Big Data – Opportunity for Innovation in Online Retailing – A Case of Amazon

Vinod N. Sambrani, B.E, MBA, Ph.D
Associate Professor,
Kousali Institute of Management Studies
Karnatak University, Dharwad
Karnataka State, India
Email: vinodsambrani@gmail.com

ABSTRACT:
Today we all are surrounded by data - data from various sources, from banking data to GPS signals on our smart phones. According to McKinsey report this is ‘Big Data’ and it is the next frontier for innovation in diverse fields. By changing what and how we see, big data is reshaping everything coming under its ‘Midas’ touch. Many companies are already using big data to create value. Big data allows organizations to create highly specific segmentations to tailor products and services precisely to meet the needs of the customers in those segments.
Big data enables companies to enhance existing business models and innovate new models, products, and services. Manufacturers are developing the next generation of products and to create innovative after-sales service offerings based on data obtained from the use of actual products. By mining big data, companies can develop insights and identify micro-markets that represent opportunities for growth. Big data will act as the source for large amount of technology based innovation.
The focus of this paper is to look at how big data is connected to technology based innovation in various sub-functions of retailing like marketing, technologies at the point of sale, merchandising, operations, supply chain and new business development.

Keywords: Innovation, Innovation Management, Big Data, Online Retailing, Amazon.

INTRODUCTION
The year to year viability of an organisation depends on its ability to innovate; this is true given the changing market expectations and competitive pressure. Organisations big and small are aware that innovation provides the only sure path to sustainable growth. This is true to retailing as well. Current innovations in retailing range from changes in business models, store formats and technologies to fundamentally new ideas and concepts for pursuing growth opportunities in global markets. Technology innovation in the form of Internet and other Information Technology tools have revolutionised the retail business. Some businesses have innovated new markets like the Apple iTunes, others have innovated multichannel business model wherein the same customer visits the retailer via different channels for different purposes (i.e., obtains information online, makes purchases offline, and contacts customer support via telephone ,like buying an IFB washing machine).
Retailers are innovating methods and models to allow them to focus from selling products to engaging and empowering customers, with the goal of creating a rewarding customer experience. Technological innovations allow retailers to expand the boundaries of their target markets and develop new ways for interacting with customers and channel members. Online retailers like Amazon.in cater simultaneously to multiple niche segments and as a result effectively exploit the “long-tail”.
The wave of innovative technologies available for selling goods and services is spreading fast, by making available many interactive and innovative systems able to support both consumers while shopping and retailers for achieving fast and update information on market trends and selling process (Eleonora Pantano, 2014).
The extensive use of sensors in technology-based innovations in retail industry is making possible a new level of automation in data collection, transmission, and analysis. The resulting information networks promise to create new business models, improve business processes, and reduce costs and
risks (McKinsey Quarterly article). As sensors become omnipresent in all domains, the amount of data generated by them will be huge, in terms of multiple petabytes and exabytes to zettabytes. This is ‘big data’ and it's too big for conventional systems to handle it. Innovative methods are being developed to collect, store, organize, analyze and share data, so that the huge data generated by sensors can be used to add value to old products and create brand new ones, and ‘the data itself offers opportunities for innovation’.

By the end of year 2016 over 1 billion people will be using smartphones and tablets (Forrester, 2013), generating zettabytes of both structured and unstructured data. If all this data is brought into play, it opens up until now unexplored avenues for innovation.

The focus of this paper is to look at how big data is connected to technology based innovation in retailing.

UNDERSTANDING BIG DATA

Data is flowing into every area of the global economy. Businesses churn out a huge volume of transactional data, capturing zettabytes of information about their customers, suppliers, and operations. As businesses go about their business and interact with individuals, they are generating a tremendous amount of digital “exhaust data,” This exhaustive data becomes so large that one has to start innovating how to collect, store, organize, analyze and share it. This is “Big Data” (Andreas.Weigend, former chief scientist, Amazon).

Edd Dumbill, Editor-in-Chief, Big Data says, Big data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn’t fit the strictures of your database architectures. To gain value from this data, you must choose an alternative way to process it. Big data is traditionally defined as comprising of data sets that are generally beyond the capability of an organisation to capture, analyse and process.

Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data (McKinsey Quarterly).

