analytics that are based on actual spending patterns, the result can be a bad outcome and an unhappy customer. As Andy Palmer, CEO of the DataOps firm Tamr puts it, "DataOps technologies look at the data itself to understand the data, rather than rely solely on declarative rules which may not be flexible enough to handle heterogeneity or ambiguity."

Clare recalls an example drawn from personal experience. In 2010, while on a business trip, he learned that an Apple store had a received a limited inventory of new iPads. He purchased the device, but by the time he had reached the door he'd received a phone call and text message from his bank notifying him that his credit card had been shut off. While this transaction may not have made sense when viewed from the vantage point of being a large electronics purchase in a foreign city, when taken in the context of his total purchase and business travel history, it wasn't out of line. DataOps presents an alternative. If the bank had engineered a system which dynamically incorporated new data to generate insights about Clare from his spending, rather than checking his transaction against static rules, it would not have needed to send a text alert, call his phone or lock his account, leading to a much better customer experience. By focusing on engineering repeatable, reliable and dynamic analytics, rather than static rules, enterprises are able to create a better picture. That engineering challenge is difficult, but embracing DataOps technologies and processes will allow firms to automate as many processes to create better analytics pipelines In this case, dynamic customer spend analytics reduce fraud detection friction with customers, in turn reducing churn.

DataOps approaches enable enterprises to derive value quickly and efficiently from their data assets. In the example of consumer spending, real-time resolution is the target outcome. As a result of more efficient and accurate data preparation and data integration, organizations can finally begin to rapidly address real consumer needs that accelerate the time in which it takes to deliver business value.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Farmers Accelerates Its Time to Impact with AI

August 1, 2018

By Thomas H. Davenport and Randy Bean

Insurance, which is an industry made up of data, transactions, and decisions, is one of the industries that is most likely to be affected by artificial intelligence (AI). Farmers Insurance, the California-based insurer, intends to apply the technology aggressively to its business. AI projects have proliferated around the company over the past few years, and now its leaders are viewing AI as an enterprise capability and are organizing to allow for greater impact.

AI is being used in a wide variety of use cases at Farmers. A recent Wall Street Journal article, for example, described the use of AI-based image recognition in the automobile insurance claims process at the company. Chatbots and other natural language processing-based tools are being used in contact centers. The potential of machine learning algorithms is being explored as part of the insurer's s usage-based insurance program, called Signal. The company is also using automated machine learning tools to develop better predictive models. Robotic process automation—although some at Farmers aren't sure it qualifies as AI—is being tested to automate back-office digital processes.

As at many organizations, the initial AI projects at Farmers were fragmented around the company, with little coordination. This led to the early adoption of multiple different tools and vendor relationships—sometimes for a similar purpose. Now, however, Farmers has developed a variety of coordination mechanisms to ensure a well-managed enterprise rollout of AI. One key mechanism is the AI Council, for which there are both business and IT versions. The business AI Council considers new use case proposals, which is the initial way to consider a project. The business units themselves do not lead with AI, but rather with the problem to be solved—if AI surfaces as a means to accomplish the desired business objective then the use case is routed to the council.

In the IT AI Council, the primary focus has been on identifying enterprise tools, vendors, and data. In some cases, Farmers works with "insurtech" vendors on vertical solutions for use cases like analyzing customer verbatims with machine learning models. In other cases, for example in the development of chatbots, they use their own resources to develop solutions using widely available platforms like Amazon's Lex. Katie Meyers, Farmers Chief Data Officer, told us that Farmers plans to develop a series of rationalized platforms in key areas of AI with an objective of progressing individual AI initiatives to a mature aggregate capability.

The IT organization's AI Council has categorized AI projects in terms of their stage of implementation—with an eye toward moving beyond pilots and proofs of concept to production. The projects are in three categories:

- Discovery—considering the project or technology, assessing viability, light experimentation;
- Pilot—formal experimentation;

- Production (or on the highway to it)—full implementation of system and related processes.

Meyers notes that there are quite a few projects in the production category, although some involve relatively simple AI technologies like chatbots. At least 50% of their projects are in the discovery phase, with the expectation that the pipeline will be continuously replenished as additional business use cases are brought forward and technology improves.

In terms of the impact of AI on human workers, Farmers has an enterprise philosophy of augmentation, rather than employee displacement. Meyers stated,

We see the power of AI as being the machine working in tandem with employees. How do we use human skills where they are of most benefit to the customer? We are not focused on replacing people. Human interaction plays a huge part in our business and customer relationships.

There are already some examples at Farmers of how AI is enhancing human jobs rather than replacing them. In contact centers, for example, chatbots are freeing up representatives from answering the most repetitive and basic calls. Supervisors in contact centers have historically used relatively simple tools to support the development and performance management of people who take calls. They listened in on calls for opportunities for improvement but only had time to review a fraction of a service advocate's calls. Their goal in the QA exercise would be to figure out what reps needed to do differently—when are they making offers effectively, or using high-effort language requiring work by the customer.

Now an NLP system converts all speech recordings from calls to text. Supervisors can look at 100% of calls if necessary, but the AI system automatically identifies problematic language and interactions. As a result, supervisors can do much more effective and informed coaching with the call reps.

Dexter Johnson, who heads strategy for Farmers Service Operations, says that supervisors now have jobs that are substantially more data-oriented than in the past. He notes:

We want our service advocates to make it easy for customers to do business with Farmers. So we are particularly focused on avoiding high-effort language. The supervisors have a dashboard that monitors high effort language for them. They can tell how many of those high effort calls there are, compare their team to other teams, look at individual members of the team, and listen to either problematic or particularly good calls with the rep to give them some coaching. It's already had a big impact on our service levels.

Johnson doesn't envision a time when AI-based agents handle all customer interactions. Humans, he notes, are better at making a personal connection with customers than machines—and he expects that will be true for a long time. For relatively simple transactions, he says, customers seem to prefer machines for their speed and 24/7 availability. But for more complex issues like understanding coverage, or explaining why a premium changed, there will need to be an element of discussion with people. That's a division of labor among humans and machines that benefits both employees and customers.

Forbes

Will Blockchain Transform Healthcare?

August 5, 2018

By Randy Bean

The Wall Street Journal recently noted that the United States "will soon spend close to 20% of its GDP" on healthcare. While it might not be possible to address the rising costs of healthcare in the immediate future, there are steps that can be taken to address issues of customer service and efficiency to improve the overall healthcare experience, while ensuring the protection of customer privacy. In recent months, there has been a flurry of excitement about the role that blockchain technology might play in the long-term transformation of U.S. healthcare.

I recently spoke with a few individuals who are deeply conversant in the challenges facing healthcare and how solutions like blockchain can be brought to bear. John Halamka is Chief Information Officer of Beth Israel Deaconess Medical Center in Boston, a Harvard University teaching hospital, a position that he has held since 1998. He also held the position of CIO for Harvard Medical School from 2001 through 2012. Halamka recently assumed responsibility as Editor-in-Chief of the new academic journal, Blockchain in Healthcare Today. In the inaugural issue, published in March of this year, Halamka proclaimed his manifesto, "As the Editor-in-Chief of **Blockchain in Healthcare Today**, my goal is to publish high-quality opinion pieces and research papers about use cases that really require blockchain". Halamka continued, "Just using blockchain in healthcare because it's cool does not make sense".

