Big Data Insecurity:

The forgotten Security Landmine

December 21, 2017

In this era technology advancements are moving at a breathtaking pace with businesses looking for more innovative ways to gather insights about the massive amounts of data that is being collected from web, social media , mobile, internet of things systems and other data sources . The drive to be manage and maintain connectivity and interconnectivity to the internet, B2C (Business to Consumer), B2B(Business to Business) has driven organization to generate mountains of data that goes a long way to help build efficiencies, maintain better relationships, generate better marketing , advertising and sales capabilities etc..

These aforementioned reasons have led most mid-size to large scale organizations to look into building Big Data Infrastructure to take advantage of the explosion in connectivity as we know it today. The traditional database is unable to hold such massive amounts of data both from capacity as well a as the fact that some of the data is also unstructured. This has led to the rise of Big Data technologies to address these needs.

Big Data technologies addressed these opportunities by providing better capabilities in handling Volume, Variety, Velocity and Veracity of Data as shown below:

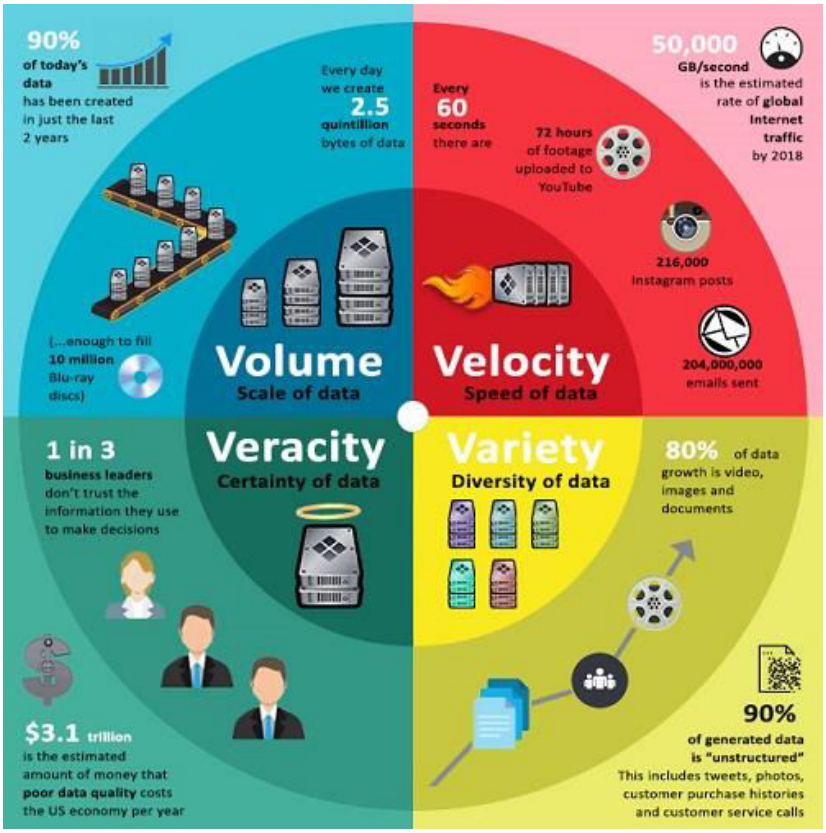
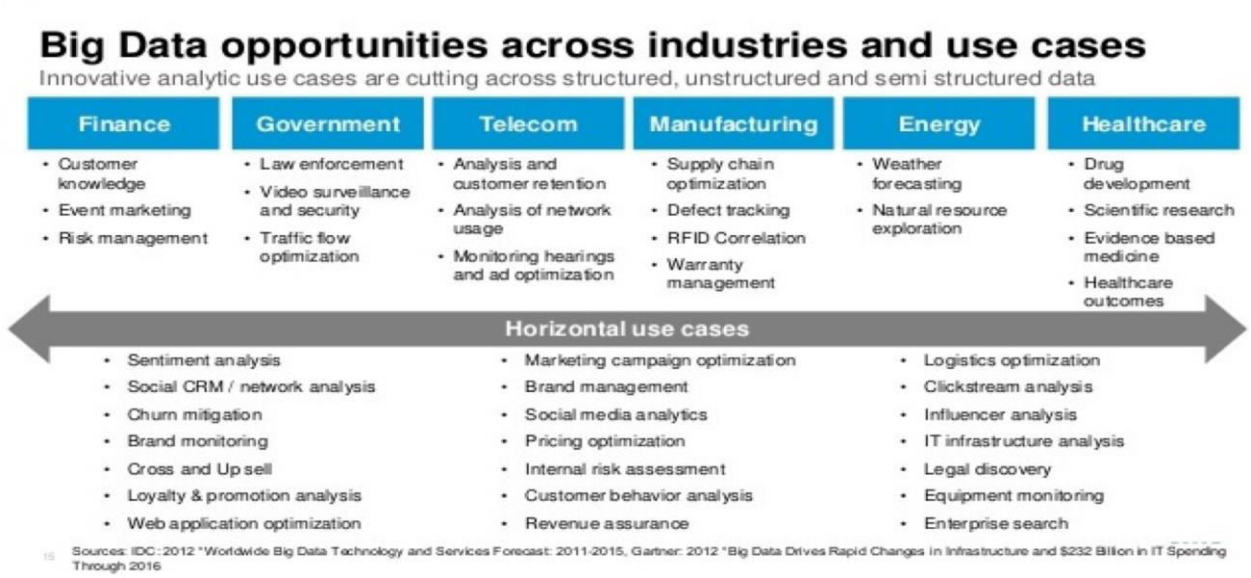


