



सिद्धिमूलं प्रबन्धनम्
भा. प्र. सं. इन्दौर
IIM INDORE

Indore Management Journal

Volume 6 Issue 2

July - December 2014

Indore Management Journal

The Indore Management Journal (IMJ) is a biannual journal published by the Indian Institute of Management Indore, with an objective to provide a diverse perspective on management functionalities to its readers as well as a medium to share experiences, knowledge and practices with the wider community. IMJ publishes empirical and theoretical investigations that enhance the understanding of various phenomena related to Business and Management. The journal is structured to include one general and one special issue every year. We welcome proposals for special issues of IMJ from potential guest editors. Please share your suggestions with us.

Special Issue : Guest Editor

Rajhans Mishra, *Indian Institute of Management Indore*

Editorial Team

Editorial Advisory Board

Pawan Budhwar, *Aston Business School, UK*

Ramesh Rao, *Oklahoma State University, USA*

Rajagopal, *EGADE Business School, Mexico*

Vikas Kumar, *University of Sydney Business School, Australia*

Atreyi Kankanhalli, *National University of Singapore, Singapore*

Rupa Chanda, *Indian Institute of Management Bangalore, India*

Tathagata Bandyopadhyay, *Indian Institute of Management Ahmedabad, India*

Editor

Sushanta K. Mishra

Associate Editors

Gaurav Singh Chauhan

Joysankar Bhattacharya

Prashant Salwan

Rajhans Mishra

Saripalli Bhavani Shankar

Shweta Kushal

Sujay K. Mukhoti

Editorial Staff

Soma Lilhare

© 2015, Indian Institute of management Indore

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical or photocopying or otherwise, without prior permission in writing from IIM Indore.

The editor, editorial team and the printer do not hold any responsibility for the views expressed in the Indore Management Journal or for any error or omission arising from it.

Editorial

We are happy to present the second issue of the sixth volume of Indore Management Journal (IMJ). This is a special issue focusing on the domain of information technology and management. Information technology is cutting across all the disciplines of management studies in various ways. The present issue has tried to explore various emerging topics of information technology and management along with its interaction with real world applications.

This issue has various papers on contemporary topics of information technology and management. Big Data is emerging as an important area for management professionals. First article has discussed the various facets of Big Data and its relevance in the present scenario. The article is supplemented by a HR practitioners' view on Big Data. Technology acceptance is playing an important role for several business models. Various technology acceptance models have been reported in the second article. Social media has emerged as an important platform for expression of views and likings of individuals. The third article has investigated the role of self-disclosure for participating in social networking platforms. Besides that there is also one article on e-Government services. Last article has discussed the role of smart devices and its linkages with data mining.

We have also included three book reviews that are highly contemporary and useful to manifest the importance of information technology and systems management.

We are confident that the articles presented in this issue will be useful to current management researchers and practitioners.

Editorial Team

Indore Management Journal (IMJ)

CONTENTS

EDITORIAL

- 1 Big Data: Road Ahead for India
Madhukar Dayal, Sachin Garg and Rubaina Shrivastava
- 15 Big Data - Intuitive Lessons for HR Managers
Debolina Dutta
- 17 A review of Evolution of Theories and Models of Technology Adoption
Rajesh Sharma and Rajhans Mishra
- 30 Participating in Social Networking Sites (SNS):
Mediating Role of Self-disclosure and the Effects of Well-being
Manoj Das
- 39 Government Process Re-engineering for an E-governance Implementation for
Motor Vehicle Registration in India
Suresh Subramoniam and Dev Twinky
- 47 Internet of Things (IoT) and Smart Technologies:
Framework for Temporal Data Mining Concerning Smart Meters
Shashi Kant Srivastava

Book Review

- 56 Think Bigger: Developing a Successful Big Data Strategy for your Business
Shweta Grover
- 59 Semi-Organic Growth - Tactics and Strategies behind Google's success
Kakul Modani
- 63 Code Halos: How the Digital Lives of People, Things, and
Organizations are Changing the Rules of Business
Suganya Balakumar

Big Data: Road Ahead for India

Madhukar Dayal, Sachin Garg and Rubaina Shrivastava

Abstract

Advancements in computing technologies make new platforms and large volumes of data available to businesses and governments to discover hidden underlying patterns in the data and creating new knowledge. While businesses need to embrace these technologies in order to stay ahead of competition, governments can reap great benefits in cost effectively delivering social services and bring about improvement in social development indices. However, before any new technology can become a powerful resource (for business or for government), there exists a fundamental need for extensive planning, such that one can chalk out a future trajectory, prepare for the changes to come, and invest prudently. Exploitation of Big Data platforms and technologies requires both corporate strategies and government policies to be in place much before the results would start pouring in. In this paper, we investigate the potential of available Big Data platforms and technologies, their current use by various governments, and their potential for use by the central and state Governments in India.

Keywords: big data, business strategy, government policy, social welfare.

1. Introduction

The world is changing: earlier understanding of the historical chain of events was viewed as knowledge but now its meaning has turned into being a capability to predict and influence the future, including the ability to diminish negative future outcomes and enhance positive ones. In one of its emerging forms, this science is known as Big Data.

There is no rigorous definition of Big Data. As pointed out by, Mayer-Schönberger & Cukier (2013, pp 7): "...the real revolution is not in the machines that calculate data but in data itself and how we use it". According to Gartner: "Big Data is high volume, high velocity and high variety information assets that demand cost-effective, innovative forms of information processing

for enhanced insight and decision making". Morton, Runciman & Gordon (2014) reflecting on the characteristics of Big Data list them as: Volume, Velocity, Variety, Veracity and Value. Such voluminous data first accumulated from the astronomical and weather data (for example, collected from various satellites). Today, such data additionally comes from a wide variety of sources, such as, sensor data, web logs, data streams on the Internet, social media, customer transactions, etc.

Big Data is perceived as comprising structured, unstructured and semi-structured data. Amongst them the unstructured data lead, with an estimated share of over 95% in Big Data. Structured data are those that are systematically stored for retrieval, manipulation and analysis, for example, as in relational databases. Semi-structured data do not reside in relational databases but have some organisational properties making them easier to analyse. With a few alterations semi-structured data can often be reorganised in relational databases. XML is an example of semi-structured data. Unstructured data, on the other hand, do not follow any specified format and are largely void of meta-data, such as data from social media, emails, videos, photos and audio files data.

1.1 Understanding Big Data

To better understand and appreciate "Big Data", we should go back to what Diebold (2012) talks about Big Data being three things: the term ("firmly entrenched"), the phenomenon ("continuing unabated") and an "emerging" discipline. Thus, Big Data is many things to different people and it is imperative to understand it deeper before it can be put to use. Towards this end, we look at how Big Data is changing the paradigms of social science research (and thus the lenses through which we perceive the world) and follow this with how Big Data tools and techniques are being used to make smart policy and business decisions. This is followed by a deeper look at Big Data in the Indian perspective and the road ahead.

Big Data and Social Science Research Paradigms: The Debate: One of the big controversies about Big Data and its use for science was started by Chris Anderson, Editor in Chief of Wired magazine when he claimed "the end of scientific theory building and hypothesis testing" was near - "faced with massive data, this approach to science-hypothesize, model, test-is becoming obsolete" (Anderson, 2008). His contention was that Google had "conquered the advertising world" without knowing anything about the "culture and conventions" of advertising, merely on the assumption that better data, analyzed better would win. A similar theme has been evoked by Mayer-Schönberger & Cukier's (2013) claim that data analysis has now shifted from using a sample size of N (where N is a subset of the population ($N \ll \text{all}$)) to the entire population ($N = \text{all}$). As it is infeasible to collect data from the entire population, it is the accepted research practice to survey a statistically significant sample and extrapolate the findings to inform decisions applicable to the entire population (Agresti & Finlay, 2009; Hutcheson & Sofroniou, 1999; Salsburg, 2001; Velleman, 1997). However, Mayer-Schönberger & Cukier (2013) contend that with fixed sample sizes, one could not be sure that all population characteristics were accounted for, and extremely small groups might not even find a place in the sample. As $N \rightarrow \text{all}$, this is no longer the case and there is greater freedom in asking relevant questions. These claims have sparked off a vociferous debate on the role of theory in social science research, Big Data's contribution(s) to social science research, and even more on what big data detracts from social science research. One of the most influential participants in this debate have been boyd and Crawford. They rebutted Anderson's claim(s) by providing a bold new definition of big data and provoking a conversation around it.

1.2 boyd and Crawford's Definition

boyd and Crawford (2012) defined "Big Data" as a cultural, technological, and scholarly phenomenon that rests on the interplay of:

Technology: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets;

Analysis: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims;

Mythology: the widespread belief that large datasets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy.

The idea that "Big Data" rests on "mythology" needs deeper digging into. Similar to the ideas put forth by Gitelman (2013) and Bollier and Firestone (2010), boyd and Crawford posit that "all researchers are interpreters of data". Data is not a "given" Gitelman (2013, ch 1), but is rather subject to a "cleaning" and "interpretive" process. Thus, their contention - just because we have more data, it is a myth to presume that our insights will be truthful, accurate or more objective. They argue that in the case of "Big Data", the four forces that regulate social systems - market, law, social norms and architecture (code in case of technology) are frequently at odds and it is necessary to provoke conversations around what it all means. To accomplish that, they list out six provocations that we use to review the current state of the "correlation-causation" debate.

Provocation 1: Big Data Changes the Meaning of Knowledge: boyd and Crawford posit that Big data profoundly changes our thinking at the epistemological levels by reframing key questions about knowledge and research processes. Numbers don't speak of themselves and others methods of studying phenomena can get blown away by the sheer force of numbers. This talks to the qualitative/quantitative divide also talked about by Manovich (2012) who talks about "deep data"-about a few people and "surface data" about lots of people. There are different and distinct questions that can be asked and answered using the two types of data, and researchers should be cognizant not to prefer one to the detriment of the other. They also point out that the tools often used to study Big Data phenomenon, for example Twitter and Facebook come with their own limitations and restrictions. They treat the study of society using the Big Data tools as analogous to "accounting tools" that "shape the reality they measure". The idea that tools limit what can be collected resonates with Vis (2013) who points out that the Application Programming

Interfaces (APIs) that are used by researchers to collect data limit what can be collected. Tufekci (2013) points out how the practice of using "hashtags" to filter tweets biases the data towards a particular demographic—those who use a particular hash tag are more "wedded" to the issue, and are thus different from the rest. Manovich (2012) cautions us against taking what is spoken on social media as authentic by telling about his personal experiences growing up in the erstwhile Soviet Union and how what was spoken out was very different from what was actually meant.

Provocation 2: Claims to Accuracy and Objectivity are Misleading: boyd and Crawford claim "all researchers are interpreters of data". Data is not a "given", existing in and of itself. Data is actively collected and sourced. As Desouza & Jacob (2014) point out that it is difficult to "recognize data in its unstructured form and then to understand how to 'connect it' to more conventional forms of data". Bollier & Firestone (2010, pg. 13) ask if the data represents an "objective truth" or are interpretations biased due to the way the data is "cleaned"? Visualizations also have judgments embedded within them (Bollier & Firestone, 2010, pp 11-12). Khoury & Ioannidis (2014) point out that "Big error" is another challenge with "Big Data". They say, "big data's strength is in finding associations, not in showing whether these associations have meaning. Finding a signal is only the first step". Thus, it is important to understand the biases and the limitations of the data.

Provocation 3: Bigger Data are not Always Better Data: Using the example of Twitter, boyd and Crawford point out that we cannot assume that Twitter users provide an appropriate sample as not everyone is on Twitter and "bots" also inhabit it. Also, results are not transferable between social networks due to the network's unique demographics (Ruths & Pfeffer, 2014). Sometimes, "smaller" data might be more relevant. This is similar to Manovich (2012)'s concept of "shallow" and "deep" knowledge. Lagoze (2014) has called out Mayer-Schoenberger and K. Cukier (2013) on their concept of $N = \text{all}$. It is an unachievable mathematical ideal as not everything can ever be measured. Sometimes, sampling the data is better as pointed out by Hal Varian who

mentioned that for the economic studies Google undertakes, a random sample was good enough (Bollier and Firestone, 2010).

Provocation 4: Taken Out of Context, Big Data Loses its Meaning: Contextual integrity of the data is extremely important to gain value from the data. boyd and Crawford contend that people's real-world 'personal' networks are different from their 'articulated' and 'behavioral' networks traced out through data. Taylor & Schroeder (2014) provide an example of wrong inferences due to lack of context. Practical Big Data analysis requires the data to pass through multiple stages through the pipeline (Jagadish et al., 2014) and at each stage the data gets "repurposed, reprocessed, retrofitted, and reinterpreted" thus losing context on the way (Schintler & Kulkarni, 2014).

Provocation 5: Just Because it is Accessible Does not make it Ethical: Big Data is changing the perception of ethics. Earlier, individual decisions had specific and knowable outcomes. With the advent of Big Data, many can take actions without realizing how their actions impact others (Zwitter, 2014). Crawford & Finn (2014) talk about the use of social media (Twitter) sourced data in the context of crisis, where people share location data and other personal information because they want help. One of the bigger challenges is anonymity and re-identification. boyd and Crawford (2012) say that researchers should focus on "accountability", which is a much broader concept than "privacy". Privacy and anonymity continue to be the biggest challenges regarding the use of Big Data in social science and public policy.

Provocation 6: Limited Access to Big Data Creates New Digital Divides: The "digital divide" regarding Big Data can be thought to exist along three axes- (a) who creates the data, (b) who accesses the data, and (c) who has the resources to analyse said data. Hilbert (2013) talks about these axes as necessary, but not sufficient conditions to harness "Big Data" and the need to have institutional mechanisms (or appropriate policies) in place. Manovich (2012) also talks about the challenges of accessing data, as much of the "social media" data is only accessible to few, trained people within the companies. A similar point was made by Taylor, Schroeder, & Meyer (2014)

in the context of income and other "sensitive" data that is only accessible to a few researchers. The challenge of who creates the data is important in a policy context to maintain equity and ensure that policy makers do not divert a disproportionate share of resources to those segments on whom data is available, at the expense of the data-poor.

Once we consider these research and perspective aspects, other questions that are needed to be asked are: *What is the data? Where is the data? Who owns it? Who has access to it? Is it technically, legally, ethically, politically feasible to run analytics on it? What does it cost to analyse this data? What are the expected benefits of this analysis? Do the benefits outweigh the costs?*

Thus, it is imperative to consider that all data is collected, stored and analysed in an environment circumscribed by policy. It is policy which sets the boundaries of what is acceptable or unacceptable in particular contexts. Various regulations like HIPPA, Children Privacy Act, FERPA etc. determine the limitations of data collection and analysis. Thus, we have an interesting paradox between data and public policy: Public Policy simultaneously bounds and is bounded by the data collection and analysis framework(s).

2. Big Data in the Public Sector

The public sector is a ripe area for applying the tools and techniques of Big Data to increase the efficiencies in the sector. This can happen in two ways: by using Big Data to improve programmatic outcomes, and to improve decision making.

Improving Programmatic Outcomes: Desouza & Jacob (2014) contend that most public sector data is of low complexity and hence organizations can improve their "programmatic outcomes". They provide an example of New York city where the mayor's office is linking together otherwise unconnected databases and mining them to identify areas of focus. As an example, the office is using predictive analytics to identify potential zoning violations and target inspections towards such potential violations. The key insight here is the idea of linking together disparate databases to get a more complex database, which may not possess large volume or

velocity. However, predictive policing in the Los Angeles Police Department is an example of using "true" big data. The data in this case is a fusion of historical and real-time data (that includes real-time city and traffic camera feeds). This data is used to identify areas where crime may occur and concentrate resources in such areas. They point out that possibly one of the biggest insights of this project was the realization that video streams are just another form of data that can be connected to other data via geocoding. The public sector in US is deploying big data technologies in Postal services, health care and human services, and internal revenue services, to name a few.

Improving Decision Making: The underlying premise behind using big data to enhance decision making is providing feedback loops that allow citizens to engage with government and thus reveal preferences that are not revealed through the traditional polling process. Desouza & Jacob (2014) identify two different mechanisms that have been proposed to assess the peoples' will: prediction markets and sentiment analysis. However, these methods require new types of data, richer data and existing data cannot be shoe-horned to fit these paradigms, and thus need greater investment.

Prediction markets, designed to take advantage of the "wisdom of the crowds" work akin to a futures market, where the commodity being traded on is an event. Though such markets are part of well developed stocks and commodities markets, they are still akin to wagers. While wagering on stocks and commodities is accepted, wagering on politically sensitive issues is not and leads to ethical concerns.

Sentiment analysis draws on messages posted on social media like Twitter and Facebook as a means to understand the populace. The use of social media has been criticized by scholars as not being representative of the society (see Tufekci, 2013; boyd and Crawford, 2012; Ruths & Pfeffer, 2014). Another aspect is that social media can be manipulated as noted by Desouza & Jacob (2014). It should also be noted that sentiment analysis depends heavily on natural language processing, which is not as well developed for languages other than English. Thus the use of this technique in

other countries and regions may not be as successful. However, despite these shortcomings, we should still attempt to use these mechanisms as "another" input into the decision-making process. Some examples of how big data is being used as input into the decision making process follow.

Boston Street Bump is a project of Boston's Mayor's Office of New Urban Mechanics. It crowd sources road condition data using the accelerometer and GPS sensors of smartphones. This data is aggregated across users and used to fix short-term problems like potholes. The application found mention in Podesta et al., (2014). It was also reported that initially the app sent repair crews to wealthier neighbour hoods where people were more likely to carry smartphones, and this bias was fixed by first deploying it to city-road inspectors, who service all parts of the city equally; the public now provides additional supporting data (Podesta et al., 2014).

Flowminder used cell phones to track people's movement during the 2010 Haiti cholera outbreak (Taylor & Schroeder, 2014) and identify areas outside the capital (Port-au-Prince) at risk of cholera. This technique used near real-time (upto 12 hour) location data of cell phones (based on the phone-tower communication) from phone companies and extrapolated the results (considering that multiple people share a phone). These results were validated by comparing with on the ground data from local and UN agencies.

Billion Prices Project at MIT has created an "inflation index" by tracking online prices (Taylor & Schroeder, 2014). The initial idea behind this was to use actual on the ground prices in Argentina and compare them with officially released figures that could be potentially affected by political demands. Decuyper et al., (2014) have attempted to use indicators derived from mobile phone data (call detail records and airtime purchases) as food security indices in an African country.

2.1 Big Data and Public Policy: Key Challenges

From the foregoing, we note that big data can be used to analyze, formulate and monitor public policy in myriad ways. However, its use continues to be fraught

with many challenges, chiefly related to privacy, discrimination and liability (Schintler & Kulkarni, 2014). Of these, however, privacy has the largest mindshare, owing in part to the Snowden affair. The primary fallout of Snowden's disclosures on the NSA's bulk collection of telecommunication metadata was that privacy concerns with big data came to the forefront in the eyes of the public. This also prompted a review of US Signals Intelligence and a mandate to look closely at the "challenges inherent in big data": "Look at how the challenges inherent in big data are being confronted by both the public and private sectors; whether we can forge international norms on how to manage this data; and how we can continue to promote the free flow of information in ways that are consistent with both privacy and security." - The White House, Review of US Signals Intelligence, Jan 17, 2014.

2.2 Privacy, Anonymity and Big Data

The researcher community's concerns over privacy issues are echoed by Lane & Stodden (2013): "privacy issues could stop bona-fide data collection and statistical research in its tracks". This does not imply that the researcher community is not concerned about privacy, but rather wants a "sensible structure for data access that ensures the goal of good science is attained while protecting confidentiality and respecting individual agency" (Lane & Stodden, 2013).

The major challenges impacting the usage of Big Data for Public Policy and Social sciences are:

Lack of data collection and management infrastructure. This challenge is even more pronounced in the developing world. The developing world does not have enough ears on the ground to collect high quality data. Most of the data that Big Data and development experts talk about is essentially digital exhaust of a very specific type-mobile phone call data records (UN Global Pulse, 2012). Though analysing these records has shown utility, there are inherent challenges in accessing these records. These records are all part of private businesses owned data systems and sharing them is fraught with competition and other issues. Though the records are collected in near real time, they are actually released after a significant amount of time (cf. D4D challenge

(Taylor & Schroeder, 2014)). Even when released, they are only accessible to a small number of researchers. Hence, even plucking the low hanging fruit in this case is non-trivial (Prydz, 2014).

Lack of Institutional mechanisms to curate the data and mediate access to it. Again, while in the US, researchers have access to resources curated and mediated by the ICPSR, such institutional arrangements are largely lacking in the developing world. Even, in the USA, presence of legislation overly focused on privacy aspects has prevented the linking together of administrative record data both across agencies and across states (Lane & Stodden, 2013; Lane & Schur, 2009; Lane & Schur, 2010).

Absence of infrastructure to support privacy preserving data mining, wherein the researchers do not need access to the raw data per se. Data enclaves (Abowd & Lane, 2004) are a possible solution, but have concerns regarding accessibility to a select few. Thus, some virtual solutions are needed.

Crisis of Reproducibility is illustrated by Google Flu Trends (Lazer et al., 2014). Both science and social science are passing through a crisis of reproducibility, partly because of entrenched notions on sharing and the incentive mechanisms, and partly because it is difficult or impossible to share. Studies like the Facebook contagion study (Kramer, Guillory, & Hancock, 2014) cannot be replicated (and thus criticized) outside the platform for which they were designed and same is the case with multiple other studies, many of which depend on data sourced from commercial entities mediated through Application Programming Interfaces (Vis, 2013; Ruths & Pfeffer, 2014).

Privacy and anonymity as Daries et al., (2014) say, the two are intimately linked and sides of the same coin. We need to clearly understand what needs to be protected and build strong policy foundations for the same. The consent framework that does not take care of end-use only instills a false sense of privacy, as it does not really protect end-use (Barocas & Nissenbaum, 2014; Mundie, 2014; Podesta et al., 2014).

Data Brokers or "omnibus information providers" are largely unregulated and hold detailed profiles on almost

all citizens (Nissenbaum, 2010; Podesta et al., 2014). Whereas the government is not able to link together its own administrative databases together because of the "big brother is watching attitude", multiple data brokers acquire data from federal, state and county governments, and link them together to form extremely rich datasets (Washington, 2014). Ansolabehere & Hersh (2012) provide an illustration when they detail how the private firms they engaged "Catalist" and "Polimetrix" shared data amongst themselves to link various voting records, and shared de-identified records with the researchers.

Lack of Data Integrity or provenance which is a "cornerstone of credible science" (Lagoze, 2014) is a major challenge to the reproducibility and applicability of results. A lot of big data lacks provenance as (i) it has not been designed for research (social media data) (Lazer et al., 2014), (ii) it has been stripped of key identifiers (Podesta et al., 2014), or (iii) it has been "repurposed, reprocessed, retrofitted, and reinterpreted" (Schintler & Kulkarni, 2014) multiple times.

Till now we have looked at the various aspects of Big Data from the data and domain perspective. Now, we turn our gaze to the tools and techniques used to analyse this data and gain value.

3. Big Data Technologies and Challenges

The open source project Hadoop (by Apache Software Foundation) is a primary Big Data analytics platform which is built to operate on large distributed (high performance) compute clusters. *MapReduce*, the most popular function, is essentially a two stage fault tolerant analytical routine which distributes the data and task at hand, first, to various compute nodes, and integrates the results obtained later. This is done using the Hadoop Distributed File System (HDFS) (adopted from Google (distributed) File System or GFS). IBM's InfoSphere BigInsights and InfoSphere Streams are commercial platforms for analysis of big data at rest and in streams respectively. A survey on Big Data describing technologies, platforms, applications, and challenges with suggestions on designing Big Data systems is presented by Chen & Zhang (2014). For more details on the platform, we refer the interested readers to *Understanding Big Data: Analytics for Enterprise Class*

Hadoop and Streaming Data by Zikopoulos, Eaton, Deroos, Deutsch & Lapis (2015). Further developments for new analytical routines to add to the Hadoop family are also underway by several organisations including the open source community.

