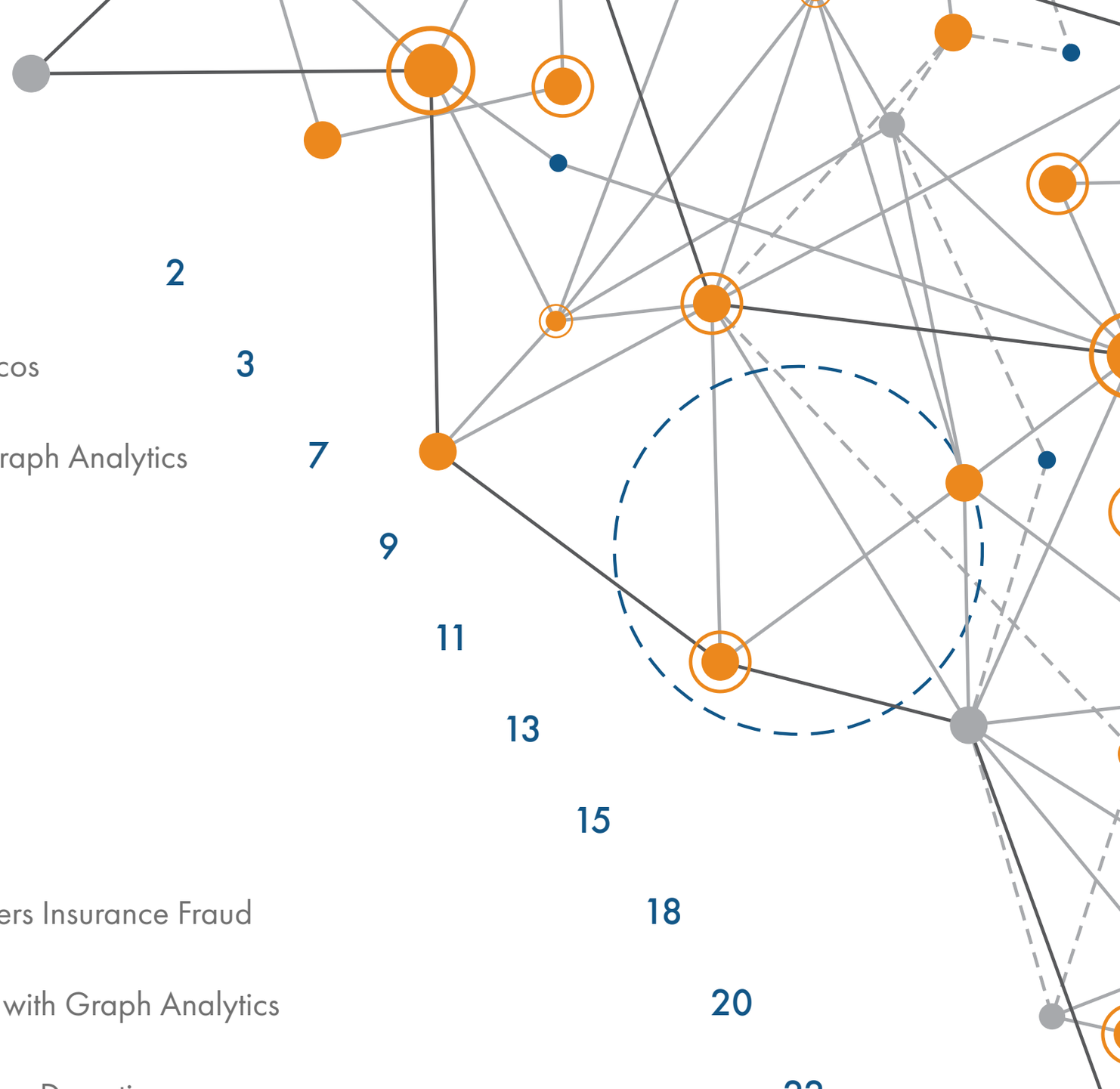
RELATIONSHIPS MATTER

# The Business Value of Connection Analytics

TERADATA.

# CONTENTS





# Introduction



Always on. Off the grid. Six degrees of separation.

It's all about relationships. That's why we call sites like LinkedIn, Facebook, and Twitter social media.

Social media is all about connections. But so is literally almost everything else: families, teams, computers, products, and even plants and animals.

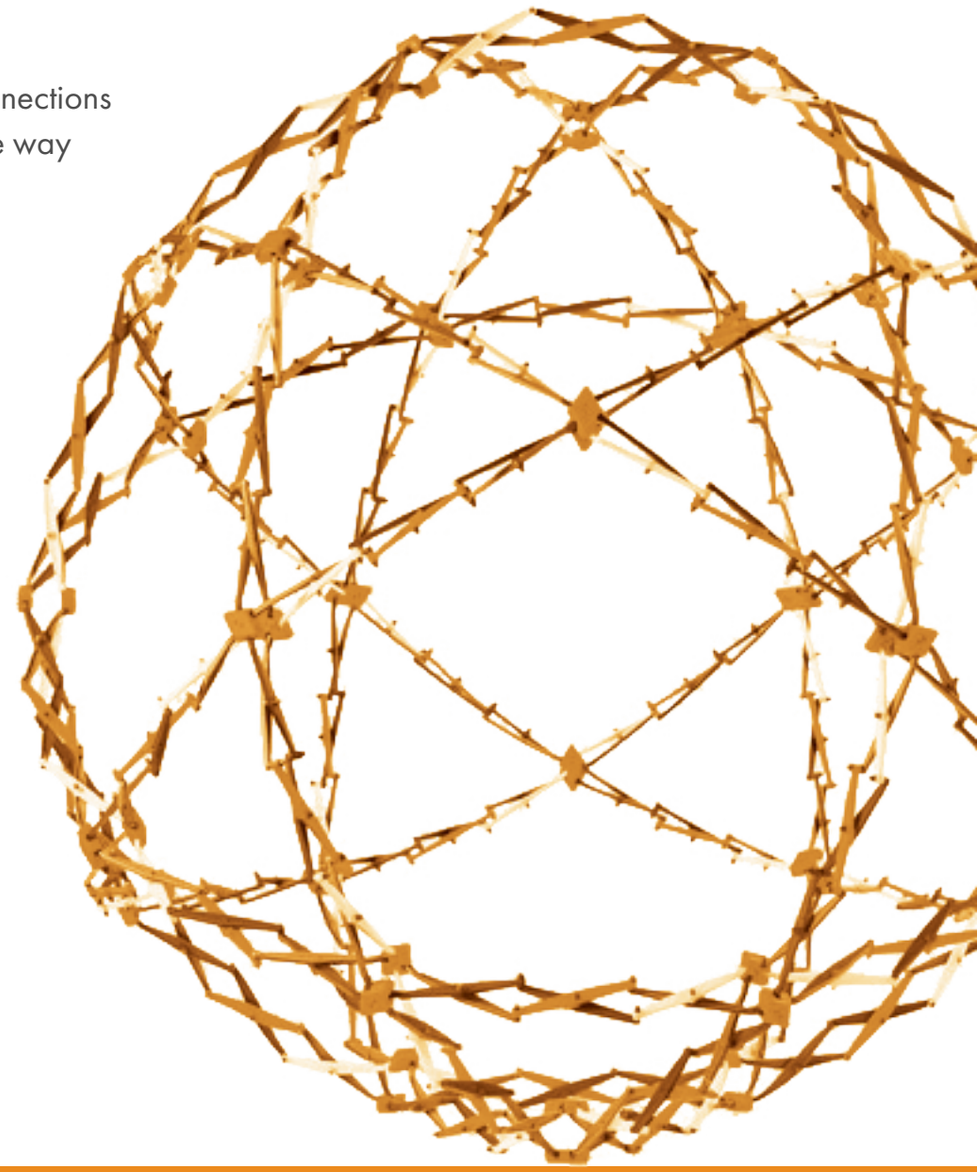
Now we're finding ways to explore those connections and see what they can tell us. And they can tell us a lot: about the phone calls we make, the people committing fraud, the products to offer together, and more.

Of course we've always had connections, but we just didn't have a good way to find the interesting connections that are lurking in our data. Now we do; it's called connection analytics and it helps us find patterns in the way people and things of all types are related.

In this ebook, we've collected some cool patterns for using connection analytics, from luring new business to preventing your network from going down just when you're working on that big proposal.

We hope you enjoy the buffet of ideas provided in this ebook and find inspiration in looking for connections in your own data.

Dan Woods
CTO and Founder, CITO Research



## About Dan Woods

Dan Woods is CTO and Founder of CITO Research. He has written more than 20 books about the strategic intersection of business and technology. Dan writes about data science, cloud computing, mobility, and IT management in articles, books, and blogs, as well as in his popular column on Forbes.com.