Fig 1.0 Big Data capabilities

These capabilities are excellent and its applications are also enormous in solving several complex problems across several industries..from healthcare, education, security(logical and physical), energy ,e-commerce , government etc. as shown below:

Fig 1.1 Big Data opportunities

With every opportunity there is a related risk. As with every aspect of technology developed for several decades, capabilities have always being developed ahead of security which seems to still be the case with Big Data. Several regulatory organizations have no guidance available as it relates to Big Data, thus



making it almost impossible for security professionals to address security concerns with Big Data. The opportunities with Velocity, Variety, Volume and Veracity can easily be represented as its risk. From a hacker's perspective you have massive Volumes of Verified Data of different Varieties being committed with great Velocity into a centralized system . Such a centralized system is what you will consider as a **hacker's goldmine**. A hacker does not have to do the hard work of figuring out which databases have sensitive data and try to attack those, this time everything is neatly represented in one place for a hacker to access what every data, what ever variety, what every classification etc..

At a recent conference with over 150 security practitioners, a quick poll showed at least 25% or organizations have a Big Data initiative. Strangely enough when I asked how many of the people in the room were involved in their organization's Big Data initiative I received a response of zero. These results went a long way to confirm my concerns about why Big Data continues to remain a major security landmine for most organizations, Big Data teams are often siloed from the rest of the organization and most often than not security teams have no visibility into the work and systems involved in such efforts.

Major security issues with Big Data:

There are several open issues with Big Data security apart from the fact that security industry offers very little guidance on the issue and has thus not created the necessary awareness to drive security people to address such gaping risks.

-Security capabilities is not native for Big Data systems. The security options and capabilities currently implemented in RDBMS systems are far advanced than currently available in Big Data systems

1. Limited natively designed security

Encryption, masking, policy control, compliance and risk management

Security controls like TDE(Transparent Data Encryption) , Data Masking, Data Redaction are controls that can be commonly found in traditional RDBMS cannot be equally found in Big Data systems. With the Variety of sensitive data from varied sources into Big Data systems presents a glaring need to ensure data is appropriately identified , properly classified and correctly secured.

2. Anonymity and Privacy

The type of data collected in big data systems has several issues of Anonymity and Privacy concerns. The reams of data and metadata collected from user activities on mobile apps, web apps, IOT systems, AI systems all leads to a huge explosion of data points about a user which has the potential of easily identifying a user with astonishing accuracies without ever collecting PII. Organizations must therefore look into extending protection of customer and user data beyond simply PII , user behavioral data and address Anonymity and Privacy concerns with both Data and Meta Data.

3. Varied and Complex

Sources: Server data, email data, cloud apps and mobile device data

Data structure: Structured and Unstructured Data

Data consumers: High-level executive, B2B, B2C

Some of the challenges organizations face is the various data sources from File systems, RDBMS, Cloud Storage systems(Amazon S3 etc..), Cloud apps and the fact that data types are so varied from driver licenses, images , text, RDBMS, files, emails etc.. makes identifying and securing such Data quite complex and almost impossible. Multiple consumers exists for Big Data systems and care will have to be taken to ensure user access rights are properly handled

4. Gold mine and target for data theft

Large amounts of data from multiple data sources pushed into a central system

Large amounts of data from multiple data sources and systems are pushed into a centralized system. A Big Data system basically centralizes all the data that a hacker will need from any organization in one place which saves an attacker an enormous amount of time and resources to launch a data breach. Instead of trying to attack multiple databases you simply have one single system to train your attacks on.

5. Low skilled individuals available & Little awareness

Few security people have big data experience

Most security people have very little knowledge and experience with Big Data systems and what they do . This issue is quite telling when you consider that most organizations running Big Data systems do not have security teams involved directly in such efforts. The challenge for most security people is that there is very little security guidance from most security regulatory bodies. Very little prescriptive guidance currently exists in the industry.

6. Data Brokers

User identity, behaviors etc..

With Big Data , there is a rich variety of consumer data that is available for Data Brokers. Consumers must push to ensure that their data is properly secured and protected by the Data Brokers as well as the consumers of such services. Always ask what is being done with your data and what protections are around data that you supply to any organization

7. Lack of products on market

Complex problem, difficult to solve

Lastly, one of the biggest challenges that organizations and security teams face is the lack of products on the market that replicate a lot of well established capabilities of traditional RDBMS as well as address the emerging security challenges of Big Data systems.

Sources: http://www.crewmachine.com/wp-content/uploads/2017/02/4Vs-of-big-data.jpg