



Big Data and Small Data: A Collaborative Affair

Is there a symbiotic relationship between big data and small data, and can it give us better insights into our business? In this eBook, we break down big data into its five V's of behavior and discuss how to harness that behavior to work for you, not against you. We also discuss the value of small data and its long-standing importance in humanizing the corporate process.

Contents:

1. Big Data and Small Data: A Collaborative Affair
An Overview
2. Big Data, Small Data or Both?: A Collaborative Affair
Volume Versus Value
3. Big Data and Small Data: A Collaborative Affair
The Human Element



1 An Overview

Before we delve into the subject, let's first define big and small data, or at least try since agreed-upon definitions for either do not exist. Big data is a term that describes the large volume of data - both structured and unstructured - that inundates a business on a day-to-day basis. These extremely large data sets may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions. Small data, on the other hand, can be defined as small datasets that can be easily stored and are capable of impacting decisions in the present. Either sourced locally or mined from big data, small data "can be organized and packaged, often visually, to be accessible, understandable, and actionable for everyday tasks."¹

In the Digital Age, big data receives all the attention. It's big, it's glamorous, and it has massive stage presence. Data analyzers have even assigned it a five-V acronym to categorize the way it behaves and differentiate it from small data: Volume, Velocity, Variety, Veracity and Value. Small data, well, its behavior has yet to be labeled. Many companies consider small data the runt of the litter, but small data can be essential to the more nuanced, strategic decisions your business will make to ensure its success. In this series, we will not only delve into the five V's of big data but also delve into its much older yet its much less talked about counterpart, small data. Our goal? Understanding the importance of both and how to parley that understanding into a strategic and successful data action plan.

But first, let's become acquainted with the five V's of big data's behavior.

Volume

No surprise, here. Volume refers to the vast amount of data available to you at any given second. Thanks to IoT and the number of connected devices, in the blink of an eye, we've skyrocketed from megabytes and gigabytes to terabytes and now up to peta/exa/zetta/yottabytes. Making sense of these big bytes requires advanced processing techniques and solutions.²

Velocity

Velocity refers to the speed at which new data is both generated and distributed.² Big data rarely even makes it to databases before it is analyzed and acted on. Data can also come in waves, depending on the circumstance, moving from a steady predictable wave stream, such as customer feedback, to an unpredictable mega tsunami, like a video gone viral that prompts a product recall.

Variety

Variety refers to the types of data available to us from structured data (i.e. date, time GPS location, database), semi-structured (i.e. clickstream), to unstructured data (i.e. text, image, voice, or video),² and big data can encompass all three. In fact, about 90% of the world's data is now unstructured,³ and many businesses are now focused on turning unstructured into actionable data.

Veracity

Veracity refers to the trustworthiness of big data, and it is quite possibly the trickiest behavioral aspect.² Quantity does not necessarily mean quality, and inaccurate or messy data often prevents our ability to create value from it.

Value

Value refers to our ability to change big data into a commodity that adds value to our business.² Of all the tera and yottabytes flying at us from different sources at different speeds, only a third or so of it contains valuable information that is actionable.² Luckily, resources exist to mine big data and find the valuable "ore."

If big data tells you the "who," "what," "where," "when," and "how"; small data tells you the "why."

As previously stated, there are no categories nor is there even an agreed-upon definition for small data. Most agree that this type of data is small enough to be "accessible, understandable, and actionable,"¹ basically pulling us back into our comfort zones of mega and gigabytes. Unlike big data, small data can often be housed in one database. Although its size is smaller and its speed is slower than big data's, its role in driving your business decisions can be just as vital. As an old German proverb once said, "God is in the detail"; small data allows us to dissect structured or unstructured data to provide us with its hidden meaning.

So, how can you approach big and small data collaboratively, ensuring both data sets are communicating effectively with each other and with you? Ultimately, our recommendation is to acquire the tools to garner both data sets. And partner with a data-savvy firm, like Copley, to ensure your data provides you with the true customer story. Knowledge is power, so stay tuned.

1. Banafa, A. (2016, July 25). Small Data vs. Big Data: Back to the Basics. OpenMind BBVA.

2. Marr, B. (2015, March 19). Why Only One of the 5 Vs of Big Data Really Matters. IBM Big Data & Analytics Hub

3. Marr, B. (2019, October 16). What Is Unstructured Data, and Why Is It So Important to Businesses? An Easy Explanation for Anyone. Forbes



2 Volume Versus Value

Mining Big Data for Value, Literacy and Sense

Think of big data as a rocky mountain, looming over your business, housing tons of valuable ore. You know the ore is there, but not how to find it. Mining big data for valuable and actionable small data poses a challenge for many SMBs and manufacturing companies, and there are many “mining packages” out there promising they are the best at data exploration and extraction. So, when it comes to approaching your mountain, how do you proceed? The best approach is the same approach a mining company utilizes to extract ore from a mountain: in well-planned phases.

