

# Big Data Handbook

*How to Unleash the Big Data Opportunity*

7/8/2013

**By Lora Cecere**  
Founder and CEO

**Supply Chain Insights LLC**

# Contents

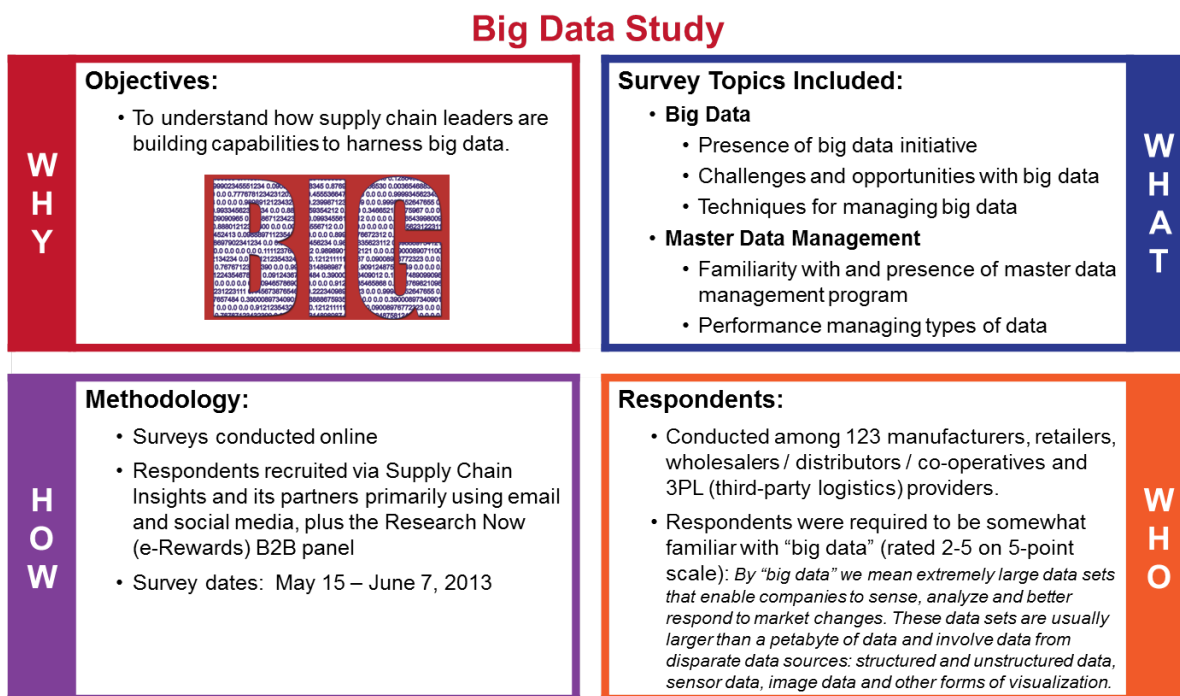
Research.....	2
Disclosure.....	2
Executive Overview .....	3
Is It a Big Data Problem?.....	5
What Is Big Data?.....	6
Companies Focused on Clean Data Today Are in a Better Place.....	7
The Opportunity.....	10
Recommendations.....	12
Conclusion.....	13
Appendix .....	14
Other Related Reports.....	20
About Supply Chain Insights LLC .....	21
About Lora Cecere .....	21

# Research

Data is piling up on corporate doorsteps. It is growing in volume, velocity and variety. New forms of data, in combination with advanced analytics, offer a wealth of new possibilities. How do companies get started?

We want to help. The goal of this report is to break through the hype and understand where companies are on their *Big Data journeys*. The recommendations in this report are based on survey data conducted online during the period of May-June, 2013. The quantitative analysis is augmented by interviews with manufacturers and retailers working on advanced analytic concepts. The overview of this study is shared in figure 1. For more detailed respondent demographics, reference the additional charts in the Appendix.

Figure 1. Overview of the Big Data Study



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)

# Disclosure

This report was solely funded by Supply Chain Insights. The analysis is an independent and objective read of the current market.

As a research firm, we are committed to Open Content research. These reports are intended for you to read, share and use to improve your supply chain decisions. Please share this data freely within your company and across your industry. As you do this, all we ask for in return is attribution. We publish under the Creative Commons License [Attribution-Noncommercial-Share Alike 3.0 United States](https://creativecommons.org/licenses/by-nc-sa/3.0/) and you will find our citation policy [here](#).

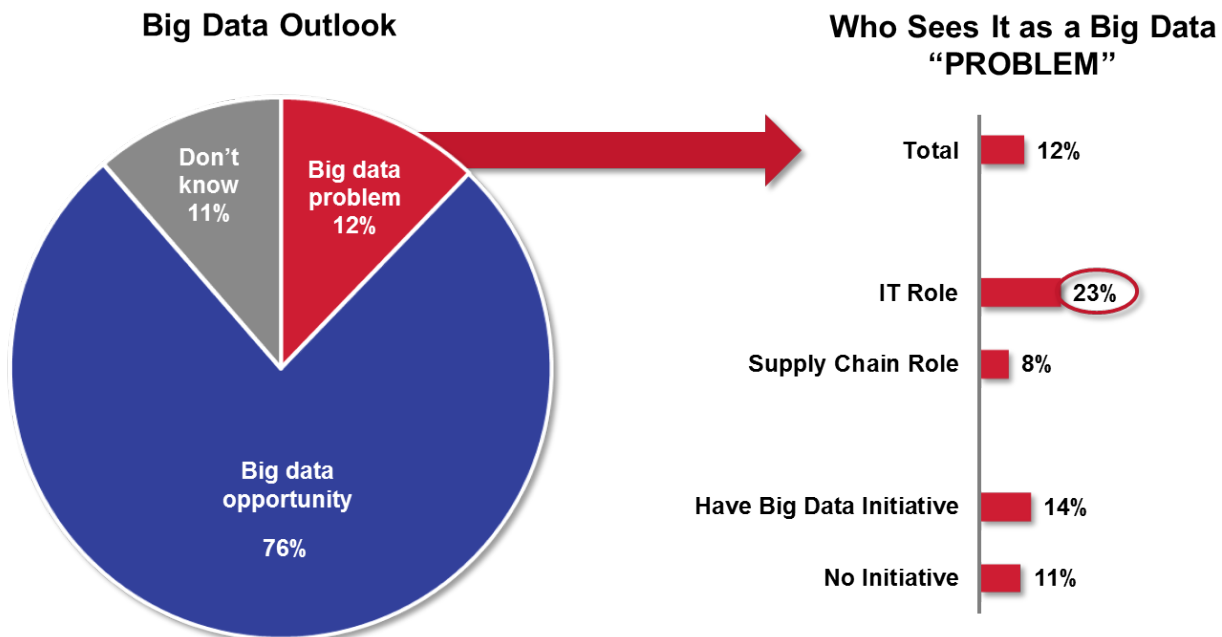
# Executive Overview

The use of new forms of data is not an evolution. Instead, powering big data supply chains, and innovating through new forms of analytics, is a step change.

New forms of data do not fit traditional architectures. Traditional supply chains were architected to use structured data with software using relational databases. The big data era will make many of the investments from the last decade obsolete.

Big data offers the opportunity to redefine supply chain processes from the outside-in (from the channel back) and define the customer-centric supply chain. This is in stark contrast to the inflexible IT investments installed over the last decade to respond inside-out based on order shipments. These traditional investments in Enterprise Resource Planning (ERP), Advanced Planning Systems (APS) and traditional Business Intelligence (BI) for reporting, improved the supply chain response, but did not allow the organization to sense, shape or orchestrate outside-in. New forms of data (e.g., images, social data, sensor transmission, input from global positioning systems (GPS), the Internet of Things, and unstructured text from email, blogs and ratings and reviews) offer new opportunities. They also require new techniques and technologies.

Figure 2. Is Big Data a Problem or an Opportunity?



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123), IT Role (n=31), Supply Chain Role (n=38), Have Big Data Initiative (n=35), No Big Data Initiative (n=88)  
Q7A. When you think about big data, do you think of it as a big data "problem" or a big data "opportunity?"  
○ Higher than other group at 90% or higher level of confidence

Big data offers new opportunities for the corporation to listen, test and learn, and respond faster. In this study, companies see the greatest opportunity to use big data for "demand" (to better know the customer and improve

the response); however, actual investments are in “supply” not “demand.” Respondents view supply-centric projects like product traceability (involving product serialization and traceability), supply chain visibility and temperature controlled handling as important.

Is big data a problem or a new market opportunity? Like the respondents of this survey, we believe that big data represents an opportunity for all. In the study, one-fourth of respondents currently have a big data initiative. However, interest is growing. Sixty-five percent have or plan to have a big data initiative in the future. Despite the hype, and the intensity of marketing rhetoric in the market, in our year-over-year studies on big data we see very little change in activity.

Despite the fact that the IT group is more likely to see big data as a problem, 49% of those with a big data initiative report that it is headed by an IT leader.