Halamka knows of what he speaks. In his Journal call to action, Halamka goes on to note, "In 2017, I worked on several production blockchain applications, so I have a sense what works and what does not. Blockchain is not meant for storage of large data sets. Blockchain is not an analytics platform. Blockchain has very slow transactional performance. However, as a tamperproof public ledger, blockchain is ideal for proof of work. Blockchain is highly resilient". I asked Halamka about what he sees as the greatest healthcare challenges where blockchain could make a difference. Noting that blockchain is ideal for ensuring data integrity where control is decentralized, Halamka cites three prominent opportunities:

- **Medical Records.** When a medical record is generated and signed, it can be written to the blockchain, which will provide absolute proof and confidence that a medical record cannot be changed. The integrity of the medical record is ensured. The same concept can be applied to clinical trials. This has impact in legal cases as well where the integrity of the medical record is pivotal.
- **Consent management**. In the current healthcare environment where every state has different privacy and consent regulations, blockchain could be used to record patient consent for purposes of data sharing. Any party seeking to exchange medical data about a patient could check the blockchain for permission to do so.
- **Micropayments**. The idea that patients might be incented is gaining traction. If a patient follows a care plan, keeps their appointments and stays healthy, there might be rewards offered through

the blockchain. Similarly, patients might be rewarded for contributing their data to clinical trials and clinical research using the same approach.

Tory Cenaj, founder and publisher of Blockchain in Healthcare Today, adds, “Blockchain technology can elevate care excellence, and enhance the participation of owning one's health and data”. Greg Matthews, whose mission is data-centered innovation in healthcare, and is creator of MDigitalLife, a platform for tracking digital trends in healthcare, offers an additional perspective, “Blockchain could make the biggest impact in healthcare in enabling health outcomes that take a 360° view of the patient’s genetic profile, their demographic and socioeconomic status, the behaviors that impact their health, and their response to different treatments or combinations of treatments”. Matthews continues, “This data exists today in one form or another, but can be tremendously difficult to stitch together at an individual level. Blockchain can enable “profile stitching”, and do so in such a way that the patient’s identity is protected”.

A Blockchain Future?

Halamka observes how blockchain is ideally suited to addressing the challenge of decentralization of medical data. “Most healthcare data is centralized at the level of a corporation, healthcare facility or government registry” notes Halamka. “Blockchain is decentralized and therefore not impacted by the behavior of any one organization. In the future we might see blockchain as a component of a system in which patients serve as stewards of their own data, rather than relying on any central source”. Matthews concurs, “We haven’t been able to aggregate patient data in one place and secure it so that only the patient has control of it and can make decisions who they want to share it with”.

Matthews envisions a future where blockchain would play an integral role in healthcare improvement. He observes, “By using blockchain in combination with AI and machine learning, we should be able to discover potential solutions to health problems that are devastating to us today”. Matthews continues, “The dream of personalized medicine looked like an almost-insurmountable problem 10 years ago because of technical challenges in linking data types and using them to find patterns across massive amounts of data. Today, the dream is more threatened by the harm that personalized medicine could do if the data and insights it yields were improperly used”. He concludes, “Blockchain could be at the foundation of the solution, with the patient having ultimate control over their data and how it’s used”.

Halamka remains cautious however. He notes that technical challenges pose obstacles to the adoption of blockchain initiatives in healthcare. “It’s slow, it’s awkward to use, the number of steps required to get and put data to blockchain are numerous and complex”. There is hope though. “There are emerging “blockchain-as-a-service” products that attempt to solve these problems, but they are very early” observes Halamka.

Matthews and Cenaj note that, in addition to these technical challenges, there are significant cultural obstacles that stand in the way of blockchain adoption as well. “Regulation, policy, and legacy practices hinder the US from assuming a leadership role. Shareholder value does not equal patient value. It may take 10-15 years unless policy changes are implemented rapidly”, comments Cenaj. Matthews remarks, “Until we have a policy change at the highest levels of government, I don’t think that blockchain will be more than a point-solution for data security. I am convinced however that when we do finally have clarity on who owns patient data, transformation in personalized medicine could happen fast”.

In spite of his pragmatism and caution, Halamka is optimistic about the future of blockchain in healthcare. “There are production applications in healthcare using blockchain now, and they will become more commonplace over the next year.

Like any innovation, we'll go through a hype phase, a disappointment phase and eventually achieve broad adoption. Expect three years before there is universal adoption of blockchain related products". If Halamka is correct, we could see blockchain driving healthcare transformation sooner than expected.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

How Fannie Mae is Creating a Modern Data Environment

August 30, 2018

By Randy Bean

It seems like it was only yesterday that a new breed of data-driven businesses arrived on the scene. In the interceding years, a number of these firms – Amazon, Apple, Google, Facebook – have gone on to achieve extraordinary market capitalizations, while challenging incumbent businesses whose market dominance has lasted generations. The impact of this disruption is just beginning to be fully realized.

According to NewVantage Partners' 2018 Executive Survey, published earlier this year, an overwhelming majority (79.4%) of Fortune 1000 executives who were surveyed reported that they fear disruption from agile, data-driven competitors. In addition, 54.4% of these executives stated that the inability to be nimble and compete on data and analytics represented the biggest competitive threat facing their businesses today. It is evident that Fortune 1000 companies have come to the realization that to compete in the 21st Century they must adapt to a new order where data-driven competitors with highly agile, data-savvy business cultures are seizing the competitive initiative.

Fannie Mae's Modern Data Initiative

Scott Richardson is Chief Data Officer of Fannie Mae, which is formally known as The Federal National Mortgage Association (FNMA). Founded in 1938 as part of the New Deal, the mission of Fannie Mae has been to provide liquidity, stability, and affordability to the U.S. housing market. Like most longstanding corporations, Fannie Mae has legacy environments characterized by data silos. This circumstance makes it challenging and expensive to access high quality data in a timely manner. In a world of increasingly data-driven competitors, firms like Fannie Mae face the challenge of making the transition to agile, more flexible and responsive data environments or running the risk of competitive disadvantage.

Hired by Henry Cason, Fannie Mae's visionary first Chief Data Officer, Richardson joined their recently formed Enterprise Data organization in 2014 after spending over a dozen years at Capital One. Unlike its traditional banking competitors, Capital One forged a name for itself since its founding in 1988 by employing advanced data and analytics techniques to revolutionize the credit card industry and demonstrate how a leading financial services firm could compete on data and analytics. Data and analytics had been in the founding DNA and lifeblood of Capital One from its inception. Richardson, who had been instrumental in creating Capital One's first enterprise data warehouse, hoped to bring that same data-driven mindset to transform Fannie Mae into becoming a more data-driven competitor.

Along with his colleague Kevin Bates, a Fannie Mae veteran and neuroscientist by training, and a highly capable data governance, development, and operations team, they sought to undertake a wholesale data transformation to create what they describe as a "modern data infrastructure" supported by "industrial scale data". To achieve this ambitious objective, Fannie Mae needed to build a centralized data operation. And to realize such an ambitious undertaking, they asked for and received CEO and Management Committee sponsorship. The goal of this effort would be to establish a centralized data function where

they could manage data to a high-level of quality, consistency, and timeliness across the enterprise. “The right data, with the right quality, at the right time”, as Richardson puts it.

Reducing the Cycle Time to Data Analytics

Borrowing from the Development Operations (DevOps) movement, which was created to ensure automation and monitoring that shortens development cycles and deployment frequencies, Richardson and Bates are following a DataOps methodology to reduce the cycle time for the firm’s data analytics, from data preparation through reporting. From a process and methodology perspective, DataOps applies Agile software development, DevOps, and the statistical process control used in lean manufacturing, to data analytics. As they describe it, Fannie Mae’s new data platform is now “managing data with the same level of criticality as developing applications”. Foundational to their approach is a common data model, an Agile methodology aligned to a single governing model, and modular architecture designed to “automate everything that you can”.

One step Fannie Mae has taken to achieve their DataOps capability has been to deploy a dynamic data platform that accelerates data flow through the automation of data movement and self-service. By accelerating the testing of a major common securitization platform environment, Fannie Mae could refresh and restore its test data environment in a matter of minutes with an immediate reduction of 28 days of cycle processing. This drove significantly faster testing cycles for a very complex technology and process transformation effort, and has enabled Fannie Mae to support loan origination and servicing activities for its partners and customers with greater flexibility. Data flow solutions from firms such as Delphix have been deployed by Fannie Mae and a number of leading financial services firms as they seek to transform and virtualize their data development environment.