3.1 Challenges

Kambatla, Kollias, Kumar, & Grama (2014) point out that due to the scattered nature of Big Data it is difficult to store, process and analyse it at one place. Hence, it needs to be segregated and processed over different servers. But with such distributed databases there arises the complexity of privacy, fault-tolerance, security and access controls. Chen & Zhang (2014) highlight that the lack of awareness pertaining to Big Data poses serious threats to the nation's cyber security and is also a barrier to country's socio-economic development.

Big Data poses a serious challenge in regard to data complexity, large scale data integration, sheer volume and lack of availability of supporting high performance computing cluster (HPCC) hardware and software platforms to tackle the aforesaid challenges. For more detailed discussion on involved issues, we refer the interested readers to *Big Data: Opportunities and Challenges* by Morton, Runciman & Keith (2014).

3.2 A Techno-Legal Perspective

Big Data: Information Security Panorama: Secure cyberspace has become an indisputable need. In the context of Big Data, all organisations involved in its life cycle must have robust information security frameworks, incorporating at least:

1. Limiting access through segregation and separation of duties with defined access rights restrictions and strict authentication and authorization parameters.
2. Use of data anonymization and a control on de-anonymization techniques while storing personal identifiable information (PID) or other sensitive information.
3. Establishing a trust boundary between data owners and data storage owners.
4. Implementation of sound access control policies and customized firewall configurations in parlance to the

value and sensitivity of the information/data.

5. To conduct periodic internal and external security audits.
6. Real-time security monitoring to detect and respond to any alarming event.
7. Use of Fully Homomorphic Encryption (FHE) method, in order to keep a balance between need to perform operations on encrypted data packets and also keeping it secure while in transmission.
8. Cyber threat intelligence mechanisms.
9. Hosting critical information only in hardened host servers.

Mayer-Schönberger & Cukier (2013, pp 27) cite an example of 'Xoom', which is a firm holding a big name in the context of Big Data. Xoom analyses its transactions in totality and triggers an alarm if any suspicious behaviour is detected. Usually to detect malicious behaviour it works on pattern based detection techniques, which implies that whenever any suspicious behaviour is detected which appears to be against the 'normal' behaviour pattern of the firm, the software will raise an alert/alarm. Xoom provides 128 bit encryption protection for securing transactions on its website whether the user is logged in or not. Xoom is a Verisign certified site and a certified licensee of TRUSTe (www.xoom.com).

Big Data: Legal Panorama: There exist variety of sensitive information, such as, confidential organisational information, intellectual property (e.g., trade secrets), healthcare information (e.g., patient records or insurance information), personal financial information (e.g., employee salary details, social security details) which need to be protected from unauthorized disclosure, access, alteration or damage. Several nations have enacted laws to protect personally identifiable information, for example, European Union Data Protection Directive, Enhancing Privacy Protection Act, Health Insurance Portability and Accountability Act (HIPAA), and for protecting personal financial information the Gram-Leach-Bliley Act (GLBA). In India, there is a pressing need to frame and enact suitable data protection legislation and incorporate compliance mechanisms. As of now, personal data protection is

covered by provisions in the Information Technology (Amendment) Act, 2008 under Sections 43, 66, & 72 and/or under the provisions of Indian Penal Code, 1860.

Big Data legal requirements and its sector wise applicability: The multiple layers of regulations can be implemented jointly or independently depending upon the case facts. In the context of Big Data, the applicability of the relevant law is ascertained on the basis of various factors, such as, the type of data (personal health, financial or corporate information). The applicable laws to select from are:

- (1) Health Information Technology for Economic and Clinical Health Act (HITECH Act) - applicable to health care providers, health care clearing houses storing, processing, and exchanging electronically protected health information (e-PHI). HITECH Act widens the scope of privacy and security available under HIPAA. It further increases the potential liability in case of non-compliance and bestows better enforcement.
- (2) Children's Online Privacy Protection Act - applicable to organizations collecting personal information of children (the age limit varies from country to country).
- (3) CAN-SPAM Act 2003 - protects customers from targeted marketing campaigns of companies which results into unsolicited bulk emails.

Google's 'Usage of Big Data: A Strategic Business Purpose' example (in book *Big Data: A Business and Legal Guide* by Kalyvas & Overly): The analyses of Big Data is often for a purpose different than the one for which it was collected. Although, the commercial use of Big Data is apparent but organisations need to be transparent regarding its business purpose while using Big Data and ensuring that it does not exceed the defined purpose(s). Google narrowly escaped in a litigation, *Authors Guild, Inc. v. Google Inc.*, 770 F.Supp.2d 666 (S.D.N.Y. 2011) in which Google had successfully avoided legal liabilities by clearly defining well in advance a business purpose for use of Big Data. In this case, Google was sued for violating copyright by creating a copy of authors books in the form of e-books (using optical character recognition technology) and then

responding to the users queries/searches on the basis of matching keywords and thereby increasing Google books sales. The Honourable court favoured Google by stating that Google's usage does not fall into the category of 'massive copyright infringement' (as claimed by the plaintiff) or adversely impacts the rights of copyright holder, as it had followed the required due diligence and also is eligible to fall under the 'fair use' category. Google through its security measures didn't allow the users to have a complete view of the books but only snippet views, thus, giving respectful consideration to the author's rights and creativity. Not only that, the court held that Google had incorporated better research tool (in form of data mining), ease of access, efficient mechanism for identifying and locating books and quick search results for end users. On the whole, we may say that an organisation can mitigate the risk of litigation arising out of using Big Data with well-defined business purpose (inclusive of transparency to users regarding usage of data collected, protections from any competing commercial interests that may arise, and above all serving the public good).

Big Data in Healthcare Industry: Mckinsey Global Institute estimated that healthcare analytics will generate more than \$300 billion in business value per year. Big Data can make significant changes and developments in reshaping public health. Google published a paper in the scientific journal 'Nature' estimating the likelihood of rapid spread of the H1N1 virus, just few weeks before the virus actually hit (Mayer-Schönberger & Cukier, 2013). As pointed out by Bill Hamilton (2012), "If a group of patients is discussing quality of care about a provider, there will likely never be 100% consensus. Patient experiences will be different, and there will be biases based on accidents, misunderstandings and other factors. The challenge will be to create useful information out of this collection of data to provide information such as provider ratings and improvement guidance" (Hamilton, 2012).

For more examples and a detailed discussion on legal aspects we refer the reader to *Big Data: A Business and Legal Guide* by Kalyvas & Overly (2014). Criticism of real life Big Data application has also surfaced (Lazer et al., 2014). However, we observe that Big Data

applications and development are at a nascent stage and we envisage that over a period of time the technology, platforms, and applications will mature proving their utility.

4. Big Data Opportunities

Big Data being voluminous allows us to explore new information avenues with better granularity and without the risk of blurriness. Immense volumes of data lie around us needing to be collected and processed to extract value. One of the major benefits of creating and using Big Data is that it highlights and spots such points of concern which otherwise may be entirely undetectable when using sample data (Mayer-Schönberger & Cukier, 2013).

Governments and PSUs: Governments are increasingly adopting digital technologies. USA.gov and 'Digital India' are notable examples of this trend. The 2012 presidential election campaign in U.S has seen one of the remarkable uses of Big Data for better decision making. President Barack Obama's campaign team conducted Big Data analysis to target voters and identify the most responsive regions for campaigning and then allocating the resources to the destined areas. The winning of Obama and his getting re-elected as president of U.S.A demonstrated and unfolded a new strategic step in making sense of Big Data (Jin, Wah, Cheng & Wang, 2015).

Law Enforcement Agencies: Big Data can be used by Law enforcement agencies in order to analyse voluminous data and impede crime and terrorist attacks. The case of a notorious Chinese serial killer 'Zhou Kehua' is an example of usage and summary analysis of various information obtained from Big Data. The Big Data consisted of video data, photographs and some other related content and on the basis of it Zhou Kehua was tracked, investigated and captured. In this case, Big Data analysis played a decisive role for the law enforcement agency. Big Data may also prove to be applicable in identifying potential criminals (Jin et al., 2015).

Business and Economic Systems: Big Data studies can be applied to raise the economic value and to bring significant societal and scientific impact. Forecast, the

air fare predictions website for best buy price helps consumers based on the Big Data analyses on earlier air fare data and thus giving substantial economic benefit to passengers (Mayer-Schönberger & Cukier, 2013, pp 4-5).

4.1 International Big Data Initiatives

United States of America: In September 1993, the 'Information Highway' program was launched in USA. Similarly, in March 2012 the 'Big Data Research and Development Initiative' was launched. The project envisions to improve and facilitate use of Big Data by extracting valuable information insights for better development. It primarily focuses on healthcare, emergency response and disaster recovery, cybersecurity, education and employability, transportation and energy sector (Jin et al., 2015).

United Kingdom: COSMOS (see *What is COSMOS?*) aims to be an open platform for social data analysis that can harvest, archive, analyze and visualize social media streams. In due course, the platform is expected to link to other social data and is currently linked to the UK Police API, harvesting crime statistics. Collectively, the European Union has also started partnering through the program 'Horizon 2020'.

Japan: Aspires to be the World's Most Advanced IT Nation by year 2020. 'The Integrated ICT Strategy for 2020' has already been launched with a mission to develop Japan as a leader of Information Technology with Big Data at its centre stage. The aforesaid IT strategy focuses at implanting the highest level of standards in Big Data technology and IT infrastructure (see *Declaration to be the World's Most Advanced IT Nation*).

Germany: The German Government has announced a Big Data research initiative namely 'production intelligence'. The aim is to perform real time analytics on all manufacturing data. This Big Data analysis will help to evaluate, improve, and enhance the manufacturing capacities and processes, to automate, and in effective decision making, and to achieve optimal manufacturing scenarios (see German government announces "Production Intelligence": funding for Jedox's Big Data project).

Australia: The Australian Public Service ICT Strategy 2012-2015 aims to use Big Data for better service delivery, efficient and effective mechanisms for e-governance, preserve national information assets, improve health service offerings and better emergency response mechanisms. Australian government uses Patient Admission Prediction Tool (PAPT) (in collaboration with Australian e-Health Research Centre Queensland Health, Griffith University and Queensland University of Technology) software for Big Data analytics in health industry. PAPT aims to achieve predictions for number of patients that hospital may expect in the near days, emergency cases, hospital staff's case(s) handling capacities, available and required labor pool as and when need arises, and balanced workload. These predictions can achieve timely service delivery, better disaster resilience and a far better quality care offering.

United Nations: UN recently launched a project 'Global Pulse: Harnessing Big Data for Development and Humanitarian Action'. Global Pulse is intended to ascertain and predict the societal issues like unemployment, disease outbreaks, and likewise. It aspires to achieve proactive approach in handling alarming events arising out of humanitarian grounds. It works for creating awareness and development in regard to Big Data opportunities and its value addition for society (<http://www.unglobalpulse.org/>).

5. Big Data: The Road ahead in India

Substantial Big Data is being generated (and stored) by Government departments in India already. Department of Science and Technology, GoI has announced plans to take Big Data research forward in the Indian context, including financial support for teams taking up such projects (<http://dst.gov.in/scientific-programme/bigdatainitiative.html>). However, continuous effort shall be needed for a long period of time before some success stories of big data studies and their results are visible.

More efforts to tap the potential of big data analytics, especially in the social welfare sphere, are needed. The bottlenecks needed to be overcome are: (i) not much (big) data is being collected and stored in India (leaving a few segments, such as, scientific community with

space and weather data), (ii) accessibility to expensive platforms (hardware and software needed, though open source can be deployed) is limited, (iii) efforts are needed in the direction of preparing policies and legal frameworks covering issues such as responsibility for collection, storage, and preservation, protection from illegal use, ownership of the data and (extent of) freedom to share with others, etc.

Prime Minister's farmer soil health card is an initiative which could provide extremely valuable data in future contributing to the nation's food security. Similar schemes are also needed for our other national natural resources, such as: (i) monitoring (underground and surface) water availability, usage, and its preservation in India, (ii) rainfall harvesting activities and potential, (iii) land and its (current and possible) usage across the country, (iv) forest areas monitoring, (v) wildlife data, (vi) air quality data from cities (recent media reports indicate Delhi as the most polluted city in the World, now since a few years), (vii) wind farming potential, and so on.

On the human development index front, monitoring of diet and health data (including disease spread and control, vaccinations, etc), levels and adequacy of nutrition intakes in society, education (availability and usage), transport (needs, trends, and consumption), electricity (generation, distribution, shortages, losses, etc), provide scope for exploiting big data applications for big gains.

The central and state Governments in India stand to gain a lot by joint planning, collection, sharing, and analysis of big data to develop appropriate talent development plans for future, planned farming to avoid over and under production in a season leading to excess or shortage, and similar other schemes.

Effort is needed to tap the potential in big data starting with: to identify, support (such as through fully funded academic scholarships), develop, and employ special talent to tap the potential of big data. Simultaneously, to set up big data analytics centres with necessary infrastructure, and accessibility to the scientific and academic community, supported by series of funding

for incorporating future developments in these technologies and ensuring their immediate availability to the talent pool for productive deployment.

As the currently available big data talent pool is small, it may not yield the critical mass necessary to push a series of big data projects ahead rapidly. Hiring talent from other countries, if available, may also be an expensive alternative. Hence, progressively expanding Government support appears to be the good road ahead. The first among these steps could be identifying and supporting scholars and students in acquiring the necessary knowledge and skills.

Being a very large nation with equally large variety of data generated (and additionally, challenges of different languages, notations, formats used etc), it is desirable for India to develop and follow a schema designed and suggested by scholars, researchers, various data users, and Governments collectively. Though, Big Data technology is designed to address analysis of humongous volumes and variety of data, bringing even a partial order to its collection and storage by pursuing an established policy can ease analytical loads (by somewhat reducing data's dimensionality), help in ensuring that it meets legal frameworks and manage changes brought to it from time to time.

Organisations who share their data with public should do so with an explanatory data dictionary, including Government organisations. Development of a data storage standard (such as Data Collection and Storage ISI standard) is also recommended, though both, its compliance by all involved agencies and retrofitting old data can become extremely challenging (if not impossible) undertakings. However, this is expected to be helpful in better collection of data in future, including its integration among several parties for meaningful analysis and insights.

Additionally, a central repository for organisations to contribute semi-structured and unstructured data (with its description, even if brief) should be sponsored by central and state Governments (a data-history-pedia). Such data may prove to be useful for hitherto unforeseeable analysis.

6. Conclusion

Big Data platforms and technology have crossed the chasm of mere interest. Across the world, scientific, academic, research, business, as well as, government communities are aggressively charting plans and paths to benefit from developments in the big data field. The issues pertaining to policy and existing frameworks developed over the last few years in some advanced countries have been identified and critiqued to identify unresolved issues. We anticipate much action in the business and government domains in the years to come, and one such potential arena would lie in big data which spans national boundaries.

Summarizing learnings from recent applications of big data across the world, and identifying the initiatives embraced by governments of prosperous nations, we underscore the huge potential that big data holds and can unfold for the central and state governments in India. We also propose steps that can be taken in India during policy formulation, legal frameworks enactment, infrastructure upgradation and talent pool creation, for the nation to benefit from Big Data platforms and technology.

References

- Abowd, John M. and Lane, Julia. (2004). New Approaches to Confidentiality Protection: Synthetic Data, Remote Access and Research Data Centers. English. In: *Privacy in Statistical Databases*. Ed. by Josep Domingo-Ferrer and Vicenc, Torra. Vol.3050. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp.282- 289. ISBN: 978-3-540-22118-0. DOI: 10.1007/978-3-540-25955-8_22. URL: http://dx.doi.org/10.1007/978-3-540-25955-8_22.
- Agresti, A. and Finlay, B. (2009). *Statistical Methods for the Social Sciences*. Pearson Education. Pearson Prentice Hall, New Jersey.
- Anderson, Chris. (2008). The end of theory: The data deluge makes the scientific method obsolete. In: *Wired* 16.7, pp. 106-129. URL: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.
- Ansolabehere, Stephen and Hersh, Eitan. (2012). Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate. In: *Political Analysis* 20.4, pp. 437-459. DOI: 10.1093/pan/mps023. eprint: <http://pan.oxfordjournals.org/content/20/4/437.full.pdf+html>. URL: <http://pan.oxfordjournals.org/content/20/4/437.abstract>.

- Australian Public Service Better Practice Guide for Big Data.* (2015). Retrieved from: <http://www.finance.gov.au/sites/default/files/APS-Better-Practice-Guide-for-Big-Data.pdf>
- Barocas, Solon and Nissenbaum, Helen. (2014). Big Data' a End Run around Anonymity and Consent. In: *Privacy, Big Data, and the Public Good*. Cambridge University, 44-75.
- Bollier, David and Firestone, Charles M. (2010). *The promise and peril of big data*. Aspen Institute, Communications and Society Program Washington, DC, USA.
- boyd, Danah and Crawford, Kate. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. In: *Information, Communication & Society* 15.5, pp. 662-679. DOI: 10.1080/1369118X.2012.678878. eprint: <http://www.tandfonline.com/doi/pdf/10.1080/1369118X.2012.678878>. URL: <http://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.678878>.
- Bryant, Antony and Raja, Uzma. (2014). In the realm of Big Data In: *First Monday* 19.2. ISSN : 13960466. URL : <http://firstmonday.org/ojs/index.php/fm/article/view/4991>.
- Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314-347.
- Crawford, Kate and Finn, Megan. (2014). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. English. In: *GeoJournal*, pp. 1-12. ISSN : 0343-2521. DOI : 10.1007/s10708-014-9597-z. URL : <http://dx.doi.org/10.1007/s10708-014-9597-z>.
- Daries, J. P., Reich, J., Waldo, J., Young, E. M., Whittinghill, J., Ho, A. D., ... & Chuang, I. (2014). Privacy, Anonymity, and Big Data in the Social Sciences. In: *Commun. ACM* 57.9, pp. 56-63. ISSN: 0001-0782. DOI: 10.1145/2643132. URL: <http://doi.acm.org/10.1145/2643132>.
- Declaration to be the World's Most Advanced IT Nation.* (June 14, 2013). Retrieved on 20-Apr-2015 from: http://japan.kantei.go.jp/policy/it/2013/0614_declaration.pdf
- Decuyper, A., Rutherford, A., Wadhwa, A., Bauer, J. M., Krings, G., Gutierrez, T., ... & Luengo-Oroz, M. A. (2014). Estimating Food Consumption and Poverty Indices with Mobile Phone Data. Version 1. In: preprint arXiv:1305.3212. arXiv: <http://arxiv.org/abs/1412.2595v1> [cs.CY, physics.soc-ph].
- Desouza, Kevin C. & Jacob, Benoy. (2014). Big Data in the Public Sector: Lessons for Practitioners and Scholars. In: *Administration & Society*. DOI: 10.1177/ 0095399714555751.eprint:<http://aas.sagepub.com/content/early/2014/11/06/0095399714555751.full.pdf+html>. URL: <http://aas.sagepub.com/content/early/2014/11/06/0095399714555751.abstract>.
- Diebold, Francis X. (2012). On the Origin(s) and Development of the Term 'Big Data'. PIER Working Paper No. 12-037. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2152421>
- German government announces "Production Intelligence": funding for Jedox's Big Data project.* (April 13, 2015). Retrieved from: <http://www.protext.cz/english/zprava.php?id=22879>
- Gitelman, Lisa. (2013). "Raw Data" is an Oxymoron. *MIT Press*.
- Halevy, A., Norvig, P. and Pereira, Fernando. (2009). The Unreasonable Effectiveness of Data. In: *Intelligent Systems, IEEE24.2*, 8-12. ISSN:1541-1672. DOI: 10.1109/MIS.2009.36.
- Hamilton B. (2012). Big Data is the Future of Health Care. *Cognizant 20-20 insights*. pp-5. Retrieved from: www.cognizant.com/.../Big-Data-is-the-Future-of-Healthcare.pdf
- Hilbert, Martin. (2013). Big Data for Development: From Information-to-Knowledge Societies. Available at SSRN 2205145. DOI: 10.2139/ssrn.2205145. URL: <http://ssrn.com/abstract=2205145>.
- Hutcheson, Graeme D and Sofroniou, Nick. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. Statistics Series. SAGE Publications.
- Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big Data and Its Technical Challenges. In: *Commun. ACM* 57.7, pp. 86-94. ISSN : 0001-0782. DOI : 10.1145/2611567. URL : <http://doi.acm.org/10.1145/2611567>.
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research* (In press).
- Kalyvas, J. R., & Overly, M. R. (2014). *Big Data: A Business and Legal Guide*. CRC Press.
- Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, 74(7), 2561-2573.
- Khoury, Muin J. and Ioannidis, John P. A. (2014). Big data meets public health. In: *Science* 346.6213, pp. 1054-1055. DOI : 10.1126/science.aaa2709. eprint:<http://www.sciencemag.org/content/346/6213/1054.full.pdf>. URL : <http://www.sciencemag.org/content/346/6213/1054.short>.
- Kramer, Adam D. I., Guillory, Jamie E., & Hancock, Jeffrey T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. In: *Proceedings of the National Academy of Sciences* 111.24, pp. 8788-8790. DOI: 10.1073/pnas.1320040111. eprint: <http://www.pnas.org/content/111/24/8788.full.pdf+html>. URL: <http://www.pnas.org/content/111/24/8788.abstract>.