Big data represents a new opportunity, but seizing it requires a new form of leadership. It can ignite new business models and drive channel opportunities. However, it cannot be big data for big data itself. Instead, the initiatives need to be aligned to business objectives with a focus on small and iterative projects. It requires innovation. To move forward, companies need to embrace new technologies and redesign processes. It is not the case of stuffing new forms of data into old processes. In the survey, as shown below in the sample of quotes from respondents, business leaders have a range of emotions. The path forward requires leadership. People need a nudge to think beyond traditional data types and applications. Here are some quotes from respondents in the survey:

*Big data is a distraction at the moment. [Our] key challenge is how to automate useful insight.*

*We have not explored this opportunity. Skillsets in supply chain and IT are limited.*

*This information can give companies a “heads-up” to help sense, analyze, and better respond to market changes.*

*Data equals information and information equals profit.*

---

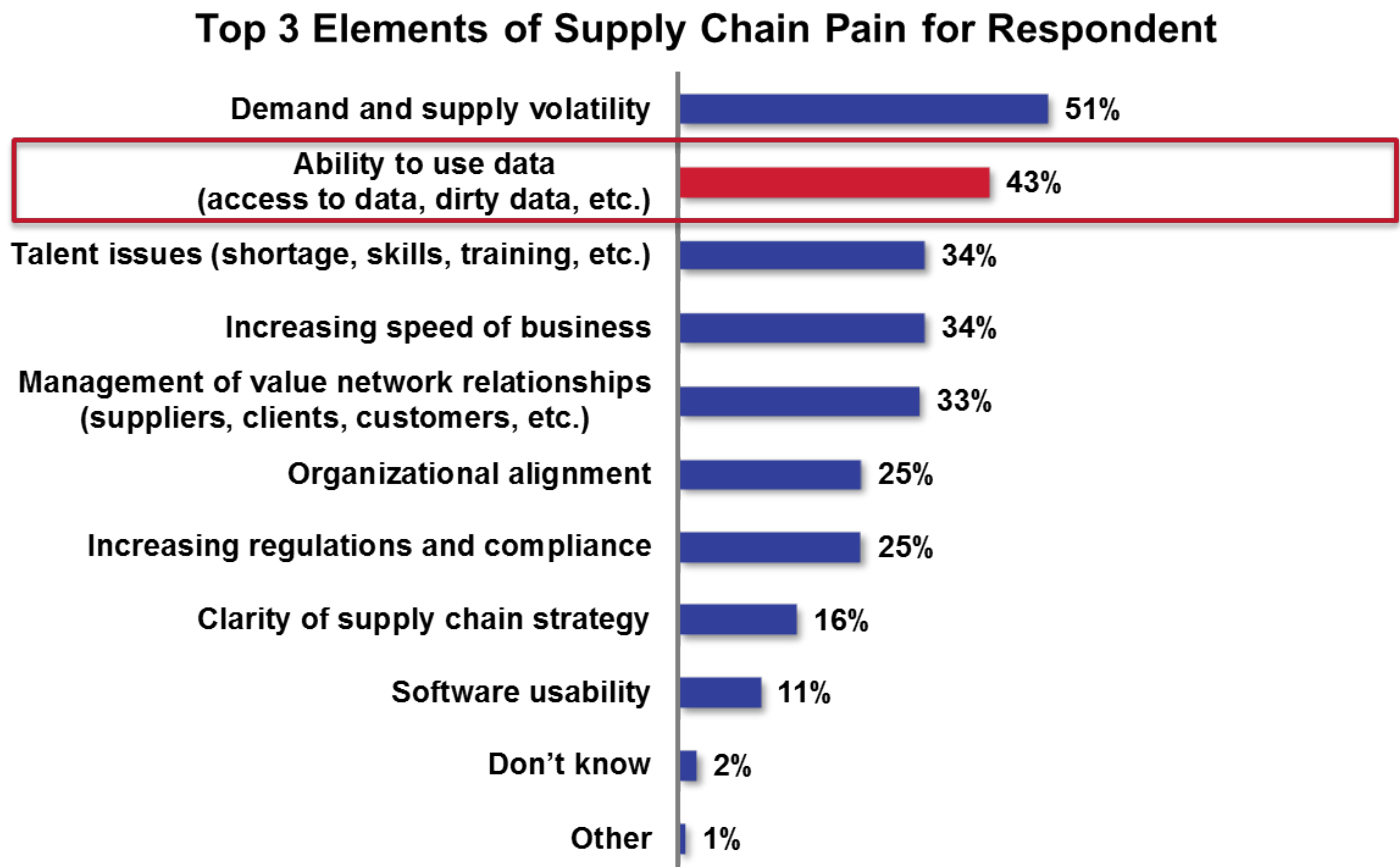
In capitalizing on the opportunity, data-driven companies—focused on the use of analytics and clean data—have an advantage. One of the first steps is to get good at managing data. In this study, we find that companies that are more advanced in Master Data Management (MDM) are more likely to have a big data initiative today.

Unfortunately, the term “big data” is the buzzword du jour. Due to the size of the prize, many consultants and technology partners are hyping the concepts. We find that most technology providers and consultants use it in their marketing pitches. Confusion reigns. To drive the opportunity, companies must bypass the hype. To help readers, we share insights on how to navigate the hype and maximize the opportunity in the recommendations section of the report.

## Is It a Big Data Problem?

Today, most companies feel that they do not have a big data problem. While data volumes are growing, the velocity of data is accelerating and the variety of data is increasing; the larger top-of-mind issue (as shown in figure 3) is that companies cannot use the data that they have today. While they are intrigued about how to use new forms of data to solve the issues of demand and supply volatility, few know how to get started.

Figure 3. Top Levels of Supply Chain Pain for the Respondents



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)

Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)

Q4B. When it comes to supply chain management, which of the following are the top 3 elements of business pain for you personally? Please select no more than three.

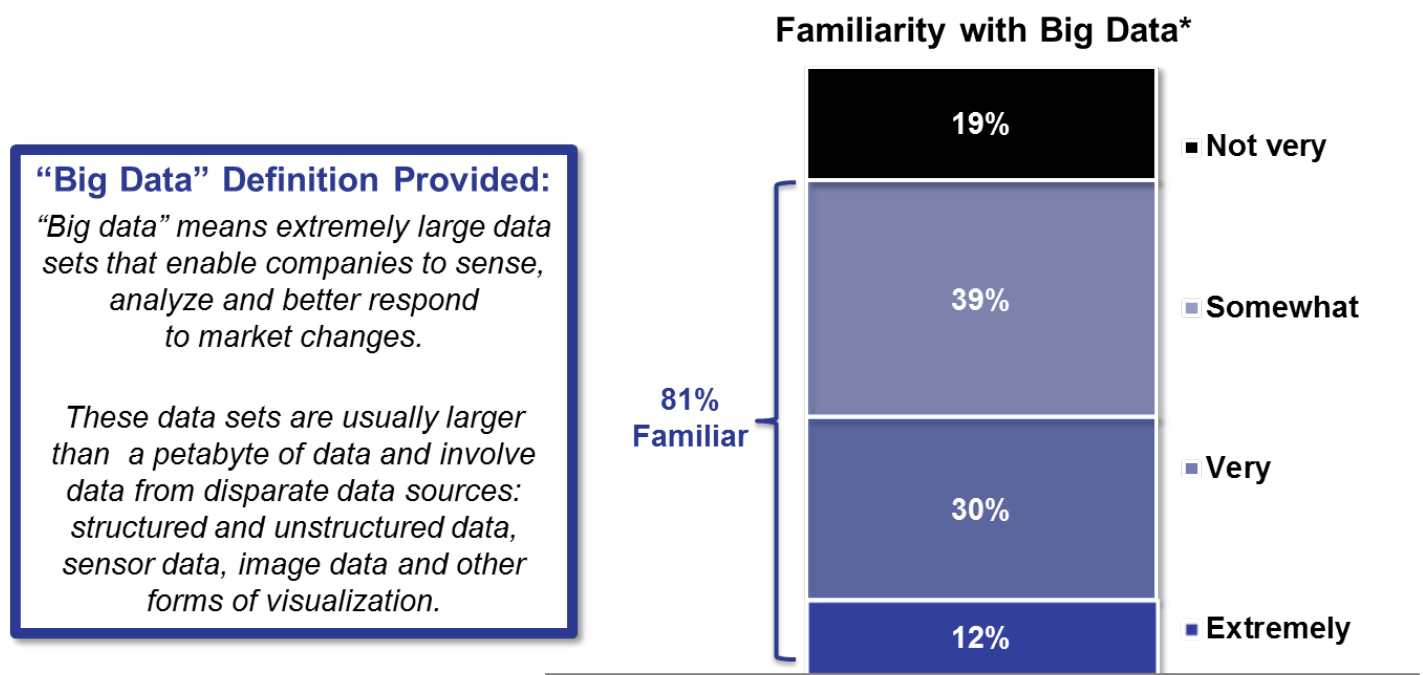
Different forms of data offer a realm of new business challenges and opportunities. For companies new to analytics, these data forms are often called "dirty" when they are really just "different." They do not fit into conventional relational database applications. The data often has different context and attributes. So while people want to use new data forms, they struggle with how to get started.