Business Adoption Key to Success

For firms that undertake ambitious data transformation initiatives, there is a common recognition that changing behaviors in a legacy culture requires time and patience. The process should be thought of as a journey. Companies that are successful focus on quick wins and demonstrate incremental progress. Richardson and Bates cite the value of “pushing value into production at least monthly”. They describe their effort as the culmination of a multi-year process. Noting initial wins in the early days of the transformation effort, they cite these wins as a critical factor in establishing positive momentum. One of the innovations that Fannie Mae instituted was to establish a community of Business Data Officers, to ensure that data is fully owned and cared for by business leaders, and that new initiatives consider the creation, ongoing quality, and effective usage of data from the outset.

Today, Fannie Mae has established a modern data environment characterized by near real-time updates, which results in a richer and more granular real-time customer experience. “Data is now viewed as a business asset” noted Richardson. He continues, “We are engaged in thinking about business strategy through the lens of furthering our mission and improving the customer experience with data”. Bates adds, Fannie Mae is now “encouraging experimentation at scale with data”, with the result that the company now has the means as well as an “appetite for reimagining new businesses and business processes”. Iteration and learning fast -- this is what it means to compete on data and analytics in the competitive Era of Big Data.

Forbes

The State of Machine Learning in Business Today

September 17, 2018

By Randy Bean

Artificial Intelligence (AI), Machine Learning, and Deep Learning are all topics of considerable interest in news articles and industry discussions these days. However, to the average person or to senior business executives and CEO's, it becomes increasingly difficult to parse out the technical differences which distinguish these capabilities. Business executives want to understand whether a technology or algorithmic approach is going to improve business, provide for better customer experience, and generate operational efficiencies such as speed, cost savings, and greater precision. Authors Barry Libert and Megan Beck have recently astutely observed that Machine Learning is a Moneyball Moment for Companies.

State of Machine Learning

I met last week with Ben Lorica, Chief Data Scientist at O'Reilly Media, and a co-host of the annual O'Reilly Strata Data and AI Conferences. O'Reilly recently published their latest study, The State of Machine Learning Adoption in the Enterprise. Noting that "machine learning has become more widely adopted by business", O'Reilly sought to understand the state of industry deployments on machine learning capabilities, finding that 49% of organizations reported they were exploring or "just looking" into deploying machine learning, while a slight majority of 51% claimed to be early adopters (36%) or sophisticated users (15%). Lorica went on to note that firms identified a range of issues that make deployment of machine learning capabilities an ongoing challenge. These issues included a lack of skilled people, and ongoing challenges with lack of access to data in a timely manner.

For executives seeking to drive business value, distinguishing between AI, machine learning, and deep learning presents a quandary, as these terms have become increasingly interchangeable in their usage. Lorica helped clarify the distinctions between machine learning (people teach the model), deep learning (a subset of machine learning characterized by layers of human-like "neural networks") and AI (learn from the environment). Or, as Bernard Marr aptly expressed it in his 2016 article What is the Difference Between Artificial Intelligence and Machine Learning, AI is "the broader concept of machines being able to carry out tasks in a way that we would consider smart", while machine learning is "a current application of AI based around the idea that we should really just be able to give machines access to data and let them learn for themselves". What these approaches have in common is that machine learning, deep learning, and AI have all benefited from the advent of Big Data and quantum computing power. Each of these approaches relies upon access to data and powerful computing capacity.

Automating Machine Learning

Early adopters of machine learning are findings ways to automate machine learning by embedding processes into operational business environments to drive business value. This is enabling more effective and precise learning and decision-making in real-time.

Firms like GEICO, through capabilities such as their GEICO Virtual Assistant, have made significant strides through the application of machine learning into production processes. Insurance companies, as an example, may implement machine learning to enable the offering of insurance products based on fresh customer information. The more data the machine learning model has access to, the more customized the proposed customer solution. In this example, an insurance product offer is not predefined. Rather, using machine learning algorithms, the underlying model is "scored" in real-time as the machine learning process gains access to fresh customer data and learns continuously in the process. When a firm employs automated machine learning, these models are then updated without human intervention since they are "constantly learning" based on the very latest data.

Real-Time Decision Making

For businesses today, growth in data volumes and sources -- sensor, speech, images, audio, video -- will continue to accelerate as data proliferates. As the volume and speed of data available through digital channels continues to outpace manual decision-making, machine learning can be used to automate ever-increasing streams of data and enable timely data-driven business decisions. Today, organizations can infuse machine learning into core business processes that are connected with the firm's data streams with the objective of improving their decision-making processes through real-time learning.

Businesses that are at the forefront in the application of machine learning are using approaches such as creating a "workbench" for data science innovation or providing a "governed path to production" which enables "data stream model consumption". Embedding machine learning into production processes will help ensure timely and more accurate digital decision-making. Organizations can accelerate the rollout of these platforms in ways that were not achievable in the past through techniques such as the Analytics Workbench and a Run-Time Decision Framework. These techniques provide data scientists with an environment that enables rapid innovation, and helps support increasing analytics workloads, while leveraging the benefits of distributed Big Data platforms and a growing ecosystem of advanced analytics technologies. A "run-time" decision framework provides an efficient path to automate into production machine learning models that have been developed by data scientists in an analytics workbench.

Driving Business Value

Leaders in machine learning have been deploying "run-time" decision frameworks for years. What is new today is that technologies have advanced to the point where machine learning capabilities can be deployed at scale with greater speed and efficiency. These advances are enabling a range of new data science capabilities including the acceptance of real-time decision requests from multiple channels while returning optimized decision results, processing of decision requests in real-time through the execution of business rules, scoring of predictive models and arbitrating among a scored decision set, scaling to support thousands of requests per second, and processing responses from channels that are fed back into models for model recalibration. Firms are deploying run-time decision frameworks with embedded and scalable machine learning, and are achieving notable results:

- Establishment of a demand funnel of 90 use cases for a next-generation advanced analytics platform
- Reduction of run time for analytics by more than 90%
- Support for millions of decision requests per day
- 46% more leads measured versus a control group
- $100M+ incremental annual revenue measured versus a control group.

This may sound like data science jargon, but firms are experiencing quantifiable business results. For organizations seeking to compete on data, machine learning has reached the stage of providing a critical business edge.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Every Company Is a Data Company

September 26, 2018

By Jedediah and Randy Bean

In July 2018, Facebook's market value suffered the biggest one-day fall in US stock market history—$120 billion. After alleged election tampering by the Russians and the Cambridge Analytica scandal, Facebook's costs to manage data security and privacy have skyrocketed, negatively impacting earnings.

To put the loss in perspective, that's more than the market cap for GE, UPS, or American Express.

Many companies interpreted the news as a cautionary tale. But that's a misread of the bigger story, which has played out over the last decade and will likely predict the rise and fall of industry leaders over the next several years:

Today, every company is a data company.

Whether they know it or not.

A decade ago, Internet companies began to fulfill the promise of the 90s dot-com boom. Companies such as Apple, Amazon, Alphabet (Google), and Facebook began an unchecked ascent, taking full advantage of the Internet's frictionless reach to rewrite the rules of nearly every major industry.

They also crowned themselves the most valuable companies in the world.

In 2011, Marc Andreessen wrote that every company must be a software company to survive. Many companies spent billions on software, a binary arms race. Yet the data paints a bleak picture of the results. Over the last decade, nearly three-quarters of Fortune 1,000 companies have been replaced—despite aggressive investments in software.