# Improve Profits Using Social Network Analysis in Telcos



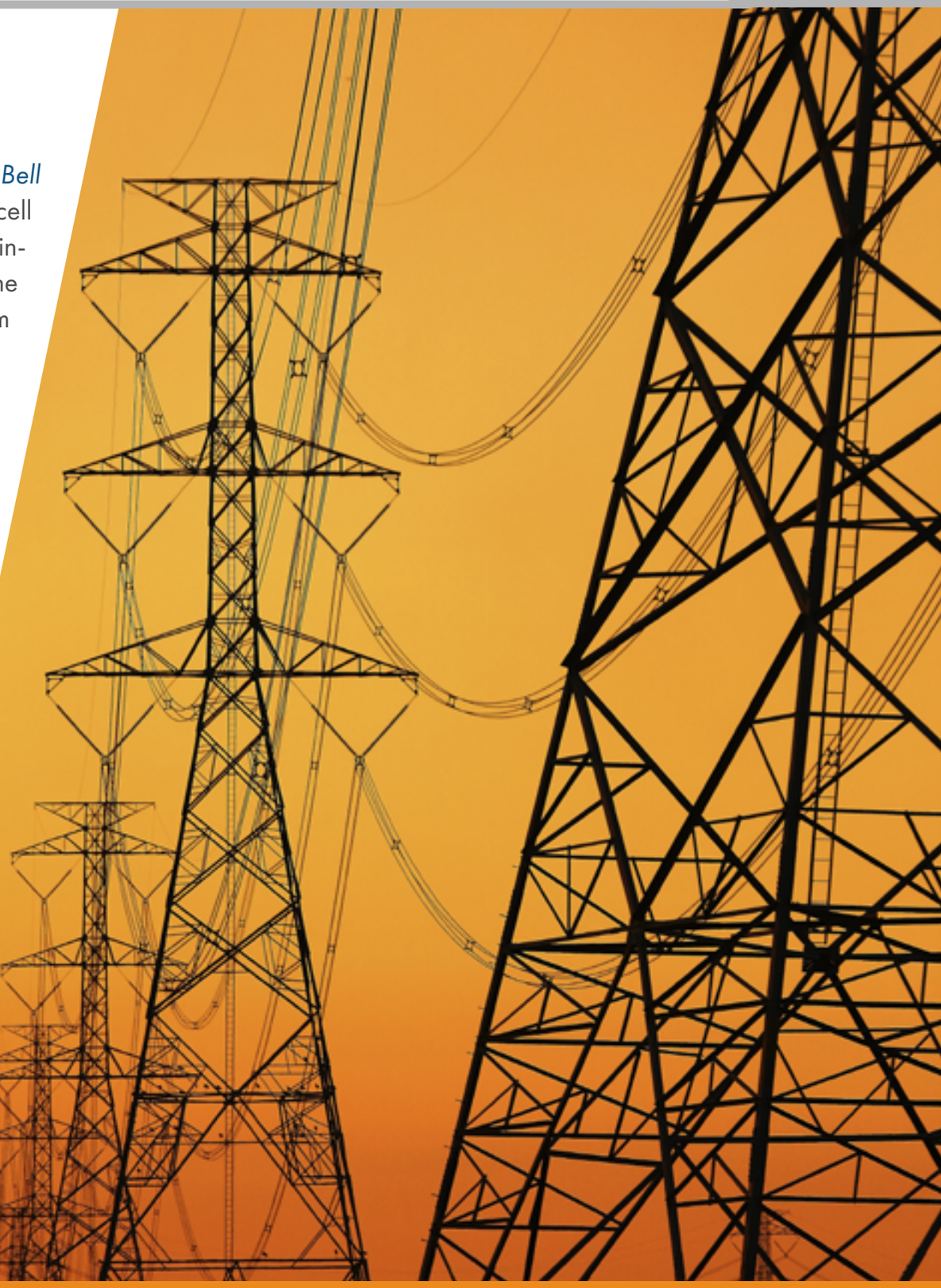
## The Problem of Customer Retention

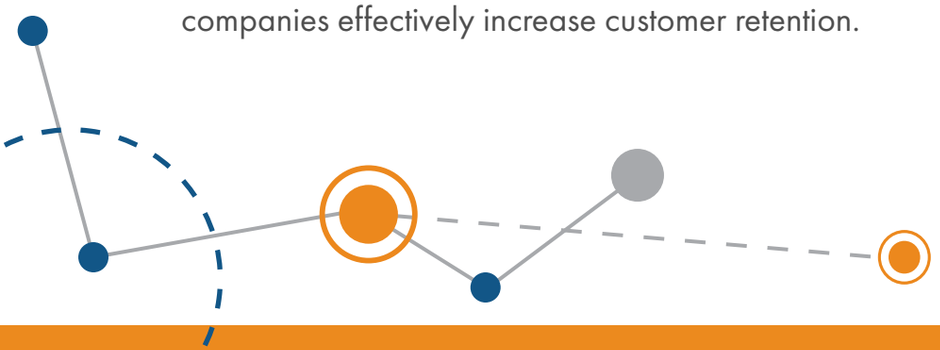
For children of the early 90s, the gigantic cell phone Zach Morris used on *Saved By the Bell* is iconic because it was so novel: at the time, only a select number of individuals had cell phones. Twenty years later, 95% of the world's population has a cell phone. We are all increasingly connected. The rising volume and number of ways we are connected created the field of **connection analytics**, which seeks to examine our connections and analyze them to understand the interdependencies between us.

For telecommunications companies (telcos), exploring the content and context of connections is both an opportunity and a challenge. One of the biggest problems facing businesses across industries today is customer retention: it costs far less to retain existing customers than to attract new ones. Telcos face a high rate of customer turnover, with consumers switching from one provider to another. In such a saturated market, losing customers can decimate a company because there are so few new customers to attract.

It's an even bigger problem if your customers leave and take their friends and family with them. When influential customers leave, customers connected to them are six times more likely to cut service as well.

In order to increase retention, telcos need analytics tools to identify customers likely to leave (especially key subscribers), understand why they're thinking of leaving, and intercede quickly with an offer to retain them. And in cases where an influential customer does leave, the company needs to plug the dam before the river drains. Traditional analytics methods have not been up to the job, but connection analytics techniques, one of which is graph analytics, can help companies effectively increase customer retention.





## Why hasn't the problem of telco customer retention been solved already?

The problem of telco customer retention has existed for years, but only now have tools become available that any company can utilize without specialized expertise. The two main reasons for this lag are the complexity and volume of telco data.

Telco customers are highly interdependent—we have phones so that we can call other people. We generate incredible amounts of complex data through our phone: who we call, how often, and when, in addition to our texts and browsing history. There is data generated through activities like customers calling the customer care center and making payments. And behind the scenes, there are a lot of operational systems kicking out data—network switches, servers, and firewalls. Without help, data warehouses and traditional analytics tools could not handle this data volume and variety. Memory and processing constraints, along with varying source types, made it impossible (or at least prohibitively expensive) to represent the complexity of the relationships in the data. Systems could not scale to pinpoint influential customers across a network of millions of subscribers, tens of millions of calls, and billions of interactions. Often, to get adequate performance, you needed to plan the schema to support a specific type of query. This meant you had to know what you were looking for before you looked for it. As a result, telcos struggled to identify influential customers and became increasingly skeptical about the ability of big data and various analytics tools to improve customer retention.

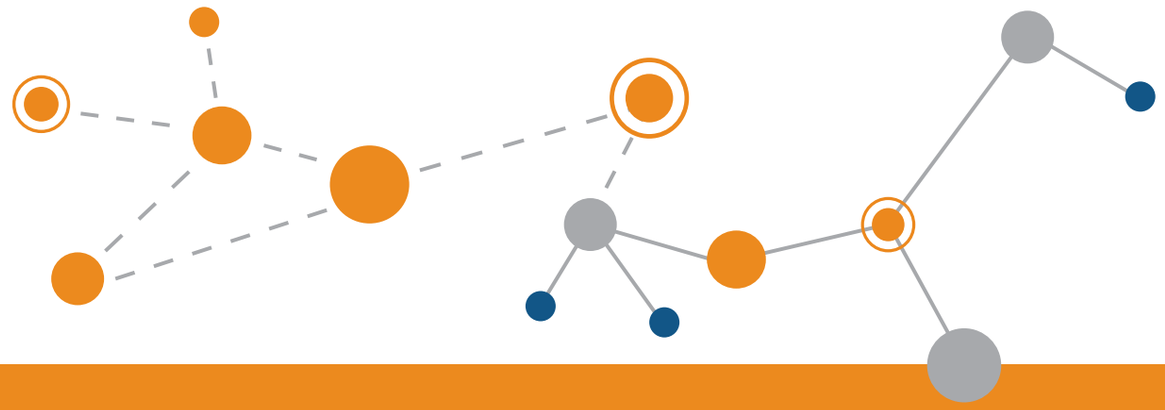
To improve retention, companies needed a solution that can synthesize this data tsunami, recognize key influencers and their connections as well as customer sentiment, and use data from all sources, all in near real time and in an easy to use interface. Enter graph analytics.

## How can graph analytics help telcos retain customers?

Graph analytics is predicated on the larger idea of connection analytics, in which graphs are used to identify the relationships between anything, from telco users to products often purchased together. Graph analytics can help telcos find ways to retain customers because the relationships being analyzed are naturally represented by a graph of nodes and edges. Each person is a node. Each call or other interaction is an edge. By tracking who is connected to whom and how often one person calls another, it is easier to spot influencers and their networks and gauge the satisfaction of the influencers as a first line of defense in retaining customers.

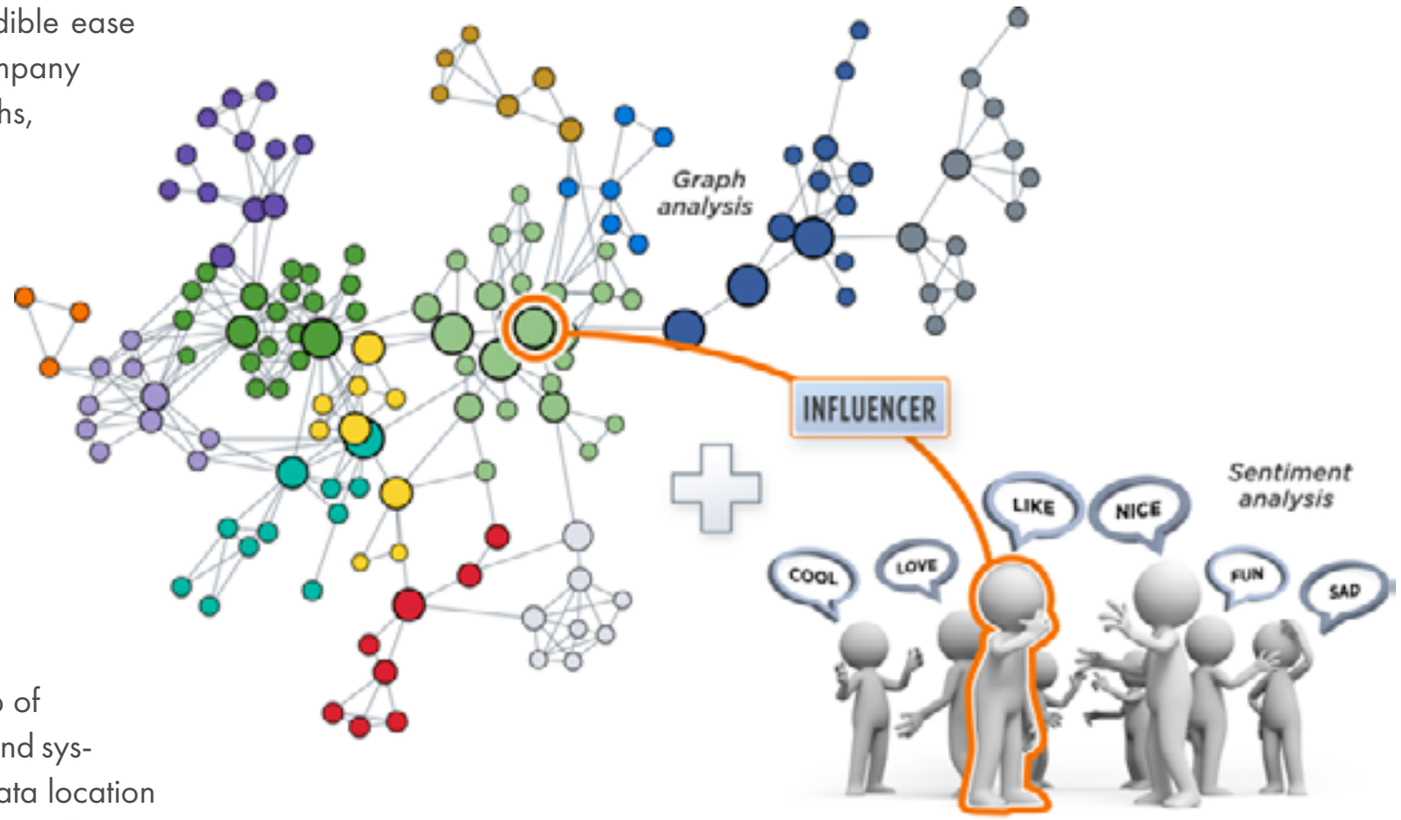
Graph analytics presents data in terms of relationship touch points, allowing telcos to find early indicators that customers may leave or influence others to leave, and enabling them to isolate behaviors that indicate customers are thinking of leaving and preventively reach out with an offer to build loyalty.