# What Is Big Data?

Let's start with a definition. For the purposes of this study, big data is defined as data with a volume greater than a petabyte coupled with a growing variety of data. A petabyte of data sounds like a lot of data; but, how much is it really? A petabyte of data is 1024 terabytes. A terabyte of data is 1024 gigabytes. In more graphical terms, a petabyte of data represents twenty million four-drawer filing cabinets filled with text or the storage of 13.3.years of HD-TV video. Twenty petabytes represents the amount of data processed by Google on a daily basis. It also represents the total hard drive space manufactured in 1995.<sup>i</sup>

As shown in figure 3, companies cannot effectively use data today within their organization; and as will be shown later in this report, they also do not have the levels of data volume that would qualify for the big data definition used in this study. They are just beginning to understand and use the concepts. Those with a big data initiative are more likely to consider their ability to use data to be one of their top business pains: 57% versus 38% for those without a big data initiative. As shown in figure 4, the concepts are emerging.

**Figure 4. Currently Level of Understanding of the Concepts**

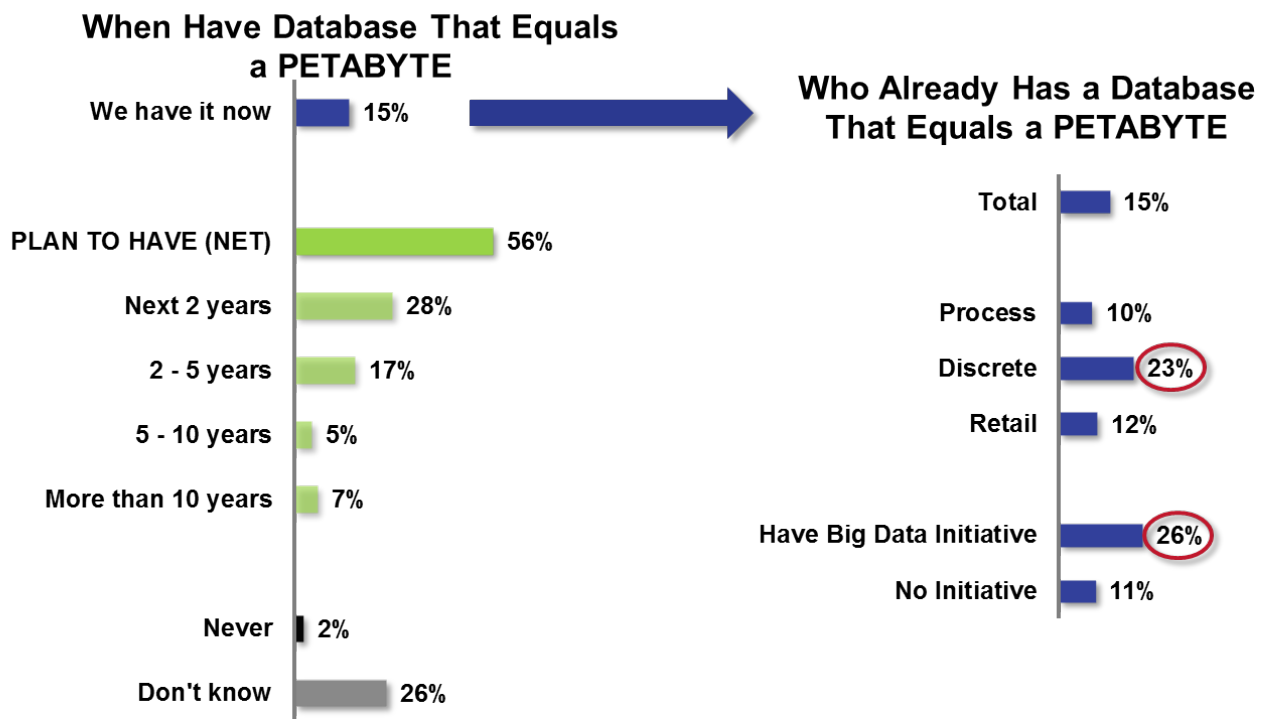


Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with “big data” – Total (n=123)  
\*NOTE: Respondents had to be at least “not very familiar” to qualify for the survey  
Q1D. In general, how familiar are you personally with “big data?” SCALE: 1=Not at all familiar, 5=Extremely familiar

So, where are we in the growth of databases? Do we currently have a big data problem? In our survey, 15% of respondents report having a database that equals or is greater than a petabyte of data. In addition, another 28% of respondents expect to have a database this size within the next two years. Based on our interviews with clients, these large databases are usually associated with customer and product data and the mining of

customer insights. It is less likely to be ERP. While ERP databases are growing, only 18% of the study respondents currently have an ERP database that is more than 6 terabytes.

**Figure 5. Current Database Size within the Corporation**



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)

Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123), Process Industry (n=40) Discrete Industry (n=48), Retail Industry (n=26- SMALL BASE SIZE), Have Big Data Initiative (n=35), No Big Data Initiative (n=88)

Q6a. How soon do you expect your company will have a database that equals a petabyte?

○ Higher than other group at 90% or higher level of confidence

## Companies Focused on Clean Data Today Are in a Better Place

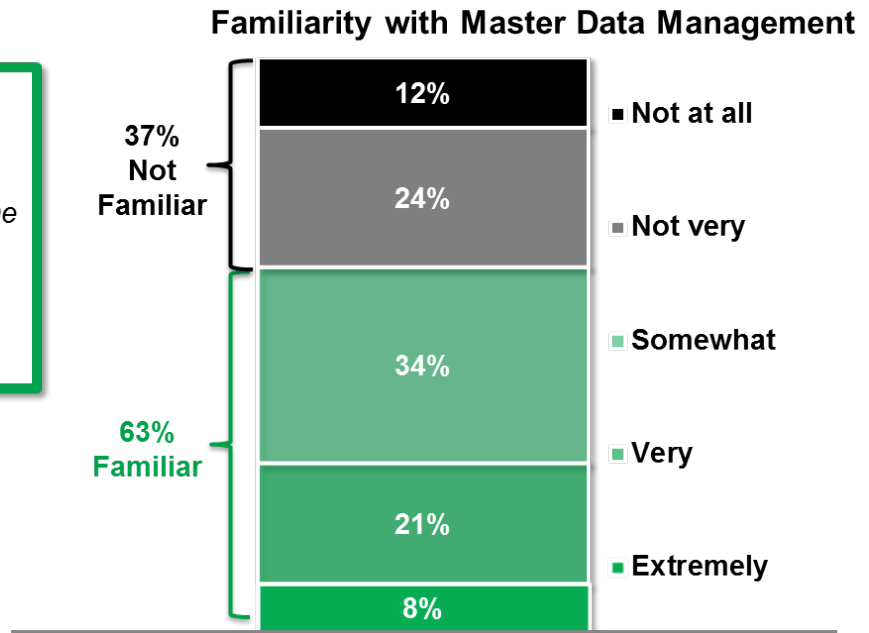
Respondents with a master data initiative are nearly three times more likely to have a big data initiative. For most companies, Master Data Management (MDM) challenges are not new. MDM initiatives are ongoing. Thirty-two percent of survey respondents currently have an MDM initiative, and 30% plan to have an MDM program in the future.

In working with clients, we can clearly see that the understanding of MDM is more mature than the evolution of big data concepts. Companies with maturity in MDM have a leg-up in the use of "different data": the explosion of data volume and types.



Figure 6. Survey Respondent Familiarity with Master Data Management

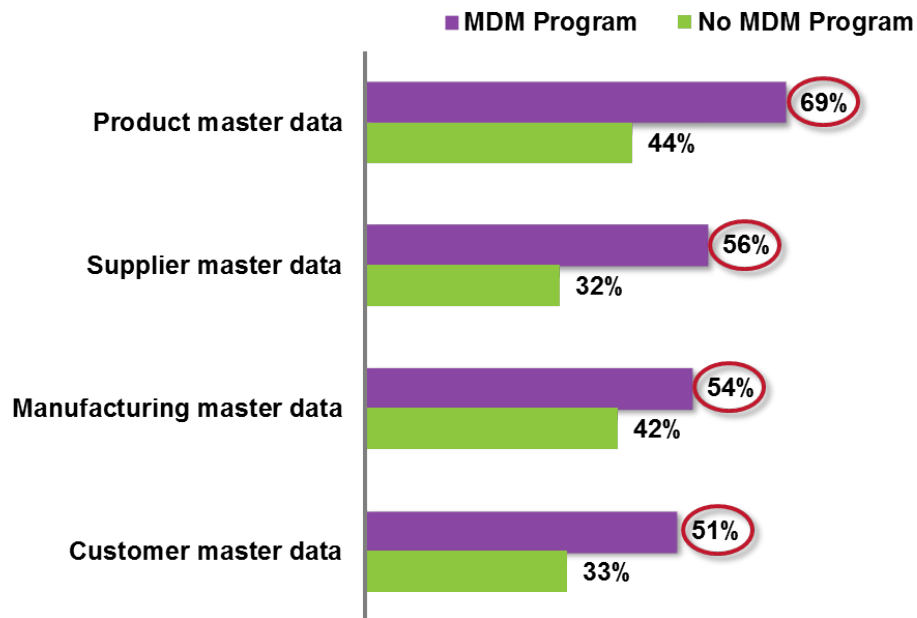
**“Master Data Management” Definition Provided:**  
 Master data management (MDM) programs are designed to improve the management of structured “master data,” which includes data about customers, products, employees, materials, suppliers, etc.



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with “big data” – Total (n=123)  
 Q22A. In general, how familiar are you personally with master data management?