The *real* story of the Facebook decline is how the social network amassed value so quickly it could lose $120 billion and *still* be one of the ten most valuable companies in the world. It's a story of the power of data.

Facebook is a data company.

NewVantage Partners recently released the results of their annual Big Data Executive Survey. The survey featured responses from C-level executives representing the largest bellwether blue-chip firms—American Express, Bank of America, Bloomberg, Capital One, Charles Schwab, Farmers Insurance, Fidelity Investments, Ford Motors, Goldman Sachs, MetLife, and Verizon, among others.

The findings underscore the increasing urgency around data. 79.4% of executives report fear of disruption by data-driven competitors.

Over the last two decades, we have both participated in the fast-evolving world of data management. At NewVantage Partners, Randy Bean has advised and helped companies to leverage data as an asset and to become data-driven.

Jedidiah Yueh has invented data products that have generated over $4 billion in revenue as founding CEO of Avamar (acquired by EMC) and the founding CEO of Delphix. At Delphix, he has been privileged to work on innovative programs for many of the most advanced companies in Silicon Valley and around the world—a chance to peek under both tents.

And the differences between big tech companies and mainstream legacy companies are proving to be pivotal for the global economy.

We've both learned a great deal from our customers. The very first customer to ever put Delphix into production was the former CIO of Facebook, Tim Campos. He showed how it was possible to increase project velocity by accelerating data flow between systems. He also showed how to live by the Facebook ethos emblazoned on the company's walls: "Move Fast and Break Things."

And, wow, did they ever move fast and break things.

Under a darkening regulatory cloud, Facebook has begun to mature, but only after their data practices catapulted them to more than half a trillion in market cap.

Even in IT, they use data to innovate, creating Facebook CRM (customer relationship management) and Audience Insights—products that helped them scale from a billion to over $40 billion in revenues.

Legacy companies are burdened by regulatory constraints and decades old systems. They constantly struggle with what products to build and if they have the right digital transformation strategy. They scrutinize programs and budgets with an accountant's attention to details and risk.

But at Facebook, Campos quickly learned that "data wins arguments." They built their products *first* and then proved the value with results. They have adopted the fail fast, learn faster mantra.

Herein lies a difference between data companies and many traditional companies.

Data companies aggressively leverage data as core assets. They drive continuous returns by purposely instrumenting their companies to collect data and then experiment to develop value.

Legacy enterprises, on the other hand, often treat their data as liabilities. They obsess over risk, costs, and management—especially the ones in heavily regulated industries. Too often, they want to know the results *a priori*—before they invest.

Take financial services firms, for example. Too many collect data as a *byproduct* of their transactions—not the core product that makes up their business and differential value in the world. They respond to new regulations and news of security breaches by adding ever more belts and suspenders (that continue to fail)—aggressively conservative.

We may be standing at the brink of Ray Kurzweil's singularity, a moment when artificial superintelligence changes the trajectory of human history forever. As we accelerate into that future, tech companies in Silicon Valley have embraced a data gold rush.

Yet too many legacy companies are still optimizing for pennies when facing a potentially existential crisis.

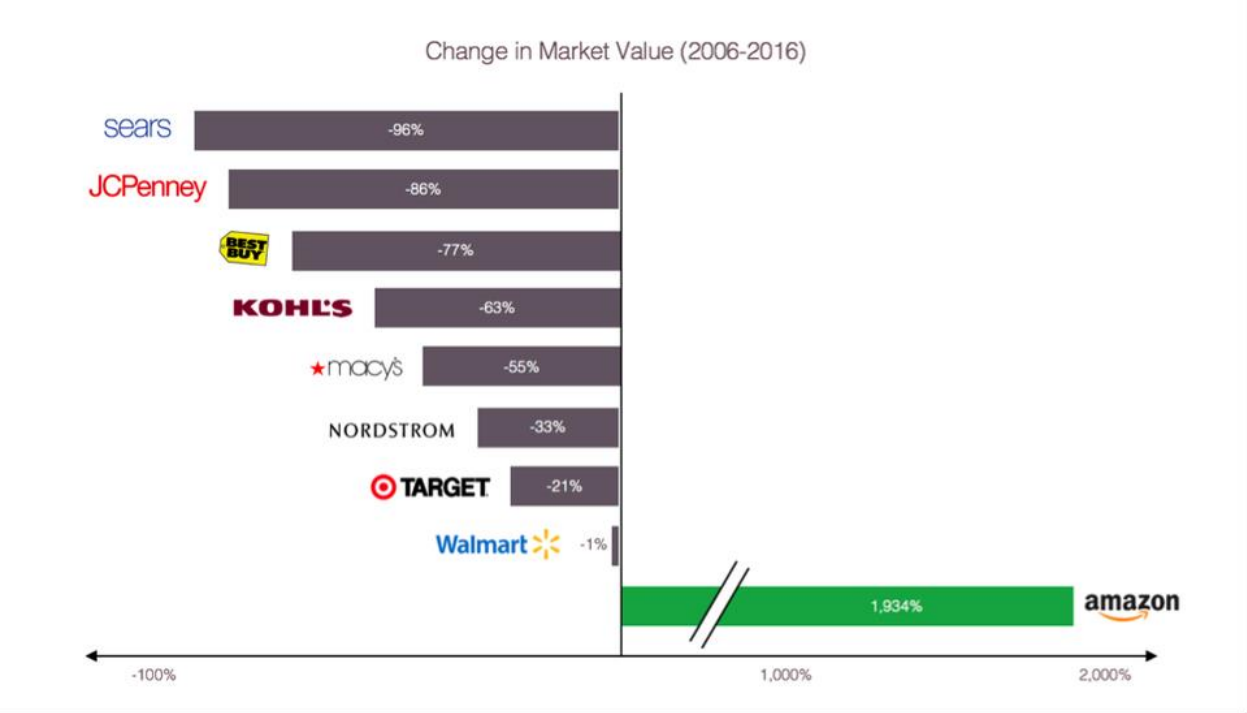
There's an AI maturity model at work in the world today.

At level one, companies run AI programs that drive operational efficiency. These are the "dabblers," companies that drive tens of billions in revenues a year and save a couple million using AI to automate tasks previously done by employees.

At level two, companies run AI programs to drive significant earnings or revenue impact. These are the "practitioners," like Farmers Insurance and PayPal. They layer machine learning through their businesses and use it to transform user experience and customer value.

At level three, companies run AI programs that drive industry change and transformation. This is the domain of big tech—the "experts." Facebook determines what we see in our feeds with AI. Apple uses AI and AI chips to power marquee iPhone features like Face ID and Siri. Google, Amazon, and Microsoft use a range of AI services such as machine learning, deep learning, and image recognition to run their businesses and also sell them as *products* in their clouds to arm the rest of the world.

In a world where every company is a data company, the fast eat the slow. Level three AI companies, like Amazon, are modern predators, consuming market value by feeding on legacy prey.



Change in market value.| Jedidiah Yueh

The timid too often obsess over data liabilities and risk bankrupting their companies, while the bold—who pan for gold and strike it data rich—work aggressively to inherit the industries of the future.

You can't focus on defense and expect to win the future. Leveraging data—in an ethical manner—has to be at the heart of your company and product strategy. Collect, sift, then monetize.

Because in the end, data companies will reinvent the world.