To perform graph analytics, you need a graph analytics engine. Graph analytics engines have no memory or processing limits, don't require specially designed schemas, and can handle the full spectrum of telco data, all with unparalleled speed. Graph analytics engines are built to absorb data from a wide variety of sources at tremendous volumes. In this way, graph analytics engines exceed the potential of graph databases, which assume the information being analyzed resides in the database in a particular format. Graph databases are also often bound by memory and processing constraints.



Another important difference between graph databases and graph analytics engines is that graph databases are good for finding the answer to a particular question (such as who is in a particular person's network of connections), whereas graph analytics engines can be used to explore patterns that answer particular questions (such as finding all the influential customers) and, if those customers were to leave, who connected to them is also likely to leave. This is a subtle but important distinction. Graph databases answer specific questions whereas graph analytics engines enable discovery by helping you figure out what questions to ask.

The ability of graph analytics engines to create a single graph from numerous sources provides analysts with the ability to ask questions that are not possible to answer in other ways. This flexibility leads to incredible ease of use, not just for data scientists, but for anyone in a company who knows how to use SQL. Unlike other ways of using graphs, Teradata® Aster Discovery Platform allows analysts to query graphs with a form of SQL that has built-in analytics functions for common graph queries and analytics that allow organizations to identify complex relationships for improving customer retention, promoting products, improving sales, reducing fraud, and solving other common business problems. Companies don't have to waste time re-creating these algorithms and can instead focus on extracting information from their data.




Graph analysis identifies influencers, bridges, and other roles. Companies can then target high sentiment influencers for viral marketing campaigns

The graph analytics capabilities in Teradata® Aster Discovery Platform support all standard interactive BI tools as well, from Tableau to Tibco. Essentially, the engine sits on top of and seamlessly integrates with a company's other data tools and systematizes the way users can probe this data, regardless of data location or format. Not only does this mean telcos don't have to replace their existing BI infrastructure, but it also means they can take advantage of the unique benefits of an ensemble of analytic tools.

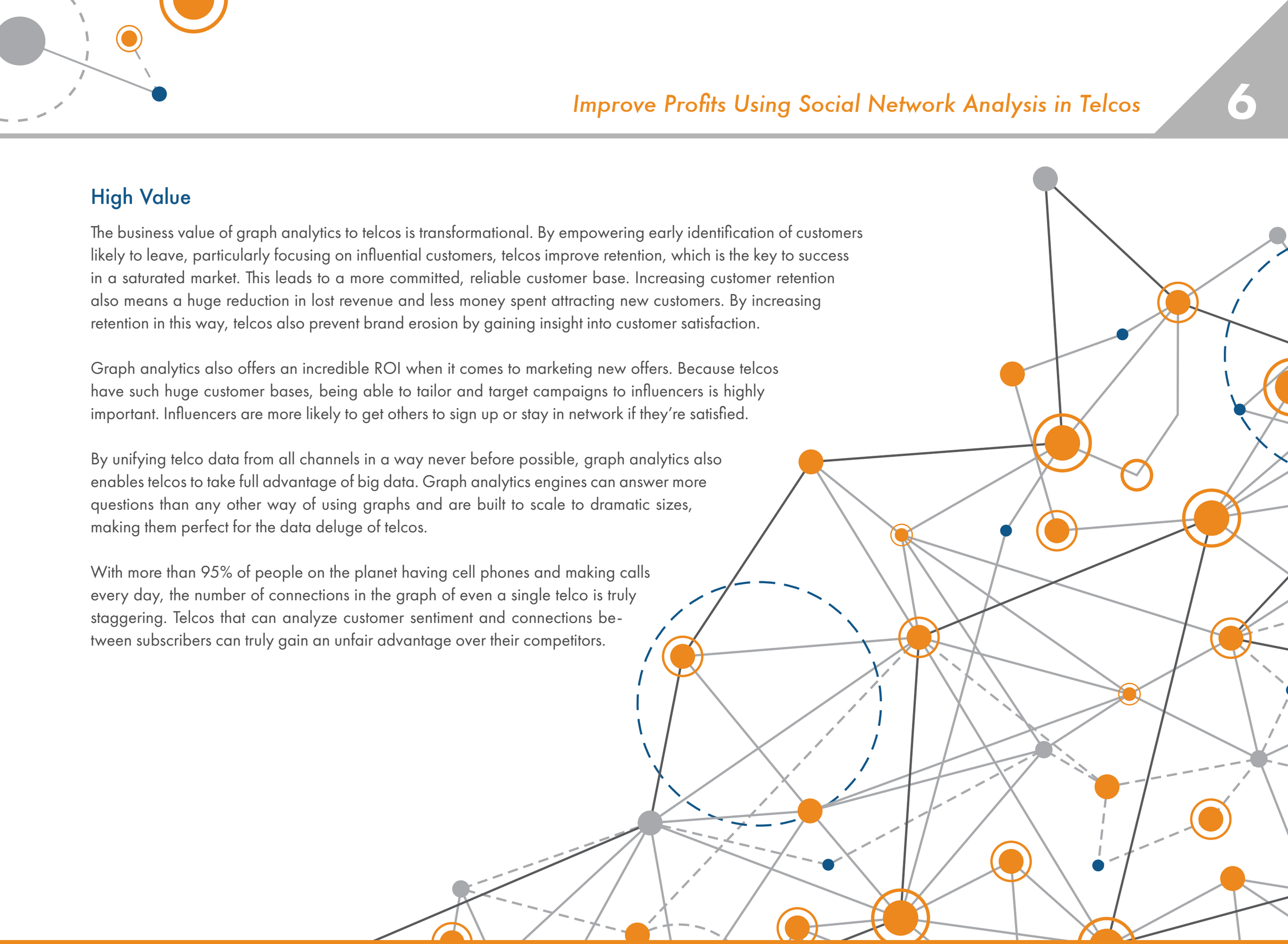
## High Value

The business value of graph analytics to telcos is transformational. By empowering early identification of customers likely to leave, particularly focusing on influential customers, telcos improve retention, which is the key to success in a saturated market. This leads to a more committed, reliable customer base. Increasing customer retention also means a huge reduction in lost revenue and less money spent attracting new customers. By increasing retention in this way, telcos also prevent brand erosion by gaining insight into customer satisfaction.

Graph analytics also offers an incredible ROI when it comes to marketing new offers. Because telcos have such huge customer bases, being able to tailor and target campaigns to influencers is highly important. Influencers are more likely to get others to sign up or stay in network if they're satisfied.

By unifying telco data from all channels in a way never before possible, graph analytics also enables telcos to take full advantage of big data. Graph analytics engines can answer more questions than any other way of using graphs and are built to scale to dramatic sizes, making them perfect for the data deluge of telcos.

With more than 95% of people on the planet having cell phones and making calls every day, the number of connections in the graph of even a single telco is truly staggering. Telcos that can analyze customer sentiment and connections between subscribers can truly gain an unfair advantage over their competitors.



# Would You Like Fries with That? Cross-Selling with Graph Analytics



## The Challenge: How To Find Customer Behavior Patterns

Less than $30. That's the amount of the average grocery-shopping trip. Given the price of groceries and the tight margins in the sector, that's not much per transaction.

The cross and upselling so common in fast food and high-touch marketing is much harder to carry out when you have about 60,000 items in your store, many of which have low margins.

What are the items that link low-margin purchases to high-margin purchases? To find this out, you'd need to analyze the items purchased in thousands of shopping trips, looking for a pattern that leads to particular kinds of high-margin items, like wines or meats.

## Knowing Customers Better Than They Know Themselves

In the past, companies have had to throw darts at the wall when attempting to inspire reliable low-margin customers to buy high-margin products. Today, shifting buying behavior is crucial to the bottom line. Especially in industries with tight margins, like grocery retailers, the difference between a customer who buys a bottle of wine with a 2% margin and a 5% one can be staggering. And in terms of cross-sells, companies have had to rely more on assumptions about their customers than on data.

With big data, we have vast amounts of information on buying habits. Armed with this information, what if companies could identify bridge products that inspire purchases in high-margin areas? A retailer could determine that given the right enticements, buyers will opt for a more expensive wine or a high-margin microbrew. Or for cross-sells, that same business might find that customers who buy a certain brand of salad dressing also tend to buy high-margin meat and wine when they purchase fancy cheese. With the right technology to harness the full power of big data, retailers can both earn more and achieve higher customer satisfaction.

## Using Graph Analytics to Find Patterns

But big data doesn't offer up this type of insight on its own. Traditional BI and analytic tools have been overwhelmed by big data, unable to provide 360° views integrating all data sources in one place and analyzing millions of daily transactions to uncover subtle behavioral patterns. These older tools simply don't have the processing, memory, and analytic chops to support the kind of analysis being discussed here.

To coax the profit drivers out of all the raw data, companies need analytic tools that can form coherent narratives. Graph analytics, a form of connection analytics, is one of the most powerful big data analytic tools now available: it shows the relationships between anything, from people to products. For retail, graph analytics presents products as nodes and transactions as edges and then shows these relationships in graphs.

By seeing these links, retailers can better identify bridge products and increase cross-sell and up-sell opportunities. The potential is staggering. Graph analytics run on a graph analytic engine. Graph analytic engines have no memory or processing limits, don't require specially designed schemas, and can handle the full spectrum of product and transaction data, all with unparalleled speed. In this way, graph analytic engines exceed the potential of graph databases, which assume the information being analyzed resides in the database in a particular format. Graph databases don't have the requisite analytic chops and are bound by memory and processing constraints.

Graph analytics engines need to be able to tap into all of a company's data sources. Many retailers have data spread across numerous databases and warehouses. A graph engine has to be able to unify all of these sources so that analysts know they are getting the complete picture. And, in many cases, one insight leads to another—analysts begin to probe data differently, asking questions they have never posed in the past.