Figure 7. MDM Performance by Data Type

**Performance in Using Types of Master Data – by Presence of MDM Program (Rated 5-7 on 7-Point Scale)**



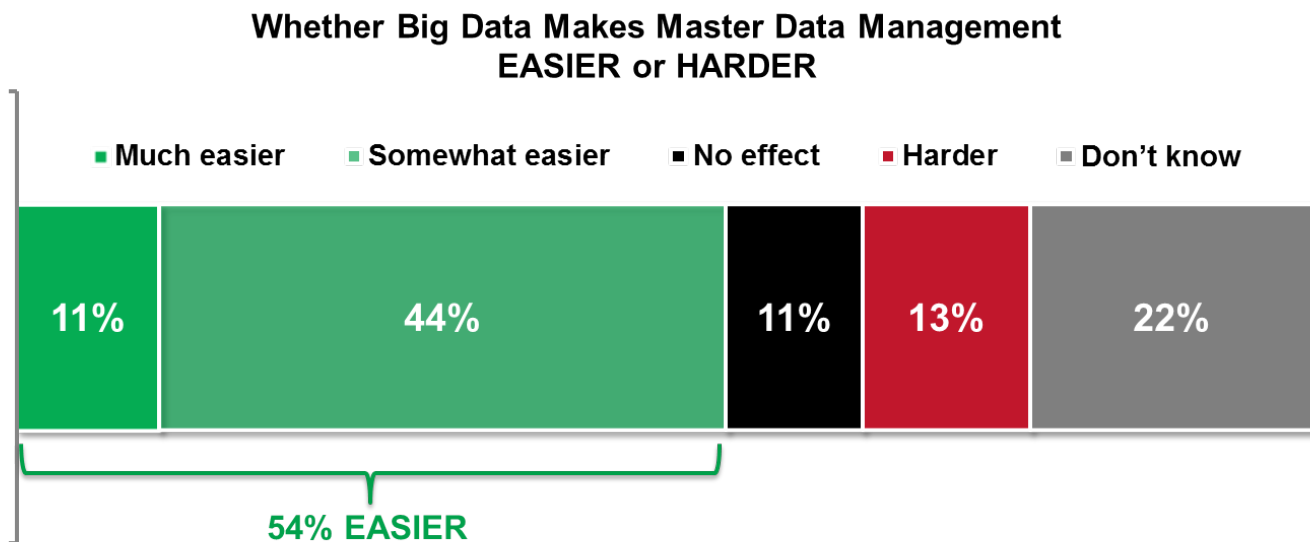
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with “big data” – Have Master Data Management Program (n=39), No MDM Program (n=84)  
 Q26. How would you rate your company’s current performance in terms of using the following types of master data? SCALE: 1=Poor, 7=Excellent

When companies focus on understanding data today and cleaning data for corporate use, as shown in figure 7, they self-assess their ability to use reference data significantly higher than those who do not have an MDM program.

Companies with well-established MDM programs understand the intricacies of data management better than those that don't. Traditional MDM techniques are problematic. Packaged approaches in ERP have not met organizational requirements.

Companies with experience understand the needs for data context and attribution. They understand that the problem is more complex than what can be delivered in one-to-one rules engines (Business Activity Monitoring (BAM)) or the harvesting of unstructured data through Natural Language Processing (NLP). Instead, it requires organizational discipline in the management of data and the combination of techniques found in emerging big data analytics to meet the challenge. One of the barriers is that 49% of those with an MDM program report it is headed by IT leaders. They struggle to get line-of-business vision and support.

**Figure 8. The Relationship between Big Data and MDM Programs**



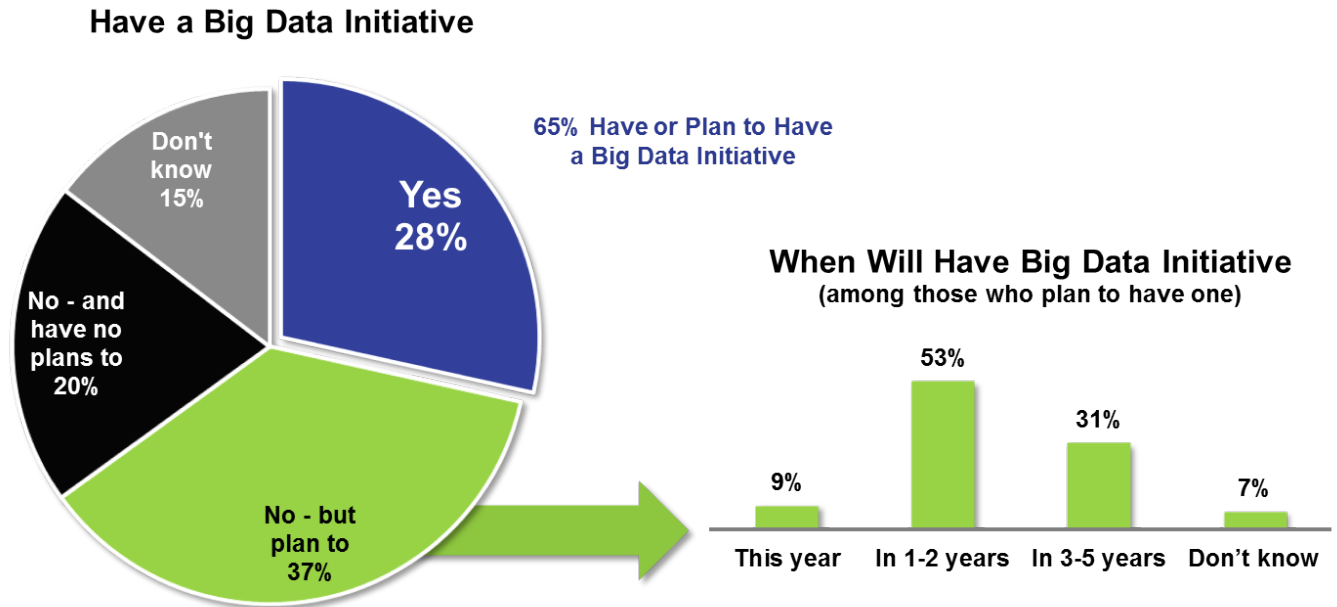
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
Q27. In your opinion, does big data tend to make the management of master data easier or harder?

When we hear the leaders of e-commerce speak on the use of data and their big data programs, they never talk about the pain of MDM. Their work with cognitive learning and advanced analytics support the belief statement in figure 8. We believe, based on interviews with big data leaders, that the approaches with big data help to solve some of the historic and legacy issues of MDM.

# The Opportunity

Big data investments are just starting. Over one-third of respondents in the study consider it one of the top three supply chain trends. Today, as shown in figure 9, 37% percent of respondents plan to have a big data initiative and 28% percent have a big data program today.

Figure 9. Current State of Big Data Initiatives



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)

Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)

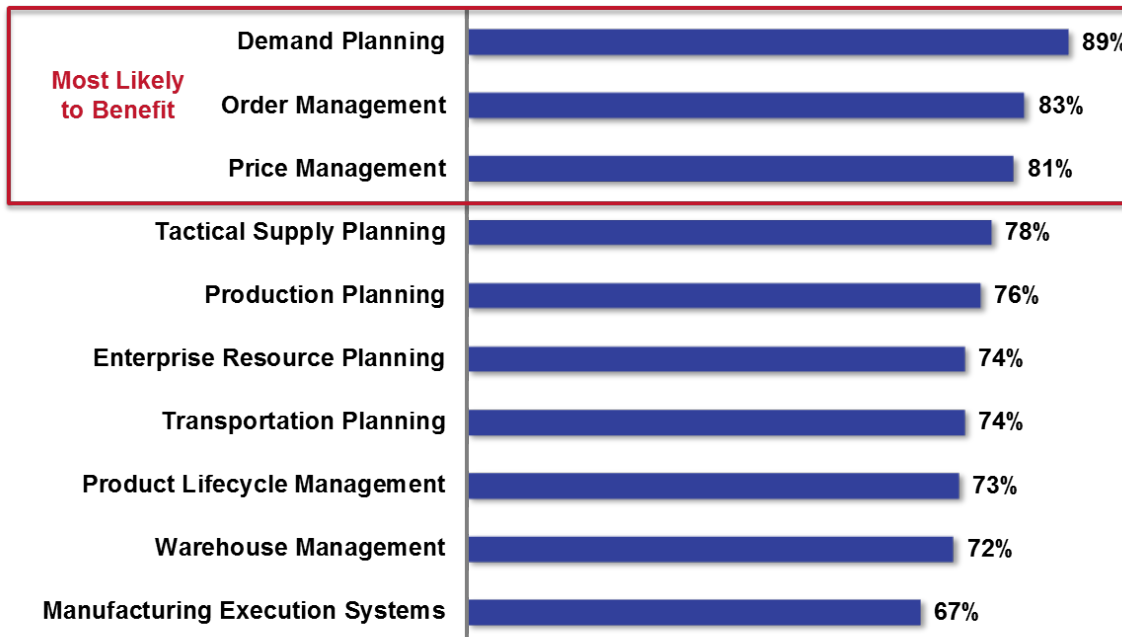
Q7. Does your company currently have an initiative in place to evaluate how to use big data?

Base: Plan to have initiative (n=45) Q7B. How soon do you expect to have a big data initiative at your company?