Jedidiah Yueh is the bestselling author of Disrupt or Die. Prior to his book, he led two waves of disruption in data management, first as founding CEO of Avamar (acquired by EMC in 2006 for $165 million), which pioneered data de-duplication. After Avamar, he founded Delphix, which provides fast, secure data pipelines for over a third of the Fortune 100 for critical programs such as leveraging AI services in the cloud. Yueh was designated a US Presidential Scholar, graduated Phi Beta Kappa from Harvard, and has over 30 patents in data management. You can contact him at jed@delphix.com and follow him @JedidiahYueh.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

A Rising Crescendo Demands Data Ethics & Data Responsibility

October 29, 2018

By Randy Bean

Data ethics is not a subject that you would have expected to be a centerpiece of conversation among Chief Data Officers and senior business leaders in the recent past. However, times are changing. Just this past week, Apple CEO Tim Cook condemned what he called the "data-industrial complex". In calling for accelerated data protection regulation, Cook declared, "Our own information—from the everyday to the deeply personal—is being weaponized against us with military efficiency. Today, that trade has exploded into a data-industrial complex."

Cook's call comes among increasing attention to issues of data privacy and data protection, fueled in part by enactment earlier this year by the European Union of the General Data Protection Regulation (GDPR). Cook and others are calling for legislators in the United States to adopt similar federal privacy laws.

It is against this backdrop that data ethics has rapidly moved to the forefront of any meaningful discussion about data. A spate of recent articles -- Never Heard of Data Ethics? You Will Soon, It's Time to Talk About Data Ethics, Data Ethics: The New Competitive Advantage, Will Democracy Survive Big Data and Artificial Intelligence – underscore the increasing urgency and highlight the ethical considerations that organizations must address when managing data as an asset, and considering its impact on individual rights and privacy.

I recently convened two thought-leadership roundtables of Chief Data Officers and executives with responsibility for data initiatives within their organizations. The increased focus and concern for the ethical use of data is born out of widespread reaction to recent and highly publicized misuses of data that represent breaches of public trust -- whether this be unauthorized data sharing by social media platforms, reselling of customer information by businesses, or biased algorithms that reinforce social inequalities.

It is within this context that we are now witnessing increased corporate attention to data for good initiatives. Companies are increasingly recognizing and acknowledging that ethical action and doing well can be synonymous with doing good. A few corporations, notably Mastercard through their Center for Inclusive Growth, Bloomberg through Bloomberg Philanthropies and the Data for Good Exchange, and JP Morgan through the JP Morgan Institute have been among those corporations at the forefront of ethical data use for public good.

JoAnn Stonier is Chief Data Officer for Mastercard, and in this capacity works closely with Mastercard's Center for Inclusive Growth. She notes how the work of the Center has contributed to positive change in the overall culture of Mastercard. Ms. Stonier remarks, "The Center for Inclusive Growth has taken a completely new approach to philanthropy - using the best of Mastercard's assets for social impact". She observes that "doing well by doing good" is not a new concept. However, she observes that "Having a philanthropic hub at the heart of the company reminds all of us employees about the value and impact we can have in the communities we serve".

She credits the Mastercard executive team for creating the vision and passion which has permeated every corner of the company. “Every day, we think about how we can make a broader impact, from delivering aid in a digital way so tsunami victims can get needed food and equipment faster to helping transit run faster in developing markets”, says Stonier.

Increasingly, corporations are focusing on issues of data ethics, data privacy, and data philanthropy. An executive representing one of the nation’s largest insurance companies noted, “We are spending more hours on legal and ethical review of data than we are on data management”. He cited this as a positive and constructive development. A number of leading universities have established collaborations with major corporations to collaborate on ethical data initiatives. One example is Washington University in St. Louis, which recently hosted a Data for Good Conference, featuring a discussion by Mastercard executive Chris Merz on “Personalizing Experiences and Protecting Customers”. A number of organizations are undertaking initiatives to develop data codes of conduct. Investment firms like Nuveen, a division of TIAA, have launched responsible investing initiatives that rely on responsible use of data.

Organizations are also turning their attention to algorithmic integrity and its consequences for local communities. Virginia Eubanks, a professor of political science at the University of Albany, SUNY, recently published a study on the impact of data algorithms, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. Other recent articles, such as Algorithm Blues in the Economist, discuss the work of social organizations such as the Human Rights Data Analysis Group. Increasingly, community groups and corporations are thinking about how data can be used to effect positive social outcomes.

Those who are most closely studying the social consequences of data see a growing disparity between what they characterize as the “data haves and have nots”. Ms. Stonier of Mastercard observes, “One of the major issues of our time is information equality. Often, those most in need of data and insight are the least likely to have access”. She cites an example. “Communities trying to recover after a manufacturing plant closes or after a natural disaster strikes can benefit greatly from information insights that can help them rebuild or restructure or understand how to reinvest. The private sector is astute at data insights and analytics, but too few are pioneering data philanthropy”.

One organization that has been a pioneer in this effort is DataKind, whose mission is to “harness the power of data science in the service of humanity”. Through 18,000 volunteers, 6 worldwide chapters, and 300 projects, DataKind is working with corporations and communities to use data to achieve a positive social impact. Ms. Stonier notes that, “These skills need to be combined with the organizations on the ground who need the data to solve the problems and develop the interventions and do not always have the tools or expertise to make informed decisions or changes. We need to close that divide to help tackle society’s biggest problems together”.

Firms like Mastercard, Bloomberg, JP Morgan, and Nuveen are among a growing list of corporations that are taking steps to tackle ethical data use and socially responsible applications of data. Through community and corporate partnerships, individuals and organizations can begin to think about data in creative and positive ways that can move the needle on humanitarian challenges. The crescendo of voices that are calling for data ethics and greater data responsibility is growing. Data ethics and data responsibility are not just good citizenship. They are good business.

Forbes

Munich Re: How Data & AI Reduce Risk from Global Calamities

November 4, 2018

By Randy Bean

Technology is at its most impactful when it is applied to addressing big problems. Perhaps there are no bigger problems than the occurrence of calamities, whether in the form of natural disasters, epidemics, or other catastrophic events. It is in response to seismic events, that fast actions can ameliorate acute conditions and mitigate potentially greater disaster. Such was the case in the wake of the recent super Hurricane Florence which had a massive regional impact when it struck the continental U.S. and devastated the Carolinas with severe flooding and widespread damage.

It was in the wake Hurricane Florence that Munich Reinsurance Company, commonly known as Munich Re, stepped into the aftermath to help devastated homeowners and business owners get back on their feet, as it has in response to past calamities. The reinsurance business may seem like a mystery to those not familiar with the role that reinsurance firms play. Munich Re was founded in 1880 and is headquartered in Munich, Germany. Warren Buffett was the single largest shareholder of Munich Re for many years, and remains a significant shareholder today. As one of the largest reinsurance firms in the world, Munich Re provides insurance products and services to insurance companies to help mitigate the risk of high impact crises. These crises generate severe capital demands on insurers to cover payouts and property losses in a timely and responsive manner. By helping assure part of the risk, reinsurers like Munich Re can help primary insurers in times of crisis.

Munich Re Chief Data Officer Wolfgang Hauner expresses it aptly -- “For the average person, you only value insurance when you have a loss”. And, it is in the event of the biggest of losses where Munich Re rises to the occasion. It is then that technology and data innovation come to the forefront. Munich Re has been hard at work developing and enhancing its data engineering and data analytics capabilities in anticipation of calamitous events. These efforts have included the application of image classification algorithms which rely on Artificial Intelligence (AI) to assess the severity of damage, produce immediate and automated damage estimates, and accelerate badly needed damage payouts. Munich Re employs remote sensing devices to capture high resolution images of property damage that are fed into an AI engine which is supported by advanced Machine Learning algorithms. Munich Re is then able to help customers through calamities with a rapid response to damage assessment that transforms what had been a 100% manual process into an automated process that reduces costs, increases productivity and efficiency, and results in enhanced customer satisfaction and loyalty.

It should not come as a surprise that Munich Re has embarked on initiatives such as image classification using AI. Munich Re CDO Hauner describes the journey that Munich Re has undertaken to leverage data and analytics to enable business innovation. He describes this journey as an evolutionary process that has been formally underway for several years now. In 2015, Munich Re launched an Advanced Analytics team to create a central hub of data and analytics skills and capabilities. In early 2016, a Chief Data Officer function was established under the leadership of Hauner.