- Lagoze, Carl. (2014). Big Data, data integrity, and the fracturing of the control zone. In: *Big Data and Society* 1.2. DOI:10.1177/2053951714558281.eprint: <http://bds.sagepub.com/content/1/2/2053951714558281.full.pdf+html>. URL: <http://bds.sagepub.com/content/1/2/2053951714558281.abstract>.
- Lane, Julia and Schur, Claudia. (2009). Balancing Access to Data And Privacy. A review of the issues and approaches for the future. Working Paper Series of the German Council for Social and Economic Data 113. German Council for Social and Economic Data (RatSWD). URL: <http://ideas.repec.org/p/rsw/rswwps/rswwps113.html>.
- Lane, Julia and Schur, Claudia. (2010). Balancing Access to Health Data and Privacy: A Review of the Issues and Approaches for the Future. In: *Health Services Research* 45.5p2, pp. 1456-1467. ISSN: 1475-6773. DOI: 10.1111/j.14756773.2010.01141.x. URL: <http://dx.doi.org/10.1111/j.14756773.2010.01141.x>.
- Lane, Julia and Stodden, Victoria. (2013). What? Me Worry? What to Do About Privacy. *Big Data, and Statistical Research*. URL: <http://magazine.amstat.org/blog/2013/12/01/bigdatastatresearch/>.
- Lazer, D. M., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. In: *Science* 343.6176, pp.1203-1205. DOI: 10.1126/science.1248506.eprint: <http://www.sciencemag.org/content/343/6176/1203.full.pdf>. URL: <http://www.sciencemag.org/content/343/6176/1203.short>.
- Manovich, Lev. (2012). Trending: The Promises and the Challenges of Big Social Data. Debates in the Digital Humanities. In: *Debates in the Digital Humanities*. Ed. by M K Gold. University of Minnesota Press. Chap. 27, pp. 460-475.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A revolution that will transform how we live, work, and think*. John Murray, London.
- Morton, J., Runciman, B. & Gordon, K. (2014). *Big Data: Opportunities and challenges*. BCS Learning & Development Limited, Great Britain.
- Mundie, Craig. (2014). Privacy Pragmatism; Focus on Data Use, Not Data Collection. In: *Foreign Aff.* 93, p. 28.
- Nissenbaum, Helen. (2010). *Privacy in Context: technology, policy, and the integrity of social life*. Stanford, Calif: Stanford Law Books. ISBN: 978-0804752367.
- Online Security and Privacy at Xoom. (n.d.). Retrieved on 20-Apr-2015 from: <https://www.xoom.com/india/security-center>
- Podesta, J., Pritzker, P., Moniz, E.J., Holdren, J., & Zients, J. (2014). Big Data: Seizing Opportunities, Preserving Values. *Tech. rep. Executive Office of the President*. URL: http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.
- Prydz, Espen Beer. (2014). Knowing in time: How technology innovations in statistical data collection can make a difference in development. *PARIS 21*. URL: <http://www.paris21.org/sites/default/files/PARIS21-DiscussionPaper2Knowing.pdf>.
- Ruths, Derek and Jürgen Pfeffer. (2014). Social media for large studies of behavior. In: *Science* 346.6213, pp.1063-1064. DOI:10.1126/science.346.6213. 1063.eprint: <http://www.sciencemag.org/content/346/6213/1063.full.pdf>. URL: <http://www.sciencemag.org/content/346/6213/1063.short>.
- Salsburg, David. (2001). *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. Macmillan.
- Schintler, Laurie A. and Rajendra Kulkarni. (2014). Big Data for Policy Analysis: The Good, The Bad, and The Ugly. In: *Review of Policy Research* 31.4, pp. 343- 348. ISSN: 1541-1338. DOI: 10.1111/ropr.12079. URL: <http://dx.doi.org/10.1111/ropr.12079>.
- Taylor, Linnet and Schroeder, Ralph. (2014). Is bigger better? The emergence of big data as a tool for international development policy. English. In: *Geo Journal*, pp. 1-16. ISSN: 0343-2521. DOI: 10.1007/s10708-014-9603-5. URL: <http://dx.doi.org/10.1007/s10708-014-9603-5>.
- Taylor, Linnet, Schroeder, Ralph & Meyer, Eric. (2014). Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same? In: *Big Data & Society* 1.2. DOI : 10.1177/2053951714536877. eprint: <http://bds.sagepub.com/content/1/2/2053951714536877.full.pdf+html>. URL: <http://bds.sagepub.com/content/1/2/2053951714536877.abstract>.
- Tufekci, Zeynep. (2013). Big Data: Pitfalls, Methods and Concepts for an Emergent Field. In: *SSRN* (March 7 2013). DOI: 10.2139/ssrn.2229952. URL: <http://ssrn.com/abstract=2229952>.
- United Nations Global Pulse White Paper. (2012). *Big Data for Development: Opportunities and Challenges*. Retrieved on 15-Apr-2015 from: <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf>
- Velleman, P F. (1997). The Philosophical Past and the Digital Future of Data Analysis: 375 Years of Philosophical Guidance for Software Design on the occasion of John W. Tukey's 80th Birthday. In: *The practice of Data Analysis: Essays in Honor of John W. Tukey*. Princeton Legacy Library, pp. 317-337.
- Vis, Farida. (2013). A critical reflection on Big Data: Considering APIs, researchers and tools as data makers. In: *First Monday* 18.10. ISSN : 13960466.
- Washington, Anne L. (2014). Government Information Policy in the Era of Big Data. In: *Review of Policy Research* 31.4, pp. 319-325. DOI: 10.1111/ropr. 12081.

What is COSMOS? (n.d.). Retrieved on 12-Apr-2015 from <http://www.cs.cf.ac.uk/cosmos/>

Zikopoulos, P., Eaton, C., Deeros, D., Deutsch, T. & Lapis, G. (2015). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media.

Zwitter, Andrej. (2014). Big Data ethics. In: *Big Data & Society* 1.2. DOI : 10.1177 / 2053951714559253. eprint: <http://bds.sagepub.com/cgi/reprint/1/2/2053951714559253>. URL : <http://bds.sagepub.com/cgi/content/abstract/1/2/2053951714559253>.

Madhukar Dayal has served as a Gazetted officer in Indian Railways (Indian Railways Service of Mechanical Engineers) for over twenty years. His experience spans Railway operations, information technology projects, R&D, as well as, teaching. He has travelled widely across the country conducting seminars, delivering speeches, advising IR in IT projects, and in applications of emerging technologies in Indian Railways. He was also the editor of IRIMEE, Jamalpur newsletter and conference proceedings. He earned his Fellow (Computers & Information Systems) degree from IIM Ahmedabad. He has taught C, C++ (OOP), Industrial Management, computer hardware, software, and networking courses at the graduate level and in Management Development Programs. Currently he teaches Spreadsheet Modelling, Information Systems for Managers, and Modern Computing Applications for Businesses in the post-graduate programs, and DBMS & OLTP in the Fellow (PhD) program. His research interests include high performance compute cluster algorithms and business and government applications of technology.

Sachin Garg is a Ph.D. student at George Mason University's School of Policy, Government & International Affairs. He is currently researching how Big Data can be used to analyse and inform in Public Policy, and how policies impact the generation, collection and dissemination of such Data. Prior to joining the doctoral programme, Sachin was working as a Software Architect in Yahoo!. He has extensive and varied experience with Open Source Software, especially Linux. He holds a Masters in Computer Science from the University of Allahabad, India.

Rubaina Shrivastava is an Academic Associate (Information Systems) at Indian Institute of Management, Indore (IIMI). Earlier, she worked as a Research Fellow (Information Security) at National Law Institute University (NLIU, Bhopal). She has received Post Graduate from NLIU, Bhopal in the stream of Cyber Laws and Information Security. Also, she holds two globally respected certification viz. CCNA and RHCE. She is a compassionate enthusiast for addressing cybersecurity techno-legal challenges and to inform others about the good that can be done in line.

Big Data - Intuitive Lessons for HR Managers

Debolina Dutta

"Big Data" has emerged as the new buzz word in management and industry, who look to tame the volume, velocity and variability of massive data to reveal hitherto insights, which would provide organizations with a competitive edge. After all, analysis and insights from big data is what has been the genesis of large number of service and product organizations. Trend patterns emerging from user actions have resulted in crafting highly personalized user experience and the new wave of advertising. This has been leveraged successfully by the large web based startups like Amazon, Facebook, Google and Yahoo.

The buzz around big data is enough to make sensible people head for the hills! Terms in the Hadoop bestiary like HDFS, Mapreduce, Pig, Zookeeper, Hive, Mahout, Sqoop, Flume are enough to deter most HR professionals from testing these waters. However, just as driving a car requires some overall coordination, sensitivity to traffic behavior and a clear sense of what is the intended destination and not a degree in automobile engineering or a minute understanding of the workings or mechanism of the vehicle, grappling with big data needs skills which are not so difficult.

So what competencies are required to become a true big data master? Since the differentiation emerges from the speed of the insights provided, agility is a key requirement. However, what to look for necessitates a high level of domain understanding, an ability to build correlation with strategic outcomes for the organization, an innate curiosity and an entrepreneurial bent of mind. If retail analytics have established a correlation between diaper and beer sales, resulting in retail outlets bundling these products together, can HR not find similar patterns? Surely, it is not too much to ask for from HR practitioners? Not really. Most HR functions are sitting on goldmines of large employee data and each function within HR can independently explore what insights could be gleaned from the volume, velocity and variety of data accessible to them. The variety aspect of the data

would refer to the multiple sources of data, which does not necessarily fall into neat relational structures. Moreover, experienced practitioners have an intuitive understanding of which levers in their operations could produce strategic outcomes for the business and the organization. All it requires is a questioning bent of mind, to challenge existing paradigms of efficiency and effectiveness and evaluate the impact of these initiatives through the arsenal of big data.

It is also important to understand the critical difference between correlation and causality. Therefore, a word of caution on the tendency to find "spurious correlations". For instance, a phenomenal correlation of 99.79% exists between US spending on science, space and technology with suicides by hanging, strangling and suffocation ! Or how about a 66.6% correlation between films Nicholas Cage appears in with the number of people who drown by falling in a pool! It would be extremely imprudent to establish causality in these cases and take strategic decisions driven by these correlations. ! However, it is equally not important to establish or even explain the causality to appreciate the correlation . This is where a rich domain understanding becomes relevant to connect the dots- prudently!

Too much to digest? Let me share a few illustrative examples. A discussion with a recent assessment firm indicated that they were offering services to a large number of organizations for filtering the large volume of campus recruits applying to organizations. A sharp campus recruitment head then took it upon himself to link the assessment scores (available from a different data source), to performance on the job and retention of the campus hires (available in the organizations performance management system). Some extremely interesting insights emerged from the data trend, which were useful for the organization to sharpen its recruitment strategy for better organizational outcomes.

I could share a personal example of how a Talent Acquisition function I was part of, was actually working

blindly against non-validated perceptions and biases which were actually counter-productive to the organizations' interests. The business head had given us a mandate to improve quality of hire, with a clear directive to increase the intake from top tiered organization. The rationale behind this mandate seemed intuitive, with talent at top tiered organization having gone through sufficient screening process to ensure high caliber talent. Additionally, having worked at these top tier organizations, the assumption was that these hires would produce non-linear outcomes in terms of quality of performance and productivity. The Talent Acquisition function took affirmative action and congratulations were generously awarded by all, when the intake of talent from these top tier organizations was increased from 23% of lateral hires to 78% of lateral hiring. Naturally, this came at a significant increase in the salary cost. However, reviewing the hiring data of over 5000 hires over a couple of years, correlating with performance and retention indicated that there was no significant correlation of pedigree of company hired from with the performance. What this meant was that with no significant productivity increase, we had, in effect, increased the cost base for the organization by 5% and actually got appreciation for it! Naturally, corrective strategies were applied and hiring was then focused on increasing Tier hire, but within the same salary grids applicable to the larger organization and no deviations in cost allowed.

An interesting example from another top IT services firm has been how they have been using the data from their pre-boarding gamification platform. Insights from how offered candidates engaged with the organization through the pre-boarding gamification activity provided a heads-up to the Talent Acquisition function about which candidates were likely to abort. Proactive recruitment was therefore initiated only for these high risk cases and this contributed to overall utilization level increasing and bench strengths reducing,

making this firm a frontrunner on these metrics, with some of the best statistics in the industry.

Fundamental to all these examples is the strong feedback loop used to link data streams to build powerful and actionable insights. The purpose of Big Data is not to generate more data but to produce actionable outcomes. Examples cited here are not necessarily restricted to the domain of talent acquisition. For instance, HR Managers could be encouraged to look at trends within performance management systems, link with reward and recognition or correlate with training interventions and retention statistics. Depending on the business problem critical to the specific organization, HR managers should start use the "drivetrain approach" in conjunction with a model assembly line and this can be used to link a predictive model to actionable outcomes. Can HR managers consider performance metrics of certain teams or business units to provide recommendations to other business functions on HR interventions which will result in improved business outcomes, much akin to the recommendation engine of Ecommerce sites? So instead of waiting for best practice sharing on a knowledge site, can some proactive data mining throw up these nuggets? A lot of HR decisions are based on lag data and maybe the time has come for us to get into predictive indicators for our decision making. We can all start with asking simple questions around "What objectives is our HR function trying to achieve? What levers do we have at our disposal to achieve this objective?"

All of us would be amazed at the possibilities which emerge from these simple questions.

Debolina Dutta has 23 years of work experience in HR & Sales domain. She had a degree in Electrical Engineering from College of Engineering, Pune, a Post Graduate Degree in management from IIM, Bangalore and completed her FPM-Industry program from IIM Indore. She is working as Director and Head of HR, VF Brands (I).

A Review of Evolution of Theories and Models of Technology Adoption

Rajesh Sharma and Rajhans Mishra

Abstract

Technology adoption is one of the mature areas of IS research. Research in this domain has evolved over time by conceptualizing new factors which can better explain the phenomena of technology adoption resulting in development of several theories and models. This process of evolution has been primarily driven by rapidly changing technology scenario and has led to new factors which are grounded in theory from other disciplines. In this paper, we strive to trace the evolution of various theories and models of technology adoption over the years with the objective of presenting an overview of this important domain to the researchers who intend to apply these models in their research. We also look at need for identification of new factors in the technology adoption models in wake of rapid technology changes in future.

1. Introduction

Technology adoption is one of the mature area of research in information systems. Carr (1999) has defined technology adoption as the 'stage of selecting a technology for use by an individual or an organization'. With rapid strides being made in technology innovations in every conceivable domain, the issues related to technology adoption have gained increasing prominence in recent times. Huge investments are made by organizations and governments for introducing new technologies that have the potential of bringing a paradigm shift in the life-style of the users. However, these investments may not yield results if the innovations are not adopted by the intended users. Initial failure of diffusion of Electronic Health Record (EHR) systems in US (Simon et al. 2007) and Enterprise Resource Planning (ERP) systems (Addo & Helo 2011) are some of the examples of the technologies that failed to take off in spite of promising start. More recent examples are that of cloud computing (Low et al. 2011) and e-

Government (Venkatesh et al. 2012) that were promising in respect of the advantages they offered to the users but still have not been adopted by the users to the extent expected.

Several studies have revealed that technology adoption is not related to the aspects of technology alone but has evolved as a much more complex process involving dimensions of user attitude and personality (Venkatesh et al. 2012), social influence (Ajzen and Fishbein 1975), trust (Gefen et al. 2003) and numerous facilitating conditions (Thompson et al. 1991). It is necessary to understand the evolution of this research area in Information Systems and look at future research opportunities.

In this study, we trace the evolution of research in the area of technology adoption over the years by means of a review of the existing literature on the subject. The study is undertaken with two objectives. The first objective of this paper is to present an overview of this important domain to the researchers who intend to apply these models in their research. Second objective of the study is to look at the need to identify new constructs that may possibly be used in explaining adoption of emerging technologies such as e-government, cloud computing, mobile government etc. and further build the theoretical foundations of the existing body of knowledge.

2. Research Methodology

Methodology of literature survey was followed for this paper. Research papers with relevant keywords (such as technology adoption, technology adoption theory, technology adoption model etc) were downloaded from online databases like EBSCO, Google Scholar, Proquest, INFORMS etc. The papers were scrutinized to identify and classify them on the basis of themes on which they were focussed. Details of methodology, data set and major findings of the papers were tabulated for extracting

their common and differentiating features. References of latest papers were once again searched online to include any missing papers left out due to new keywords. Since the objective was to look at evolution of theories and models over the years rather than their application, we picked up a survey paper by Sarkar (2009), review papers by Chuttar et.al. (2009) and Long (2010) that provided an overview of some of the important theories that have evolved over the years. Seminal papers related to these theories were then downloaded and studied in depth to understand the constructs and the context of the study.

3. Technology Adoption Theories and Models

Literature review revealed interchangeable use of the terms adoption and diffusion although these terms are quite distinct from each other. Therefore, noting the difference between these two terms is in order. Adoption refers to "the stage in which a technology is selected for use by an individual or an organization" (Carr, 1999) while the term diffusion refers to "the stage in which the technology spreads to general use and application" (Rogers, 2003). Therefore, while the term adoption is used at individual level, diffusion can be thought of as adoption by the masses. From the point of view of our study, both the terms are important because adoption will generally lead to diffusion. Hence, while looking into the evolution of research of technology adoption we take into account the diffusion studies as well as adoption studies.

Two major streams of research have evolved on adoption of technologies. One stream pertains to adoption at individual and the other at organizational level. If the intention or usage by an individual is the subject of study, it is considered as adoption at individual level (e.g., Compeau and Higgins 1995; Davis et al. 1989). On the other hand, if success of implementation by the organization is under study, it is considered as adoption at organization level (Leonard-Bartonan and d Deschamps 1988). Adoption at individual and organizational levels leads to mass adoption which is termed as the diffusion of technology.

The theories and models that have evolved for explaining adoption of technology are summarized in chronological order as follows:

A. Diffusion of Innovation Theory (Roger, 1960): Research in diffusion can be traced back to the epic work by Everett Roger's in 1960 named as the Diffusion of Innovation Theory which has been widely applied by the researchers over the years.

The main idea of the theory is that there are four elements that influence the spread of a new idea: the innovation, communication channels, time and social system. The process of diffusion consists of five stages, namely, knowledge, persuasion, decision, implementation, and confirmation. It results in six categories of users: innovators, early adopters, early majority, late majority, laggards and the leap-froggers. The theory can be depicted as shown in Figure 1.

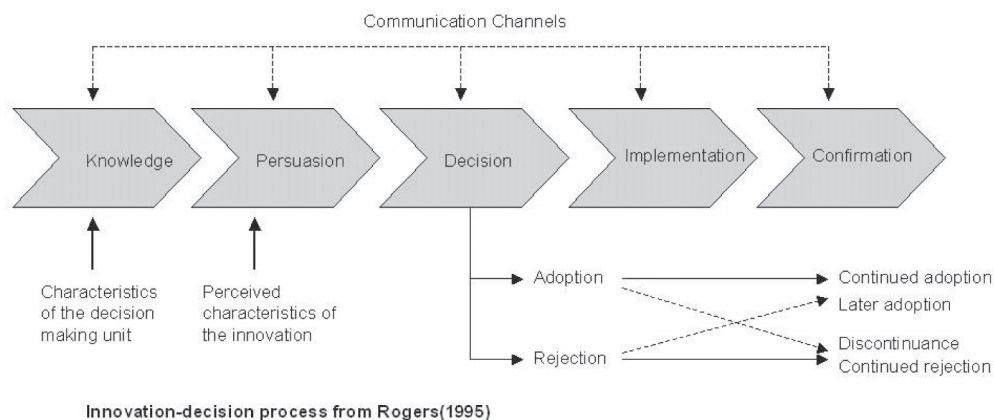


Figure 1: The Diffusion of Innovation Theory (Rogers, 1960)

The diffusion innovation theory provided the concept of S-shaped curve of adoption which was also called as the epidemic model of adoption. According to this curve, spread of infections among the population can be held as an analogy to the pattern of spread of a new technique or idea. According to this analogy, initially the rate of spread is slow. In the mid range of the graph, the rate of spread accelerates and finally the rate of spread tapers off resulting in an S-shaped curve depicted in Figure 2.

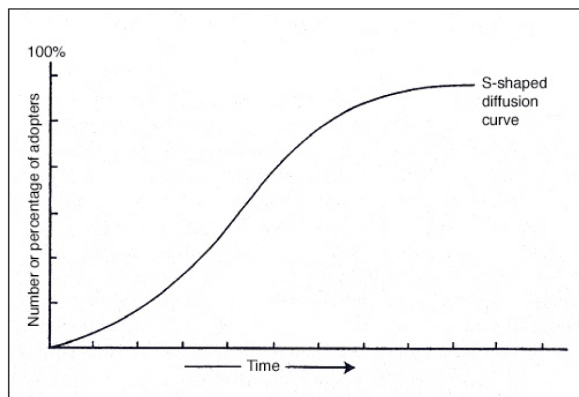


Figure 2: S-shaped Adoption Curve (Rogers, 1960)

The reasoning for such S-shape curve is that initially the innovation has to come from outside the boundaries of the social system prevalent at that time. This implies that number of people that are exposed to the innovation are few in the beginning. As these people in the social system start accepting the innovation, they bring it in contact with more and more people. Therefore the rate of spread keeps on increasing. Eventually, the innovation is accepted by most of the members of social system and the rate of spread declines. As there are no more members left for accepting the innovation, the spread stops completely.

The S-shaped curve depicted in Figure 2 illustrates that there is a critical "take off point" at which the

slope of the growth curve becomes positive and number of members who have adopted the innovation becomes so large that there are hardly any new members left for adopting it. According to Rogers (1960), this point occurs when nearly 10% to 20% of the members of the social system have adopted the innovation.

The S-shaped adoption curve described above applies to most of the innovations that come up from time to time. However its application is of special significance for adoption of communication technology where it is referred to as Metcalfe's law (Gilders 1993). In this case, value of the innovation is enhanced for existing users of the communication system as more and more people adopt the innovation. Each addition of user has a positive effect on existing users of the system which results in acceleration of the adoption curve. Phenomenal growth of the Internet over last one and half decade is often interpreted by this law.

B. Theory of Reasoned Action (Fishbein and Ajzen, 1975): Theory of Reasoned Action (TRA) has its roots in social psychology setting. The theory proposes three general constructs, namely "behavioural intention (BI), attitude (A), and subjective norm (SN)". According to TRA behavioural intention of a person depends on his attitude and subjective norms. Mathematically, it can be interpreted that behavioural intention is the summation of attitude and subjective norms. Moreover, intention of a person likely to convert to action if there is the intention to behave in a specific manner is strong enough.

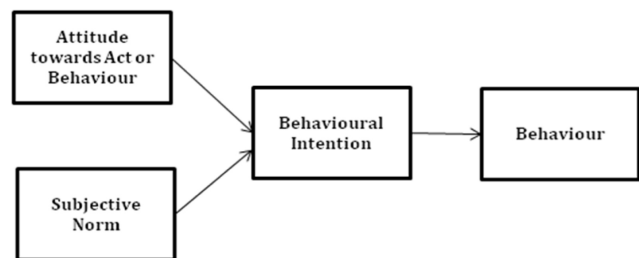


Figure 3: Theory of Reasoned Action, (Fishbein and Ajzen, 1975)

The definition of various constructs used in the theory are as given in Table 3.

Table 1: Constructs used in Theory of Reasoned Action

Construct	Definition
Attitudes	"Sum of beliefs about a particular behaviour weighted by evaluations of these beliefs".
Subjective norms	"Influence of people in one's social environment on his behavioural intentions; the beliefs of people, weighted by the importance one attributes to each of their opinions that will influence one's behavioural intention"
Behavioural intention	"Function of both attitudes toward a behaviour and subjective norms toward that behaviour which has been found to predict actual behaviour"

Source: Fishbein and Ajzen (1975)

C. Theory of Planned Behaviour (Ajzen, 1991): The

Theory of Planned Behaviour (TPB) was proposed by Icek Ajzen in 1991 and was developed from the Theory of Reasoned Action (TRA) which was proposed by Martin Fishbein and Ajzen in 1975. TPB adds the concept of Perceived Behavioural Control (PBC) to the constructs attitudes and subjective norms which make the TRA. Perceived behavioural control refers to "people's perception of the ease or difficulty of performing the behaviour of interest". It differs from Rotter's (1966) concept of perceived locus of control because it is not constant and varies with different situations faced by the individual. Locus of control is considered to be a more generalized expectancy of the individual that remains fairly stable across situations. In this way, the criticism faced by TRA that it is based on relatively static construct of attitude and thus cannot be used for prediction of behavioural outcome has been addressed by TPB. The roots of concept of PBC are grounded in the Self-Efficacy Theory (SET) proposed by Bandura (1977) which in turn came from the Social Cognitive Theory. Bandura (1986) defined self-efficacy as "the judgments

of how well one can execute courses of action required to deal with prospective situations". According to the theory, self-efficacy is the most important determinant for behavioural change since it leads to building up of coping behaviour.

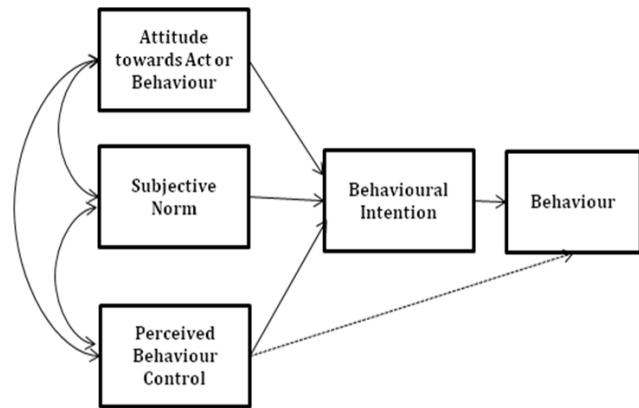


Figure 4: Theory of Planned Behaviour (Ajzen, 1991)

The definition of additional construct (PBC) is given in Table 2.