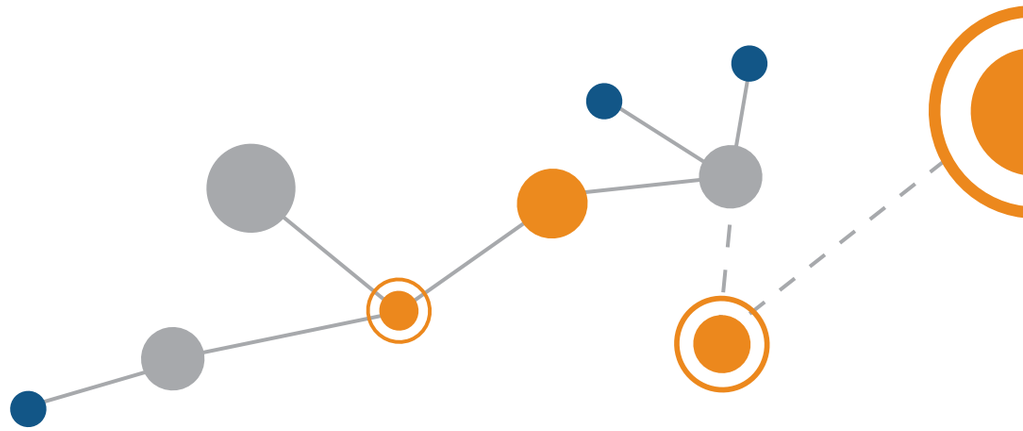
Another important difference between graph databases and graph analytic engines is that graph engines are good for finding patterns that answer particular questions, such as finding one or more bridge products between multiple clusters of low-margin and high-margin products. Graph databases answer specific questions (which products have been purchased with this particular SKU) whereas graph analytic engines enable discovery by helping you figure out what questions to ask (show me all transactions that include products with a margin of 4% or more).

Because of the sheer volume of product and transaction data and the potential for connection analysis to identify new cross-sell and upsell opportunities, graph analytics is ideal for retail. Using graph analytics, retailers can identify high-margin products that appeal to customers who already buy low-margin ones. But companies can also do exploratory research, conducting market basket analysis of sales with more high-margin products to see what results in increased sales and what does not. Just as importantly, graph analytics is the best tool for determining bridge products that connect two different types of high-margin products.

## Conclusion

Graph analytics is already changing the way retailers do business. At checkout, whether in-store or online, retailers can offer products they know the customer, based on analytic expertise, is likely to buy. This is the souped-up version of candy in the checkout line, except that, with graph analytics, retailers know exactly which brand of candy to offer that will return the highest margin.

But graph analytics has applications for retailers beyond targeted advertising and product placement. Because graph analytics reveals connections, retailers can also use it to improve the effectiveness of their social media outreach. They can trace the connections between customers and allow them to share their purchases, opinions, and wish lists with others in their networks, resulting in sales from other linked customers, as well as better data about buying patterns. Retailers won't have to guess where to put those new high-margin gluten-free items: they'll have data to help make their decision and revenue to show it worked.



# Retention: The Best Hiring Strategy



## The Vicious Cycle of Employee Turnover

It's no secret that collaborative, supportive relationships among employees are foundational for effective companies. When employees trust their coworkers and genuinely like them, it is easier to unite a workforce behind a corporate vision and have them work together to meet their targets.

But this type of camaraderie also has a flip side—when one employee leaves, other employees in that person's network are likely to leave, too. Those interpersonal relationships upon which companies base their success also can lead to mixed loyalties, with some employees more loyal to their coworkers than to their employer. What can result is a cascading effect in which an entire department is quickly filled with vacancies following the departure of one leader. This obviously leaves companies at risk of mass flight that can impact the integrity of the entire business. It's like the game Jenga—pull out too many blocks and everything collapses on itself.

On top of this, losing employees and hiring new ones is expensive. Experienced employees exit along with their institutional knowledge that is incredibly difficult to replace. The money and energy put into training is also lost. And finding replacements is perhaps the greatest cost of all, both in terms of time and money for retraining and onboarding. Numerous studies indicate that companies spend one-fifth of an employee's salary to replace a worker. As many HR directors will tell you, a company's best hiring strategy is effective retention.

Yet, there's a difference between recognizing retention as a problem and being able to do something about it. Businesses should not just accept employee turnover as a given or a factor that is out of their control. Now, more than ever before, the big data that companies accumulate about their staff can be a powerful component of their retention strategy. With the right analytics tools, companies can identify their most integral employees and proactively seek to keep them. When employees do leave, companies can identify other employees closely aligned with them and work to prevent a domino effect of departures.

When many companies think of big data analytics, they think only of what they can do for their customers, especially in terms of marketing, advertising, and sales. In truth though, big data analytics can be just as important for the internal operations of a business.

Google is famous for its use of big data to help gauge Googlers' happiness (referred to as GoogleGeist). Google's HR department does big data analytics in the service of improving life at Google, making changes, and quantifying the results. And it's working—Fortune has named Google the best place to work for five years running. One direct result of such analytics: offering longer maternity leave to retain female employees, along with flexibility in how they can take that leave.

## The Power of Graph Analytics

Good managers are well aware who their most vital employees are. These workers are not only adept at their jobs, but indispensable in terms of team cohesiveness. However, companies have never before had analytic tools that can accurately map this web of relationships and key players using hard data. That is, until now.

More and more frequently, companies are collecting data about their employees. This includes the frequency and target destinations of their emails, meeting requests, phone calls, and texts. Used in the right way, this information can be an ally in the fight to improve retention.

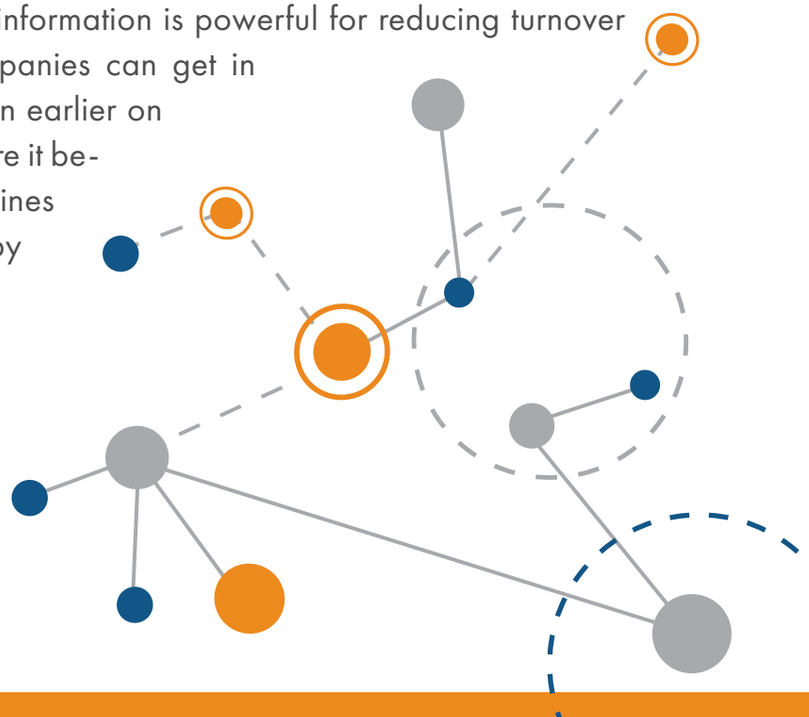
As with all data, big data must be properly analyzed and contextualized to be relevant. Traditional BI and analytics tools, while proficient at portraying trends, have not had the ability to depict relationships. Graph analytics is a type of connection analytics that can do what these previous tools could not, accurately outlining the connections between employees. In graph analytics, employees are nodes, and the communications between staff members are the edges. With graph analytics, companies can visualize workforce relationships and interdependencies, detecting clusters of affiliations. This can give companies an advantage in predicting the effects of any single departure and lead to retention solutions.

Because graph analytics, run on a graph analytics engine, can track all communications, both work- and non-work-related, employers can see how their workers are connected both inside and outside of the office. For instance, one manufacturing company is using Teradata® Aster Discovery Platform, a leading graph analytics engine, to keep track of the intermural activities of its workforce—the employees playing in the softball league and those playing on the basketball team. The friendships formed through these activities greatly assist the bottom line by building loyalty among staff. Employers need to be able to identify those loyalties however, in order to get in front of retention should any single employee leave, especially those with the largest networks.

The key here is that, with graph analytics, employers can dramatically improve their understanding of their staffs. This understanding is not limited to identification of corporate and personal relationships—it also helps employers to determine their most valuable employees and how influence flows through the networks. With graph analytics, employers can know who to reward and whose job satisfaction to constantly monitor. This results in smarter, more targeted preventative retention practices and a more rationally compensated workforce.

## Retaining the Best

Graph analysis allows companies to peer behind the curtain of their workforces. With these analytic tools, the value of individual employees can be mapped in insightful, suggestive ways, allowing companies to identify influencers and the influenced, the network of relationships and alliances that make any business thrive. This information is powerful for reducing turnover and its associated costs. Companies can get in front of employee dissatisfaction earlier on and do something about it before it becomes a problem that undermines morale and productivity. And, by boosting retention, graph analytics helps businesses keep their best employees, the ultimate strategy for success.



# Breaking Up Fraud Rings with Graph Analysis



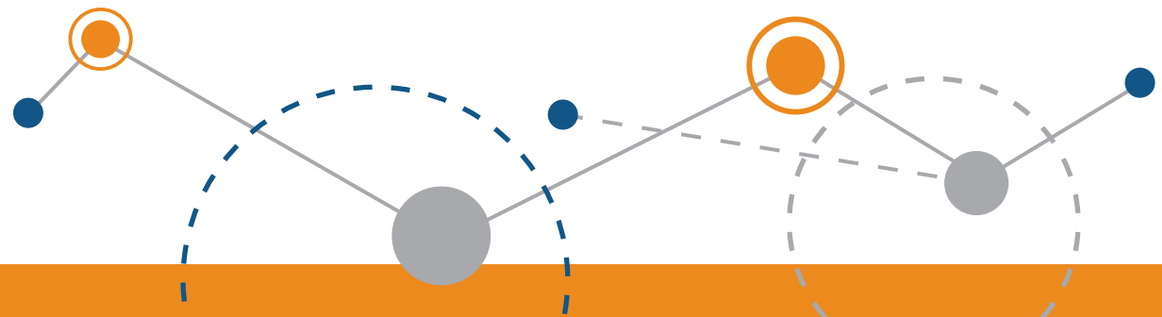
Stories hit the news nearly every day, stories of ATM skimming, point-of-sale credit card breaches, compromised gas pumps, and restaurants where credit cards have been compromised. Financial fraud is hitting consumers hard. In 2012 alone, credit and debit card fraud amounted to $11.3 billion dollars worldwide. In the US alone, 42% of all cardholders have been affected by fraud at some point—the highest rate of any country in the world.

Yet, while consumers think most about fraud, they are often shielded from its financial impact, as retailers, credit card companies, and banks absorb the losses. Banks, retailers, and other companies are finding themselves all too familiar with other less publicly reported types of financial fraud. First-party fraud affects banks and occurs when individuals or rings of fraudsters apply for lines of credit such as loans or credit cards, without ever intending to repay the bank. Fraudsters often open a series of accounts with assumed identities and act like normal customers until suddenly maxing out all the credit lines at once and then disappearing. Banks have no recourse but to absorb the losses. Conservatively, banks lose 15 cents of every dollar received to first-party fraud annually.