Big data initiatives are in their infancy. We are at the starting line. The big data opportunity is seen in "demand." The areas of demand management, order management, price management and channel sensing, as shown in figure 10, rate the highest in terms of areas to benefit from big data. However, as shown in figure 11, the highest importance is placed on "supply-based initiatives."

Figure 10. The Big Data Opportunity

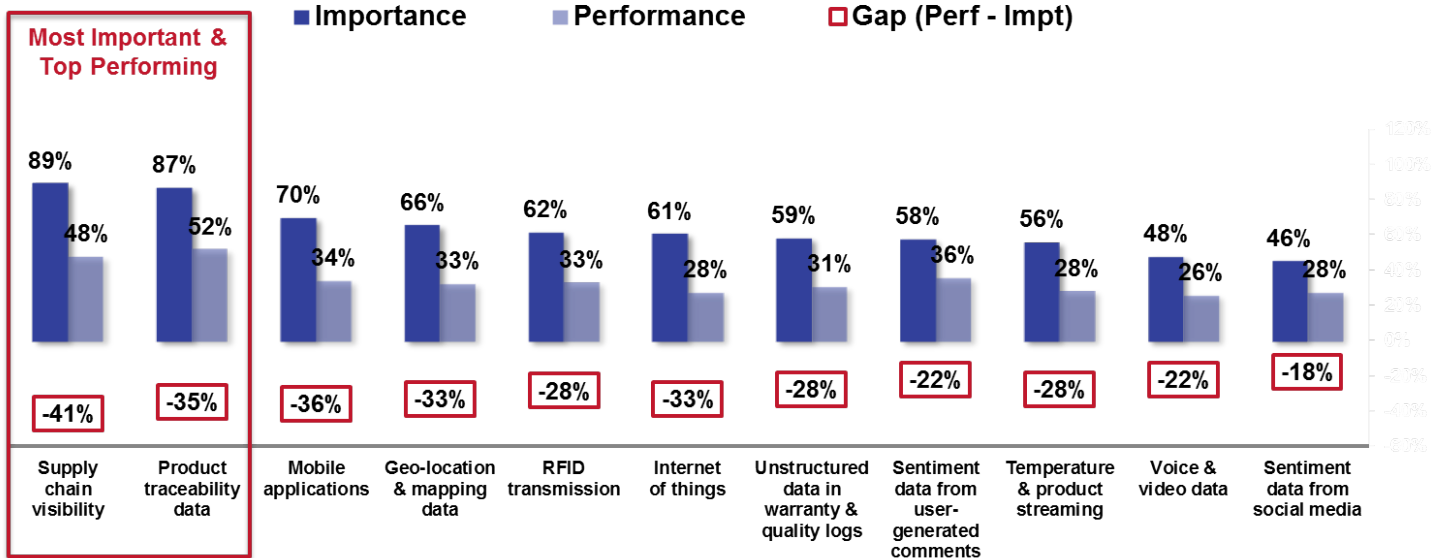
### IT Systems to Benefit from Big Data Strategies (Rated 5-7 on 7-Point Scale)



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q16. In general, how much do you think each of the following IT systems could benefit from big data strategies at your company?  
 SCALE: 1=No benefit, 7=Benefit greatly

Figure 11. Importance versus Performance of Current Types of Data

### Types of Data: Importance vs. Performance (Rated 5-7 on 7-Point Scale)



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q10. How important do you think it is to have each of the following types of data? SCALE: 1=Not at all important, 7=Extremely important  
 Q11. How would you rate your company's current ability to use each of these same types of data? SCALE: 1=Poor, 7=Excellent

One of the barriers to capturing the big data opportunity is the “definition of supply chain.” When the supply chain is seen as a “function”—versus an end-to-end process from the customer’s customer to the supplier’s supplier—the organization is more likely to move forward with big data projects as a functional program.

## Recommendations

The term *big data* is “hot,” and will get “hotter,” as companies realize the opportunity. As companies mobilize for the future, we offer these recommendations:

**Avoid the Hype.** If a vendor or a technology consultant comes in and shares a slide on how their solution helps with big data, ask them, “*What do you mean by big data? And how does your solution help to solve the problems to harvest the big data opportunity?*” In parallel, drive change in the organization by facilitating and elevating discussions on analytics and the use of data. Any time there is an internal discussion on “big data,” ask for the definition and try to “ground the discussion in reality.” Is it a big data opportunity because the data is different? Or is it a big data opportunity due to the volume and velocity of data?

**Focus End-to-end. Build a Cross-functional Team.** Big data offers a wealth of opportunities. It enables market sensing, consumer listening, test and learn strategies, new forms of visibility and safe and secure supply chains. It is an opportunity to use new data forms and emerging analytics to build processes outside-in from the customer back. This can best be accomplished when there is a team of IT and line-of-business leaders that can work cross-functionally with a focus on end-to-end processes. This team is best led by a line-of-business leader and is guaranteed a higher level of success if it does the following:

- **Sidestep Religion.** Do not get entangled in religious arguments of supply chain as a function or an end-to-end process.
- **Start Small and Iterate.** Do not get caught up in the ERP-like mindset of big projects with a series of releases. Focus on small wins and use the learning from the use of analytics to spread to other functions. For example, the use of in-memory reporting from QlikView and visualization from Spotfire and Tableau are being used by a number of our clients to improve data usage today to power organizational support of funding for big data initiatives. Organizations have many technologies and systems and need to free the decisions for analytics from the decisions being made on their systems of record implementations (ERP). ERP is only one source of data, and over time, will become a less significant contributor to the overall response.
- **Provide Innovation Funding.** Give these teams innovation dollars to innovate for the company. Some companies have had success with having departments submit requests for consideration for innovation spending on analytics and the use of different data forms by cross-functional teams working on big data initiatives.

- **Consolidate Business Intelligence Centers of Excellence and Master Data Efforts into Big Data Initiatives with a Goal to Serve the Business.** Organizations that are good at using data will win in driving big data opportunities. The goal needs to be on solving business problems through the harnessing of the cross-functional efforts of knowledgeable people working on teams to solve data and analytical problems and opportunities.

## Conclusion

Big data is more of an opportunity than a problem. Companies need to mobilize now to build new processes and differentiated services. The focus needs to be on what the end-to-end supply chain can be. To do this, companies need to free themselves from the limited definition of what it is today. As a result, some organizations are calling it the “digital supply chain.”

It is a great opportunity to build processes outside-in and harness new channel opportunities. It is also critical to meet new supply chain challenges for safe and secure and temperature compliant supply chains. It is foundational for digital manufacturing and solving the tough issues of Corporate Social Responsibility (CSR).

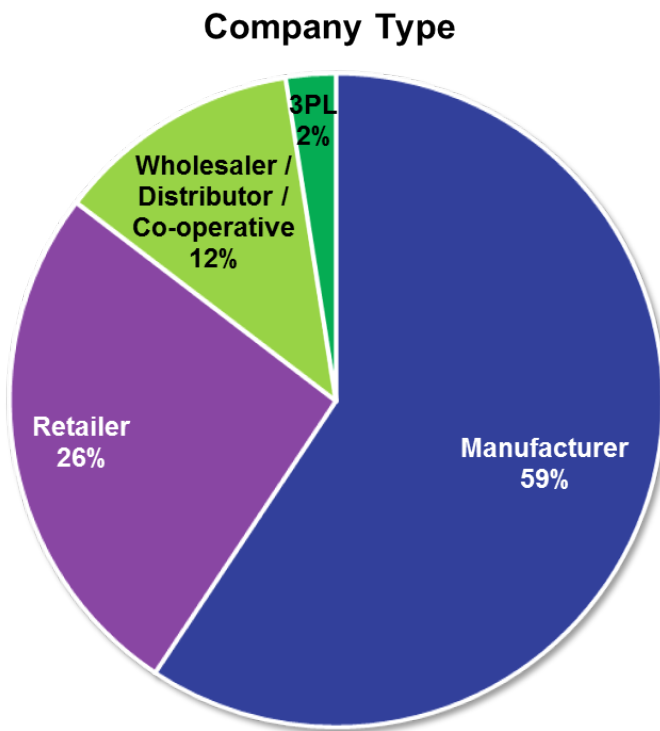
The possibilities are endless. It starts with thinking about how to use different data in a differentiated way.

## Appendix

In this section, we share the demographic information of survey respondents. The respondents answered the survey questions of their own free will. There was no exchange of currency. The only offer made to stimulate a response was to share the survey results in the form of Open Content research at the end of the study.