Phase 1: Hire the experienced miner to help you establish your goals:

First, if you aren't an experienced data miner, you need to start by hiring one, meaning, find the right consultant. Before you even begin the mining process, you need to clearly define what it is your company needs to know from big data. A consultant will guide you in asking the right questions, allowing you to better define your true business needs before you waste time exploring the wrong parts of the mountain. That big mountain contains more invaluable than valuable data, and the former can distract you from the true course. The best way to define your data mining goals is by first defining your business objectives. Once you've established your goals and objectives, your consultant can help you lay out a detailed mining plan, a blueprint so to speak, centered on finding the data that you truly need. Once you've hired your expert, you can proceed to phase two...

Phase 2: Begin your prospecting (or data exploration):

Without clear goals and objectives, it is easy “to embark on big data initiatives without a clear understanding of the business value it will bring.”⁴ Think of this phase as if you were exploring your mountain, looking for the “veins” that contain the most valuable deposits of data, the data you have targeted to help you fulfill your established needs. Data exploration, or prospecting, begins with “exploring a large set of unstructured data, looking for patterns, characteristics, or points of interest. Summarizing the size, accuracy and initial patterns in the data is key to enabling a deeper analysis.”⁴ This phase may sound daunting, but your consultant will employ key solutions designed for this very process. Once you've located where your “mineral deposits” are, it's time to build.

Phase 3: Build the mine and extract the ore:

You cannot extract the valuable deposits if you don't build the mine. There's no easy way to say this: if you don't invest in the structure and the software that can drill down and do the work, everything else is for naught. It's just like mining for minerals. After a mineral deposit has been identified through exploration, one must make a considerable investment in mine development before production begins.⁵ It's no different with your mountain of big data. Building a mine is akin to updating or upgrading your legacy system so it can deploy solutions that target and extract valuable data from the mountain. Yes, data extraction can be a costly investment, but luckily, your consultant will ensure your investment yields the best results. Once you've extracted your valuable data, you can enter the final phase.

Phase 4: “Smelt” your big data yields into small data:

“Smelting” your big data into a valuable commodity that will do the most for your business is the final step of the mining process. In mining, smelting uses “heat and a chemical reducing agent to decompose the ore, driving off other elements, such as gasses or slag, and leaving [only] the metal behind.”⁶ What a great word to translate into the technology world for extracting small data from big data. It's time to make sense of the data you've invested in retrieving and “smelt” it into valuable, actionable intelligence. Think of your analytics software tools as the reducing agents that carry out this final process. All the unnecessary slag and gasses not needed from big data is funneled away, and only the valuable remains, your small data. Once smelted, you can package it into accessible, understandable, manageable, and actionable reports, metrics, and alerts for your company's key decision makers and employees. But one thing to remember: one department's “slag” is another's treasure, so be careful what you discard. Analytics tools, such as Qlik, can assist in locating data that other query-based solutions often overlook, data known as grey data, and make sense of it to companies. But we'll delve into grey data in a future discussion.

Overall, it is easy to fall into the buzz trap and embark on big data initiatives without a clear understanding of the business value they will bring.”⁴ If organizations dig haphazardly into their mountain of big data, looking blindly for the good stuff, they can easily be buried in the rubble before striking it rich. Developing a successful data mining program before you break ground will ensure your mountain of big data is transformed into actionable information vital to your company's success.

4. Import.io. (2019, October 8). What is Exploratory Data Analysis and Why is it Important? How to use Data Exploration to Gain Insights for Your Organization.

5. Committee on Technologies for the Mining Industries. (2002). Technologies in exploration, mining, and processing in evolutionary and revolutionary technologies for mining (pp. 19-45). National Academies Press.

6. Lu, L., Pan, J., & Zhu, D. (2015). Quality requirements of iron ore for iron production. *Iron Ore*, 475-504.



3 The Human Element: Building Bridges

You may have heard the saying, “big data” is about machines while “small data” is about people, and the saying still has merit. Yes, big data can be quite distracting: it’s diverse, it’s constantly flowing, and it contains incredibly helpful information to guide and support your business objectives. But don’t let it pull you entirely away from ultimately what’s important, the human element.