MaryAnne M. Gobble in her article Big Data: The Next Big Thing in Innovation feels that data is considered big when it's too big for conventional systems to handle it. Bigness is not just about size—data may be big because there's too much of it (volume), because it's moving too fast (velocity), or because it's not structured in a usable way (variety).

Big Data refers primarily to the vast amount of data continually collected through devices and technologies such as credit cards and customer loyalty cards, the internet and social media and, increasingly, WiFi sensors and electronic tags. Much of this data is unstructured – data that does not conform to a specific, pre-defined data model.

We’re seeing data that is different from what we would call ordinary, or server-type data. Big data for us is really a collection of datasets so large and complex that you simply can’t use the existing database to manage it. (Ashton Dallsingh FCCA, Vice President and CFO – EMEA and Russia, Cisco) Big Data relates to rapidly growing, Structured and Unstructured datasets with sizes beyond the ability of conventional database tools to store, manage, and analyze them. In addition to its size and complexity, it refers to its ability to help in “Evidence-Based” Decision-making, having a high impact on business operations (CRISIL GR&A analysis).

A formal definition was given by Gartner in 2012, he defines big data as ‘big velocity, and/or high require new forms of decision making, insight optimization’ (Gartner).
THE 3V’S OF BIG DATA

What are the 3 V’s of Big Data? ‘Volume’, ‘Variety’ and ‘Velocity’ are the 3 V’s commonly used to describe big data.

There is exponential growth of data. Gartner, says the volume of data is set to grow 800% over the next five years and 80% of it will be unstructured. Voluminous data flow is from transactions, records, tables and files.

Data comes in all types of formats – structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

Velocity of data flow is in real time, near time, streams and batches from mobiles, blogs, and social media.

POWER OF BIG DATA IN ONLINE RETAILING

Retailers are realizing the power of Big Data. Today, customers buy products across multiple channels using digital technology, make buying decisions through ratings, reviews, price comparisons, and product recommendations as compared to earlier days, where the contact was limited to the physical store.

Customers leave a trail of information about their preferences and behaviour at all the touch points and across all transactions. This digital data left behind by buyers is so important, because of big data; managers can measure, and hence know, radically more about their businesses, customers and directly translate that knowledge into improved decision making and performance.

Big Data provides competitive advantage to enterprises such as extreme personalization, real-time context-aware recommendations, dynamic pricing, micro-segmentation, improved supply chain efficiencies by facilitating advanced inventory management for retailers with expansive item catalogs and analysis to improve supplier negotiations, improved operations by enabling standardization and cost savings, leading to better performance and higher return on investment and improved merchandising activity by using innovative business models to help customer get involved in product co-creation, thus optimizing production costs and creating a suitable product mix.
Big Data’s analytics helps businesses discover new trends and patterns, combines internal and external data to provide overall customer view. It enables the ability to tap opportunities unforeseen. McKinsey Global Institute has identified 16 big data levers in retail, which retailers can employ along the value chain. These levers fall into the five main categories of marketing, merchandising, operations, supply chain, and new business models.¹

Marketing levers can affect 10 to 30 percent of operating margin; merchandising levers can affect 10 to 40 percent; and supply chain levers can have a 5 to 35 percent impact. In contrast, price transparency levers will tend to cut prices and squeeze margins. As per McKinsey report, retailers who have started utilizing Big Data have experienced a 60% enhancement in business margins and a 1% improvement in labour productivity. The likes of Apple, Amazon, Google, and Wal-Mart are already using big data to create value. Big data equals big value.

### APPLICATION OF BIG DATA IN AMAZON

1. **Cross Selling**
2. **Location Based Marketing**
3. **In-store Behaviour Analysis**
4. **Customer micro-segmentation**
5. **Sentiment Analysis**
6. **Enhancing the multichannel customer experience**

1. **Assortment optimization**
2. **Pricing optimization**
3. **Placement and design optimization**

1. **Performance transparency**
2. **Labour inputs optimization**

1. **Inventory Management**
2. **Distribution and Logistics Optimization**
3. **Informing Supplier Negotiation**

1. **Price comparison services**
2. **Web-based markets**

¹ Src. McKinsey Global Institute Analysis
Amazon uses big data also to offer a superb service to its customers. It has approximately 1 million hadoop clusters to support their risk management, affiliate network, website updates, machine learning systems and more.

Approximately 30% of Amazon’s sales are generated through its recommendation engine. The company leverages strong analytics and real-time decision making to proactively create compelling cross sell offers based on customer needs.