Today, the CDO office has primary responsibility for three principal activities:

1. **Data Engineering**, including management of the data lake, data governance, and ownership for the global data and analytics platform
2. **Analytics**, which is primarily focused on the development of sophisticated analytics, many of which are employed for one-time use, while others lead to analytical software products
3. **Artificial Intelligence**, which is where leading edge technologies are applied to problems of unstructured data such as image classification and text analytics.

Hauner describes a fourth area which is being developed to focus on Analytics Operations, which will be the process by which models are maintained through their life cycle. The role and responsibility of Analytics Operations will be to run and maintain these models in a highly operational fashion based out of Munich.

Being a global business, Munich Re recognizes the necessity of decentralization. To avoid business bottlenecks, Munich Re is undertaking efforts to provide greater regional proximity to its clients. As an example, Munich Re is engaging in parallel recruiting efforts for data science and analytics talent in regional markets, recognizing that it is easier to compete and hire scarce analytic talent on a decentralized basis, rather than hire many people within a single market.

Reinsurance is a complex and highly impactful business. Hauner describes the characteristics of Munich Re clients, and levels of servicing that are required to meet their needs. High-value, high-frequency customers benefit from Munich Re's AI image classification initiative because the need is great, the timeframe is immediate, and the magnitude is widespread. The result is that the recovery process for property owners shrinks from months to days. Munich Re also seeks to bring these same high-end AI capabilities to enable smaller insurers to realize the same positive customer impact without having to fund and develop these sophisticated data and AI technology capabilities on their own. Munich Re provides this functionality to insurers in a services model. Lastly, Munich Re supports low-value, high-frequency initiatives such as identifying and compensating customers for flight delays, through the detection of weather patterns that correlate to the risk and probability of flight delay. An example is the flight delay insurance developed by Munich Re and offered by a client of Munich Re in China, which is helping to insure almost 50 million air travels -- out of around 500 million air travels annually in China at an average premium of 20 Yuan (less than $3) per policy. As climate change accelerates, AI predictability of weather patterns and their risk impact will become increasingly critical.

Munich Re is proceeding on its data and innovation journey, but recognizes that further progress is required. As Hauner notes, "We are so far perhaps 40% of the way there". To manage the data lake, Munich Re has created a data ecosystem supported by software tools such as those provided by Alation. The Alation tools enable Munich Re to navigate their data lake so that analysts and data scientists can work collaboratively with their data. The tools provide a central place to search for and access data. Hauner observes, "This helps us break down organizational silos. We are able to think differently. Data sources provide value across the organization. We are able to understand who is using data, in addition to our own units".

Thinking differently! Indeed, it is not hard to appreciate how Munich Re is relying on data and AI to meet the needs of a changing world while also assuming the risks. Hauner sums up, "Catastrophic events tend to be cyclical. We and our customers are encountering more frequent and evolving risks. These risks

require great solutions. We look to leverage great technologies, while we focus on what we do best – the management of risk taking and providing policy solutions".

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

Allstate's Data-Driven Business Transformation Initiative

November 11, 2018

By Randy Bean

Like the majority of legacy firms today, Allstate Insurance, founded in 1931 as the insurance arm of Sears Roebucks, must find new ways to compete in an increasingly competitive and digital marketplace. Luckily, Allstate was spun out of Sears in 1993, and unlike Sears, has continued to grow and prosper. While Allstate is the 2nd largest provider of property and casualty insurance in the U.S. according to A.M. Best, Sears has become the poster child for how data-driven tech giants, in this case Amazon, can thoroughly disintermediate and destroy a once thriving and iconic company.

In stark contrast to Sears, Allstate has demonstrated a commitment to leveraging data as an enterprise asset that can be used to transform business processes and services. Under the leadership of Allstate's Chief Data and Analytics Officer Eric Huls, the firm has launched the ambitious "D3" organization. Within Allstate, D3 refers to the 3 cornerstones of the firm's data-driven analytics journey – Data, Discovery, and Decision Science. The 300 person D3 organization is organized around core horizontal activities – Data Strategy, Technology, and Analytics Center of Excellence – to support core vertical business activities – product, claims, marketing, agencies, amongst others.

"We've taken an aggressive approach to integrating data into business, which ultimately gives us an edge when it comes to innovative pricing solutions and knowing the why or what our customers need even before they know it," said Huls. "We continue to grow our data team and almost every day we're seeing a new or different way the information we already have can be beneficial to our customers."

The Analytics Center of Excellence function is led by Rick Bischoff, Director of Data Science, whose mission it is to deploy analytics to drive business transformation. Bischoff has undertaken a multi-front effort to help enable transformational analytical improvements. Through the establishment of an Analytics Center of Excellence and through university partnerships, Bischoff is working to educate the business community within Allstate on the power of analytics techniques and applications of Artificial Intelligence (AI) to critical business problems. One area of notable business innovation is in the application of video footage to enhance analytics prediction. Working in partnership with Northwestern University, Allstate is also employing AI techniques to better understand the history of policy holder interaction. By analyzing photographic images and mountains of text data, Allstate is able to detect additional signals that can predict policy renewal and result in a better customer experience.

Data and analytics are integrated into almost every operation within Allstate. From claims handling, to billing inquiries, agent interactions, and the technical help-desk that Allstate employees call when their desktop computer doesn't work, Allstate analyzes mounds of free-form text and voice data to derive insights that inform business decisions. Taking this a step further, Allstate is continuing to develop machine learning and natural language processing capabilities which use this data to deliver an enhanced customer experience.

Allstate's QuickFoto Claim, or QFC, initiative, allows customers to provide collision photographs without relying on claims adjusters to be physically at their location. To quote Allstate, "The QuickFoto Claim feature of the Allstate[SM] Mobile App helps you get paid quicker so you can get back to normal faster". This represents transformational business change, notes Bischoff. No travel is required and advanced analytics can eventually be applied to photographic images to help adjusters resolve insurance claims.

Allstate is also deploying advanced analytics to drive product and pricing decisions with industry-leading sophistication. The Analytics Center of Excellence operates as a shared service which offers best practices and helps business constituents accelerate the use of analytics within their lines of business. Through the deployment of "analytics sandboxes", Allstate can explore the art of the possible, while sourcing new data and expanding data collection capabilities.

Allstate's D3 organization relies on a robust data ecosystem. One of the solutions that the Decision Sciences team has evaluated as part of their analytics ecosystem is Domino Data Lab's data science platform for empowering model-driven organizations. Domino has built a model building facility for data scientists that presents an easy way for data scientists to build and govern their own data environment. The Domino platform delivers productivity gains by enabling data scientists to run multiple models in parallel. The model management platform enables Allstate's newly recruited data scientists to hit the ground running, resulting in improved efficiency.

In his role as head of data science for Allstate, Bischoff sees the D3 organization helping Allstate quantify the business benefits and prioritize data activities that deliver the greatest business value to the firm. Bischoff refers to this as "transformational analytical improvement". His goal is to demonstrate the power that analytics can deliver to transform business initiatives across Allstate. In the face of competitive threats of encroachment from tech giants such as Amazon, and FinTech solutions such as Lemonade, Allstate is committed to building out rich data and analytics capabilities within a nimble, highly agile environment.

Allstate is proving that an 87-year-old insurance company can adapt and compete on data and analytics. Where its one-time parent Sears missed the window of opportunity, Allstate recognizes that complacency is the enemy of innovation and customer service. To quote Allstate," The Good Hands Are Doing More Than Ever Before".