While advanced algorithms can lead to better machine learning or AI, better customer insights and better forecasting, they can also remove extensive human understanding into core data sets and create a divide between those who embrace big data and those who don’t. It’s important to bridge the gap between humans and machines in a world that is being driven by powerful data.

Building a Bridge Between Your Data and Your Company

The continued adoption of big data technologies, cloud services and machine learning or AI have provided an unprecedented opportunity for businesses to experiment at scale and cost effectively,⁷ but often at the cost of human judgment and thorough understanding. Often, businesses find themselves reliant on analytics software powered by learning algorithms they don’t understand. As a result, companies leave these algorithms to their own devices and often accept their insights and conclusions as gospel.⁸ It’s important to find and employ the talent, whether it be consultants or data scientists, to not only translate big data into value but also to apply ethics and morality to data-driven dilemmas, something algorithms cannot currently do. Ultimately, these teams should be able to build a bridge between data and human understanding.

Another bridge that employers must build within the company is the one between the data believers and the data skeptics. The most successful enterprises “foster a culture designed to promote collaboration and data analytical skills” by building teams that consist of both IT and business execs.⁸ Each employee brings different skill sets to the table whether it be data analytics or organic customer/employee experience, and collaboration between this diverse culture results in decisions that have a high percentage of buy-in.

Building a Bridge Between Your Company and Your Customer:

When we say small data is about “people,” we are generally referring to users, customers, and their behaviors. Small data highlights the reason why behind the trends of big data and therefore can be very insightful and actionable. But again, small data should not only be mined from big data, it should also be mined from genuine human interaction.

New York Times best-selling author and public speaker Martin Lindstrom explains small data in its most human form. Many years ago, to interview IKEA founder Ingvar Kamprad (now deceased), he went into one of his stores in Stockholm, Sweden to meet up with him. Kamprad was nowhere to be found in the main offices, so Martin asked the staff, “Where is he?”

“Well, he’s probably at the usual spot,” they replied.

“Where is that?”

“...The checkout.”

Lindstrom went down to the checkout, and there stood Kamprad, sitting behind one of the cash registers and ringing up his customers’ purchases. Lindstrom said to him, “Why are you doing that?”

He replied, “Because this is the cheapest and the most efficient research ever. I can ask everyone why they choose it and why they didn’t choose it.” Kamprad’s response is the essence of how good business leaders can continue to embrace the human element of their businesses.⁹

There is only so much the data and an algorithm can determine for you. Speaking directly to the consumer or having more intimate knowledge will help one understand why someone purchased or chose one product over another. Was it based on size? Color? Or some other differentiating factor that AI or machine learning may not pick up. Yes, small data can be gathered by mining big data, but that doesn’t allow for true customer interaction. An algorithm cannot compute a satisfied smile, it cannot monitor a pivotal conversation between sales associate and customer, and it cannot “hang” around the water cooler” to gauge employee satisfaction. Let’s face it, there are still some sacred places where human interaction dominates.

Ultimately, building bridges between your data and the human element builds trust and understanding. And finding the right data scientists and consultants to interpret that data frees you up to leave the computer screen to visit the water cooler, the shop floor, even the checkout line for a more organic assessment of your company’s success...and the occasional chocolate-covered donut.

7. Lane, A. (2020, August 9). Why analytics still needs the human element: Analytics can technically be used by itself, but needs the human element. Innovation Enterprise.

8. Lindstrom, M. (2016, March 24). Why small data is the new big data. Knowledge@Wharton: Wharten@University of Pennsylvania.

9. Mazzei, C. (n.d.). Addressing “the human element” in data analytics. CIOReview.



About the Copley Consulting Group:

For over 30 years The Copley Consulting Group has delivered consulting and business analytics, to more than 400 enterprises. From Fortune 1000 companies to start-up operations, Copley has provided education, training and technical services melded with a focus on best practices. As one of Infor's premier Gold Level Channel Partners and a Qlik Elite Partner, Copley has assembled a team of dedicated professionals committed to increasing the productivity and profitability of our customers.

For more information about our Infor Cloudsuite Industrial (SyteLine) or our Qlik Analytics Solution, contact a Copley Consultant at [855.884.5305](tel:855.884.5305) or sales@copleycg.com.

Follow us on social media, click on the icon below:



facebook.com/copleycg



twitter.com/CopleyCG



linkedin.com/company/copley-consulting-group



youtube.com/user/TheCopleyConsulting