Every time a user logs into his or her Amazon account and makes purchases or browses various products on the site, Amazon collects this data and the next time the customer’s returns, they offer them products based on their previous purchase and browsing history. This also helps Amazon identify various trends amongst people who make similar purchases.

Amazon provides people with more choices by segmenting the customers based on their interests and purchase patterns, thereby tempting them to make additional purchases, even if they are not looking to buy other products.

Amazon’s ‘one-click’ purchase has made it simpler for customers to buy from them across a number of channels (e.g. Kindle, web, mobile) without having to re-enter credit card details.

Amazon Price Check Mobile App helps to scan the products in store, take a picture of the product or perform a text search to find the lowest prices. The app also prompts the customers to submit the in-store price for the products through the app which helps Amazon know if it’s offering the best prices to its customers ensuring that it is ahead of its competitors.

Amazon is collecting intelligence and valuable pricing information (big data) from its competitors. Amazon uses big data to operate effectively in a fast-paced and competitive e-commerce environment where price and online advertisements dominate.

Amazon has recently obtained a patent on a system designed to ship goods to us before we have even decided to buy it – predictive despatch based on predictive analytics.

The Amazon Web Services (AWS) cloud accelerates big data analytics allowing them to use big data to monitor, track and secure its 1.5 billion items in its retail store and perform various analytics in different areas to provide excellent customer service.

**HOW AMAZON DOES IT?**

Amazon uses the Matrix factorization - A process to decompose one large matrix into smaller matrices whose product approximates the original matrix. Methods to do this are old; applying this to recommendations is relatively new.

Create the item (I) matrix and the user (U) matrix to approximate seen data (figure 2), to predict unseen data (figure 3).

![Matrix Factorization Diagram](image)
By this factorization method Amazon is able to show its user the things they like (figure 4).

The goal is to find the optimal matrices $U$ and $I$ such that we can approximate our original matrix. Such that

$U'$ contains each user and a set of parameters and $I'$ contains each item and a set of parameters. The process is to update user parameter matrix $U'$, update each user parameter vector $u$ to better predict each item preference for that user. Update item parameters matrix $I'$, update each item parameter vector $i$ to better predict each user preference for that item. Repeat the above ‘N’ number of times.

Amazon uses the Elastic MapReduce (Amazon EMR) web service that makes it easy and cost-effectively process vast amounts of data. Amazon EMR simplifies big data processing, providing a managed Hadoop framework that makes it easy, fast, and cost-effective to distribute and process vast amounts of data. The result – Amazon can now partition its data and run at Amazon scale!

For instance

a. Runs on 21 m2.2xlarge
b. 3 hours for 8MM customers, 200 parameters
c. ~90% of time is spent on computation
CONCLUSION

“You can’t manage what you don’t measure”.

There’s so much wisdom in the above saying, and it explains why the recent explosion of digital data is so important. Simply put, because of big data, managers can measure, and hence know, radically more about their businesses, and directly translate that knowledge into improved decision making and performance. Big data, is far more powerful than the analytics of the past.

Big data is the next frontier for innovation in retailing, transforming business to deliver high value and accumulate higher margins and profits. Online Retailers can tap into the large volume and high velocities of unstructured data being collected at various touch points as customer go shopping across multi-channels. Online Retailers can use this data for effective targeted messaging, product creation, and supply chain planning.

Big data analytics help in effective customer micro-segmentation, assortment optimization, in-store analysis, location-based pricing, inventory management, and predictive demand forecast. Big Data provide opportunities for innovations and transform businesses to deliver value.

BIBILOGRAPHY

1. 5 Big Data Use Cases- How Companies Use Big Data, 06 Aug 2015
3. Big data: The next frontier for innovation, competition, and productivity, McKinsey Global Institute
5. Big Data: The Next Big Thing in Innovation, MaryAnne M. Gobble
7. Big Data in Big Companies, Thomas H. Davenport , Jill Dyché, May 2013,International Institute for Analytics
8. What is Big Data? Sas.com
9. How Amazon Uses Big Data to Boost Its Performance, June 2014, Bernard Marr
10. Innovations in Retail Business Models, Journal of Retailing
12. Big Data- The Next Big Thing, Nasscom
14. IBMbigdatahub.com
16. What is Big Data? Edd Dumbill