Just as painful for companies are insurance and ecommerce fraud. With insurance fraud, accidents or services that either never happened or the severity of which are exaggerated, are billed to insurance companies. By using a web of people, some of whom are providers (such as lawyers and doctors), coupled with a network of participants (such as victims and witnesses) who falsely report incidents, fraudsters can scam large sums from insurance companies or from the liability insurance of other businesses. The insurance industry estimates that the impact of fraud on its operations runs to at least $80 billion a year.

Ecommerce fraud occurs when individuals or syndicates acquire the credit card or banking information of individuals and use it to purchase products online. Due to US laws and regulations, banks and credit card companies are on the hook, rather than consumers, when such crimes are perpetrated.

Despite the steady increase of all three types of fraud, businesses should not concede the fight with fraudsters. In fact, new data analysis technologies are providing banks, insurers, and credit card companies with more powerful, incisive tools to combat fraud than they have had at any time in history. And the most potent tools of all are using graph analytics to prevent and stop fraud.



## Finding the Connections

How can graphs, seemingly straightforward visual representations that we've all used since high school math, be such formidable weapons in the fight against fraud? The very nature of graphs makes them ideal for exposing the interconnections of any fraudulent scheme: graphs, by definition, are designed to show the connections between people and things. Graph analytics operates on the principle of connection analysis. In the case of fraud, connection analytics helps to first identify that fraud is occurring and then find out the people and businesses that are colluding to enact the crimes.

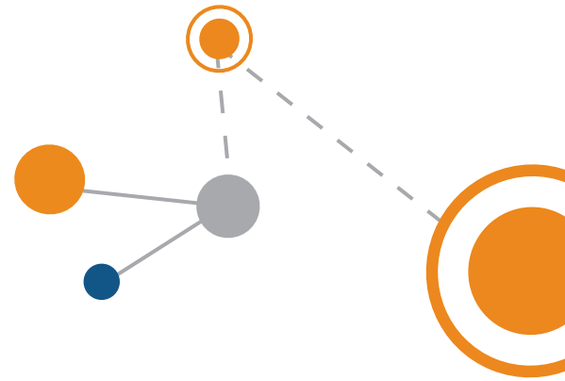
Using a graph analytics platform, such as the Teradata® Aster Discovery Platform, companies can look to find suspicious patterns that indicate fraud is occurring. This is especially important because companies need to detect fraud as quickly as possible, as perpetrators often disappear immediately after committing the act. Graph analytics can pinpoint these patterns because it synthesizes vast amounts of data and makes it searchable. The graph data structure mirrors the way fraud occurs: through a series of interconnected people and businesses.

In graphic analytics, people and vendors are nodes and the transactions are the edges. This is perfect for depicting fraud rings. To detect fraud, all an analyst has to do is walk the graph. In the case of first-party fraud, an analyst can quickly trace contested transactions to the same businesses or see how webs of criminals share identities. With graph analytics, companies can also see similarities between victims, such as determining whether those with compromised credit cards used the same ATM or gas pump in the recent past. Armed with that information, businesses can more easily deduce whether a skimmer designed to steal credit card information might be installed at one of those locations, ultimately leading them to the perpetrators.

In one recent case, a global credit card company noticed a suspicious fraud pattern in which customers' credit card data was being used illegally at online retailers. In each case, the card would be used for a small "trial" transaction before being followed by a larger transaction that often maxed out the card. The company used graph analytics to pinpoint where the fraudsters were acquiring the credit card numbers: the victims had all used one of three ATMs within a one-mile radius of downtown Miami. The credit card company was thus easily able to identify ATM skimmers as the source, and, by monitoring those ATMs, they eventually found the people behind the fraud. The company was able to reach this resolution in a matter of days. Simply following the connections shown using graphs between the purchases and activities of the victims led to solving the crime.

## Graphs: The Anti-Fraud Analytics

At the end of the day, reducing fraud helps not just banks and credit card companies, but consumers too. Companies pass on the losses they have to absorb through higher interest and insurance rates to customers. And, with recent credit card scams, we all need to be more vigilant in the fight against fraud. As the ultimate tool to illustrate the connections between the perpetrators and victims of fraud, graph analytics is a bold new step forward in the fight to keep financial information safe.



# 20/20 Vision for the Internet of Things

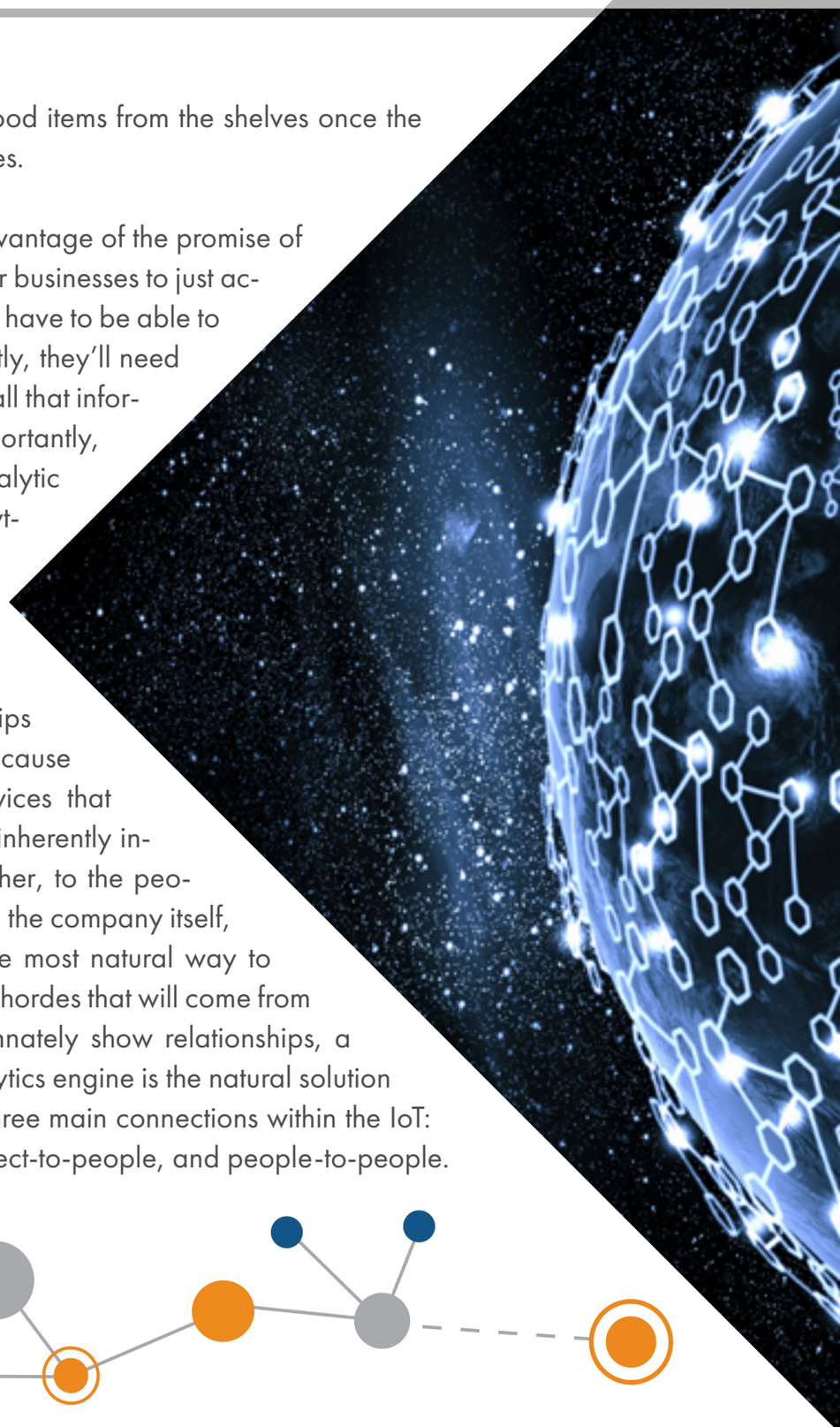
According to astrophysicists, we live in an ever-expanding universe. All of the other galaxies, planets, stars, and matter are racing away from us, and, as they do so, are increasing the size of the cosmos. Even if you're not a Carl Sagan disciple, trying to comprehend the vastness of the universe leads to a state of awe.

But we don't have to look beyond the Milky Way to be wowed by a universe that is expanding on a vast scale. Big data, in all its forms, is creating a data universe here on earth unlike any we have ever experienced. And, just like our own universe, the realm of big data is in a state of constant expansion. And its growth is only going to accelerate with the onset of the Internet of Things (IoT).

The IoT is predicated on the idea that we will increasingly make smart objects rather than static ones. We're all accustomed to the idea that our phone can provide us with directions or that we can connect to the Internet through our Xbox. But we're not just talking about our mobile technology devices. The IoT takes this idea of connectedness a step further, connecting every object in our lives, as well as the infrastructure used to build those objects. That could mean everything from our house key to our medicine is embedded with sensors that provide status information about their usage and condition.

The estimates about just how many smart objects we'll have in our lives in the near future are utterly staggering: Cisco estimates that by 2020, there will be 50 billion connected devices; Intel calculates that there will be 200 billion smart objects or 26 for every person by that same year.

Regardless of the true number, obviously, all those connected devices and sensors will produce incredible amounts of data, redefining our current definition of big data in the process. For companies, the IoT's big data production offers previously unheard of potential to gain granular understanding and control of their complete supply chains, inventories, and operations. Imagine a bike manufacturing company that can use IoT feedback to improve the durability of their models or a baby food manufacturer than can automatically alert retailers to remove its baby food items from the shelves once the expiration date passes.



However, to take advantage of the promise of IoT, it's not enough for businesses to just accumulate data—they have to be able to act on it. Consequently, they'll need a repository to store all that information, but more importantly, they'll need an analytic platform and analytics engine that can help their employees make sense of it by finding the patterns and relationships that drive value. Because the sensors and devices that make up the IoT are inherently interrelated to each other, to the people using them and to the company itself, graph analytics is the most natural way to analyze the big data hordes that will come from the IoT. As graphs innately show relationships, a powerful graph analytics engine is the natural solution to capitalize on all three main connections within the IoT: object-to-object, object-to-people, and people-to-people.

## Smart Cities and the Internet of Things

Just how well will graph analytics empower people to take advantage of the IoT revolution? Consider a smaller universe: the modern city. A city that fully embraced the IoT would have benefits for the people, government, and companies living and operating within it. With the IoT, every object and piece of infrastructure within the city could be communicating with each other. Apartment and office buildings would be constructed with devices that could monitor occupancy and adjust heating, lighting, and air conditioning instantaneously, resulting in significant energy savings. Buildings could alert owners to the need for preventive maintenance, which might mean replacing a $10 part that, if it failed, would cause damage that costs thousands to repair. The lives of the city's residents would also become easier. For instance—the exact arrival time of buses, subways, and trains would no longer be a guessing game; historic sites could automatically send information to tourists passing by; city planners could help program traffic lights to improve traffic flow based on in-the-moment feedback. The possibilities are limitless.