Table 2: Additional Construct used in Theory of Planned Behaviour

Construct	Definition
Perceived Behaviour Control	"People's perception of the ease or difficulty of performing the behaviour of interest which in turn depends on the self efficacy which is the judgments of how well one can execute courses of action required to deal with prospective situations."

Source: Ajzen (1991)

D. The Social Cognitive Theory (Bandura, 1986): Focus

of the Social Cognitive Theory (SCT) is on the concept of self-efficacy which is defined as "the judgment of one's ability to use a technology to accomplish a particular job or task" (Compeau and Higgins, 1995). According to SCT, behaviour of the user is influenced by expectations of outcome related to personal as well as performance-related gains. Self-efficacy, in turn, influences the expectation of outcome of both types. While esteem of the person and his sense of achievement relate to personal outcome expectations, outcome expectations related to performance on the job

lead to performance related expectations. According to SCT, there are two opposing factors that influence behaviour of the users. Positive contribution is made by the factor "affect" which is the extent to which an individual likes his job. On the other hand, negative contribution to desired behaviour is made by the factor "anxiety" which is the anxious reaction of the person while performing a job such as trying to use a computer with which the person is not very familiar. This theory has been widely used in adoption studies.

E. Technical Adoption Model (Fred D Davis, 1989):

Technology Adoption Model (TAM) has been widely used in technology adoption studies. The strength of the model lies in its simplicity as it has only two constructs, namely, "perceived usefulness" and "perceived ease of use" for predicting extent of adoption of new technologies at individual level as shown below

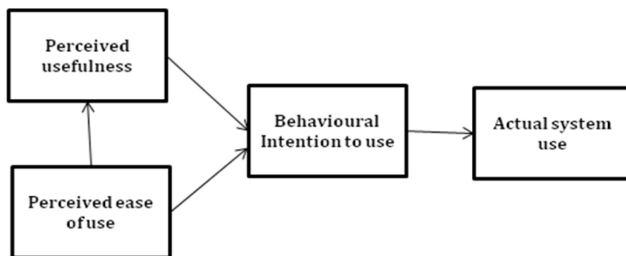


Figure 5: Technology Adoption Model (Davis, 1989)

These constructs are derived from Bandura's Self Efficacy Theory (1982) which defines perceived ease of use as "the judgments of how well one can execute courses of action required to deal with prospective situation" and from Rogers and Shoemaker (1971) paper which defines complexity (interpreted as ease

of use) as "the degree to which an innovation is perceived as relatively difficult to understand and use". The definitions of these constructs are depicted in Table 1.

TAM was originally tested in the context of adoption of email service and file editor at IBM Canada with 14 items on each of 2 constructs. The results of the survey on sample of 112 users validated the model with the finding that perceived usefulness is a stronger factor than perceived ease of use that drives technology adoption. In next ten years, TAM became well-established as a robust, powerful, and parsimonious model for predicting user acceptance. King and He (2006) presented a meta analysis of TAM and found that it is a valid and robust model with applications in a wide range of areas. Dwivedi et al (2010) carried out a comparison of TAM and UTAUT (Venkatesh et al. 2003) and found that focus is now shifting away from TAM to UTAUT while citing in the research articles. In another study, Benbasat & Barki (2007) have criticized TAM especially on the grounds of its limitations in the fast-changing IT environment.

F. The Model of PC Utilization (Thompson et. al. 1991):

The model is based on the Theory of Human Behaviour by Triandis (1977) which differs in some ways from the Theory of Reasoned Action because it makes a distinction between cognitive and affective components of attitudes. Beliefs belong to the cognitive component of attitudes. According to this theory "Behaviour is determined by what people would like to do (attitudes), what they think they should do (social norms), what they have usually done (habits), and by the expected consequences of

Table 3: TAM Model by Davis (1989)

Construct	Definition	Theoretical Background
Perceived usefulness	The degree to which a person believes that using a particular system would enhance his or her job performance	Bandura's Self efficacy theory (1982)
Perceived ease of use	The degree to which a person believes that using a particular system would be free of effort	Rogers and Shoemaker (1971)

Source: Davis (1989)

their behaviour". This theory primarily deals with extent of utilization of a PC by a worker where the use is not mandated by the organization but is contingent on the option of the user. In such a setting, the theory posits that the use of computer by the worker is likely to be influenced by several factors such as his feelings (affect) toward using PCs, prevalent social norms regarding use of PC at the workplace, general habits related to use of the computer, consequences expected by the user by using the PC and extent of conditions that are present at the work place for facilitating use of PC. These constructs are depicted in the figure 6.

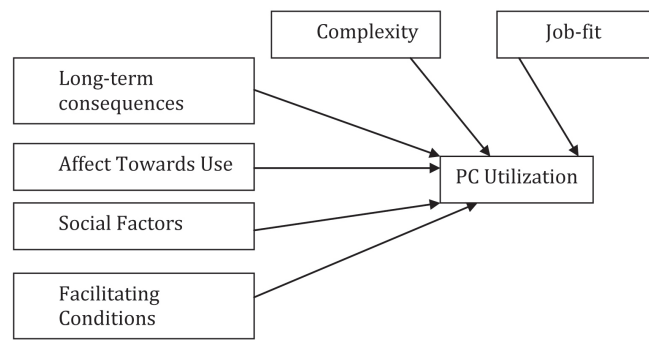


Figure 6: The Model of PC Utilization (Thompson et al. 1991)

The definition of the constructs used in the model are given in Table 4.

G. The Motivation Model (Davis et al., 1992): Davis applied the motivational theory to study information technology adoption and use. The main premise of the Motivation Model is that there are extrinsic and intrinsic motivations that shape the behaviour of the user. Extrinsic motivation is defined as the perception that users want to perform an activity "because it is perceived to be instrumental in achieving valued

outcomes that are distinct from the activity itself, such as improved job performance, pay, or promotions" (Davis et al., 1992, p. 1112). Examples of extrinsic motivation are perceived usefulness, perceived ease of use, and subjective norm. On the other hand, if performing an activity leads to a feeling of pleasure and results in satisfaction for the individual, such behaviour can be classified as intrinsic motivation. (Vallerand, 1997). Users want to perform an activity "for no apparent reinforcement other than the process of performing the activity per se" (Davis et al., 1992,

Table 4: Constructs used in the Model of PC Utilization (Thompson et al. 1991)

Construct	Definition
Job-fit	"The extent to which an individual believes that using a technology can enhance the performance of his or her job."
Complexity	"The degree to which an innovation is perceived as relatively difficult to understand and use."
Long-term consequences	"Outcomes that have a pay-off in the future."
Affect Towards Use	"Feelings of joy, elation, or pleasure, or depression, disgust, displeasure, or hate associated by an individual with a particular act."
Social Factors	"Individual's internalization of the reference group's subjective culture, and specific interpersonal agreements that the individual has made with others, in specific social situations."
Facilitating Conditions	"Provision of support for users of PCs may be one type of facilitating condition that can influence system utilization."

Source: Thompson et al. (1991)

p. 1112). An examples of intrinsic motivation is the extent of enjoyment that a person derives from playing with a computer (Davis et al., 1992; Venkatesh, 2000).

H. Extended TAM2 model (Venkatesh and Davis, 2000): Venkatesh & Davis modified TAM to include additional key determinants of TAM's perceived usefulness and usage intention constructs in their extended TAM model. The additional constructs included social influence processes (subjective norm, voluntariness and image) and cognitive instrumental processes (job relevance, output quality, result demonstrability and perceived ease of use) which are depicted in Figure 7.

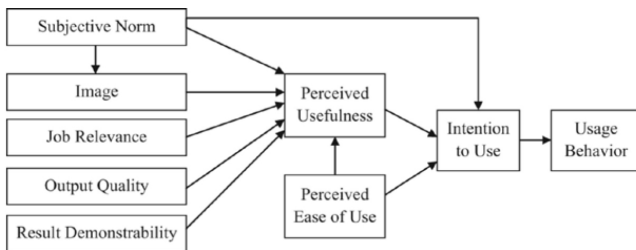


Figure 7: Extended Technical Adoption Model (Venkatesh & Davis, 2000)

The definitions and theoretical basis of the constructs are summarized in Table 5.

I. Unified Theory of Acceptance and Use of Technology (Venkatesh, 2003): This theory, popularly referred as UTAUT was postulated in 2003 by Venkatesh et.al. by a systematic review and consolidation of the constructs of earlier eight models (TRA, TAM, MM, TPB, TAM2, DOI, SCT and model of personal computer use). It is meant to serve as a comprehensive model that can be applied across a range of applications. It has four key constructs namely "performance expectancy, effort expectancy, social influence and facilitating conditions" which are depicted in Figure 8.

For developing the unified model, the authors have compiled and tested all the constructs that were used in previous models and theorized that out of the seven constructs used earlier, four constructs shown above are most significant as determinants of intention to use information technology. They have hypothesized

that remaining three constructs, namely, attitude toward using technology, self efficacy, and anxiety are theorized not to be the direct determinants of intention as they are fully mediated by ease of use which has been considered in the unified model as performance expectancy. Therefore, these three constructs have been removed from the UTAUT model. The constructs in the unified model are defined as in table 6.

The unified theory is proposed to be superior as it is able to explain 70% of the variance while the earlier theories were explaining only 30-40% variance in the adoption behaviour (Venkatesh et al. 2003). However, it is criticized on the grounds of being overly complex, not being parsimonious in its approach and its inability to explain individual behaviour (Casey & Wilson-Evered 2012; Van Raaij & Schapers 2008). A comprehensive review of 450 articles that have cited UTAUT was carried out by Williams et al (2011) and they found that only a small number of articles have actually used the constructs of UTAUT in their study-rather, it has been used more for theory-building.

J. Model of Acceptance with Peer Support, (MAPS, Sykes et al., 2009): Model of Acceptance with Peer Support (MAPS) provides an integration of earlier research that was focussed on individuals with relevant constructs of social network in a way that helps to extend the scope of earlier theories. The authors propose that there are two types of social ties. First tie between employees relates to obtaining help from employees that can result in extension of knowledge for using the system. Another kind of tie between employees is related to providing assistance and help to co-workers for enabling better understanding of configuration and deployment of the system. These ties are labelled as "get-help" and "give help" ties. The authors propose two new constructs, namely, "network density" and "network centrality" that relate to the concept of "get-help" and "give-help" respectively. The theoretical backing for these constructs is drawn from earlier research carried out in social network and are posited as key predictors of system use. These constructs are further extended as "valued network density" and "valued network cen-

Table 5: Extended TAM Model by Venkatesh and Davis (2000)

Construct	Definition	Theoretical Background
Subjective Norm	Person's perception that most people who are important to him think he should or should not perform the behaviour in question".	Theory of Reasoned Action (Fishbein and Ajzen 1975) and the subsequent Theory of Planned Behaviour (Ajzen 1991).
Voluntariness and Compliance with Social Influence.	Voluntariness is the extent to which potential adopters perceive the adoption decision to be non-mandatory.	Hartwick and Barki (1994) found that even when users perceive system use to be organizationally mandated, usage intentions vary because some users are unwilling to comply with such mandates.
Image	Moore and Benbasat (1991) define image as "the degree to which use of an innovation is perceived to enhance one's . . . status in one's social system".	Moore and Benbasat (1991)
Job relevance	Defined as an individual's perception regarding the degree to which the target system is applicable to his or her job. Regarded as cognitive judgment that exerts a direct effect on perceived usefulness, distinct from social influence processes.	Cognitive instrumental theoretical underpinnings come from three main areas: work motivation theory (e.g., Vroom 1964), action theory from social psychology (e.g., Fishbein and Ajzen 1975), and task-contingent decision making from behavioral decision theory (e.g., Beach and Mitchell 1978).
Output quality	Output quality measures perception of how well the system performs the job related tasks.	
Result Demonstrability	Defined by Moore and Benbasat (1991) as the "tangibility of the results of using the innovation". This implies that individuals can be expected to form more positive perceptions of the usefulness of a system if the co-variation between usage and positive results is readily discernible.	

Source: Venkatesh & Davis (2000).

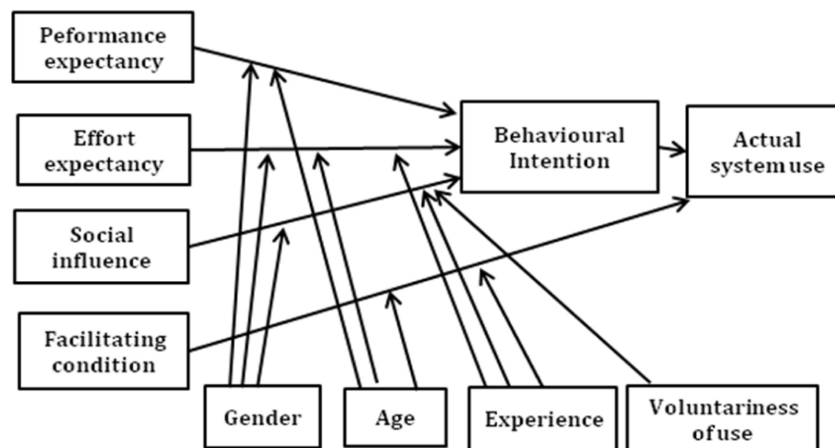


Figure 8: The UTAUT Model (Venkatesh et. al. 2003)

Table 6: Constructs used in UTAUT (Venkatesh et.al. 2003)

Construct	Definition	Root source of the construct from earlier models	Moderators
Performance expectancy	Performance expectancy is defined as the degree to which an individual believes that using the system will help him or her to attain gains in job performance.	The five constructs from the different models that pertain to performance expectancy are perceived usefulness (TAM/TAM2), extrinsic motivation (MM), job-fit (MPCU), relative advantage (IDT), and outcome expectations (SCT).	Gender, Age
Effort expectancy	Effort expectancy is defined as the degree of ease associated with the use of the system.	Three constructs from the existing models capture the concept of effort expectancy: perceived ease of use (TAM/TAM2), complexity (MPCU) and ease of use (IDT).	Gender, Age, Experience
Social influence	Social influence is defined as the degree to which an individual perceives that important others believe he or she should use the new system.	The three constructs related to social influence: subjective norm (TRA, TAM2/IDTPB, TPB), social factors (MPCU), and image (IDT).	Gender, age, voluntariness and experience
Facilitating conditions (no effect on use intention but direct effect on use behaviour)	Facilitating conditions are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system.	Three different constructs used in earlier models are: perceived behavioural control (TPB, DTPB, C-TAM-TPB), facilitating conditions (MPCU) and compatibility (IDT).	Age and experience

Source: Venkatesh et al. (2003)

trality" by taking into account the extent of the resources, information and knowledge available in the system. It is inferred that these constructs can act as additional predictors.

The definitions of these constructs are given in Table 7. The authors conducted a social network study on 87 employees of a supplier-focused business unit of a large multinational company in Finland. With this empirical study they were able to obtain support to their proposed model. They found that the new social network constructs which augmented the previously

established determinants of system use, namely behavioural intention and facilitating conditions, were able to explain about 20 percent additional variance.

To summarize, Table 8 presents the summary of the various theories and models of technology adoption that have evolved over the years.

4. Discussions & Findings from the Study

The study was carried out with two objectives. The first objective was to present an overview of technology

Table 7: Model of Acceptance with Peer Support (Sykes et al., 2009)

Construct	Definition
Behavioral intention	Behavioral intention is defined as "a person's subjective probability that he will perform some behavior" (Fishbein and Azjen 1975, p. 288).
System use	System use is defined as the frequency, duration, and intensity of an employee's interactions with a particular system (Venkatesh et al. 2003).
Facilitating conditions	Facilitating conditions is defined as the "degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system" (Venkatesh et al. 2003, p. 453).
Network density	Network density describes the connectedness of a network and is defined as the actual number of ties in a network as a proportion of the maximum possible number of ties.
Network centrality	Network centrality is defined as the extent of an individual's involvement in assistance exchanges with co-workers (Mossholder et al. 2005; Sparrowe et al. 2001). An individual's centrality has been linked to influence (Burkhardt and Brass 1990), involvement in innovation (Ibarra 1993a, 1993b), and attitudes toward new technology (Rice and Aydin 1991).
Valued network centrality	Valued network centrality refers to peers' perceptions of the level of system-related resources controlled by a focal employee.
Valued network density	Valued network density refers to the connectedness of a focal employee to others, weighted by the perceived strength of the tie and control of system-related information (such as system features, upcoming releases, demo dates), knowledge (such as tips and tricks, short cuts, process sequences), and other tangible resources (such as training resources, manuals, tutorials) that are needed for effective use of a system.

Source: Sykes et al 2009.

adoption theories and models to the researchers who intend to apply these models in their research. This objective has been achieved as evolution of various models has been presented in Table 8. In section 3, we have presented definitions of the constructs and provided literature references to the criticism and applications of various models. It is felt that the paper provides sufficient knowledge of this evolving field in a concise form to the researchers who intend to probe deeper into the field of technology adoption.

Second objective of the study was to look at the necessity to identify new constructs that may possibly be used in explaining adoption of emerging technologies such as e-government, cloud computing, mobile government etc. Need for a sound theoretical model in the fast changing information technology environment has been articulated by Benbasat & Barki (2011). They suggest

that researchers should "redirect their focus toward examining different antecedents (e.g., IT artifact and design) and different consequences (e.g., adaptation and learning behaviors) in order to reach a more comprehensive understanding of what influences adoption and acceptance in different IT use contexts and to provide more useful recommendations for practice" (pp 216). This view of the authors clearly hints towards probable inadequacy of present models in explaining adoption of upcoming technologies such as Internet of Things (IOT), cloud computing and m-Government among others. This also points towards a future research opportunity that needs to be probed by the researchers and fulfils the second objective of the research.

It is therefore posited that the systematic overview of technology adoption contributes to the area of IS/IT

Table 8: Evolution of Theories and Models of Technology Adoption

Year	Theory/Model	Developed By	Constructs/ Determinants of adoption
1960	Diffusion of Innovation Theory	Everett Roger	The innovation, communication channels, time and social system.
1975	Theory of Reasoned Action	Ajzen and Fishbein	Behavioural intention, Attitude (A), and Subjective Norm.
1985	Theory of Planned Behaviour	Ajzen	Behavioural intention, Attitude (A), and Subjective Norm, Perceived Behavioural Control.
1986	Social Cognitive Theory	Bandura	Affect, anxiety.
1989	Technical Adoption	Fred D Davis	Perceived usefulness and perceived ease of use.
1991	The Model of PC Utilization	Thompson et al.	Job-fit, Complexity, Long-term consequences, Affect Towards Use, Social Factors, Facilitating Conditions.
1992	The Motivation Model	Davis et al.	Extrinsic motivation (such as perceived usefulness, perceived ease of use, and subjective norm) and intrinsic motivation (such as perceptions of pleasure and satisfaction).
2000	Extended TAM2 model	Venkatesh and Davis	Social influence processes (subjective norm, voluntariness and image) and cognitive instrumental processes (job relevance, output quality, result demonstrability and perceived ease of use).
2003	Unified Theory of Acceptance and Use of Technology (UTAUT)	Venkatesh et al.	Performance expectancy, effort expectancy, social influence and facilitating conditions.
2009	Model of Acceptance with Peer Support (MAPS)	Sykes et al.	Behavioural intention, System use, Facilitating conditions, Network density, Network centrality, Valued network centrality, Valued network density.

adoption and diffusion research by highlighting various theories and models, including their criticism. The paper identifies the possibility and the need for further evolution in this domain so that more holistic models are developed that can take care of fast-changing technology environment. Therefore it is evident that research has to continue in this domain because the environmental factors driving the adoption process are dynamic and can reveal new constructs of adoption.

5. Conclusions

This review paper provides an overview of theories and models which have evolved over the years for better understanding of adoption process of technologies by individuals and organizations. It is intended to present a bird's eye view of various constructs related to this relatively mature area of information system research and will be of immense value to the contemporary researchers who are planning future work in adoption

of new technologies in the present dynamic environment. The review provides identifies several new areas of research where the existing models of technology adoption may prove to be inadequate. The paper concludes with the view that further evolution in this area of study may be necessary. Future research in this domain may require that existing theories are augmented with constructs from other disciplines of management in order to explain the adoption process of new technologies.

References

- Addo-Tenkorang, R., & Helo, P. (2011, October). Enterprise resource planning (ERP): A review literature report. In Proceedings of the World Congress on Engineering and Computer Science, 2, 19-21.
- Ajzen, I.(1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211
- Ajzen, I. (1985). From intentions to actions: a theory of planned behaviour, in Kuhl, J. and Beckmann, J. (Eds), *Action-Control: From Cognition to Behaviour*, Springer- Verlag, Heidelberg, pp. 11-39.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.
- Bandura, A., & Cervone, D. (1986). Differential engagement of self-reactive influences in cognitive motivation. *Organizational Behavior and Human Decision Processes*, 38(1), 92-113.
- Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review*, 3(3), 439-449.
- Benbasat, I., & Barki, H. (2007). Quo vadis TAM?. *Journal of the Association for Information Systems*, 8(4), 7.
- Carr Jr, V. H. (1999). Technology adoption and diffusion. The Learning Center for Interactive Technology.
- Casey, T., & Wilson-Evered, E. (2012). Predicting uptake of technology innovations in online family dispute resolution services: An application and extension of the UTAUT. *Computers in Human Behavior*, 28(6), 2034-2045.
- Chuttur M.Y. (2009). Overview of the Technology Acceptance Model: Origins, Developments and Future Directions, Indiana University, USA. Sprouts: Working Papers on Information Systems, 9(37). <http://sprouts.aisnet.org/9-37>
- Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS quarterly*, 19(2), 189.
- Davis, F. A.(1989). Perceived usefulness perceived ease of use and user acceptance of information technology, *MIS Quarterly*, 8, 318-339.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace1. *Journal of Applied Social Psychology*, 22(14), 1111-1132.
- Dwivedi, Y. K., Mustafee, N., Carter, L. D., & Williams, M. D. (2010). A Bibliometric Comparison of the Usage of Two Theories of IS/IT Acceptance (TAM and UTAUT). In AMCIS (p. 183).
- Fishbein M., Ajzen, I.(1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesley, Reading, MA.
- Gallivan, M.J.(2001). Organizational adoption and assimilation of complex technological innovations: development and application of a new framework, *The DATABASE for Advances in Information Systems*, 32(3), 51-85.
- Gefen, D., & Straub, D. W. (1997). Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model. *MIS quarterly*, 21(4), 389.
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated model. *MIS quarterly*, 27(1), 51-90.
- Gilder, G. (1993). Metcalfe's law and legacy. *Forbes ASAP*, 13.
- Igbaria, M., & Parasuraman, S. (1989). A path analytic study of individual characteristics, computer anxiety and attitudes toward microcomputers. *Journal of Management*, 15(3), 373-388.
- King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755.
- Leonard-Barton, D., & Deschamps, I. (1988). Managerial influence in the implementation of new technology. *Management Science*, 34, 1252-1265.
- Low, C., Chen, Y., & Wu, M. (2011). Understanding the determinants of cloud computing adoption. *Industrial Management & Data Systems*, 111(7), 1006-1023.
- Moore, G., & Benbasat, I.(1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3) 192-222.
- Rogers, E. M. (2003) *Diffusion of Innovations*, Free Press, New York, Fifth Edition 2003.
- Rogers, E. M., & Shoemaker, F. F. (1971). *Communication of innovations: A cross-cultural approach*. Free Press, New York.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1.