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

UBS Asset Management Taps Alternative Data to Increase Alpha

November 18, 2018

By Randy Bean

Asset management firms are in the business of managing money on behalf of institutional clients -- typically pension funds and mutual funds. They operate as investment advisors that purchase and sell financial instruments, with a fiduciary responsibility to generate the best possible returns for their clients. To accomplish this, asset managers seek to "increase alpha", meaning generating results that outperform market indices. In today's highly competitive investments markets, asset managers are increasingly turning to so-called "alternative data" sources to deliver superior returns to the institutions that entrust them with their financial assets.

UBS was founded in 1862 in Switzerland. The firm has operated under the UBS brand since the 1998 consolidation of Union Bank of Switzerland and Swiss Bank Corporation. UBS, which is headquartered in Zurich and Basel, and is publicly traded on the Swiss and New York Stock Exchanges is one of the largest global asset management firms. Thomas Heinzl is Chief Operating Officer for UBS Asset Management, and Suvrat Bansal is Chief Data Officer and Head of Innovation. I spoke with Messrs. Heinzl and Bansal about how they are leveraging Big Data and Alternative Data to increase alpha for their clients.

Heinzl and Bansal are operating in a new data-rich investment world, characterized by massive availability of data in traditional and new varieties and formats. Massive availability dictates different ways and methodologies for managing and analyzing data. UBS Asset Management seeks to identify market and investment signals that are contextually relevant, and which will have a material impact on the performance of investments in a portfolio. The focus is on long-term performance, and the sustainability of these asset's performance.

The new sources that constitute alternative data are typically non-financial data elements that can be used to gain better insights to assess the future price performance of invested assets. As an example, UBS is leveraging card payment information to monitor sales data against earnings estimates and potential share price impacts. These alternative data signals help asset managers minimize risk while ensuring the delivery of superior investment performance on behalf of their clients. For public pension funds, minimizing risk, and understanding a firm's diversity and sustainability are central tenets of socially responsible investing.

UBS Asset Management is engaged in capturing alternative data to expand the volume and variety of net new data sources that can be fed into advanced analytics models that are leveraging machine learning capabilities. Some asset managers use geospatial information and satellite imagery; satellite imagery can be used to count the number of cars in store parking lots as a metric for retail sales activity. Geospatial analysis can be used to identify the geographical proximity of competitors, or in which neighborhood new stores have been opened. These beneficial indicators can help paint a picture of the health and prognosis for a business over time.

UBS is not alone in turning to analytics and alternative data to achieve sustainable value creation and increase alpha. A recent study was conducted by the New York headquartered data and analytics consultancy Element-22, with sponsorship from UBS Asset Management. The study featured the participation of 20 asset management firms with a combined $14 trillion in assets under management. Study results reaffirmed the perception that leading asset management firms are turning to advanced analytics and alternative data sources to fuel investment performance, grow their client base, and improve operating efficiency. Among key findings of the study:

- Alternative data usage is exploding, with sources including transaction data, satellite imagery, weather analysis, sentiment analysis, geolocation, and video capture. AlternativeData.org projects spending on Alternative Data sources to exceed $1.7 billion by 2020.
- Asset management firms are increasingly looking to advanced analytics and alternative data to gain an investment advantage. The study reported that 85% of firms were employing advanced analytics and 55% were using Alternative Data.
- Advanced analytics capabilities such as machine learning are being used to reduce costs and increase efficiency. Approximately 75% of firms reported using advanced analytics to aid in business operations.
- Comprehensive data capabilities are being used to manage new regulations such as GDPR, fueling investments in foundational data capabilities. Firms like UBS operate under strict data protection policies, which ensure that data is certified and compliant with local laws.

The investment industry is following suit. Earlier this month, The Wall Street Journal featured a story Alternative Data is Valued on Wall Street. The story references the firm Thasos, co-founded by MIT Professor and computer scientist Sandy Pentland, noting that "Thasos is at the vanguard of companies trying to help traders get ahead of stock moves using so-called alternative data". According to the story, Thasos will set offer its data through Bloomberg terminals.

Heinzl and Bansal of UBS envision a rich new world of informed investment, fueled by data, with machine learning playing a central role. "We are starting to test, or apply machine learning to fund flows, product innovation, alpha generation, and risk and middle/back office operations", they note. "Machine learning enables us to identify opportunities and make investment decisions faster". Big Data and Alternative Data will be at the core of machine learning algorithms as asset managers seek to identify and manage risk. For Heinzl and Bansal, faster data access, richer data sources, and more robust data analytics are translating into the delivery of superior client performance. This is the front line of innovation in the asset management world – applying alternative data sources as the next frontier of a data-driven world.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.

Forbes

A 2019 Forecast for Data-Driven Business: From AI to Ethics

December 17, 2018

By Thomas H. Davenport and Randy Bean

It should come as no surprise that 2018 continued to mark another year in the progression of data adoption in business. Companies are pushing forward with efforts to become increasingly data-driven. Firms are investing in transformation initiatives to establish a "data culture" within their organizations. Early adopters are focused on data-driven business innovation.

As we look ahead to 2019, we reflect on a year of accomplishments and emerging areas of focus – from AI through Ethics (listed alphabetically):

- **AI/Machine Learning**—AI continued to grow in popularity over the past year, becoming well-institutionalized within many large enterprises. We argued in a previous post, however, that too many companies employed AI pilots and prototypes, and not enough firms had implemented production deployments. As with analytics, the use of AI is increasingly being democratized through automated machine learning (AutoML). Several contributors to KD Nuggets' review of AI and ML trends for 2019 suggested that AutoML would become more popular over the next year. It will make machine learning models easier to create for business analyst types, as well as dramatically increasing the productivity of data scientists—that is, if they can be persuaded to use it. We also predict that deep learning, which has been the fastest-growing and most popular AI technology over the past several years, will continue to advance in power and prevalence for several years. However, we also expect that deep learning will increasingly be augmented by other approaches to AI. NYU professor Gary Marcus has argued, and we agree, that artificial general intelligence—or even generally useful AI—will have to employ various techniques beyond deep learning in order to be successful.

- **Automation**—One trend we noticed in 2018 is that there are a variety of automation technologies for organizations wishing to employ "digital labor" to perform structured work tasks. Robotic process automation, workflow, business rules, process mining, and some forms of AI all have the goal of automating human labor, or at least freeing up humans to do higher-level work. We see increasing numbers of companies embracing these technologies and determining how best to design work to maximize the respective capabilities of humans and machines. Given the proliferation of automation options, it's important to begin identifying, prioritizing, and categorizing automation use cases so that the right technology is used for each application.

- **Blockchain —** 2018 represented a year of major advancement for blockchain solutions as firms sought to ensure that data can be trusted, particularly when managing data in a distributed fashion. The need to ensure data trust received heightened attention in 2018 due to the adoption of the European General Data Protection Requirement (GDPR), resulting in greater focus on developing trusted frameworks for data sharing.

Healthcare has taken a lead in the adoption of blockchain capabilities, forging ahead with initiatives to ensure the protection of patient data and electronic medical records. Notably, leading healthcare providers such as Massachusetts General Hospital are evaluating initiatives to store patient data using blockchain. 2019 should be an active front for blockchain in healthcare.

- **Cloud Computing**—The cloud continued its march toward domination in 2018. Two Deloitte surveys, for example, indicated that 90% or more of global executives are adopting, considering, or already using the cloud. Amazon Web Services, Microsoft Azure, and Google Cloud are all growing rapidly. They are increasingly adding software and data management capabilities to their clouds, including enterprise data warehouses, advanced analytics, various forms of AI, Internet of Things, blockchain, and robotics applications. Some companies have converted their computing architecture to 100% cloud. The only question still to be resolved is what large organizations will do with all that well-cooled data center space they are vacating.