Perhaps more importantly, the IoT would also offer substantial assistance in combating emergencies. All fire, police, and emergency response vehicles could be linked and centrally monitored. If a fire broke out in a restaurant, dispatch would know the precise location of every vehicle in the vicinity and the employees operating the restaurant would know when help would arrive. When a crime occurred, police could use sensors in sidewalks or buildings to track fleeing suspects. Gas and water lines could alert city officials to their own deteriorating conditions so that breaks could be prevented.

Yet, while the IoT and the big data it produces will ensure that all these improvements are possible, they will not become a reality without the right analytic tools. All that connected data is best analyzed by a solution designed for finding patterns in related data: graph analytics. Graph analytics is designed to represent relationships. Graph analytics separates things into nodes and edges: in the case of smart cities, the nodes are the devices and the relationships between them are the edges. Graph analytics shows the interdependence of things. Thus, it can show how development in a neighborhood is straining the capacity of its transportation infrastructure, or reveal how putting more cops on the beat in one area affects crime rates in that location as well as in surrounding locations. For businesses, graph analytics on IoT data can provide insights into the most efficient route for the delivery trucks or the best location to establish a retail store.

Since the first sensor was created, the problem for the IoT has always been the scale and proliferation of the data. Smart devices will emit a stream of data, and all those devices and all that data are connected. The real payoff for all of this intelligence comes in analyzing the data for new and interesting patterns that will drive further optimization and innovation. Graph analytics is the most effective tool for the flood of data from the IoT.

## Conclusion

The IoT will soon be an integral part of all our lives—and the operations of all businesses and the cities where they are located. Graph analytics are a natural fit for IoT data and applications. Graph analytics engines can find the kinds of patterns that will help everyone—whether a city or a corporation—profit and learn from the IoT.

# Managing Cyber Threats with Graph Analytics



## Introduction

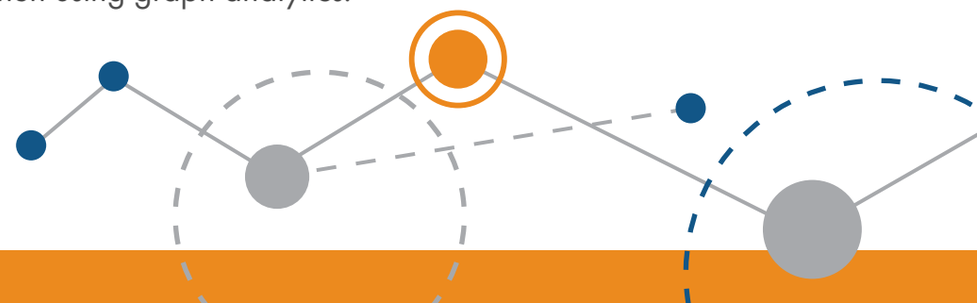
For businesses, governments, and individuals alike, web security has become an increasingly serious problem. According to a 2014 report by security provider McAfee and the Center for Strategic and International Studies, a Washington DC think tank, the annual global cost of cyber crime is between $400 billion and $575 billion. Think about that: It's nearly $100 for every man, woman, and child on the planet. Obviously, for companies, protecting the data of their customers is crucial: it has a negative impact on consumer confidence, brand perception, and revenue as well. But, in both the public and private sectors, IT departments have to be in a state of constant vigilance that extends beyond just consumer data, monitoring external threats like advanced persistent threats, distributed denial of service attacks, zero day exploits, and malware infestations. Internally, those departments have to be just as attuned to overseeing users and their access rights, ensuring that employees only have access to information relevant to their jobs.

While hackers are becoming ever more sophisticated, organizations have a new tool to ensure the security of internal resources and to ward off and prevent outside attacks: graph analytics. Unlike traditional relational databases and analytics, graph analytics shows patterns in data, enabling organizations to detect abnormal behavior and recognize attacks sooner. In an era of ever-evolving and emerging threats, graph analytics is a potent weapon that organizations cannot afford to leave out of their cyber-security arsenals.

## How It Works

While the technology behind graph analytics is sophisticated, its fundamental premise is not. The world is made up of relationships—between people, companies, and things. Nowhere is this more true than online. Graph analytics allows you to find patterns in interconnected data in a way that traditional analytics cannot.

Being able to chart these relationships becomes increasingly difficult in a world of big data, where large companies are collecting between 10 and 100 billion security events per day. The only way for a business to keep on top of this type of data flow is to have an analytics engine that can quickly spot patterns in the data that indicate anomalies. Graph analytics engines offer this, making the information immediately actionable when using graph analytics.



## Security and Access Management

Any CTO will acknowledge that, when it comes to security, being able to fend off internal threats is just as important as thwarting external ones. Though extreme, Edward Snowden's revelations about US security offer a prime example of the insider threat organizations face. His disclosures stemmed from a lack of internal controls and checks and balances around internal access that enabled him to review and then remove classified information.

Graph analytics allow companies to mitigate this problem by managing access to content and monitoring the relationships between users, groups, assets, and collections. In the case of internal security, people, users, and content are the nodes, and the access to and between them are the edges. Graph analytics offers potent summaries of who is accessing what and when, and who they are communicating with about that information—both inside and outside the organization. Armed with these insights, organizations can detect whether there are problems and intervene to stop them.

## Looking for Attack Patterns

Graph analytics is just as powerful in securing organizations against external threats. Companies do not suffer from a lack of information about their networks. For most organizations, the main issue with identifying security risks has been insufficient analytical power to make all that information useful. The key to understanding these risks is to identify patterns that allow security professionals to distinguish between normal and abnormal behavior.

For example, consider advanced persistent threats (APTs). These attacks often involve a web page downloading dropper malware, a very small executable that then runs to download a tool that the attacker can use to control the system in question. Often the user is tricked into clicking on a link or opening a document that then executes the malware. Such attacks are not easy to recognize because the dropper malware is so small. After the attacker takes control of the system, however, the effects become more noticeable as a particular system starts trying to access many connected systems and files and begins exfiltrating data to the attacker. By establishing normal usage patterns as a baseline, such radical changes stand out. Graph analytics is helpful both for establishing what a normal pattern is for a given user and for monitoring departures from that profile.

Graph analytics helps find patterns at scale. Thus, graph engines are a great fit to monitor the relentless stream of data generated from IP, network, server, and communication logs with which IT departments have to grapple. And, because graph analytics can synthesize all this data rapidly, threats can be detected in near real time. What becomes clear through this type of pattern recognition is that graph analytics exposes not only what companies know about their systems, but also, perhaps more importantly, what they don't know. Such analytics can help organizations set a strategy for how to address these gaps.

## Creating an Early Warning System for Cyber Threats

Graph analytics' ability to recognize patterns also greatly assists organizations in setting up early warning systems for cyber problems. For instance, Georgia Institute of Technology is using graph analytics to combat security threats through a system called Black Forest that helps to anticipate attacks by aggregating information from hacker sites. By collecting and then analyzing information of this kind from across the Web, graph analytics can beat malware developers and hackers at their own game by offering organizations a counterintelligence strategy that allows them to stay one step ahead. Additionally, not only can graph analytics help companies track domains that have been problematic in the past, but also, domains related to those already identified that could pose a risk in the future. Networks are made up of connections, so it only makes sense to use a technology specifically designed to finding patterns among connections.

## Conclusion

Graph analytics is a natural fit for finding patterns in big, connected data. Armed with graph analytics, IT staff can keep internal resources accessible to the right people and help them identify anomalies by quite literally connecting the dots in an attack.

# Debunking the Claims: How Graph Analytics Uncovers Insurance Fraud

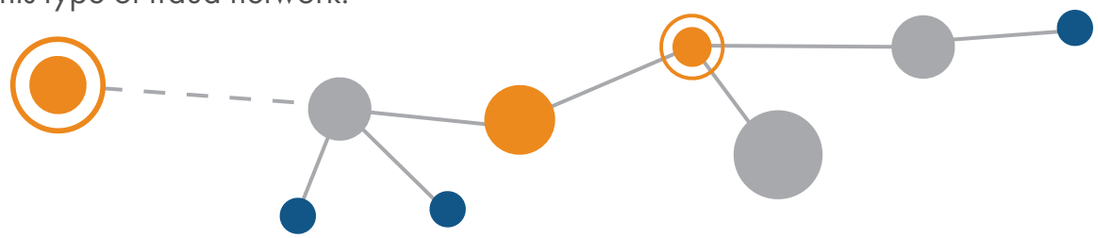
Ask people what they think is the most serious type of nonviolent crime and they'll probably say drug use, burglary, robbery, or arson. If they follow finance, maybe they'll mention white collar crimes like insider trading. Few, if any, would mention insurance fraud. And yet the magnitude of insurance fraud in the US is staggering and it's costing businesses and governments, as well as taxpayers, billions of dollars every year.

Though technical and somewhat wonky in nature, given the amount of money involved, it's surprising that insurance fraud does not dominate the headlines. The FBI estimates that non-health insurance-related fraud in the US runs to $40 billion annually, costing the average family $ 400 to $ 700. Factor in health insurance fraud and the numbers get even more shocking, with Attorney General Eric Holder estimating that Medicare and Medicaid fraud cost the US $120 to $180 billion each year.

And that doesn't even include fraud with private health insurance providers. For government agencies, fraud in the wake of natural disaster recoveries, such as hurricanes, is a huge problem, with as much as 10% of claims following a disaster classifying as fraudulent. This type of corruption not only is a significant theft from businesses, governments, and taxpayers, but it also undermines the faith all of us have in our support systems and drives up premiums for everyone.

The most common types of fraud include the embezzlement of premiums, fee churning, and asset diversion. All of these crimes generally share something in common—they require networks of individuals or resources to carry out. As a result, as companies and government agencies ramp up their efforts to combat the insurance fraud epidemic, they're increasingly turning to graph analytics to help them detect and prevent fraud.

Graph analytics is a natural solution for the fraud problem because it is predicated on relationships—between people, accounts, and funding streams—and patterns in those relationships. Insurance fraud is often highly sophisticated, involving many criminals, who each play a small part in a larger scheme. This can include everyone from doctors who diagnose false injuries or write claims for patients they never saw to automobile body shops that claim to have performed work they did not do. For those trying to fight fraud, graph analytics can quickly and efficiently link the diaspora of criminals involved in this type of fraud network.