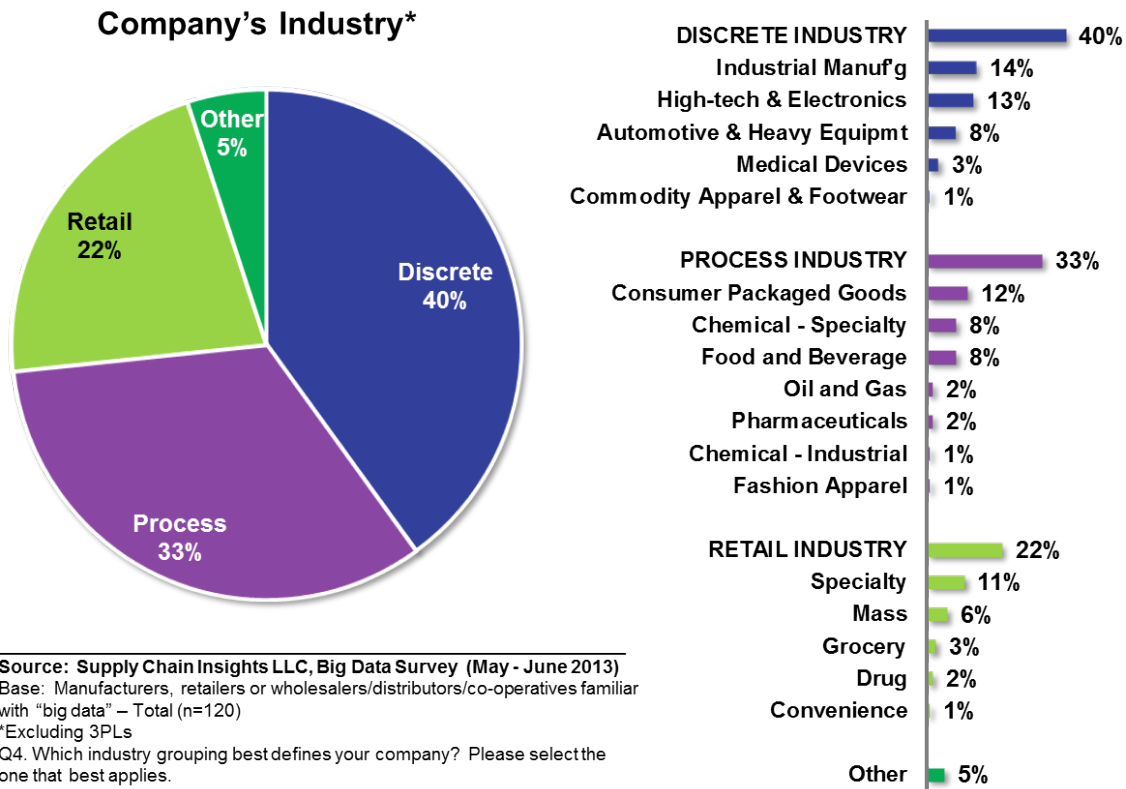
The names of those that completed the surveys are held in confidence, but the demographics are shared to help the readers of this report gain perspective on the respondents. The demographics supporting these figures are found below in Figures A-L.

**Figure A. Survey Respondents by Type of Company**

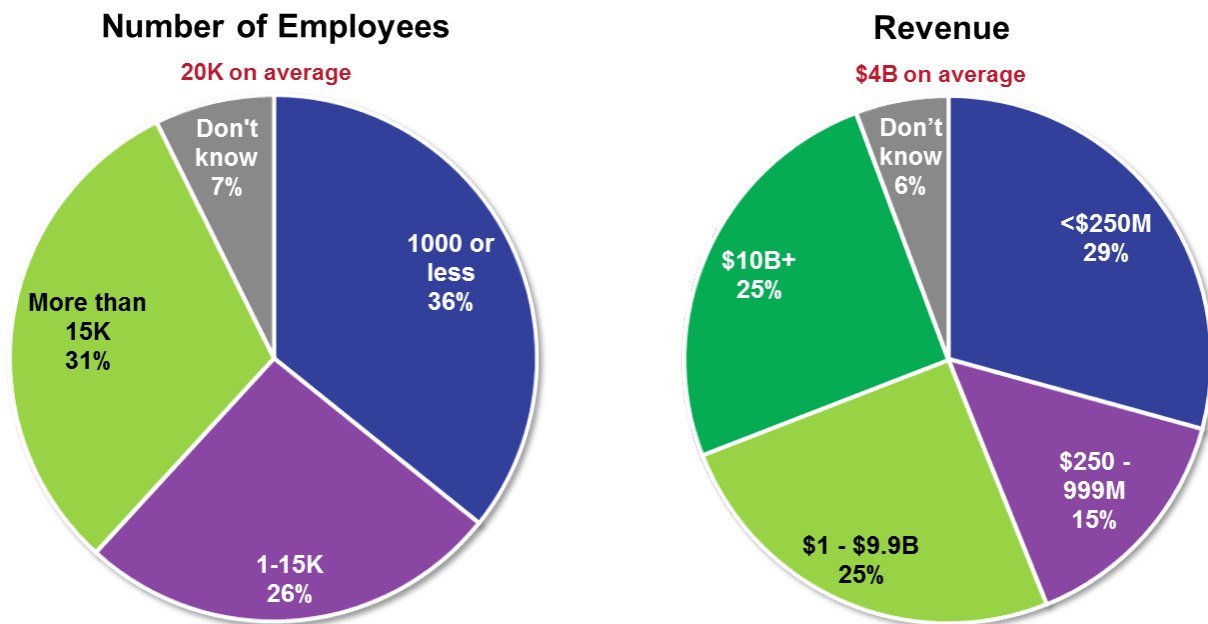


**Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)**  
Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
Q1A. Which of the following best describes you or your company? Please select the one that fits best, even if the terminology isn't quite right.

**Figure B. Breakout of Respondents by Industry**



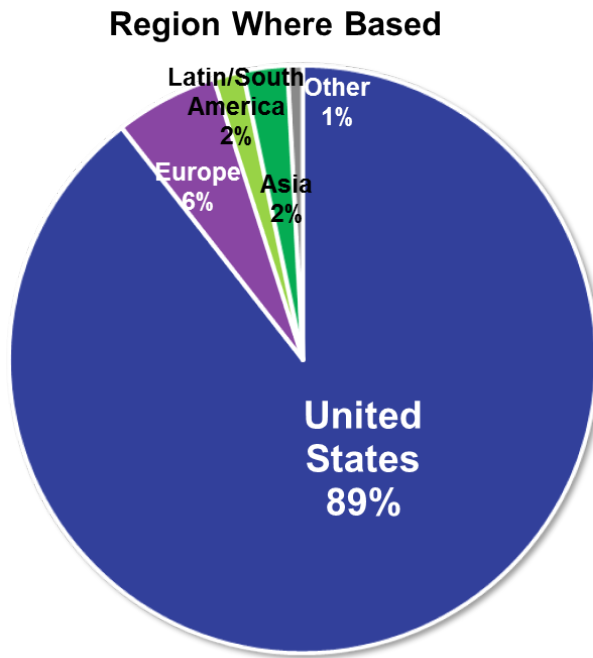
**Figure C. Survey Respondent Demographic Data on Company Size**



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q3. What is the size of your company, in terms of number of employees? Your best estimate is fine. NUMERICAL RESPONSE  
 Q3A. Approximately, what was your company's last fiscal year revenue? Your best estimate is fine.