- **Cybersecurity**—The serious cybersecurity events of 2017—WannaCry and NotPetya—led to many attempts to emulate them in 2018. As data-related activity by good guys grows, data breaches, hacks, and ransomware from bad guys seems to grow even faster. The latest McAfee Lab's Threats Report suggests that malware exploiting software vulnerabilities grew by 151% in the second quarter of 2018. The volume of these attempts leads us to believe that the only way to address them is through the use of AI/machine learning for cyber-threat intelligence, detection, and resolution. At the moment, however, those technologies don't seem to reduce the burden on humans much. In fact, they generate way too many false positives. It will be interesting—and important for the world's data security—to see whether the good guys or the bad guys master AI first.

- **Data Analytics**—One of the most prominent and durable trends in analytics is the rise of their use by amateurs or "citizens." As Tom wrote in a co-authored post for the International Institute for Analytics, graphical and search-based interfaces to analytical programs are increasingly making it possible for those without analytical skills to find data and specify the analytics they need. This opens-up the possibility of data-driven decision-making to many more parts of organizations. This trend started several years ago and will, we believe, continue for many more. If there is data available on a topic within an organization, there will be no excuse for not using it. Analytics in many organizations was also augmented by artificial intelligence, and in many cases a single group supported both technologies. In its simplest forms, machine learning and predictive analytics are basically the same.

- **DataOps –** Data Operations (DataOps) is rapidly emerging as a discipline for organizations that continue to struggle with the management of data as a shared business asset. DataOps brings a set of data engineering principles which borrow from the DevOps software development movement. The intent is to deliver "rapid, comprehensive, and curated data" to business analysts and decision-makers. We expect 2019 to be a breakthrough year for DataOps approaches as firms strive to derive value quickly and efficiently from their data assets. We also believe that companies will increasing use machine learning to integrate and improve their data environments, as we described in a post about GlaxoSmithKline.

- **Ethics** – Lastly, but by no means to be forgotten, data ethics emerged in 2018 as one of the single most important priorities for leading businesses, stung by security breaches and highly publicized misuses of customer information that represented breaches of public trust. 2018 was in some ways the year that data received a black eye. Now organizations must rebuild that trust. 2019 can be expected to be a year in which corporations step up efforts to ensure ethical data use and ethical data practices. As we have noted, the demand for corporate data ethics and greater data responsibility is increasing. Data ethics is not just good citizenship – it is good business practice. We expect that more companies will add new roles and governance approaches to address this issue over the next year.

And, on this note, we wish everyone a prosperous and data rich 2019. May you progress on your data journey with excitement, accomplishment, and responsibility. Here's to a promising New Year.

Thomas H. Davenport is the President's Distinguished Professor in Management and Information Technology at Babson College, a research fellow at the MIT Initiative on the Digital Economy, and a senior adviser at Deloitte Analytics. Author of over a dozen management books, his latest is Only Humans Need Apply: Winners and Losers in the Age of Smart Machines. You can follow him at @tdav.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can follow him at @RandyBeanNVP.

Forbes

How Data Became Mainstream: To Chief Data Officer 4.0

December 26, 2018

By Randy Bean

Peter Serenita recalls a time when to work with data was considered a solitary pursuit that was shunted off to the backroom. This was long before the notion that working with data was considered the "sexiest job of the 21st Century". Serenita holds the distinction of being one of the very first executives appointed as a Chief Data Officer – a position that he was named to as Chief Data Officer, Worldwide Securities Services at JP Morgan in 2006. By the time that Tom Davenport and D.J. Patil heralded the era of the data superstar, Serenita had learned many valuable lessons which have continued to stand him in good stead as we move toward the era of the Chief Data Officer 4.0.

Serenita is now U.S. Chief Data Officer for Scotiabank, the 3rd largest bank in Canada, serving over 25 million customers. I caught up with Serenita at the recent 2018 Chief Data Officer Summit, organized by Evanta/Gartner Group, and held on November 26-27 in New York. This summit is one of a series of industry events that now confirm the prominence of the Chief Data Officer (CDO) as a senior executive decision-maker. This year's CDO Summit comprised over 150 data executives and CDO's, featuring a range of presentations and panel discussions on the state of data and the evolution of the CDO role within leading companies. One of the keynote presentations was delivered by Serenita on the history of data management and the march to managing data at scale. This was a history lesson that few have experienced as directly, and from its infancy, as has Serenita.

Organizations are at very different stages of maturity in their adoption of data management. This is something that Serenita understands well. While some organizations have yet to reach the stage where they have appointed a Chief Data Officer, the most ambitious organizations are moving to the stage of Chief Data Officer 4.0, characterized by using data as a source of innovation and for purposes of business monetization. Most organizations sit somewhere along this continuum. For data-native businesses like an Amazon or Google, the notion of appointing a Chief Data Officer seems inherently redundant – every business leader should have data in their DNA. The role of the Chief Data Officer is most needed for mainstream, legacy companies who are undertaking a vast business transformation that leads to becoming data-driven organizations.

Serenita describes the evolution of the Chief Data Officer role over the past dozen years, as each organization has embarked on its own individual implementation of the data management journey. He reflects on the early days of data management, where data was largely an Information Technology (IT) function. He notes that the early wave of CDO's (CDO 1.0) were focused on establishing data governance policies and practices relating to data quality, data lineage, remediation, and data ownership roles and responsibilities. Then came the 2008-2009 financial crisis, and with it the emergence of regulatory demands, standard data management practices in the form of the Data Management Capability Assessment Model (known as DCAM), and an emerging focus on the use of analytics (CDO 2.0). Serenita also notes that, "The data profession is very collaborative. Organizations can learn from one

another. Data executives can benefit from the experiences of others who have been in the data trenches. But the 4.0 data executive also brings a new way of thinking that will further advance the data practice".

In recent years, Serenita has seen a growing movement to integrate data and analytics responsibilities and functions. Whereas firms had appointed Chief Data Officers (CDO) and Chief Analytics Officers (CAO) with often parallel responsibilities, there has been a sharp movement toward the establishment of the updated role of the Chief Data and Analytics Officer (CDAO). This is CDO 3.0, which is characterized by a combined focus on data governance and analytics, coupled with a growing adoption of AI and machine learning practices, along with advanced analytics techniques.

The progressive adoption of Big Data technology solutions is enabling businesses to process and analyze whole data sets in their raw form, rather than relying upon individual data domains and tables. Data does not have to be refined in advance either. Machine learning capabilities can be employed to understand and transform data into its most usable state. Subject experts can teach the data or algorithmic model initially, and then rely on the machine to learn along the way. Serenita remarks, "We have gone from the Stone Age to space travel in the course of a decade".

So what lies ahead? Serenita envisions a future where data is now understood to be a business enabler and a revenue generator, and where the CDO reports to the head of business strategy (CDO 4.0). This is a far cry from the days when data was relegated to the back room. Yet, Serenita is cautious as well as realistic based on his many years of experience. In the data business, you need to be. "One size does not fit all", notes Serenita. "Factors such as stage of maturity, capability, organization, centralized versus federated management, risk, and business enablement must all be taken into consideration".

Serenita believes that to be successful a CDO must bring together within a single executive a diverse set of skills. He describes the CDO function as being a "multi-disciplinary approach" where CDO's should refocus on the "art of the possible". Serenita sums up, "data has become mainstream". He advises that organizations need to move rapidly to adopt agile data processes that enable the business to move quicker and be more responsive to internal and external customers, and notes that organizations need to leverage technology through AI and machine learning techniques. "Let the machine be the heavy lifters; let the data analyst guide the machine", Serenita concludes.

Randy Bean is an industry thought-leader and author, and CEO of NewVantage Partners, a strategic advisory and management consulting firm which he founded in 2001. He is a contributor to Forbes, Harvard Business Review, MIT Sloan Management Review, and The Wall Street Journal. You can contact him at rbean@newvantage.com and follow him at @RandyBeanNVP.