## Case In Point: Health Insurance

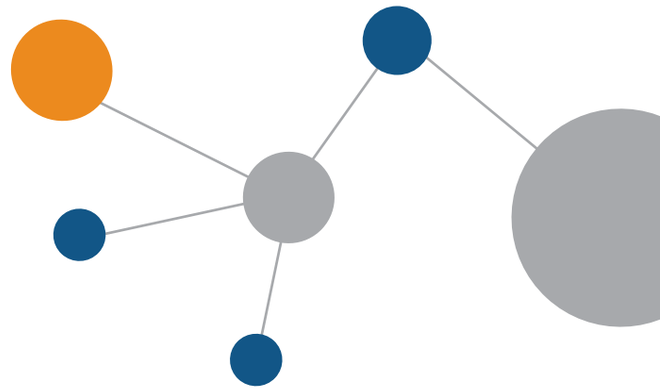
The insight and power graph analytics can provide in this battle is well illustrated by its use in health insurance fraud cases. Unraveling health insurance fraud is centered on detecting patterns in suspicious behavior and knowing where to look to find these patterns. First, it's important to recognize that 80% of health insurance fraud is perpetrated by providers. Providers often recruit people to operate in widespread networks to hide the fraud. For instance, doctors may refer patients to other doctors they know will reciprocate and then overbill for services never rendered. Or providers will identify certain patients with chronic conditions and then overcharge or overbill. As many patients pay nothing other than their co-pay, insurance companies are often left on their own to determine whether or not fraud has been committed.

By using graph analytics and walking the graph from provider to claim to patient, insurance companies and agencies can identify providers perpetrating fraud regardless of the complexity of the scam. Graph analytics can show the relationship between a provider, his entire patient base, and the claims he's submitted for them, or providers with similar claim patterns, with a depth and clarity that relational databases cannot match. The reason for this is that humans think in terms of graphs and connections, not rows and columns. Analysts can examine how much other providers with similar patients bill and then target the outliers. In cases where multiple doctors and fake patients are involved, insurance companies can weave a coherent narrative, depicting the fraud by using graph analytics to find the connections between all the parties. What might previously have been regarded as a singular, unique instance can suddenly be traced to a chronic pattern of fraud.

Graph analytics makes the fraud investigatory process easier and more efficient. By prioritizing and recognizing the most suspicious claims earlier, insurance companies spend less time chasing false leads. Additionally, health insurance fraud cases, where doctors are providing kickbacks to patients involved, are particularly susceptible to the power of graph analytics because graphs portray social networks so accurately. Rather than viewing each patient or provider in isolation, once a problematic claim has been flagged, graph analytics can show all those who are in frequent contact with that individual and the claims they've filed. So, if an entire family were committing fraud with separate providers, for example, analysts could use graph analytics to detect this pattern. Graph analytics makes it much more difficult for criminals to escape detection and can find the larger network of fraud, rather than just part of the fraud ring.

## Conclusion

Insurance fraud is a fact of life. As long as there are insurance companies, there will be dishonest criminals trying to game the system that all of us depend on as our stopgap when unexpected injuries or problems beset us. Yet, while fraud may be inevitable, accepting it is not. Graph analytics are a natural fit for finding insurance fraud. It is the ultimate analytic tool to explore relationships and the connections between people, resources, and claims. It can greatly help companies or governments more deftly identify erroneous claims and connect the dots between all the actors in a fraud network. As a result, investigators spend more time on the cases that cost insurers the most, and catch criminals more quickly.



# Free Goodies Inside: Finding Profitable Loss Leaders with Graph Analytics



Everyone has heard the adage that, in order to make money, you have to spend money. Many businesses apply this mantra to their product base, offering free or heavily discounted products to customers in the hopes that, down the road, the customers will purchase higher profit-margin services. These types of services or products are generally referred to as loss leaders.

Anyone who has received a free checking account from a bank is familiar with the way the financial services industry has used loss leaders. The goal with the free checking account is to make you a customer in the short-run so that, when you advance to more profitable products, like a mortgage or credit card, you turn to that bank to provide it. But, there are examples of loss leaders in all sectors. Amazon purportedly sold its Kindle at a loss in order to attract more visitors to its site and because it made significant profits off the e-books and other electronic media used on the devices. Often, when companies have excessive supply of a product that hasn't moved, companies will drop the price to get it off the shelf. Or, in some cases, companies will push out a loss leader just to raise brand awareness—liquor companies are famous for using this technique, giving out free shots of new liquors at bars.

## Happy Hour—The Right Kind of Customer Relationship?

Loss leaders are an excellent example of how price elasticity plays out in the business world and consumer marketplace. We're all fundamentally aware that companies offer these types of incentives and promotions to lure us in—that people pay different prices for the same products depending on when or where they buy it, or who they are. Happy hours at bars are one of the most iconic examples. And yet, while these deals can often seem too good to be true for consumers, there's been very little analysis on the business side to determine a fundamental question about loss leader campaigns: Do these techniques actually work? Do they actually create the type of long-lasting relationships with entrenched brand loyalty that companies are seeking? Or, are customers only looking for the best deal and their loyalty extends only as far as their pocketbook? These questions have remained difficult to answer, especially for new campaigns. But with graph analytics, companies now have more insight into the effectiveness of their loss leaders than ever and can use that insight to power smarter, more targeted campaigns to get the results they want.

## A Graph of Products

Graph analytics help make companies smarter about their decisions because they show the relationships between any network—of people, products, services, purchases. In the case of measuring the effectiveness of loss leaders, the products are the nodes and the transactions involving those products are the edges.

In the past, using traditional BI tools, when companies have tried to determine whether the total value of their relationship with a customer has made up for the loss leader provided initially, they've essentially been engaging in a guessing game of causation versus correlation. They could see the end result, but had no way to determine whether that was brought on by the loss leader or not. Graph analytics changes that equation.

## Validating the Graph

Graph analytics allows companies to dig deeper into the outcomes of their campaigns. Graphs are predicated on revealing relationships and, therefore, are the perfect tool to monitor and measure how well loss leaders are working. With graph analytics, businesses can essentially walk the graph, tracing the impact of campaigns on their customer base. They can quickly see if a particular product is more popular with a specific set of customers they have been trying to target, analyzing demographics about where those customers are from, how frequently they move on to other, more profitable products, as well as other demographic characteristics. The entire network of purchases and profits can be analyzed more easily because humans think in terms of relationships, not rows and columns.

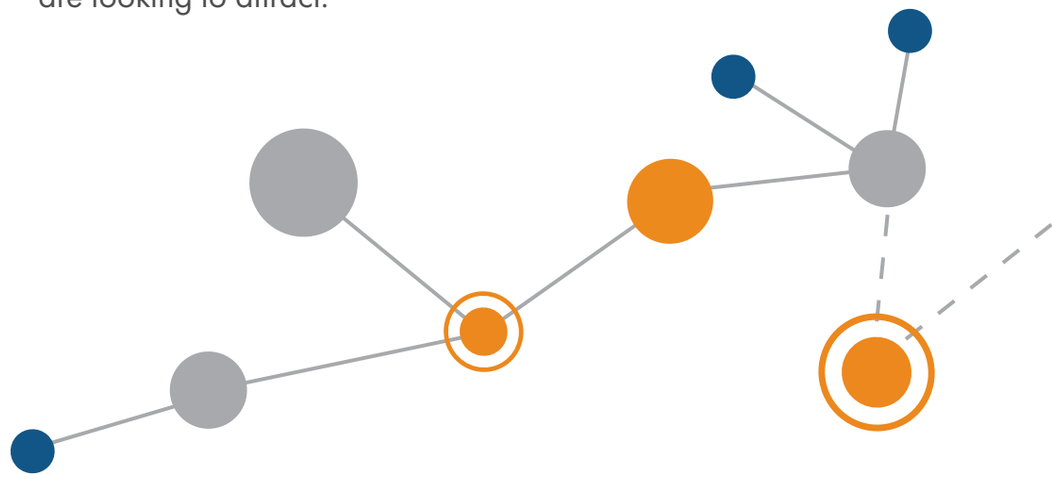
Using this type of information, companies can make more reasoned, informed decisions about whether their loss leaders are really leading at all—or if they're just costing the company money. Using graph analytics, companies often find that certain loss leading products work much better with certain types of customers. Consequently, they will offer these products to a smaller, more targeted set of potential customers, rather than to everyone. The result is better tuned advertising and marketing campaigns, pinpointing the customers who will actually generate higher profits, while avoiding giving free or discounted offerings to customers who are looking for a one-time deal.

This strategy also works for a company's existing customer base and not just in attracting new customers. In the case of free checking accounts, banks can avoid offering free checking accounts to existing customers who keep a balance of near zero, as they obviously do not have the type of income to support a mortgage or consistently pay off a credit card. In the long run, using graph analytics, companies can not only make more money off their loss leaders, but they can also save significant waste by not making offers to the wrong customers.

By validating this strategy with existing data, companies can validate the loss leader strategy before ever making an offer. They can leverage their data to determine whether there might be a more innovative, attractive, and also profitable offer to make.

## Conclusion

Offering loss leaders is a risky strategy for any business. Using graph analytics can reduce the risk and provide companies with the visibility to see if the offer will bring in the business they are looking to attract.



# Peak Productivity: How Graph Analytics Can Minimize Network Downtime



Give a network engineer a whiteboard and a marker and you'll find that they can't help themselves. They represent networks, and all changes to networks, as graphs. This is only logical—the very word network invokes a web of connections.

This almost inherent human tendency to portray networks as graphs would seem to make graph analytics a natural fit for network analysis. And yet, few companies have taken the time to model their networks in a graph data structure. However, given how complex and myriad the networks are becoming, how costly even minutes of downtime are to companies, and how all businesses are becoming ever more dependent on their networks, there are many reasons it makes sense for companies to use graph analytics to obtain the most accurate, real-time assessment of their networks possible.

## The Scourge of Downtime

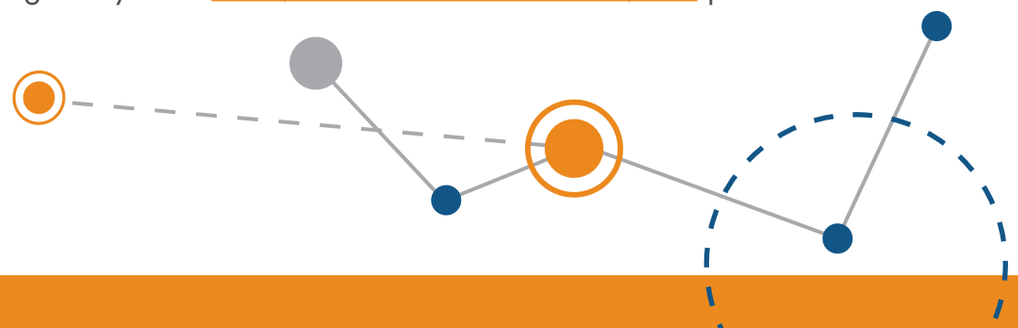
When it comes to a business's network, transparency of and information about all network operations are the ultimate preventive medicine for the scourge of downtime. Companies need to know as much about their networks as they can, in order to avoid the high costs, lost productivity, and wasted time that result from outages.