- Sarkar, J. (1998). Technological diffusion: alternative theories and historical evidence. *Journal of economic surveys*, 12(2), 131-176.
- Simon, S., Kaushal, R., Cleary, P., Jenter, C., Volk, L., Poon, E., Oray, E., Lo, H., Williams, D. & Bates, D. (2007). Correlates of electronic health record adoption in office practices: a statewide survey. *Journal of the American Medical Informatics Association*, 14(1), 110-117.
- Sykes, T. A., Venkatesh, V., & Gosain, S. (2009). Model of acceptance with peer support: A social network perspective to understand employees' system use. *MIS quarterly*, 33(2), 371-393.
- Thompson, R. L., & Higgins, C. A. (1991). Personal Computing: Toward a Conceptual Model of Utilization. *MIS quarterly*, 15(1), 125.
- Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation in *Advances in Social Psychology* (29), M. Zanna (ed.), Academic Press, New York,, pp 271-360.
- Van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society*, 19(4), 315-326.
- Van Raaij, E. M., & Schepers, J. J. (2008). The acceptance and use of a virtual learning environment in China. *Computers & Education*, 50(3), 838-852.
- Venkatesh V., & Davis F. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science*, 46(2), 186-204.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.
- Venkatesh, V., Sykes, T. A., & Venkatraman, S. (2014). Understanding e-Government portal use in rural India: role of demographic and personality characteristics. *Information Systems Journal*, 24(3), 249-269.
- Williams, M. D., Rana, N. P., Dwivedi, Y. K., & Lal, B. (2011, June). Is UTAUT really used or just cited for the sake of it? a systematic review of citations of UTAUT's originating article. In ECIS.
- Yousafzai, S. Y., Foxall, G. R., & Pallister, J. G. (2007). Technology acceptance: a meta-analysis of the TAM: Part 1. *Journal of Modelling in Management*, 2(3), 251-280.

Rajesh Sharma is a participant in the Fellow Program in Management (FPM) at Indian Institute of Management Indore (India) in the Information System area. His interest areas in research include e-Government, technology adoption, business process management and marketing of services. He belongs to the Indian Telecom Service (ITS) and is currently on study leave from the Department of Telecom, Ministry of Communication & IT, Government of India.

Rajhans Mishra is an Assistant Professor in Information Systems Area at Indian Institute of Management Indore. He has also served as a visiting faculty at Indian Institute of Management Ahmedabad and Indian Institute of Management Lucknow. His research interest includes recommendation systems, web mining, data mining, text mining, e-Governance and business analytics. He has completed his doctoral work from Indian Institute of Management Lucknow.

Using Social Networking Sites (SNS): Mediating Role of Self Disclosure and Effect on Well-being

Manoj Das

Abstract

Social networking sites (SNS) have acquired the fascination of academicians for quite some time now due to its unique advantages and outreach. This research aims at examining the antecedents of SNS usage from the perspective of people who fear or prefer avoiding face-to-face communication as well as those who lack meaningful relations in their life. Drawing upon uses and gratifications theory (U & G) and self-determination theory, we identify the motives for SNS usage and its impact on individual's well-being.

Previous studies focused on investigation of these motives individually, in isolation of the other attributes with most of the studies being experimental in nature. However, the fact remains that, most of these attributes are co-existing within an individual, and hence it would be interesting to study them simultaneously, which would help in identifying their relative strengths and interactive effects. Hence while addressing this gap in the literature; we seek to investigate the simultaneous as well as the interactional effects of loneliness, social connectedness, social isolation, social anxiety, and the mediating role of self-disclosure. It would help practitioners in designing SNS more effectively to ensure enhanced user satisfaction with their usage.

Keywords: social networking sites, uses and gratifications theory, self disclosure, social anxiety, loneliness

1. Introduction

As per Ishii and Ullmer (1997), "we live between two realms: our physical environment and cyberspace." (p. 234). The "physical environment" represents the physical environment in which we, humans, live (Gibson, 1979). "Cyberspace" hereafter referred to as the "digital world" made up of bits & bytes and distinct due to its intangibility. It remains uncontested that these worlds

are "parallel but disjointed" (Ishii & Ullmer, 1997). With the proliferation of digital technologies in our everyday lives and the increased usage of networking sites, these two realms are slowly converging into one another. We are slowly merging our physical bodies with our online selves and moving from "you are what you wear" to "you are what you post".

Computer mediated communication (CMC) has emerged as an indispensable part of interaction in our everyday lives (Stone, 2009). A social networking site (SNS) represents an online community wherein members share their personal information by creating "profiles" with the purpose of communicating with other users of SNS in different ways and an array of topics (Mahajan, 2009; Pempek, Yermolayeva, & Calvert, 2009). They serve a variety of purposes like 'social searching'- with the purpose of finding information regarding offline contacts, and 'social browsing'- developing new connections with different individuals (Lampe, Ellison, & Steinfield, 2006; Joinson, 2008; Trusov, Bucklin, & Pauwels, 2009), befriending new people (Golder, Wilkinson, & Huberman, 2007), increasing social capital (Ellison, Steinfield, & Lampe, 2006), providing a platform for computer-mediated communication (Bennett, Maton & Kervin, 2008, Ross et.al, 2009).

Despite the immense usage of SNS, there has been a scant amount of research related to it, in particular with reference to the motives of usage and the kind of benefits derived. (Ross et al., 2009). We draw upon uses and gratifications theory (U & G), which specifies 'how and why' a particular media is used with underlying motivations and derived satisfaction associated with it (Stafford, Stafford, & Schkade, 2004). U & G theory explains usage of different media by focusing on the individual motivations driving the decision (Katz, Blumler & Gurevitch, 1999). It identifies the specific gratifications, which individuals seek to satisfy notably their psychological and sociological factors (Rubin,

2002). On the other hand, self determination theory, states that needs for competence, relatedness and autonomy (Ryan & Deci, 2000) play an important role in facilitating ones need and fulfilling need of social development and personal well-being.

This study contributes to the literature by drawing upon uses and gratifications (U&G) theory (Katz, 1959) and self-determination theory (Ryan & Deci, 2000) for identifying the motives behind SNS usage. This would help in understanding how people behave and seek gratification when using SNS. We particularly look at this from the perspective of people who fear or prefer avoiding face-to-face communication as well as those who lack meaningful relations in their life.

Previous studies focused on investigation of these effects individually, in isolation of the other attributes with most of the studies being experimental in nature. However, the fact remains that, most of these attributes are co-existing within an individual, and hence it would be interesting to study them simultaneously, which would help in identifying their relative strengths and interactive effects. Hence while addressing this gap in the literature; we seek to investigate the simultaneous as well as the interactional effects of loneliness, social connectedness, social isolation, social anxiety, and the mediating role of self-disclosure on well-being of an

individual. It would help practitioners in designing SNS more effectively to ensure enhanced user satisfaction with their usage. The structure of remaining paper is as follows: First of all literature review-highlighting motives of SNS usage, followed by hypothesis development and the conceptual framework.

1.1. Social Networking Sites and Motives for their Usage

Boyd & Ellison (2007) have defined SNS as "web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system" (p. 211).

Typically, SNS users create a profile on a SNS with a picture and personal details and interests like place of work, schools attended, favorite sports and sitcoms. After creation of profile, they search for profiles of friends and colleagues and even join groups based on common interest by sending them a request. The most common motive is maintaining existing social networks effectively (Boyd & Ellison, 2007) with sometimes sending invitations to "friends of friends" (Staples, 2009). (Table 1).

Table 1: Motivation for Using Social Media

Author	Motivation for using social media
Kollock (1999)	Anticipated reciprocity, increased recognition, and sense of efficacy
Dholakia, Bagozzi, & Pearo(2004)	Purposive value ,self discovery, maintaining interpersonal connectivity, social enhancement, and entertainment
Lakhani & Wolf (2005)	Creativity, intellectual stimulation, and improving professional skills
Ryan et al. (2006)	Autonomy, competence, and relatedness
Yee (2007)	Achievement, social, and immersion
Joinson (2008)	Social connection, shared identities, photographs, content, social investigation, social network, surfing and status updating
Park, Kee, & Valenzuela (2009)	Socializing, entertainment, self-status seeking and 'information'
Rafaeli, Hayat, & Yaron (2009)	Getting information', 'sharing information' and 'entertainment'.
Zhou, Jin, Vogel, Fang, & Chen, 2011	Functional, Experiential and Social motivations

1.2. Uniqueness of CMC

Text based communication in CMC warrants special attention due to its special features like emoticons to express varied emotions and its importance for people experiencing difficulty in face-to-face communication (Reid & Reid, 2007). (Table 2).

2. Usage of SNS and Well-Being

Researchers have established the existence of a positive relationship between SNS's usage and the psychological well-being of an individual (Kim & Lee, 2011; Liu & Yu, 2013). Having a large number of friends on face book increases an individual's sense of well-being due to a favorable representation of oneself to others (Kim and Lee, 2011), this can also be mediated by self-disclosure and friendship quality rather than have a direct influence on well-being (Valkenburg & Peter, 2009; Wang & Wang, 2011). Huang (2010) has reported that social communication via internet is a positive predictor of psychological well-being. Due to the unique features of CMC, users are comfortable in making self-disclosure, which they would have otherwise avoided in case of face-to-face interactions or online interactions, which leads to high quality friendship (Valkenburg & Peter, 2007). Disclosure is also dependent on the timing, appropriateness of the particular context and comfort level with the partner (Valkenburg & Peter, 2009). Based on displacement hypothesis, spending more time on SNS reduces the quality of real life relations while augmentation hypothesis maintains that usage of SNS increases the quality of their real life relations. (Morgan & Cotton, 2003; Shaw & Gant, 2002; Valkenburg & Peter, 2007)

2.1. Loneliness

Loneliness can be defined as "the cognitive awareness of a deficiency in ones social and per-sonal relationships and the ensuing affective reactions of sadness, emptiness or longing" (Asher & Paquette, 2003 p.75). It can also be alternatively defined as, "being alone, felt deprivation of closeness"(Parkhurst & Hopmeyer, 1999). As social beings, we become vulnerable when we en-ter in social relationships.

Loneliness has been classified into two distinct types based on the kind of relationship deficit they arise from (Weiss, 1974). It can be either emotional loneliness due to the lack of emotionally close relationship with a caring and understanding person like a romantic partner. On the other hand, social loneliness arises from the absence of group ties with friends and other peers.

Loneliness is both a cause of as well as effect of internet usage among college students (Kim, LaRose & Peng, 2009; Tokunaga & Rains, 2010). Individual feeling depression and loneliness perceive themselves to be lacking in skills with regard to social competence (Caplan, 2003). McKenna, Green, and Gleason (2002) have argued that lonely individuals are "somewhat more likely to feel that they can better express their real selves with others on the Internet than they can with those they know offline" (p.28). This would lead them to believe that they are more successful in online social interactions leading to increased usage of networking sites and feeling of well-being (Baker and Oswald, 2010).

Additional support for usages of SNS sites by lonely individuals comes from social augmentation hypothesis, which posits that an additional opportunity for

Table 2: Unique Features of CMC

Author	Feature
Mckenna & Bargh, 2000	Anonymity, it allows internet users to share even intimate conversations with anyone they desire without actually sharing identifying in-formation
Mckenna & Bargh, 2000	Absence of non-verbal cues or vocal cues
Nguyen, Bin, &Campbell, 2012; Tidwell & Walther, 2002	Increased self-disclosure and intimacy
Caplan, 2005	Control over self-presentation

interacting with others exists due to the amplification of individual's total social resources (Bessi re, Kiesler, Kraut, & Boneva, 2008). It also opens up an avenue for meeting and befriending people thereby fulfilling their need of social interaction and relationship thereby eliminating both types of loneliness by satisfying the individual's needs of relatedness as well. This is also in line with the assumptions of stimulation theory and social compensation theory, which state that lonely individuals seek to compensate for lack of meaningful relations by using the self disclosure feature of SNS (Valkenburg & Josen 2007, Skues, Williams & Wise, 2012).

Hence, we can hypothesize that:

Hypothesis 1: There exists a positive relationship between avoiding loneliness and usage of SNS and well-being of individual.

2.2. Social Isolation

Although "loneliness" and "social isolation" are often used interchangeably, but they are in fact distinct concepts. (Bernard & Perry, 2013, de Jong Gierveld, Van Tilburg & Dykstra, 2006). People can be socially isolated without feeling lonely, or feel lonely amongst others.

Social isolation refers to being ignored or excluded with or without explicit declarations (Williams, 2007). Commonly used terms that are used interchangeably are ostracism, social exclusion, and rejection. Some of the reasons for occurrence of social isolation are increasing social cohesiveness, punishment for deviant behavior, getting rid of unwanted members and maintaining order in group (Gruter & Masters, 1986). Humans as social beings are very sensitive towards social isolation as they crave the interpersonal relations and group memberships to feel connected with others (Williams, 2007). It acts as an inhibitor from realizing our basic needs (i.e., belonging, self-esteem, control, and meaningful existence) (Williams, 2009). The absence of these social bonds can lead to deterioration of well-being along with physical and mental morbidity (Cacioppo, Hawkley & Thisted, 2010; Hawkley, Burleson, Berntson & Cacioppo, 2003; Williams, 2007).

We have evolved to develop pain as alarm signal for social isolation as it activates the neural systems, which

shares our affective component of physical pain (Eisenberger, 2012; MacDonald & Leary, 2005). We tend to compensate for the loss of these social bonds by using these SNS's. Drawing upon augmentation hypothesis it can be inferred that individuals generally use SNS for developing their social relations (Walther, 1996), and usage of media enhances the existing social relations of users (Valkenburg & Peter 2007, 2009). The embedded social cues within these media foster a sense of belonging, which acts as a surrogate of having social relations (Derrick, Gabriel, & Hugenberg, 2009).

CMC helps in overcoming the barriers of spatial and temporal constraints, which impede face to face communication for individuals (Wellman & Haythornthwaite, 2002). It provides social support due to its unique features like higher anonymity, option of editing and crafting messages along with access to varied information sources (Walther & Boyd, 2002).

Hence, we can hypothesize that:

Hypothesis 2: There exists a positive relationship between avoiding social isolation and usage of SNS and well-being of individual.

2.3. Social Anxiety

Social anxiety disorder refers to a marked and persistent fear of social or performance situations in which embarrassment may occur, resulting in significant distress and difficulties in functioning (American Psychiatric Association, 2013). Drawing upon the cognitive model of social anxiety, we can infer that individuals experiencing social anxiety strong negative beliefs about themselves (Clark & Wells, 1995; Liebowitz, 1987, Mattick, & Clarke, 1998, Rapee & Heimberg, 1997).

Due to these negative beliefs, such individuals tend to evaluate most social cues as negative evaluation, which tends to increase their anxiety. This increases their inability to process face-to-face transactions due to their concerns with self-presentation (Clark & McManus, 2002; Leary, & Kowalski, 1995; Schlenker & Leary, 1982). Most of these problems are effectively resolved in computer-mediated communication due to ease of control over personal information, and decreased scrutiny in terms of physical appearance and vocal

sounds, anonymity and absence of nonverbal cues there by enhancing it usage and well-being of user (High & Caplan, 2009, Kang, 2007, Madell & Muncer, 2007, Subrahmanyam & Lin, 2007).

Hence, we can hypothesize that:

Hypothesis 3: There exists a positive relationship between social anxiety and usage of SNS and well-being of individual.

2.4. Mediating Role of Self-Disclosure

Self-disclosure occurs when a person provides (i.e., discloses) information about himself or her-self to another person (Cozby, 1973). Based upon social penetration theory, when reciprocal disclosure occurs between individuals it results in development of strong relationships along with intimacy and strong interpersonal bonds (Altman& Taylor, 1973).Sharing of private thoughts, experiences, and emotions is a common and widespread practice in SNS (Joinson & Paine, 2007). Due to the unique features of CMC, youths prefer it to face-to-face communication for self-disclosure (Nguyen, Bin, & Campbell, 2012; Schouten, Valkenburg, & Peter, 2007; Walther, 1996). Increased self-disclosure leads to trust, friendships closeness, and satisfaction with the communication leading to higher levels of engagement. (Berger & Calabrese, 1975; Collins & Miller, 1994; Draper, Pittard, & Sterling, 2008; Jourard, 1971; Morry, 2005)

Research indicates that quality of friendship and well-being are positively related, and happy individuals have better relationships (Diener & Seligman, 2002; Michalos, Hubley, Zumbo, & Hemingway, 2001).This may be due to the feeling of being connected and a sense of belongingness associated with the usage of SNS (Morrow, Ross, Grocott, & Bennett, 2010), due to fulfillment of socio emotional needs (Goswami, 2012). Moreover, SNS users are less conscious about their self-presentation on SNS, which causes them to disclose more information as op-posed to face-to-face conversation (Walther, 1996).

Hence, we can hypothesize that:

Hypothesis 4: Usage of SNS is mediated by self-disclosure and leads to well-being of an individual.

2.5. Subjective Well-being

Subjective well-being stands for how an individual makes sense of his life through cognitive and affective evaluations (Diener, Lucas, Oishi, 2005). It is common knowledge that people share their problems with others in times of stress and the support thus obtained has significant benefit on that person's health and well-being (e.g. Uchino, Cacioppo & Kiecolt-Glaser, 1996). In fact, sharing events and discussing things with others is helpful in building one's social resources thereby allowing development of positive social interactions (Gable & Reis, 2001).Hence, sharing experiences using SNS helps to improve one's social ties with others in addition to providing psychological benefit (Gable, Reis, Impett, & Asher, 2004; Shen, Liu, & Wang, 2013; Valenzuela, Park, & Kee, 2009). A high number of friends on SNS leads to satisfaction with life and well-being since individuals compensate for the loss of meaningful relations in their real life with their friends on the SNS (Lee, Lee, & Kwon, 2011).

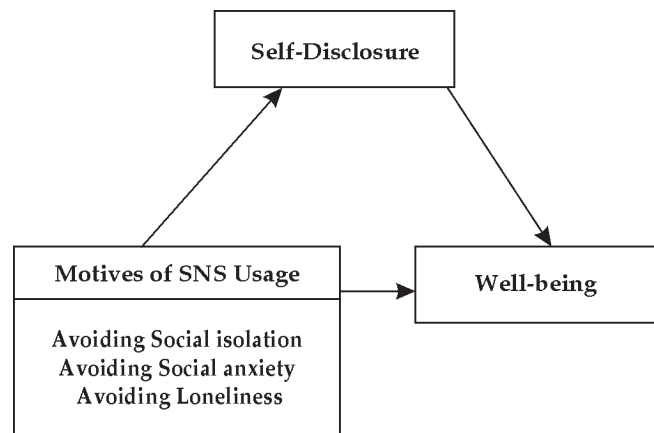


Figure 1 : Conceptual Framework

3. Discussion & Conclusion:

SNS form an indispensable part of our everyday lives; hence, there has been an increased attention towards understanding the motives for their usage. In this paper, we have tried to provide a holistic framework by identifying the antecedents of SNS usage for people who lack meaningful relationships in their lives and compensate for it by using SNS.

For such individuals SNS offers a gateway to a world where they can express their feelings without inhibition, receive social and emotional support they crave and thereby increase their sense of well-being (Goffman, 1959). Disclosing personal information is also easier due to anonymity and the absence of any kind of criticism along with feeling of being in control (Qian & Scott, 2007).

Some practical implications for companies could be targeting such individuals for targeting their products and creating virtual communities wherein they can share their ideas & comments freely (Gupta, Kim & Shin, 2010) which could act as a potential barrier for competitors. Companies would also be able to effectively promote their products and services extensively on such community pages (Algesheimer, Borle, Dholakia, & Singh, 2010). It could serve as a source for idea generation for new products and improvements in existing ones.

In addition, since the success of SNS depends on the number of members, content generation done by users, increased level of site visits and traffic which result in revenue generation (Carroll, 2007, Chen, 2013). Hence, designers of SNS need to keep in mind the motives of the users for using SNS and try to make it a pleasurable experience for using them.

Future research can empirically verify this model and identifying which are the most relevant motives for usage of SNS. It would help practitioners in designing SNS more effectively to ensure enhanced user satisfaction with their usage. Future research may also look into whether personality traits have an impact on usage of SNS. Lastly, a longitudinal study may reveal insights that are more promising.

References

- Algesheimer, R., Borle, S., Dholakia, U. M., & Singh, S. S. (2010). The impact of customer community participation on customer behaviors: An empirical investigation. *Marketing Science*, 29(4), 756-769.
- Altman, Irwin and Dalmas A. Taylor (1973). *Social Penetration: The Development of Interpersonal Relations*. New York: Holt, Rinehart and Winston.
- Asher, S. R., & Paquette, J. A. (2003). Loneliness and peer relations in childhood. *Current Directions in Psychological Science*, 12(3), 75-78.
- Bennett, S., Maton, K., & Kervin, L. (2008). The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology*, 39(5), 775-786.
- Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human communication research*, 1(2), 99-112.
- Bernard, S., & Perry, H. (2013). *Loneliness and Social Isolation Among Older People in North Yorkshire*. Social Policy Research Unit, University of York: York. Loneliness Survey Team A, 3.
- Bessièrè, K., Kiesler, S., Kraut, R., & Boneva, B. S. (2008). Effects of Internet use and social resources on changes in depression. *Information, Community & Society*, 11(1), 47-70.
- Cacioppo, J. T., Hawkley, L. C., & Thisted, R. A. (2010). Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychology and Aging*, 25 (2), 453-463.
- Caplan, S. E. (2003). Preference for Online Social Interaction A Theory of Problematic Internet Use and Psychosocial Well-Being. *Communication Research*, 30(6), 625-648.
- Carroll, E. (2007). *Success Factors of Online Social Networks*. The University of North Carolina, Chapel Hill, USA.
- Chen, R. (2013). Member use of social networking sites-an empirical examination. *Decision Support Systems*, 54(3), 1219-1227.
- Clark DM, Wells A. (1995). A cognitive model of social phobia. In *Social Phobia: Diagnosis, Assessment and Treatment*, ed. R Heimberg, M Liebowitz, DA Hope, FR Schneier, pp. 69-93. New York: Guilford.
- Clark, D. M., & McManus, F. (2002). Information processing in social phobia. *Biological Psychiatry*, 51(1), 92-100.
- Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: a meta-analytic re-view. *Psychological Bulletin*, 116(3), 457-475.
- Cozby, P. C. (1973). Self-disclosure: a literature review. *Psychological Bulletin*, 79(2), 73-91.
- Derrick, J. L., Gabriel, S., & Hugenberg, K. (2009). Social surrogacy: How favored television programs provide the experience of belonging. *Journal of Experimental Social Psychology*, 45(2), 352-362.
- Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network- and small-group-based virtual communities. *International Journal of Research in Marketing*, 21(3), 241-263.
- De Jong Gierveld, J., Van Tilburg, T., & Dykstra, P. A. (2006). Loneliness and social isolation. In A. Vangelisti & D.