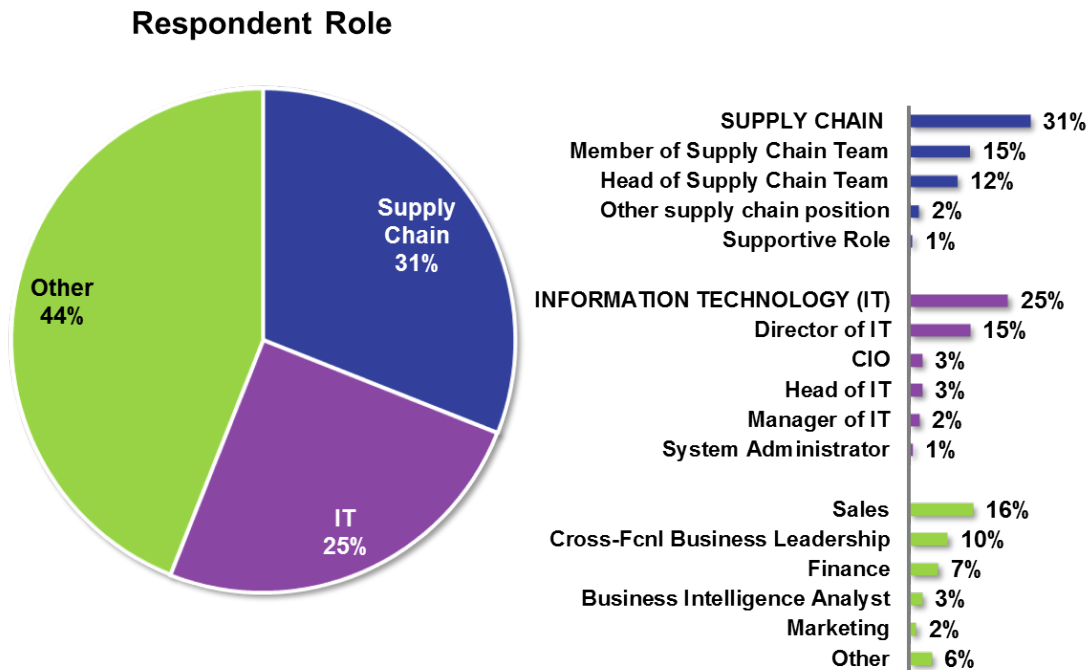


**Figure D: Geographic Respondent Information**



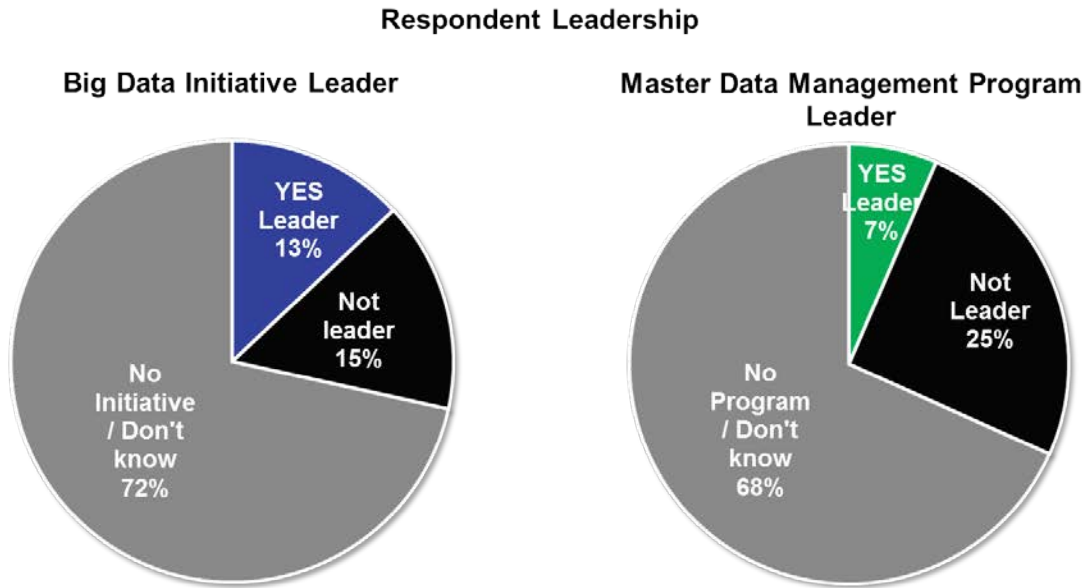
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q4A. In what region of the world are you personally based at this time?

**Figure E: Respondent Demographic Information by Role**



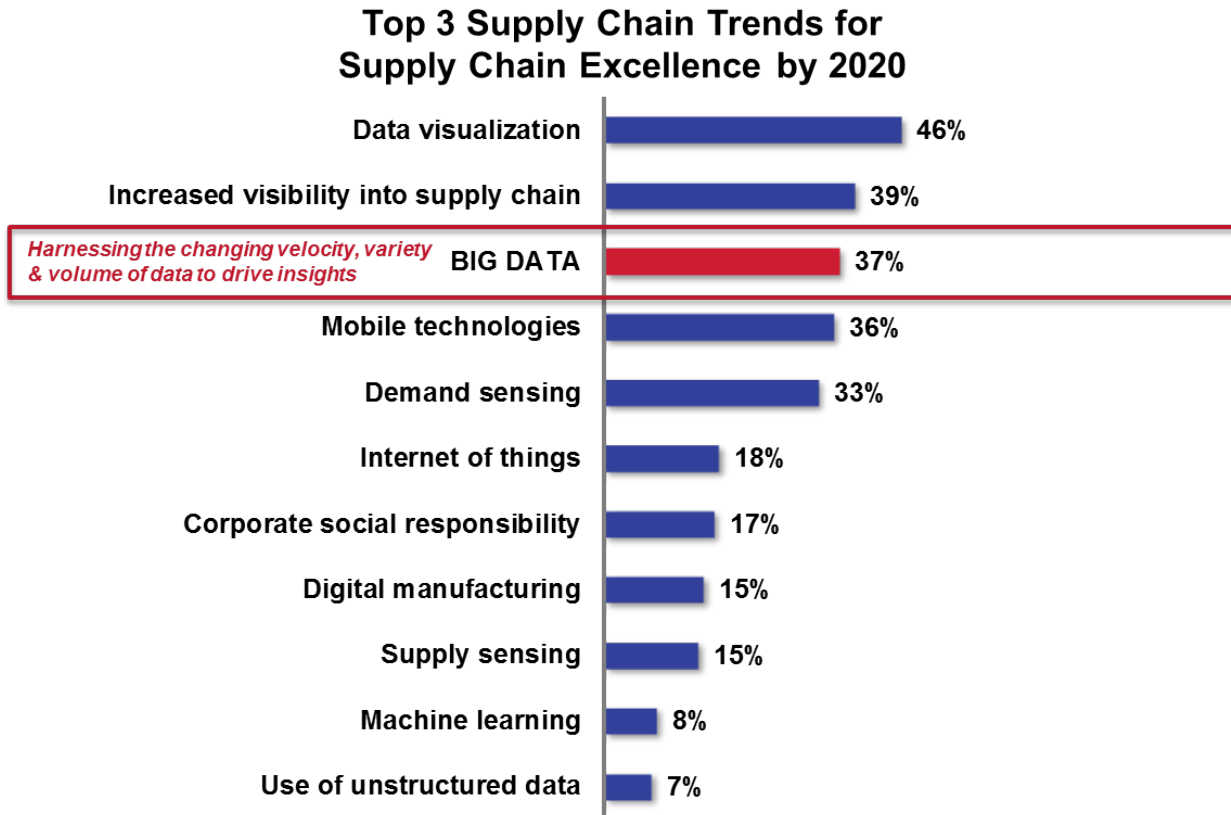
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q1B. Which of the following best describes your current functional role? Please select the one that fits best, even if the terminology isn't quite right.  
 Q1. Which of the following best describes your current title or position?  
 Q1C. Which of the following best describes your current title or position?

Figure F: Respondent Leadership of Big Data and Master Data Management Initiatives



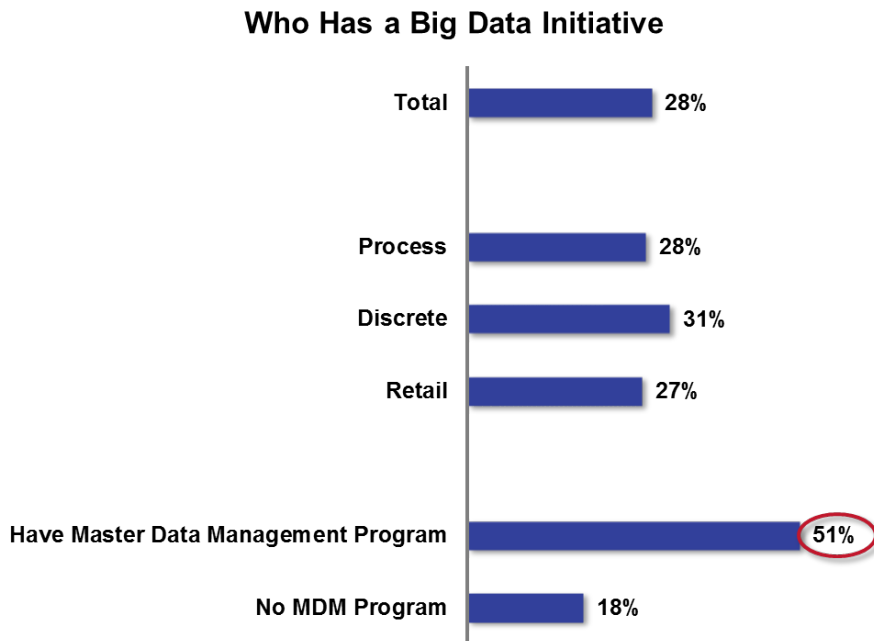
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q8. Are you personally leading the big data initiative?  
 Q24. Are you personally leading the master data management program at your company?

Figure G: Supply Chain Trends



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q19A. What are the three supply chain trends that you are most excited about, when it comes to driving supply chain excellence by the year 2020?

**Figure H: Companies with a Master Data Program are Three Times More Likely to have a Big Data Initiative**



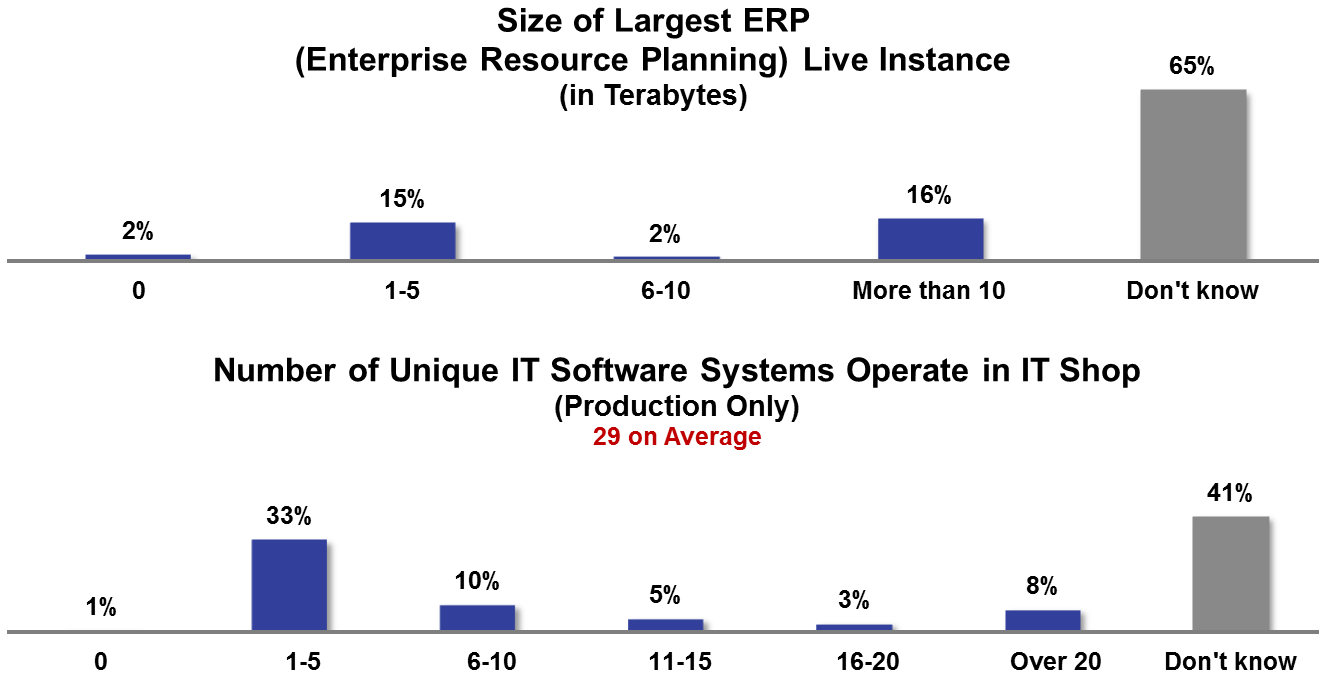
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123), Process Industry (n=40), Discrete Industry (n=48), Retail Industry (n=26—SMALL BASE SIZE), Have Master Data Management Program (n=39), No MDM Program (n=84)  
 Q7. Does your company currently have an initiative in place to evaluate how to use big data?  
 ○ Higher than other group at 90% or higher level of confidence