Just how devastating are the impacts of downtime on companies nationwide? Put together, the figures are astonishing:

- *Downtime affects all companies, regardless of sector and size*. 59% of Fortune 500 companies experience an average of 1.6 hours of downtime each week.

- *A growing problem*. Companies experienced 29% more incidents of downtime in 2013 than in 2012. As networks expand in both size and complexity, preventing downtime becomes ever more important.

- *The high tab of the unexpected*. There is no question that downtime costs companies appreciably. Estimates for the costs for an hour of downtime range from $163,674 to nearly $475,000. Estimates about the average cost of an entire unplanned outage vary from $627,000 to more than $900,000 per incident. Obviously, companies can't afford these types of losses.



What makes these numbers even more alarming is that, in most cases, businesses know the causes of the outages. Human errors during network changes are responsible for 82% of downtime incidents. Change is the enemy, because most companies do not have a full grasp of how everything in their networks is interconnected—how changing one aspect, whether temporarily taking down a server or altering the path with which certain users access databases, will impact the rest of the network. But this lack of visibility, and the daunting consequences associated with it, are not a given. Graph analytics can provide the clear insight into networks that companies need in order to deter downtime.

## The Graph Analytics Solution

Graph analytics succeeds where traditional BI tools fail. This is because they model information as it appears in the real world, as a series of connections between people, things, devices, and technologies.

Graph analytics allows companies to model their networks. Users and resources (whether routers, hubs, servers, or devices) are the nodes and the connections between them are the edges. By modeling its network, a company can assess the impact any single change or failure will have on the rest of the network. Consequently, surprises are minimized as companies experience fewer unexpected problems during network changes. Network engineers know who they'll affect if they take the system down for regular maintenance. And engineers can more easily target the root causes of problems that occur repeatedly over time.

Graph analytics offers companies insight to make informed decisions about their networks. It's not only the depiction of the interconnectedness; it's also usage frequency—how often and when users are accessing particular resources, or when certain resources are overburdened. In many cases, because companies build and grow their networks sporadically, over many years, graph analytics also helps to reveal dependencies or inefficiencies within the system engineers might not previously have been aware even existed. Information is power in the fight against downtime and graph analytics is an important resource for analyzing the information locked in log files and other data sources that companies already possess about their networks.

Ultimately, outages will occur. And most outages, due to the increasing complexity of networks, are not caused by a single problem, but rather from a cascade of events. Ultimately, the only way to fully combat these problems is through careful planning that allows companies to minimize the impact of those problems on users, and, when it comes to e-commerce, to prevent problems that can cost a lot of money, especially during peak periods like Black Friday or Mother's Day.

## Conclusion

Business is more reliant on computers and networks than ever before. That's why downtime is so costly. But it is also avoidable. By creating a graph of your network and modeling the impact on users of systems going down, companies can see how to plan for failures and where to build in redundancy to bring downtime as close to zero as possible.

# How Graph Analytics Untangles the Messy World of Compliance

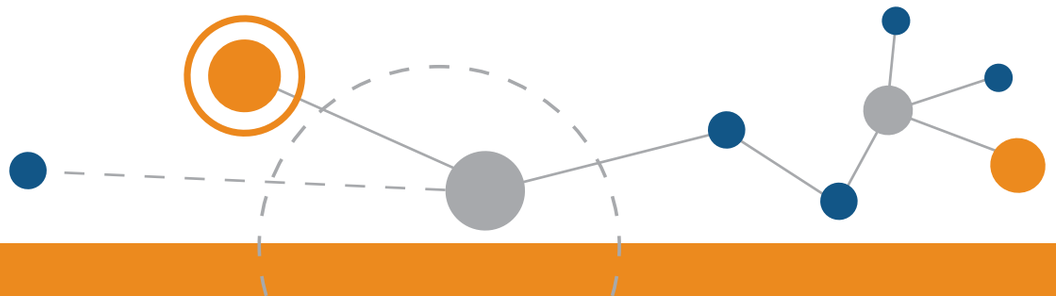
Seemingly every day, there's a new article, touting the incredible benefits of big data. And with good reason—big data certainly has the potential to radically transform and disrupt the way businesses operate in a way that no other technology has since the advent of the Internet.

And yet, as the saying goes, with great power comes great responsibility. Or, in the case of the contemporary corporate landscape, compliance issues. Even as companies fill their data warehouses with seemingly unending fountains of information about their customers, competitors, and industries, the number of laws those same businesses have to comply with in order to avoid risk and possible prosecution becomes ever larger. And the problem is particularly acute for international companies or those with international customers, as laws vary greatly by country and by industry.

But, regardless of which regulations a company must comply with, all regulations have one thing in common: they require that a business exerts tight controls over access to the data at its disposal so that only the right people have the right access to the right resources. By modeling resources and users as a graph, companies can gain visibility and insight into key patterns of permission and usage.

## Permissions Are a Graph Problem

Graphs are a logical way for companies to get a grasp of permissions, the way that computers automate who has access to what. Expressing permissions as a graph is logical because, ultimately, whether it's sensitive data or the ability to write checks as distinct from the ability to cash them, graphs are natural for expressing the relationships between users and the programs and data to which they have access. Graphs show connections and interconnectedness. They reveal dependencies and can uncover patterns. In essence, when it comes to compliance issues, graph analytics can x-ray the social network of a business, exposing the most sensitive and mission-critical resources of a company, and then linking those resources to those who have access to them and showing how frequently they use those privileges.



When companies adopt a graph analytics solution, they can the perform the following essential functions to ensure compliance:

- **Segregation of duties:** This is one of the most crucial steps any company must take toward mitigating and managing risks. Through segregation of duties, essential functions, and the access to perform them, are spread out among numerous people in a given department or team, so as to lessen the chance of fraud and errors. Companies have to be able to keep authorization, custody, record keeping, and reconciliation separated to prevent fraud. Graph analytics helps to map how these duties are divvied up, allowing companies to make changes if too much power resides in one person.

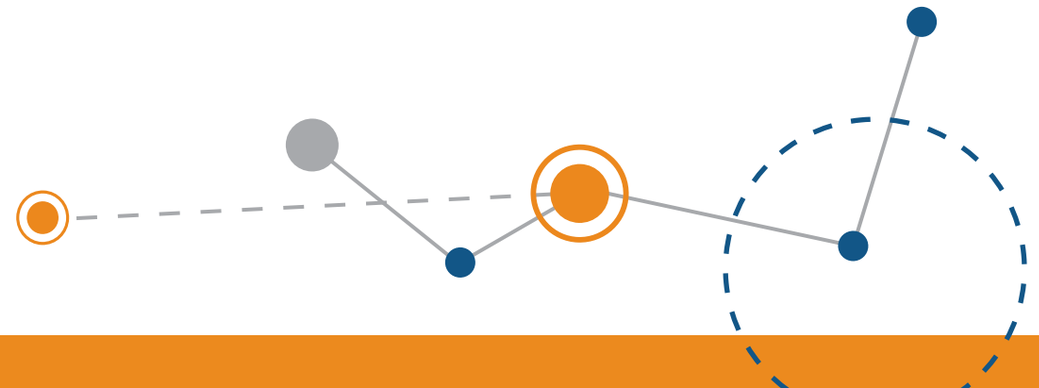
- **Identify data that is exposed or sensitive:** Companies can pinpoint their most sensitive information and then use graph analytics to see who has access to it. This helps businesses to spot weak areas where they're at the most risk of compliance failures.

- **Monitor superusers:** Every company has superusers who have greater access than most other employees. But just what are these super users accessing? And are they accessing it only from secure locations? Graph analytics can help monitor the activities of superusers.

- **Help implement governance rules, policies, and procedures:** To ensure compliance, companies have to adopt consistent rules, policies, and procedures to govern usage that are modeled and implemented from the top down. Using graph analysis, businesses can first see how these rules will affect their users before implementation, and then help with enforcement once they are adopted.

Because almost any size of dataset can be expressed as a graph, graph analytics can help companies make sense of the controls needed across millions of resources, data sources, and users. By scrutinizing patterns of access, companies can consider ways to simplify and streamline who has access to key resources, further reducing their risk.

Compliance is a critical issue for companies of all sizes. With the amount of data, users, and devices now accessing company networks, it's almost impossible to analyze compliance without the assistance of data structures like graphs. Graph analytics provides transparency into the access and control of any network, allowing businesses to make informed decisions about who should be able to see what, and from where, and who should not. And with ever-changing compliance requirements, companies need all the help they can get to analyze and simplify the maze of connections between users and resources.



# Related Resources

### Virtual Event: Data Discovery in Action

To see Data Discovery in action, check out Teradata's virtual event. You will find information about what Data Discovery does and how it can help your business. You'll also see some real-world use cases. Check it out at www.teradata.com/discovery.

### Customer Story: Verizon Wireless

America's largest wireless carrier explains how a college intern uncovered valuable new customer intelligence in just a few weeks. Hear about the unexpected customer behavior that couldn't be explained until it was brought together through Teradata's Unified Data Architecture and Data Discovery Platform. Find the story at www.teradata.com/verizon.

### Customer Story: Cardinal Health

A Fortune 500 healthcare services company optimizes a supply chain that serves over 60,000 locations. Hear how some users cut the time needed for working with raw data by 50%, leaving them time to move past reporting and do more research. Find the story at www.teradata.com/cardinal.

### Customer Story: McCain Foods

The world's biggest maker of frozen French fries motivates everyone to get involved by giving you a view of performance data. (Next on the agenda: addressing food shortages worldwide!) Go to www.teradata.com/mccain.

### Use Case: Grow Loyalty of Influential Customers

Discovering which of your customers are most influential is the start to building their loyalty. Find out more at www.teradata.com/influential.

### Whitepaper: The Rise of Data Discovery

Thomas H. Davenport, President's Distinguished Professor in Information Technology and Management at Babson College, is the author of the upcoming book *Big Data at Work* (Harvard Business Review Press, 2014). He has written many other books on IT and business, including the bestselling *Competing on Analytics: The New Science of Winning* (Harvard Business Review Press, 2007).

In a whitepaper, Davenport summarizes learning from his interviews with early adopters of Data Discovery. He outlines motivations to engage in Data Discovery and typical applications, and explains the attributes of an effective platform and process, as well as barriers that must be overcome. Find out more at www.teradata.com/tomdavenport.

### Book: *Data Discovery For Dummies*

Learn how Data Discovery gives the power of new insights to more people, with less fuss! Go to https://site.teradata.com/Microsite/data-discovery-for-dummies/landing/.ashx.

### Product Information: Teradata® Aster Discovery Platform

Go to the starting point for information on the Teradata® Aster Discovery Platform. Find it at www.teradata.com/DiscoveryPlatform.

### Animation: Next Generation Analytics With Teradata® Aster Discovery Platform

### Animation: Teradata Connection Analytics

### Talk to Teradata

Not finding the right information from these sources? Time to get on the horn and ask for what you need!
Contact Teradata at 1.888.278.3732 or www.teradata.com/contact-us.