- Perlman (Eds.), Cambridge handbook of personal relationships (pp. 485-500). Cambridge, UK: Cambridge University Press.
- Diener, E., & Seligman, M. E. (2002). Very happy people. *Psychological Science*, 13(1), 81-84.
- Diener E, Lucas RE, Oishi S. (2005) Subjective well-being: The science of happiness and life satisfaction. In Snyder CR, Lopez SJ, eds. Handbook of positive psychology. New York: Oxford University Press, pp. 63-73.
- Draper, M., Pittard, R., & Sterling, M. (2008). *Self-disclosure and friendship closeness*. Hanover, IN: Hanover College, 3(4), 5.
- Eisenberger, N. I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience*, 13(6), 421-434.
- Ellison, N., Steinfield, C., & Lampe, C. (2006). Spatially bounded online social networks and so-cial capital: The role of Facebook. In Proceedings of the annual conference of the international communication association, June 19-23, 2006, Dresden, Germany.
- Gable, S. L., Reis, H. T., Impett, E. A., & Asher, E. R. (2004). What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology*, 87(2), 228-245.
- Gable, S. L., & Reis, H. T. (2001). Appetitive and aversive social interaction. In J. Harvey & A. Wenzel (Eds.), *Close romantic relationships: Maintenance and enhancement* (pp. 169-194). Mahwah, NJ: Erlbaum
- Gibson, J. J. (1979). *The ecological approach to visual perception*. London, UK: Lawrence Erlbaum Associates.
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh Social Sciences Research Centre.
- Golder, S. A., Wilkinson, D. M., & Huberman, B. A. (2007). *Rhythms of social interaction: Mes-saging within a massive online network*. In *Communities and Technologies 2007* (pp. 41-66). Springer: London.
- Goswami, H. (2012). Social relationships and children's subjective well-being. *Social Indicators Research*, 107(3), 575-588.
- Grueter, M., & Masters, R. D. (1986). Ostracism: a social and biological phenomenon: An introduction. *Ethological Sociobiology*, 7(3), 149-395.
- Gupta, S., Kim, H. W., & Shin, S. J. (2010). Converting virtual community members into online buyers. *Cyberpsychology, Behavior, and Social networking*, 13(5), 513-520.
- Hawkley, L. C., Burleson, M. H., Berntson, G. G., & Cacioppo, J. T. (2003). Loneliness in every-day life: cardiovascular activity, psychosocial context, and health behaviors. *Journal of Personality and Social Psychology*, 85(1), 105-120.
- High, A. C., & Caplan, S. E. (2009). Social anxiety and computer-mediated communication during initial interactions: Implications for the hyperpersonal perspective. *Computers in Human Behavior*, 25(2), 475-482.
- Huang, C. (2010). Internet use and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 13(3), 241-249.
- Ishii, H., & Ullmer, B. (1997). Tangible bits: Towards seamless interfaces between people, bits and atoms. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp 234-241). New York, NY: ACM Press.
- Joinson, A. N. (2008). Looking at, looking up or keeping up with people?: Motives and use of facebook. In Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, ACM, New York, USA.
- Jourard, S. M. (1971). *Self-disclosure: An experimental analysis of the transparent self*. New York: Wiley.
- Kang, S. (2007). Disembodiment in online social interaction: Impact of online chat on social support and psychosocial well-being. *Cyberpsychology & Behavior*, 10(3), 475-477.
- Katz, E. (1959). Mass communication research and the study of culture: An editorial note on a possible future for this journal. *Studies in Public Communication*, 2, 1-6.
- Katz, E., Blumler, J., & Gurevitch, M. (1999). Utilization of mass communication by individual. In J. Hanson & D. Maxcy (Eds.), *Sources notable selections in mass media* (pp. 51-59). Guilford, CT: Dushking/McGraw-Hill
- Kim, J., LaRose, R., & Peng, W. (2009). Loneliness as the cause and the effect of problematic Internet use: The relationship between Internet use and psychological well-being. *Cyberpsychology & Behavior*, 12(4), 451-455.
- Kim, J., & Lee, J. E. R. (2011). The Facebook paths to happiness: Effects of the number of Facebook friends and self-presentation on subjective well-being. *Cyberpsychology, Behavior, and Social Networking*, 14(6), 359-364.
- Kollock, P. (1999). *The economies of online cooperation: gifts, and public goods in cyberspace*. In: Smith, M.A., Kollock, P. (Eds.), *Communities in Cyberspace*. Routledge, New York, pp. 220-239.
- Lakhani, K.R., & Wolf, R.G. (2005). Why hackers do what they do: understanding motivation and effort in free/open source software projects. In: Feller, J.R., Fitzgerald, S., Hissam, S., Lakhani, R.K. (Eds.), *Perspectives on Free and Open Source Software*. MIT Press, Cambridge, pp. 3-21.
- Lampe, C., Ellison, N., & Steinfield, C. (2006, November). A Face (book) in the crowd: Social searching vs. social browsing. In Proceedings of the 2006 20th anniversary

- conference on Computer supported cooperative work (pp. 167-170). ACM.
- Leary, M. R., & Kowalski, R. M. (1995). The self-presentation model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social Phobia. Diagnosis, Assessment and Treatment* (pp. 94-112). New York: Guilford Press.
- Lee, G., Lee, J., & Kwon, S. (2011). Use of social-networking sites and subjective well-being: A study in South Korea. *Cyber psychology, Behavior, and Social Networking*, 14(3), 151-155.
- Lewandowski, J., Rosenberg, B. D., Jordan-Parks, M., & Siegel, J. T. (2011). The effect of in-formal social support: Face-to-face versus computer mediated communication. *Computers in Human Behavior*, 27(5), 1806-1814.
- Liebowitz, M. R. (1987). *Social phobia. Modern problems of pharmacopsychiatry*, 22, 141-173.
- Liu, C. Y., & Yu, C. P. (2013). Can facebook use induce well-being?. *Cyber psychology, Behavior, and Social Networking*, 16(9), 674-678.
- MacDonald, G., & Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin*, 131(2), 202-223.
- Madell, D. E., & Muncer, S. J. (2007). Control over social interactions: An important reason for young people's use of the internet and mobile phones for communication?. *Cyberpsychology & Behavior*, 10(1), 137-140.
- Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36(4), 455-470.
- McKenna, K. Y., Green, A. S., & Gleason, M. E. (2002). Relationship formation on the Internet: What's the big attraction?. *Journal of Social Issues*, 58(1), 9-31.
- Michalos, A. C., Hubble, A. M., Zumbo, B. D., & Hemingway, D. (2001). Health and other aspects of the quality of life of older people. *Social Indicators Research*, 54(3), 239-274.
- Morgan, C., & Cotton, S. R. (2003). The relationship between Internet activities and depressive symptoms in a sample of college freshmen. *Cyberpsychology & Behavior*, 6(2), 133-142.
- Morrow, E., Ross, F., Grocott, P., & Bennett, J. (2010). A model and measure for quality service user involvement in health research. *International Journal of Consumer Studies*, 34(5), 532-539.
- Morry, M. M. (2005). Allocentrism and Friendship Satisfaction: The Mediating Roles of Disclosure and Closeness. *Canadian Journal of Behavioural Science*, 37(3), 211-222.
- Nguyen, M., Bin, Y. S., & Campbell, A. (2012). Comparing online and offline self-disclosure: A systematic review. *Cyberpsychology, Behavior, and Social Networking*, 15(2), 103-111.
- Park, N., Kee, K. F., & Valenzuela, S. (2009). Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *Cyberpsychology and Behavior*, 12(6), 729-733.
- Parkhurst J.T., Hopmeyer A.(1999). Developmental change in the source of loneliness in child-hood and adolescence: Constructing a theoretical model. In: Rotenberg K.J., Hymel S. (Eds.), *Loneliness in childhood and adolescence* (pp.56-79). New York: Cambridge University Press.
- Qian, H., & Scott, C. R. (2007). Anonymity and self-disclosure on weblogs. *Journal of Computer Mediated Communication*, 12(4), 1428-1451.
- Rafaeli, S., Hayat, T., & Ariel, Y. (2009). Knowledge building and motivations in Wikipedia: Participation as "Ba". In Ricardo, F. J. (Ed.), *Cyberculture and new media* (pp. 52-69). New York, NY: Rodopi.
- Ross, C., Orr, E. S., Sisic, M., Arseneault, J. M., Simmering, M. G., & Orr, R. R. (2009). Personality and motivations associated with Facebook use. *Computers in Human Behavior*, 25(2), 578-586.
- Rubin, A. M. (2002). The uses-and-gratifications perspective of media effects. In J. Bryant & D. Zillmann (Eds.), *Media effects: Advances in theory and research* (2nd ed., pp. 525-548). Mahwah, NJ: Lawrence Erlbaum
- Ryan, R.M., Rigby, C.S., Przybylski, A. (2006). Motivation pull of video games: a self-determination theory approach. *Motivation and Emotion*, 30(4), 347-365
- Ryan, R. M., & Edward L. D. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 55(1), 68-78.
- Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychological Bulletin*, 92(3), 641.
- Schmidt, L. A., & Fox, N. A. (1995). Individual differences in young adults' shyness and sociability: Personality and health correlates. *Personality and Individual Differences*, 19(4), 455-462.
- Schouten, A. P., Valkenburg, P. M., & Peter, J. (2007). Precursors and underlying processes of adolescents' online self-disclosure: Developing and testing an "Internet-attribute-perception" model. *Media Psychology*, 10(2), 292-315.
- Shaw, L. H., & Gant, L. M. (2002). In defense of the Internet: The relationship between Internet communication and depression, loneliness, self-esteem, and perceived social support. *Cyberpsychology & Behavior*, 5(2), 157-171.
- Shen, C. X., Liu, R. D., & Wang, D. (2013). Why are children attracted to the Internet? The role of need satisfaction perceived online and perceived in daily real life. *Computers in Human Behavior*, 29(1), 185-192.
- Sheth, Jagdish N., Sisodia, R S., & Sarma, A. (2000). The Antecedents and Consequences of Customer-Centric

- Marketing. *Journal of the Academy of Marketing Science*, 28 (1), 155-166.
- Skues, J. L., Williams, B., & Wise, L. (2012). The effects of personality traits, self-esteem, loneliness, and narcissism on Facebook use among university students. *Computers in Human Behavior*, 28(6), 2414-2419.
- Stafford, T. F., Stafford, M. R., & Schkade, L. L. (2004). Determining uses and gratifications for the Internet. *Decision Sciences*, 35(2), 259-288.
- Staples, D. S. (2009). Web 2.0 social networking sites. In M. D. Lytras & P. O. de Pablos (Eds.), *Social web evolution: Integrating semantic applications and Web 2.0 technologies* (pp. 57-75). London: IGI Global.
- Stone, B. (2009). *Breakfast can wait: The day's first stop is online*. New York Times.
- Subrahmanyam, K., & Lin, G. (2007). Adolescents on the net: Internet use and well-being. *Adolescence*, 42, 659-677.
- Tokunaga, R. S., & Rains, S. A. (2010). An evaluation of two characterizations of the relationships between problematic Internet use, time spent using the Internet, and psychosocial problems. *Human Communication Research*, 36(4), 512-545
- Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of Marketing*, 73(5), 90-102.
- Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin*, 119(3), 488-531.
- Valenzuela, S., Park, N., & Kee, K. F. (2009). Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. *Journal of Computer Mediated Communication*, 14(4), 875-901.
- Valkenburg, P. M., & Peter, J. (2007). Online communication and adolescent well-being: Testing the stimulation versus the displacement hypothesis. *Journal of Computer Mediated Communication*, 12(4), 1169-1182.
- Valkenburg, P. M., & Peter, J. (2007). Preadolescents' and adolescents' online communication and their closeness to friends. *Developmental psychology*, 43(2), 267-277.
- Valkenburg, P. M., & Peter, J. (2009). Social consequences of the Internet for adolescents: A decade of research. *Current Directions in Psychological Science*, 18, 1-5.
- Valkenburg, P. M., & Peter, J. (2009). The effects of instant messaging on the quality of adolescents' existing friendships: A longitudinal study. *Journal of Communication*, 59(1), 79-97.
- Walther, J. B. (1996). Computer-mediated communication impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23(1), 3-43.
- Walther, J. B., & Boyd, S. (2002). Attraction to computer-mediated social support. In C. A. Lin & D. Atkin (Eds.), *Communication technology and society: Audience adoption and uses* (pp. 153-188). Cresskill, NJ: Hampton Press.
- Wang, J., & Wang, H. (2011). The predictive effects of online communication on well-being among Chinese adolescents. *Psychology*, 2(4), 359-362.
- Weiss, R. S. (1974). The provisions of social relationships. In Z. Rubin (Ed.), *Doing unto others*. Englewood Cliffs, NJ: Prentice-Hall.
- Wellman, B., & Haythornthwaite, C. (Eds.). (2002). *The Internet in everyday life*. Oxford: Blackwell Publishers
- Williams, K. (2007). Ostracism. *Annual Review of Psychology*, 58, 425-452.
- Williams, K. D. (2009). Ostracism: Effects of being excluded and ignored. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 275-314). New York: Academic Press.
- Yee, N., 2007. Motivations of play in online games. *Cyberpsychology and Behavior*, 9(6), 772-775.

Manoj Das is a participant of Fellow program in Management (FPM) in Marketing Area at IIM Indore. He has done his M.Sc (Organic Chemistry) and MBA from RTMNU. He has 3 years of teaching experience prior to joining the FPM course of IIM Indore. His areas of interest in research are social media and its impact on business and individuals, personal selling, Service marketing.

Government Process Re-engineering of an E-governance Implementation for Motor Vehicle Registration in India

Suresh Subramoniam and Dev Twinky

Abstract

The study analyzes the performance of an e-governance implementation for motor vehicle registration in a state in India. Registration of new vehicles is a process which calls for detailed Government Process Reengineering study as the frequency of this process is very high, to the tune of several lakhs per year, and is increasing day by day. The objective of implementation of the e-governance project is to register a new vehicle in less than five days. However, data collection revealed that there are cases where the process took even more than twenty days, that too after the implementation of the electronic method for the same. Seven hypotheses have been stated and tested to reach a conclusion about the cases studied - whether the average time taken for registration crosses the target of five days, and also to test whether a difference exists between groups such as, transport and non-transport vehicles or between dealers. These are set as the objectives of this Government to Citizen (G2C) services. The study also comes up with a reengineered process, after eliminating non-value adding approvals and steps to minimize the delay in the process, thereby improving the overall productivity.

Keywords: government process reengineering, vehicle registration, time, hypothesis, G2C services

1. Introduction

It is stated in literature that citizen adoption of e-government systems is sluggish, particularly in developing countries (Rana & Dwivedi, 2015). An analysis on why e-government projects are prone to design-reality gap is available in literature (Guha & Chakrabarti, 2014). E-government information flow between government, intermediaries and users in varied ways impacts the effectiveness of e-government policies (Taylor et al., 2014). Another study (Alawneh, Al-Refai & Batiha, 2013) relates to the determinants of e-satisfaction with e-government. Analysis of stakeholder

expectations helps to develop e-services that offer external services and improve internal efficiency (Axelsson, Melin & Lindgren, 2013). Another research (Fogli, 2013) proposes a novel approach to the development of e-government applications for citizens and public administration employees. Also literature reveals a study on the usability of e-governance software (Kumar & Subramoniam, 2013). In developing countries with limited resources, it is vital to judiciously set e-government strategies and direct investment, giving due consideration to the risks involved (Abdallah & Fan, 2012). Another research aims to discover the quality priorities of e-government users and analyses the attitude of Greek citizens towards e-government sites (Papadomichelaki & Mentzas, 2011). Understanding the key determinants of e-government services is an important issue for enhancing the degree of the usage of services (Sharma, 2015). Nograšek & Vintar (2015) have worked to develop a more comprehensive framework that would provide better insight into the characteristics of organisational transformation of public sector organisations in the e-government era. The purpose of another research found in the literature is to assess the maturity level of the Jordanian e-government program from citizens' perspective (Anas, Hussein and Saheer, 2014). Alomari, Sandhu & Woods (2014) have explored how citizens socialise and network while using and adopting e-government. The importance of evaluation and optimization of e-government services is imperative if the government organisations consider to have an effective impact on the success and take-up, or proper buy-in of the services offered (Lee, Sivarajah, Molnar, Weerakkody & Irani, 2015). Another study focuses on a comprehensive review of the literature related to e-government satisfaction and adoption, with particular focus on the most critical factors and the manifested variables that influence user satisfaction in e-government (Weerakkody, Irani, Lee, Hindi & Osman, 2014). The primary intention of the literature review is

to establish the gap in literature as stated in the motivation section that follows.

2. Motivation for the Present Study

Literature review section has discussed the presence of a gap in the literature with respect to reengineering of existing government process flow of an already implemented e-government system, led to the establishment of the stated gap that is studied less or totally absent in the research articles, especially in the case of e-government implementations in India.

Business Process Reengineering (BPR) is defined as the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical, contemporary measures of performance, such as cost, quality, service and speed. BPR applied IT is the new Industrial Engineering (Davenport & Short, 1990; Davenport & Stoddard, 1994). BPR involves a thorough analysis of the current business processes that are redesigned to improve performance (Al-Mashari & Zairi, 2000; Davenport, 1993). Government Process Re-engineering (GPR) has evolved from applying Business Process Re-engineering (BPR) concepts to Government Services. Most organizations, both public and private, have traditional bureaucratic procedures that hinder performance and ultimately, the productivity level (Rainey, Backoff & Levine, 1976). Antiquated processes, historical Acts and Rules, and status quo continue, even with the introduction of technology to facilitate improvement of service delivery. Costly and time-consuming business processes cause inefficiency and ineffectiveness. So there is a desperate need to thoroughly analyze and reengineer the old-fashioned and obsolete business processes to improve performance (Davenport & Beers, 1995). Any e-project, at the time of implementation, will not be based on the most optimal streamlined process for various reasons such as, gaps in communication between the user and developer resulting in project failure (Bashein, Markus & Riley, 1994; Cao, Clarke & Lehaney, 2001). Firstly, jobs set aside for each section in an office or for an officer have stabilized over time due to the process of conflict over ownership of the sub-processes involved and their

resolution. Secondly, pressure on the developer for speedy implementation may result in launching the software product before it is fully ready in all aspects. Third is the time and finance constraint involved in every project. Ineffective change management, lack of awareness for training and fear among the employees about downsizing, are other contributing factors that act as barriers for communication between the client and the software supplier, preventing full understanding of the process before development.

3. Problem Statement

There is scope for dramatic improvement in performance by switching from the "as-is" process flow to the "to-be" process flow in every e-project implemented in the country. Though technology is a key enabler for Business Process Reengineering in areas other than software, like RFID or bar code for engine or chassis identification, this study focuses mainly on dramatic improvement by taking a re-look at the software process flow alone. The study is focused on the motor vehicle registration process to see whether the implemented e-project met the objective of delivering the Registration Certificate before the fifth day from application submission. Also, it seeks to ascertain whether the initially set objective of five days could be further reduced to target a lower time period through the implementation of the reengineered process.

Figure 1 shows eleven steps of the "as-is" process starting from the customer who is a citizen initiating the process and ending in the final dispatch of the Registration Certificate by post. As a first step, citizen approaches the dealer for registration, remits payments and signs papers. Then the dealer submits the details to the website. This is followed by the dealer visiting the RTO for tax token and for generation of number for remitting fees. The dealer takes the vehicle to the ground for physical verification of chassis and engine numbers by RTO, along with the originals of the submitted documents. Verification is carried out in the system and registration number is allotted for the vehicle. After this allocation, the clerk retrieves details from the implemented system and verifies. The Superintendent repeats this step. Issue of Registration Certificate is the next step. Registration

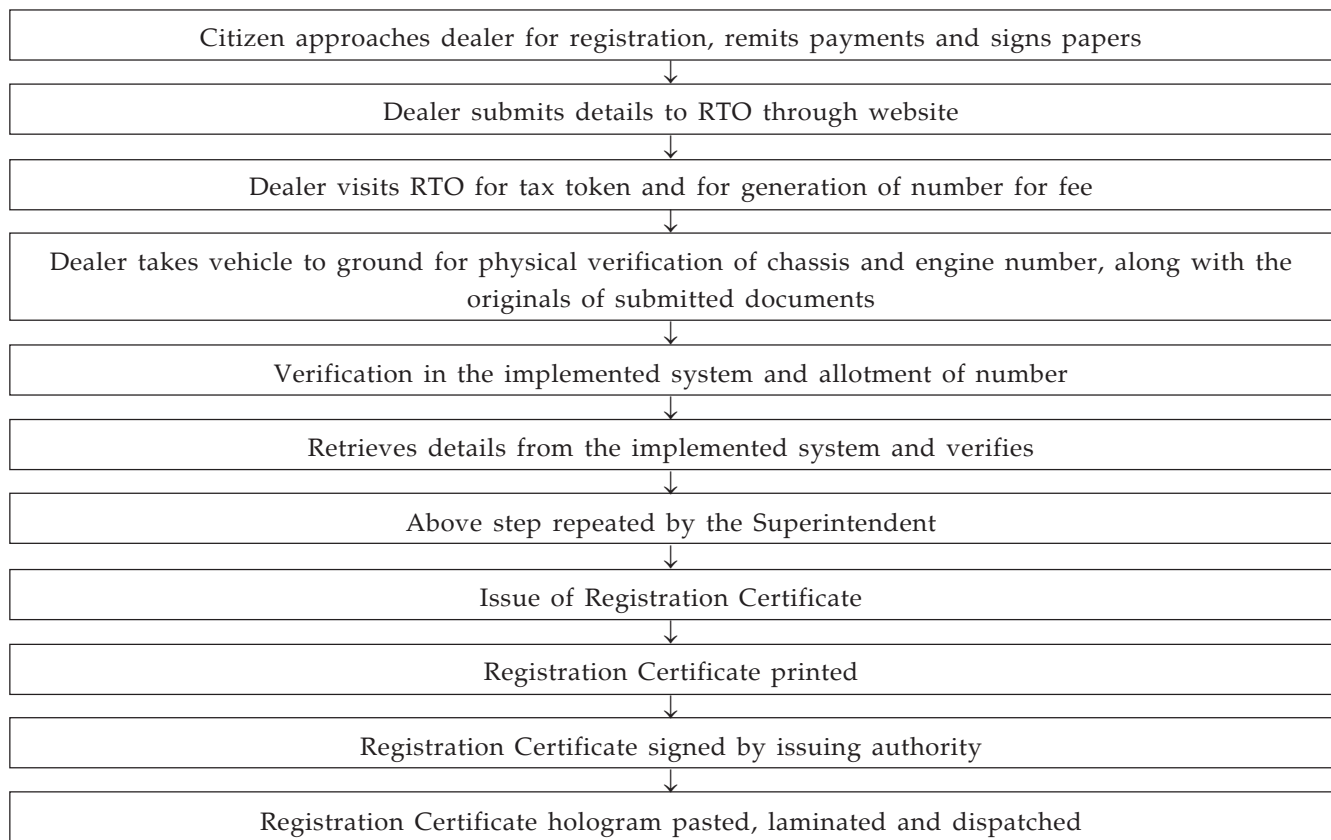


Figure 1: Schematic Representation of the "as-is" Process in Focus

Certificate is printed on a card that is later laminated. Then physical signature is affixed on the Registration Certificate by the concerned authority prior to hologram fixing, lamination and dispatch. This is the 11-step process which is the focus of this study.

4. Research Methodology

The population consists of all new vehicles registered in the state where registration is growing at the rate of ten lakh per year through nearly twenty or more Regional Transport Offices (RTO) assisted by sub RTOs numbering a little more than two and half times the number of RTOs. Three RTOs were randomly picked by Simple Random Sampling from districts where the registration rate of new vehicles is high. Twenty to thirty samples each of transport and non-transport vehicles were randomly picked from the website of the State Motor Vehicle Department for data relating to date of submission of application and date of delivery

of the registration certificate. Further, the time taken for each of the sub-processes involved as listed in the Table 1 were collected from corresponding files available in the office. Suitable hypotheses are coined to come to a conclusion about the time taken for the process of registration as discussed in the following sections. The study was done during the period January 2013 to April 2014.

The hypotheses proposed for the study are as follows:

H1: The time delay between registration on the website and data entered status of transport vehicle data is less than or equal to one day.

H2: The time delay between data entered status and verified status of transport vehicle is less than or equal to one day.

H3: The time delay between verified status and RC issued status is less than or equal to one day.

H4: The time delay between issued status and printed status of RC is less than or equal to one day.

H5: The time delay between fresh application for registration and dispatched status of Registration Certificate is less than or equal to five days.

H6: There is no difference between average time taken for registration of transport and non-transport vehicles.

H7: There is no difference between average times taken for registration of vehicles from two dealers.

5. Data Collection

The descriptive statistics of the data collected from the office files for twenty registered vehicles randomly

picked-up as sample during period of study are presented here (Table 1).