**Figure I: Big Data Initiative Leader**



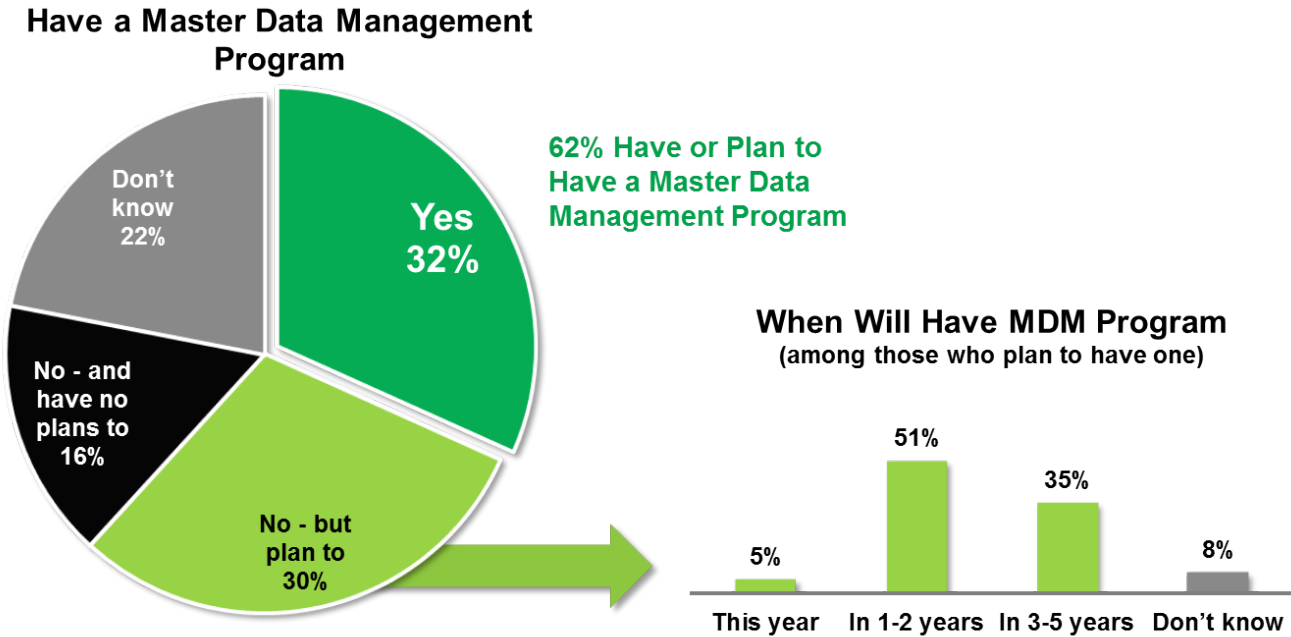
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" and have big data initiative – Total (n=35)  
 Q8. Are you personally leading the big data initiative?  
 Q9. What is the title or role of the person leading the big data initiative?

**Figure J: Size of Largest ERP Live Instance and Number of Unique IT Software Systems**



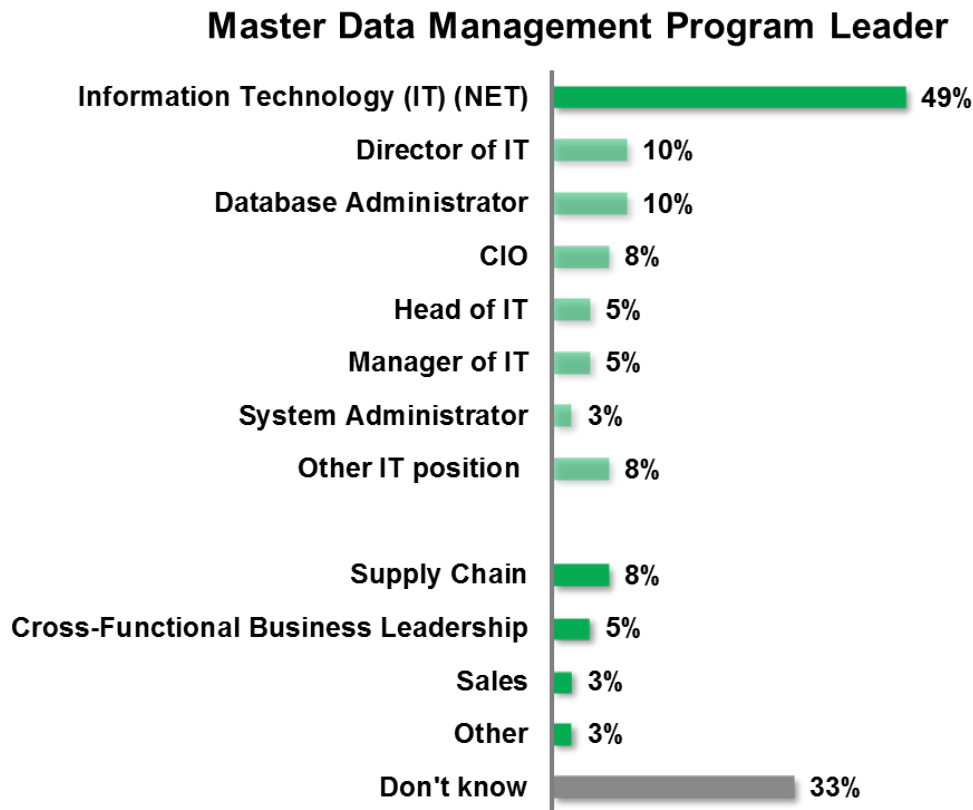
Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q6. What is the size of your largest ERP (Enterprise Resource Planning) live instance? Your best estimate is fine. NUMERIC RESPONSE  
 Q5. How many unique IT software systems does your company operate in your IT shop (production only)? Your best estimate is fine. NUMERIC RESPONSE

**Figure K: Current State of Master Data Management Programs**



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)  
 Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" – Total (n=123)  
 Q22. Does your company currently have a master data management (MDM) program?  
 Base: Plan to have MDM program (n=37); Q23. How soon do you expect to have a master data management program at your company?

Figure L: Master Data Management Program Leader



Source: Supply Chain Insights LLC, Big Data Survey (May - June 2013)

Base: Manufacturers, retailers, wholesalers/distributors/co-operatives or 3PLs familiar with "big data" and have master data management program – Total (n=39)

Q24. Are you personally leading the master data management program at your company?

Q25. What is the title or role of the person leading the master data management program at your company? Please write in title below.

## Other Related Reports

[The Art of the Possible: Actionable Analytics for Value Networks](#)

[Big Data – Go Big or Go Home?](#)

## About Supply Chain Insights LLC

Founded in February, 2012 by Lora Cecere, [Supply Chain Insights LLC](#) is focused on delivering **independent, actionable and objective advice for supply chain leaders**. If you need to know which practices and technologies make the biggest difference to corporate performance, turn to us. We are a company dedicated to this research. We help you understand supply chain trends, evolving technologies and which metrics matter.

## About Lora Cecere



Lora Cecere (twitter ID [@lcecere](#)) is the Founder of [Supply Chain Insights LLC](#) and the author of popular enterprise software blog [Supply Chain Shaman](#) currently read by 5,000 supply chain professionals. Her book, ***Bricks Matter***, (co-authored with Charlie Chase) published on December 26<sup>th</sup>, 2012, and her second book, *Metrics That Matter*, will publish in 2014.

With over nine years as a research analyst with **AMR Research, Altimeter Group, and Gartner Group** and now as a Founder of Supply Chain Insights, Lora understands supply chain. She has worked with over 600 companies on their supply chain strategy and speaks at over 50 conferences a year on the evolution of supply chain processes and technologies. Her research is designed for the early adopter seeking first mover advantage.

---

<sup>i</sup> How Much is a Petabyte of Data? Mozy Blog, <http://mozy.com/blog/misc/how-much-is-a-petabyte/>