6. Results

The results of t-tests carried out to test at 5% significance level on whether any of the sub processes crossed the time limit of one day or the whole process exceeded the target objective of five days is shown in Table 2.

For the above set of hypothesis, 't' calculated value t_{calc} is much above 't' tabulated value t_{tab} of 1.79 for both transport vehicles as well as non-transport vehicles for the corresponding one-tailed tests. Hence, hypothesis H1 that the activity is taking less than or equal to one

Table 1: Descriptive Statistics of the Samples for Transport and Non-Transport Vehicles

Average Time elapsed between the following in days	Transport Vehicle (N=20)		Non-Transport Vehicle (N=20)	
	Mean	SD	Mean	SD
Submission and Entered	4.30	4.68	12.25	15.06
Entered and Verification	0.35	0.67	1.20	2.14
Verification and Issue	0.75	1.12	2.30	2.25
Issuing and Printing	0.80	1.36	0.90	0.85
Submission and Dispatch	14.3	11.93	31.65	19.41

Table 2: Summary of the 't' Test Results Carried Out on Sub Processes Time Limit and the Whole Process

H. No.	Hypothesis for the 't' test carried out at 95% significance level for which the tabulated t value is 1.729 for 19 degrees of freedom	Transport Vehicle (N=20)		Non-Transport Vehicle (N=20)	
		t_{calc}	Test Result	t_{calc}	Test Result
H1	Time taken between application submission and data entered in the system \leq 1 day	3.15	Reject	3.34	Reject
H2	Time taken between data entered in the system and verification done \leq 1 day	4.33	Reject	1.58	Accept
H3	Time taken between verification done and issue registration certificate \leq 1 day	1.00	Accept	2.58	Reject
H4	Time taken between Issuing of registration certificate and printing \leq 1 day	0.66	Accept	0.53	Accept
H5	Time taken between submission of fresh application and dispatch of registration certificate \leq 5 days	6.14	Reject	4.99	Reject

Table 3 : Group Statistics on Time taken for Registration of Transport and Non-transport Vehicles

Time taken for registration in days	Vehicle Type	N	Mean	Std. Deviation	Std. Error Mean
	Transport	29	7.21	4.894	.909
	Non-transport	30	15.37	12.861	2.348

Table 4 : Independent Samples Test on Time Taken for Registration of Transport and Non-transport Vehicles

Time taken for registration in days		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
										Lower	Upper
	Equal variances assumed	4.406	.040	-3.199	57	.002	-8.160	2.551	-13.267		
	Equal variances not assumed			-3.241	37.47	.003	-8.160	2.518	-13.259		

Table 5 : Group Statistics on Time taken for Registration of Vehicles from Two Dealers

Time taken for registration in days	Dealer	N	Mean	Std. Deviation	Std. Error Mean
	Dealer1	20	33.35	14.240	3.184
	Dealer2	20	28.00	12.657	2.830

Table 6 : Independent Samples Test on Time taken for Registration of Vehicles from Two Dealers

Time taken for registration in days		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
										Lower	Upper
	Equal variances assumed	1.579	.217	1.256	38	.217	5.350	4.260	-3.274 13.974		
	Equal variances not assumed			1.256	37.485	.217	5.350	4.260	-3.278 13.978		

day is rejected in both the cases and it is concluded that the same is more than one day in the case of both transport and non-transport vehicles. A similar set of hypothesis have been developed. To test the hypothesis for serial numbers 2 to 5 and 't' test results on whether to accept the corresponding hypothesis or not have been listed for both transport vehicle and non-transport vehicles respectively.

It can be noted that only four of the listed sub-processes are statistically within the time limit given; rest of them exceeded the time limit. Moreover, hypothesis test on the whole process of new vehicle registration also revealed that the time limit set as 5 days has been exceeded in the case of both transport vehicles and non-transport vehicles. In order to reduce the steps in the process, a new "to-be" process is arrived at by holding internal discussions to arrive at a more simplified process, eliminating duplicated approvals and non-value adding steps as discussed in later sections of this research.

The group statistics (Table 3) and the result of the independent sample 't' test to see whether there is any difference between time taken for registration of transport vehicle and that of non-transport vehicle are presented in Table 4 for testing hypothesis H6.

Based on above results, it is inferred that hypothesis H6 which states that there is no difference between average time taken for registration of transport vehicle

and non-transport vehicle can be rejected. This is statistically significant and hypothesis H6 is rejected at 5% significance level as 0.04 is less than 0.05 or 5%. Twenty nine transport vehicle samples and thirty non-transport vehicle samples are taken randomly for the above hypothesis test.

The group statistics (Table 5) and independent sample 't' test to find out whether there exists difference between registration time taken for new vehicle registration by two different dealers are as given in Table 6. Twenty samples are taken for each dealer's case and independent 't' test is conducted to find out whether significant difference exists in the average time taken for registration between the two dealers.

A sample size of twenty from each dealer is taken for testing this hypothesis. The results showed that the hypothesis H7 could not be rejected at 5% significance level as 0.217 is much above 0.05 or 5%. Therefore, it is concluded that there is no difference in the average time taken for registration of new vehicles from two different dealers.

Figure 2 shows the "to-be" process after elimination of steps which are found no value adding so that the process can be completed within the stipulated time of less than or equal to one day instead of the previously set target of five days. The main changes which can be noted in the process flow prior to re-engineering as

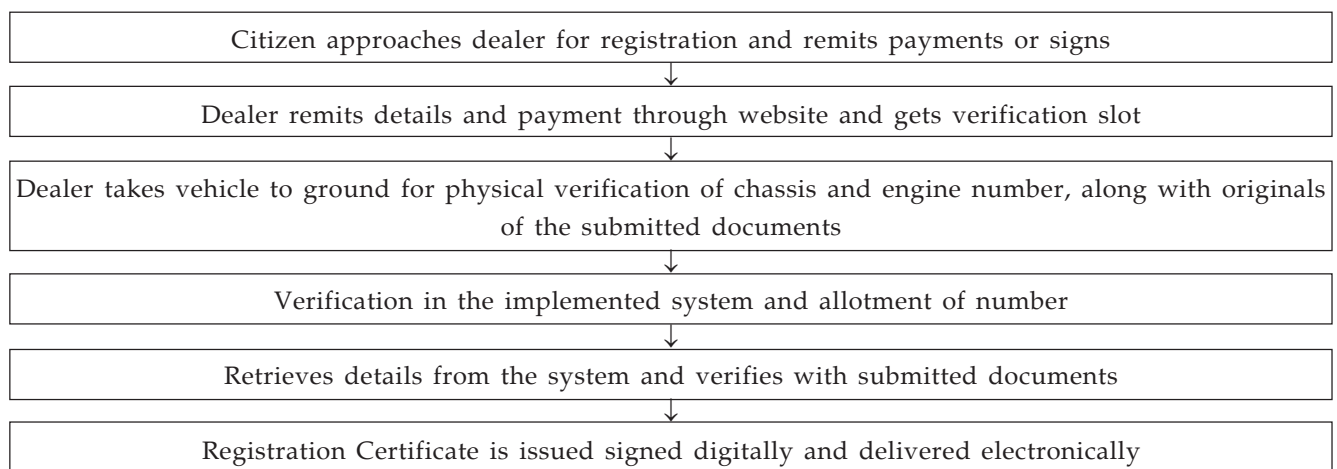


Figure 2: Schematic Representation of the "to-be" Process

in Figure 1 and re-engineered process flow as in Figure 2 are as follows: Dealer visit to RTO office for physical payment and manual receipt of registration number is avoided by online payment and electronic receipt of registration number in the re-engineered process. The verification step which is repeated by the Superintendent is eliminated in the re-engineered process. Electronic generation and delivery of registration certificate is suggested in the place of several steps which follows the step involving issue of certificate. This leads to reducing an eleven step process to a mere six step process, saving time and effort of many, improving productivity at the same time, while improving the lead time in issue of registration certificate for new vehicles.

7. Conclusion and Managerial Implications

Most-government software are designed and implemented in a project mode that results in the accidental inclusion of non-value-added steps in the process flow. Such inclusions can also be due to the organizational power play between interacting members of the system in the allocation of work in the newly introduced e-mode of functioning. Only a revisit at a later point of time with an idea of re-engineering can help to evolve a crispy set of value added tasks in the process flow, ending up with saving of resources. Though an effort is made to achieve an end-to-end process flow in every e-government project, there is a tendency for employees to fall back on the earlier manual method at least partially in duplicating the functionality already available in the software, leading to delays in the process. Revisits on the process flow at planned intervals and up-to-date application of emerging technologies widen the scope for re-engineering possibilities in every e-government project in the post implementation phase.

Seven relevant hypotheses were developed to study and test whether the steps in the process exceeded time limits set by the citizen's charter and to see whether significant difference existed between groups in the average time taken for registration of new vehicles, between transport vehicle and non-transport vehicles or between dealers. Out of four sub-processes for which hypotheses are stated and tested, two hypotheses each, for transport vehicles and non-transport vehicles,

concluded that they exceeded the set time limit of one day. The hypothesis test on whether the registration process as a whole exceeded the agreed limit of five days revealed that both transport and non-transport vehicles exceeded the set time limit of five days. Hypothesis H6 which states that there is no difference between the average time taken for registration of transport vehicles and non-transport vehicles is rejected. Further, it is concluded that there is no difference in the average time taken for registration of new vehicles from two different dealers. Every step in the existing process was re-visited to decide whether to retain it in the process flow or not. Redundant steps, if any, are removed and others are collapsed or streamlined for process simplification. In the "as-is" process eleven steps are involved, whereas the re-engineered process has only six steps. Also a reduction of nearly 45% of steps is achieved using the new process. The time limit, according to agreement with citizen's charter, for registration using the present process flow is five days. But in reality it is even more than 20 days in certain cases. It can be seen that using the re-engineered process, registration can be done in one day by the proposed process flow alone, leaving aside reduction in time which can be achieved by envisaging technologies like RFID for engine or chassis number swipe or verifications that are carried out manually at present.

References

- Abdallah, S., & Fan, I. (2012). Framework for e-government assessment in developing countries: case study from Sudan. *Electronic Government, An International Journal*, 9(2), 158-177.
- Alawneh, A., Al-Refai, H., & Batiha, K. (2013). Measuring user satisfaction from e-Government services: Lessons from Jordan. *Government Information Quarterly*, 30(3), 277-288.
- Al-Mashari, M., & Zairi, M. (2000). Revisiting BPR: a holistic review of practice and development. *Business Process Mgmt Journal*, 6(1), 10-42.
- Alomari, M., Sandhu, K., & Woods, P. (2014). Exploring citizen perceptions of barriers to e-government adoption in a developing country. *Transforming Government*, 8(1), 131-150.
- Anas R. Al-Soud, Hussein Al-Yaseen & Saheer H. Al-Jaghoub (2014). *Jordan's e-government at the crossroads. Transforming Government: People, Process and Policy*, 8 (4), 597 - 619.

- Axelsson, K., Melin, U., & Lindgren, I. (2013). Public e-services for agency efficiency and citizen benefit-Findings from a stakeholder centered analysis. *Government Information Quarterly*, 30(1), 10-22.
- Bashein, B., Markus, M., & Riley, P. (1994). Preconditions for BPR success And How to Prevent Failures. *Information Systems Management*, 11(2), 7-13.
- Cao, G., Clarke, S., & Lehaney, B. (2001). A critique of BPR from a holistic perspective. *Business Process Mgmt Journal*, 7(4), 332-339.
- Davenport, T., & Short, J. (1990). The new industrial engineering: Information technology and business process redesign. *Sloan Management Review*, summer, 11-27.
- Davenport, T. (1993). *Process Innovation*, Harvard Business School Press, Boston, MA.
- Davenport, T., & Stoddard, D. (1994). Reengineering: Business change of mythic proportions? *MIS Quarterly*, 18(2) 121-127.
- Davenport, T., & Beers, M. (1995). Managing information about processes. *Journal of Management Information Systems*, 12 (1), 57-80.
- Fogli, D. (2013). Towards a new work practice in the development of e-government applications. *Electronic Government, An International Journal*, 10(3/4), 238-258.
- Guha, J., & Chakrabarti, B. (2014). Making e-government work: Adopting the network approach. *Government Information Quarterly*, 31(2), 327-336.
- Kumar, K., & Subramoniam, S. (2013). Usability analysis of an Indian e-governance software. *Electronic Government, An International Journal*, 10(2), 211-221.
- Lee, H., Sivarajah, U., Molnar, A., Weerakkody, V., & Irani, Z. (2015). A User Satisfaction Study of London's Congestion Charge e-Service: A citizen perspective. *International Journal Of Electronic Government Research*, 11(2), 35-50.
- NograÅ;jek, J., & Vintar, M. (2015). Observing organisational transformation of the public sector in the e-government era. *Transforming Government*, 9(1), 52-84.
- Papadomichelaki, X., & Mentzas, G. (2011). Analysing e-government service quality in Greece. *Electronic Government, An International Journal*, 8(4), 290-308.
- Rainey, H., Backoff, R., & Levine, C. (1976). Comparing Public and Private Organizations. *Public Administration Review*, 36(2), 233-244.
- Rana, N., & Dwivedi, Y. (2015). Citizen's adoption of an e-government system: Validating extended social cognitive theory (SCT). *Government Information Quarterly*, 32(2), 172-181.
- Sharma, S. (2015). Adoption of e-government services: The role of service quality dimensions and demographic variables. *Transforming Government*, 9(2), 207-222.
- Taylor, N., Jaeger, P., Gorham, U., Bertot, J., Lincoln, R., & Larson, E. (2014). The circular continuum of agencies, public libraries, and users: A model of e-government in practice. *Government Information Quarterly*, 31, S18-S25.
- Weerakkody, V., Irani, Z., Lee, H., Hindi, N., & Osman, I. (2014). A review of the factors affecting user satisfaction in electronic government services. *International Journal Of Electronic Government Research*, 10(4), 21-56.

Suresh Subramoniam is an Associate Professor at the CET School of Management, College of Engineering, Trivandrum. He holds Masters degree in Industrial Engineering from Louisiana State University in USA and PhD in Management from Kerala University. He is Fellow of the Indian Institution of Industrial Engineering.

Dev Twinky is a full time MBA student at CET School of Management, College of Engineering, Trivandrum. She has special interest in Business Process Re-engineering of e-government systems.

Internet of Things (IoT) and Smart Technologies: Framework of Temporal Data Mining Concerning Smart Meter

Shashi Kant Srivastava

Abstract

This paper attempts to propose a framework to study the adoption pattern of smart electrical technologies based on the existing understanding of technology adoption. The main thrust of this research is to develop a framework that may become practical right away, on availability of data from these technologies. In the absence of existing data, this paper generates the data with the help of computer simulation, and shows how wealth and technology diffusion theories can be explored for developing the proposed technology adoption framework. It is argued that results based on the proposed framework are able to identify the trends and association in the adoption of these technologies within society.

Keywords: IoT, technology adoption, temporal data mining, smart meter, smart grid, smart home, smart devices

1. Introduction

This paper proposes a framework to mine probable and feasible business solutions from the enormous data generated from Internet of Things (IoT) devices. Smart meters are one such device which may be categorised as initial examples of IoT's that may soon become an integral part of households (Depuru, Wang, Devabhaktuni, & Gudi, 2011; Greveler, Glosekotter, Justus, & Loehr, 2012). Smart meters are capable of transferring electrical consumption metering details of a household on real time basis to electricity provider. Furthermore, this meter can also transfer equipment wise recording of the consumption pattern, in case the house is equipped with smart devices. Due to this capability of the smart meter, adoption time of various smart electrical devices installed in a residence can be

recorded very easily (Nezhad, Wijaya, Vasirani, & Aberer, 2014). This recording, if mined following data mining techniques, may provide various useful adoption patterns that remain unknown otherwise. Current electric meter does not offer this facility of observing and recording electricity consumption of different electrical devices. The importance of this research becomes evident since electric meters are part of every household numbering thousands of millions in number. Moreover, every house is equipped with several electrical devices.

The proposed topic becomes more important due to three reasons. First, IoT's are the future of the world. Second, IoT's will generate huge amount of data, the use of which is not thought of at present. Third, there is no prior research work on a similar topic. Examples of IoT's i.e., smart meter and smart devices, are to become common in the event of the present electric meter being replaced for effectiveness and efficiency. The data generated by the combination (smart meter and smart devices) will be enormous in size. This data will have record of all the operations of all the devices on 24*7*365 basis. The size of the data will be huge; conventional analytical techniques cannot handle it. Therefore, this data offers a unique opportunity to data scientists to explore techniques that may be of help to the various stakeholders involved. An example of such trends is, mining the adoption pattern of various smart electrical devices within the household. Mining these patterns will not only be of help to equipment manufacturers to understand their customers better, but also to policy makers to frame suitable policies in alignment with their long term socio-environmental objectives.

2. Literature Review

Literature review section is organized in three sub sections. First section explains the overview of data mining literature relevant to our research. Second explains the apparent research gap in the reviewed literature of data mining and the third generates research questions to be addressed based on these gaps. Data mining literature specifically focused to time trend is of interest for our research. The first mention of such an idea was observed in 1998 in a paper titled "The intelligent interface for online electronic medical records using temporal data mining" by Spenceley & Warren(1998). Since then two prominent fields have been extensively using temporal data mining; one, medical (Bellazzi, Larizza, Magni, & Bellazzi, 2005; Eriksson, Werge, Jensen, & Brunak, 2014; Svanström, Callrus, & Hviid, 2010) and another, geographic information system and urban planning (An, Zhang, Zhang, & Wang, 2014; Cheng & Wang, 2008; Kitamoto, 2002).

Besides medical and urban planning fields, research of temporal data mining is also emerging in the context of smart environment, specifically for smart homes. One research following this technique of data mining in the context of smart homes is by Cook & Jakkula, (2008). This paper uses the technique of temporal data mining for the detection of anomalies in smart homes. These anomalies are events happening within continuously observed environment that are other than day to day routine, may be accident, fire etc. Possibility of detection of such anomaly becomes possible because smart homes are residences equipped with smart devices capable of sensing the environment. Furthermore, these smart devices, in the presence of other smart technology, sequentially connect to smart meter and smart grid. So smart devices, smart home, smart meter and smart grid form interlinked chains. Combining these separate segments of the chain together results in smart environment. Out of these four sub units of smart environment, smart grid is the subject relatively more studied in research. There are quite a few papers which study application of temporal data mining in the context of smart grid (Fan, Chen, Kalogridis, Tan, & Kaleshi, 2012; Prasad & Avinash, 2013; Samantaray, 2015).

Literature review of temporal data mining clearly demonstrates that, in comparison to the field of data mining, this area is very new. Furthermore, there are relatively limited efforts by researchers to explore the field of temporal data mining. Therefore, there are very few papers pertaining to this research subject. Out of a limited number of 114 papers used for literature review, majority of efforts are presented as conference papers; publication as journal articles is limited. As a result, there are very few papers on the application of temporal data mining in smart homes (i.e. Cook & Jakkula, 2008; Jakkula & Cook, 2011). Very limited number of papers (e.g. Chou, Hsu, & Lin, 2014; Zhang, Grijalva, & Reno, 2014) and the anticipated extensive application of smart meter in future, suggest a gap in existing research efforts. Hence investigation and future research in this field is desirable. Considering the evident gap in literature, this paper aims to bridge this to an extent. The present effort of bridging the gap may be of concern to certain stakeholders of smart meters. Users, manufactures, electricity providers and policy makers within the government are few important stakeholders. Out of these, manufactures of smart devices would be the first to be concerned. Manufactures would be interested to know the adoption pattern prevailing in society. The prediction of this adoption pattern will provide a fair idea to manufactures to appropriately plan their business strategy. There are innumerable electrical devices in households at present and more or less, all are set to become smart in due course of time. A study of adoption patterns in the early stage will give added advantage to manufactures to invest on innovation and design of the appropriate product and plan its marketing accordingly. Considering these implications, this article aims to design a framework that may suggest answers of the following questions, related to manufacturers of smart devices.

1. What are the characteristics of households that adopt a particular smart device at particular time reference? User's demographic and social characteristics can be concluded, based on his electricity consumption detail (Newing, Anderson, Bahaj, & James, 2015). These characteristics are the number of occupants, presence of children, type and size of dwelling, and socio-

economic and geo-demographic characteristics of the household (Newing, Anderson, Bahaj, & James, 2015). For the purpose of simplification, this paper does not simulate the variation in power consumption of the single household. Rather only the total amount of yearly consumption is modelled. Therefore, it is not possible to mine all the possible household characteristics with the present data. But the economic aspect, significantly associated with the amount of electricity consumption, is considered as the variable for consumer segmentation. By integrating electricity consumption of the single household at the time of device adoption will help to know the present segment of customers that adopt devices. Furthermore, it gives an indication of the next appropriate consumer segment based on which customized business plan for that section can be prepared.

2. What is the sequential adoption pattern between two particular smart devices? With this analysis manufacturers may prepare for the production and marketing of the next product analyzed to be adopted by the customer segment.
3. What is the average time lag between adoptions of two particular smart devices? Knowing the time of adoption by society will give opportunity to firms to strategically plan their future products.

Remaining part of the paper is organised as follows. Next section, Section 3, describes smart meter and related technologies. Section 4 elaborates the concept of temporal data mining and its evolution. Section 5 defines the nature of temporal data availability, and Section 6 defines the problem and explains the tentative solution which temporal data mining may offer. The final section, Section 7, deals with conclusions of the paper and suggests future work to be done in this regard.

3. Smart Meter and Related Smart Technologies

In the context of electronics, smart is defined as any method or device having both sensing and control capabilities (Baz, 1996). In accordance with this definition all the smart devices are capable of sensing the surrounding environment and are able to accordingly guide and control their functions. For example, smart bulb will be capable of sensing the surrounding

illumination level and presence of human being. Based on this information the smart bulb will be able to decide whether to light itself up or not. In the context of IT, definition of smart is slightly modified; besides sensing and controlling capabilities, communication to other devices and networks is also required. This capability defines smart devices as "personal devices interacting with users, sensing their environment, and communicating with each other" (Allwright et al., 2006). Smart meters are also an example of smart devices to be deployed in future households. As defined, smart meters too are capable of sensing the electric consumption in totality, as well as partially by smart devices of the home. Smart meters may force control on consumption of electrical devices, if programmed. Smart meters are also capable of communicating this consumption data through internet to the electricity provider on near real time basis. Therefore, smart meter reaches its full capability only in the presence of smart electrical devices and in their absence offers only limited functionality. Smart grid, as one of the associated technologies, is explained below.

3.1 Smart Grid

Smart grid is defined as "the electricity grid that uses digital technology to improve reliability, security and efficiency of the electric system, from large generation, through the delivery systems to electricity consumers and a growing number of distributed generation and storage resources" (Li, Zhou, & China 2011, p.99). Definition of smart grid makes it clear that smart grid uses computer technology to integrate all parts of electricity distribution and consumption intelligently. The principle objective of this integration is to make the system more efficient. Efficiency of the smart grid is the result of availability of all consumption records in the process of electricity distribution via smart grid. Through digital technology smart grid automatically keeps the travel and consumption records of electricity, from generation till the last point of consumption. Because of its digital nature smart grid is not only capable of keeping all the data recorded in the process of electricity distribution, but it can also communicate this data to various stakeholders and help them take proper

decisions at the appropriate time. Present electricity distribution lacks this capability and hence is susceptible to several problems. The biggest problem of present day electricity transmission is energy and power shortage, power outage and electricity theft. To an extent these problems can be resolved with the help of smart grids (Seetharam et al., 2013). Smart grid offers unique solutions, which are unimaginable in present context, to these problems. Differential pricing to resolve energy shortage at peak hours and accurate data availability of power outages and theft detection are a few of them. The fully functional smart grid system does not mean just the presence of smart grid and smart meter; smart electricity devices are other essentials required to attain the full capability offered by smart grids. If smart grid is an essential requirement beside smart meter at one end, then smart electrical devices are essential at the other end.

3.2 Smart Devices

By embedding microelectronics devices to any every-day object, it can be converted into a smart object (Mattern, 2003). These smart objects or devices can communicate by wireless means and they may form networks that may rise to a world-wide distributed system network of magnitude much larger than today's Internet (Mattern, 2003). Recent developments in sensor technology has made sensors capable of detecting various environmental phenomena, including but not limited to, light, acceleration, temperature etc. Radio sensors are another interesting development in this area; they can report their observations without any energy supply within a few meters distance. These sensors obtain the required energy from the environment or directly from the measuring process itself. The capability of remote control, security alarm and sensors are utilized while converting an ordinary home into smart home with the application of these smart devices (Kadam, Mahamuni, & Parikh, 2015). An overview of smart meter, smart grid and smart devices helps us in understanding the concept of smart home and smart environment in which smart devices, in interaction with smart meter and smart grid, produce enormous amount of data. The present world has not witnessed data of such nature and quantum, and hence the potential

capability and threats posed by this data availability is still unknown. However, certain characteristics of this data may be visualized. Based on the use of smart devices, smart meters will keep recording the use with time stamp on it. Time stamping on the data from various devices puts it in the category of temporal data. Analysis of this data comes under the specialised area of data mining, known as temporal data mining. Temporal data mining may offer certain usability to this data for pattern mining, useful to stakeholders.

4. Temporal Data Mining

The previous section discussed the integration of smart meter and smart devices to produce enormous amounts of data. It is nearly impossible to handle this data without the intervention of data mining. Data produced by smart meters are of enormous volume and high velocity (Nezhad et al., 2014). Being temporal in nature, temporal data mining techniques offer better analysis of data compared to that of traditional data mining. Temporal data mining is an extension of traditional data mining. Application of temporal data mining lies in mining the sequence of activities rather than just cross sectional states of it, and thus, they offer better inferences than that of traditional data mining (Camara, Naguingar, & Bah, 2015). These inferences may result in contextual and temporal proximity of two or more subjects of study, some of which may indicate cause and effect relationship between multiple series of data (Camara et al., 2015).

5. Nature of Data

Smart devices installed in smart homes are comprised of hardware and software, with time stamping of its use recorded in a server (Mohassel, Fung, Mohammadi, & Raahemifar, 2014). Smart meters communicate consumption data equipped with time stamping to both the user and the service provider (Mohassel et al., 2014). For the purpose of this research adoption of such devices by society is assumed to be similar to other technology adoption seen historically. We further assume that once adopted by households (first time stamp recorded in the data server), the device is used on a regular basis. Cases of abandonment after adoption are not taken into account. Present study is concerned only with the first

time record of use of any new device by the consumer. Some of the information anticipated to be obtained from smart environment and used to answer the formulated research questions are:

Id<- is consumer identity

con<- per year electricity consumption by consumer

t1<- is first time in the electric meter device number one is recorded

t2<- is first time in the electric meter device number two is recorded

tm<- is first time in the electric meter device number m is recorded

The combination of this information will generate $m+2$ tuple data. The vector form of this data will be of $\langle Id, con, t1, t2, \dots, tm \rangle$ type.

6. Problem Definition

Smart meter are still uncommon in most of the underdeveloped and developing countries (Weranga, Kumarawadu, & Chandima, 2015). But the process of smart metering has started in the developed world. In Sweden and Italy most of the households have already adopted smart meter (Wehlitz, Werner, & Franczyk, 2014). In the Indian context, in the absence of deployed smart meters, data depicting consumption pattern of smart devices is not available. Considering this difficulty the data of adoption is generated through simulation. 1000 data points are generated in the R-studio statistical software. Generation of data is in the format explained in Section 5 and is supported by literature. Individual record of this data has consumption of electricity as one item and adoption times of various smart devices as the rest. In literature it has been statistically proved that electricity consumption is significantly related to the level of income of family (Francisco, Aranha, Zambaldi, & Goldszmidt, 2006). Furthermore, in many researches income is found to follow the pareto distribution (e.g. Dagsvik, Jia, Vatne, & Zhu, 2013; Persky, 1992). Therefore, following these two premises, hypothetical data in R software is generated using pareto distribution. Pareto distribution is also known as power law

distribution. In pareto distribution the frequency of observation is inversely proportional to the amount of consumption. If amount is more, frequency or probability of its occurrence will be less and vice averse. Based on average per capita electricity consumption in India as 800KWH (Chauhan & Saini, 2015; Sharma & Balachandra, 2015), variation in consumer per capita consumption is modelled with maximum yearly consumption of electricity as 12000 KWH and 100 KWH as minimum electricity consumption. Modelling of electricity consumption is followed by modelling the time pattern of smart device adoption. Rogers' diffusion of innovation (DOI) theory explains that adoption of any innovation in society follows a normal curve in time (Sahin, 2006). Innovators are the first segment of society to adopt these products, followed by early adopter. After these segments of society have adopted the product, it is adopted by early majority and late majority segments. In the end, the population characterised as laggards adopt it (Sahin, 2006). Therefore, to simulate adoption data in time, normal distribution is used. It is assumed that smart device adoption will take approximately 10 years (3650 days). Therefore the last person of the proposed 1000 people will adopt it after nearly 10 years from the first person. It has been found that there is a high degree of correlation between economic status of the individual and adoption of innovation (Al-Ghaith, Sanzogni, & Sandhu, 2010). This fact is used in simulation of data. As a result, multivariate data is generated following two distribution patterns, pareto and normal. Since there is correlation between the amount of electricity consumption by a household and adoption of smart devices, certain degree of correlation is maintained between these two in data generation.

Based on the data simulated and following the proposed framework in Figure1, the timeline of adoption of various smart devices by an individual customer can be generated. We use these schedules of adoption for temporal association mining of different devices. For the purpose of explanation the schedule of adoption of four devices by three customers are shown below:

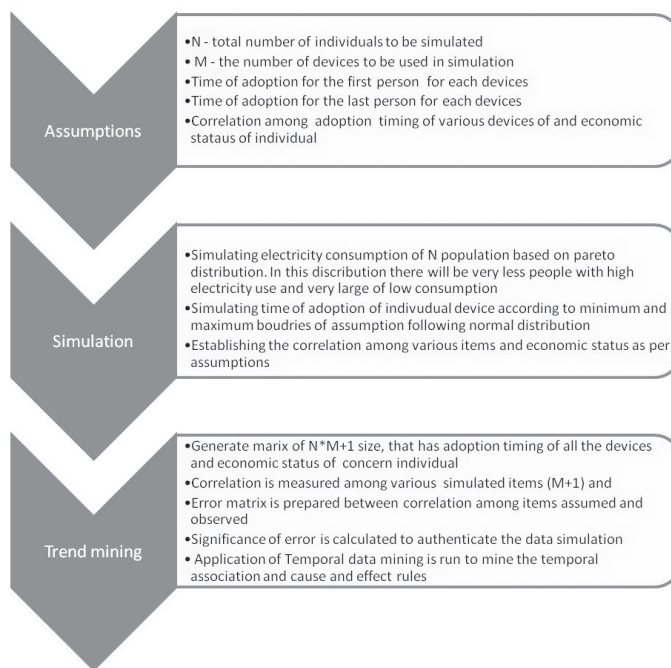


Figure 1- Framework of Adoption Trend Mining Data

Figures 3, 4 and 5 represent the time of adoption of four smart devices on the timeline by three customers. We can mine the temporal association of adoption of various devices based on this information. For example, the

association of adoption of device 2 followed by 1 is 2/3 i.e. 67%. For customers1 and 3, device2 is adopted immediately after device 1. Similarly, device 1 selected after any other device is 0%.

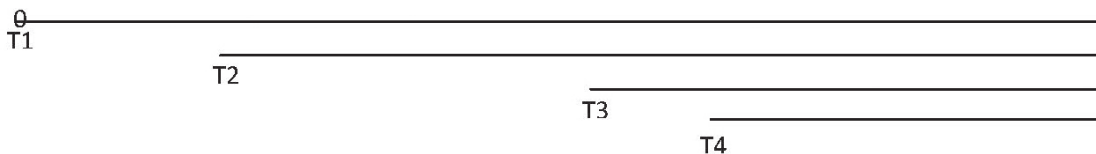


Figure 2- Adoption timing of four devices (Customer-01)

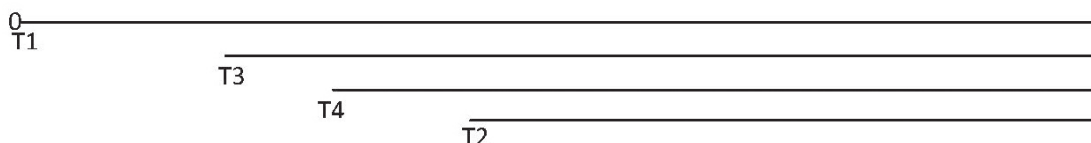


Figure 3-Adoption timing of four devices (Customer-02)

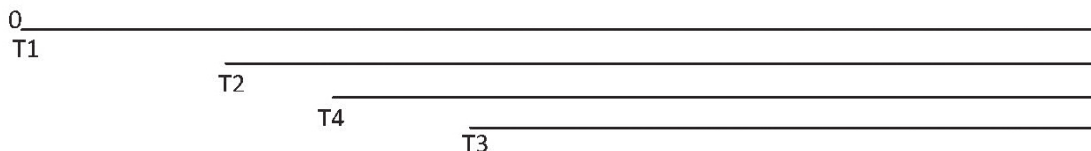


Figure 4-Adoption timing of four devices (Customer-03)

7. Conclusion and Future Work

Once smart meters are part of society huge data will be generated with the potential to mine useful insights, not only for manufacturers of smart devices, owners of electricity supply, government for policy formation, but also to the owners of the houses consuming electricity, to bring the consumption to more efficient and economical levels. The nature of data obtained from this process is big in size in comparison to the data available from say, retail stores. The number of consumers, in case of retail stores, is mostly in thousands whereas the consumers of smart meter will be millions in size. The complexity of the data is also much more compared to retail data. Retail data is mostly cross sectional in nature whereas smart meter data has time as another component. For simplification this study only uncovers one aspect of the problem, adoption mining of the smart devices in a home. There are several other further complex mining activities that can be performed with the same data. We presume that this field will mature in due course of time and researchers in future will bridge dimensions not touched in the present paper.

There are four major stakeholders for this study. First, manufacturers of smart devices who are concerned with the marketing of their product. Second, electricity distributing companies whose main concern is the elimination of existing problems associated with the distribution of electricity, especially in developing countries. Third, the government or policy makers who are enabled and are in a position to formulate appropriate policy; policy that is not only capable of eradicating existing problems of electricity generation, distribution and consumption, but is also acceptable by society. Fourth, consumers of electricity, who, in present scenario, are totally unaware of their own segregated and detailed electricity consumption. This awareness may help them to take appropriate electricity consumption decisions to optimize the level of consumption. This paper mainly focuses on manufacturers of smart devices as its subject of concern. This study will help manufacturers of smart devices to know adoption patterns and understand the consumer's segment. This understanding will help them to better formulate their marketing strategy. Furthermore, the

proposed study will help the manufacturers in two ways. First, it will mine the profiles of customers as initial adopters of their product. Second, this study will mine the association pattern between two or more launched products. Understanding of the temporal association between devices will let the manufacture know the average time lag between the adoptions of these devices. This information will guide future manufacturing and marketing strategy of the product.

Present study explains a very limited potential use of the data available from smart meters. Data simulation done for this study is also very simplistic. That is only good enough to explain a few applications. Research has found that electricity consumption variations, daily and seasonal, are capable of giving insights about the demographic and social details of households. In the present attempt, we have not covered these aspects of smart meter data mining. In the next attempt, we propose to integrate this complexity. Furthermore, in the next step of the study we aim to propose an algorithm to get insights from the simulated data.

References

- Al-Ghath, W., Sanzogni, L., & Sandhu, K. (2010). Factors influencing the adoption and usage of online services in Saudi Arabia. *The Electronic Journal Information Systems in Developing Countries*, 40(1), 1-32.
- Allwright, D., Heyes, C., Hurley, S., Kennett, M., Leese, R., Mitchell, S.,...Hulbert, P. (2006). Improving the sharing of the radio spectrum?: Final report - executive summary. Retrieved from <http://stakeholders.ofcom.org.uk/binaries/research/technology-research/final.pdf>
- An, S., Zhang, T., Zhang, X., & Wang, J. (2014). Unrecorded accidents detection on highways based on temporal data mining. *Mathematical Problems in Engineering*, 1-7.
- Baz, A. M. (1996). Method and device for active constrained layer damping for vibration and sound control. *The Journal of the Acoustical Society of America*, 101(5), 2424.
- Bellazzi, R., Larizza, C., Magni, P., & Bellazzi, R. (2005). Temporal data mining for the quality assessment of hemodialysis services. *Artificial Intelligence in Medicine*, 34(1), 25-39. doi:10.1016/j.artmed.2004.07.010
- Camara, M. S., Naguingar, D., & Bah, A. (2015). Prior data quality management in data mining process. In *New Trends in Networking, Computing, E-learning, Systems Sciences, and Engineering Lecture Notes in Electrical Engineering* (312, 299-307).

- Chauhan, A., & Saini, R. P. (2015). Renewable energy based off-grid rural electrification in Uttarakhand state of India: Technology options, modelling method, barriers and recommendations. *Renewable and Sustainable Energy Reviews*, 51(August), 662-681. doi:10.1016/j.rser.2015.06.043
- Cheng, T., & Wang, J. (2008). Integrated spatio-temporal data mining for forest fire prediction. *Transactions in GIS*, 12(5), 591-611. doi:10.1111/j.1467-9671.2008.01117.x
- Chou, J. S., Hsu, Y. C., & Lin, L. T. (2014). Smart meter monitoring and data mining techniques for predicting refrigeration system performance. *Expert Systems with Applications*, 41(5), 2144-2156. doi:10.1016/j.eswa.2013.09.013
- Cook, D. J., & Jakkula, V. (2008). Anomaly detection using temporal data mining in a smart home environment. *Methods of Information in Medicine*, 47(1), 70-75. doi:10.3414/ME9103
- Dagsvik, J. K., Jia, Z., Vatne, B. H., & Zhu, W. (2013). Is the pareto-lévy law a good representation of income distributions? *Empirical Economics*, 44(2), 719-737. doi:10.1007/s00181-011-0539-z
- Depuru, S. S. S. R., Wang, L., Devabhaktuni, V., & Gudi, N. (2011). Smart meters for power grid: Challenges, issues, advantages and status. *Renewable and Sustainable Energy Reviews*, 15, 2736-2742. doi:10.1016/j.rser.2011.02.039
- Eriksson, R., Werge, T., Jensen, L., & Brunak, S. (2014). Erratum to?: Dose-specific adverse drug reaction identification in electronic patient records?: Temporal data mining in an inpatient psychiatric population. Goodreads - User Review Google Book Preview Scimago Rank of this journal. *Drug Safety*, 37(5), 379.
- Fan, Z., Chen, Q., Kalogridis, G., Tan, S., & Kaleshi, D. (2012). The power of data: Data analytics for M2M and smart grid. IEEE PES Innovative Smart Grid Technologies Conference Europe, 4-5. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84874709434&partnerID=40&md5=e76c46b5d4866e2bd17240510eb7e24d>
- Francisco, E. D. R., Aranha, F., Zambaldi, F., & Goldszmidt, R. (2006). Electricity consumption as a predictor of household income: An spatial statistics approach. In VIII Brazilian Symposium on Geoinformatics (pp. 175-192).
- Greveler, U., Glosekotter, P., Justus, B., & Loehr, D. (2012). Multimedia content identification through smart meter power usage profiles. *Computers, Privacy and Data Protection*, 5. Retrieved from <http://www.nds.rub.de/media/nds/veroeffentlichungen/2012/07/24/ike2012.pdf>
- Jakkula, V., & Cook, D. J. (2011). Detecting anomalous sensor events in smart home data for enhancing the living experience. In *AAAI Workshop on Artificial Intelligence and Smarter Living The Conquest of Complexity* (pp. 33-37). Retrieved from <http://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/viewPDFInterstitial/3889/4212>
- Kadam, R., Mahamuni, P., & Parikh, Y. (2015). Smart home system. *International Journal of Innovative Research in Advanced Engineering*, 2(1), 81-86.
- Kitamoto, A. (2002). Spatio-temporal data mining for Typhoon. *Journal of Intelligent Information Systems*, 19(1), 25-41.
- Li, Q., Zhou, M., & China, P. R. (2011). The future-oriented grid-smart grid. *Journal of Computers*, 6(1), 98-105. doi:10.4304/jcp.6.1.98-105
- Mattern, F. (2003). From smart devices to smart everyday objects. In *Proceedings of Smart Objects Conference* (pp. 15-16).
- Mohassel, R. R., Fung, A. S., Mohammadi, F., & Raahemifar, K. (2014). A survey on advanced metering infrastructure and its application in smart grids. In *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference* (pp. 1-8).
- Newing, A., Anderson, B., Bahaj, A., & James, P. (2015). The role of digital trace data in supporting the collection of population statistics-The case of smart metered electricity consumption data. *Population, Space and Place*.
- Nezhad, A. J., Wijaya, T. K., Vasirani, M., & Aberer, K. (2014). SmartD: Smart meter data analytics dashboard. In *Proceedings of the 5th international conference on Future energy systems*. ACM, (pp. 213-214). doi:10.1145/2602044.2602046
- Persky, J. (1992). Pareto's law. *Journal of Economic Perspectives*, 6(2), 181-192.
- Prasad, S., & Avinash, S. B. (2013). Smart meter data analytics using OpenTSDB and Hadoop . Goodreads - User Review Google Book Preview Scimago Rank of this journal. In *IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)* (pp. 1-6).
- Sahin, I. (2006). Detailed review of Rogers' diffusion of innovation theory and educational technology-Related studies based on Rogers' theory. *The Turkish Online Journal of Educational Technology*, 5(2), 14-23.
- Samantaray, S. R. (2015). Special Issue on "Sensors and data analytics for smart grid infrastructure". Goodreads-User Review Google Book Preview Scimago Rank of this journal. *IET Generation, Transmission & Distribution*, 9(2), 113-114.
- Seetharam, D. P., Arya, V., Chakraborty, D., Charbiwala, Z., Ghai, S., Hazra, J., ... Kalyanaraman, S. (2013). Cyber Physical Systems for Smarter Energy Grids? *Journal of Indian Institute of Science*, 93(3), 541-551.
- Sharma, T., & Balachandra, P. (2015). Benchmarking sustainability of Indian electricity system: An indicator approach. *Applied Energy*, 142, 206-220. doi:10.1016/j.apenergy.2014.12.037
- Spenceley, S. E., & Warren, J. R. (1998). The intelligent interface for online electronic medical records using

- temporal data mining. Proceedings of the Thirty-First Hawaii International Conference on System Sciences, 5, 92328033. doi:10.1109/HICSS.1998.648321
- Svanström, H., Callrus, T., & Hviid, A. (2010). Temporal data mining for adverse events following immunization in nationwide danish healthcare databases. *Drug Safety*, 33(11), 1015-1025. doi:10.2165/11537630-000000000-00000
- Wehlitz, R., Werner, A., & Franczyk, B. (2014). SMIM-A cloud-based approach for the digitisation of smart meter installation processes. *Journal of Industrial and Intelligent Information*, 2(3), 169-174. doi:10.12720/jiii.2.3.169-174
- Weranga, K., Kumarawadu, S., & Chandima, D. P. (2015). *Smart grid and smart metering. In Smart Metering Design and Application* (pp. 1-15). Springer Berlin Heidelberg.
- Zhang, X., Grijalva, S., & Reno, M. J. (2014). A time variant load model based on smart meter data mining. In IEEE PES General Meeting / Conference & Exposition (pp. 1-5).

Shashi Kant Srivastava is a doctoral participant in the Information Systems area of Indian Institute of Management (IIM) Indore. He holds a Bachelor's degree in Architecture from University of Roorkee and Masters in Ekistics (Science of Human Settlements) from Jamia Millia Islamia, New Delhi. Prior to current engagement, he worked as an Associate Professor at Apeejay Institute of Technology, Greater Noida. He can be reached at f13shashis@iimidr.ac.in.

Think Bigger: Developing a Successful Big Data Strategy for your Business

Shweta Grover

[Mark Van Rijmenam (2014). *Think Bigger: Developing a Successful Big Data Strategy for your Business*, Amacom, Pages 288, ISBN: 978-0814434154.]

The number one benefit of information technology is that it empowers people to do what they want to do. It lets people be creative. It lets people be productive. It lets people learn things they didn't think they could learn before, and so in a sense it is all about potential".

– Steve Ballmer

Technology has been an apple of mankind's eye since the invention of the wheel. Technology in its simplest form was originally meant to make our lives easier by supporting us in various daily and routine functions. However, the expansion of civilization caused technology to not just be the backbone of society but also a pastime for purposes of entertainment, work and occupation.

Who knew that in a short span of years, the world would become a digital entity? Cloud, Data Transfer, Portable Hard Drives, Memory Cards, these developments have revolutionized the way data is stored, used and transferred. Everyday conversation is now interspersed with questions of Kilobyte capacity and the ease of storing files of considerable size. This is even more significant when it comes to businesses and entrepreneurial enterprises. Not only can such businesses or enterprises take advantage of modern technology to enhance their business and profit, but they can also benefit from the vast options provided because of the versatility of data Storage.

Businesses today are using Dashboards, Segmentation Analysis, Heat Map Analysis, Site Search Analytics, etc. to increase their productivity and revenues to gain a competitive edge and thus surpass the other participants of the industry/market.

Big Data is also one such service. Big Data i.e. the huge amount of data which has been created or collected as virtually when every choice, transaction or movement

we conduct becomes digitalized, is drastically transforming the business and its operations. Businesses can make the best use of this data to gain valuable insights about the behaviors of customers and thus, serving them the right mix of products. It has been seen that the organizations which uses the insights generated by Big Data for conducting their processes generate higher revenue and thus surpass most of their competitors and peers. However, the job doesn't end with the collection of data but it starts with it. The Company should know the right tools, techniques, algorithms and metrics to make use of it. Thus, it should have a strategy on "How to use Big Data".

Mark Van Rijmenam has written the book "Think Bigger: Developing a Successful Big Data Strategy for Your Business" to address this issue. Author, Mark Van Rijmenam, is an entrepreneur who values innovation and thus, guide and inspire others by making a difference in usual practices. He has founded Datafloq.com. Datafloq is all about Big Data which connects all the stakeholders of the global market of Big Data. He is also a strategist of Big Data and thus provides guidelines and tenets to organizations for developing strategies on Big Data. His forte includes Social Media, Strategy, Mobile and Online Marketing and Big Data. He is well versed with all the recent changes that may have an impact on the business performance. He is the co-founder of Data Donderdag. Data Donderdag is an event which is organized bi-monthly in The Netherlands to make

organizations comfortable about using and understanding Big Data. He is a highly sought-after international public speaker. He is one of the top Big Data influencers over the globe. He also writes articles and blogs for Datafloq.com and various other platforms.

Amacom has published the book, "Think Bigger: Developing a Successful Big Data Strategy for Your Business" in April/ 2014. This book gives enough practical guidelines to ensure that your organization is way ahead of its competitors. It provides the framework of possibilities which are clear and easy to understand. It also gives tenets on what your organization should do and how it should be operated and managed when there are turbulent changes in the technology. This book states that if you know the art of managing Big Data, you can get various advantages over your competitors such as personalized pricing, improvised support for customers, identification of risks, market development, etc.

The more important question the book asks is, however, that given the rise of Data usage and storage, why should the IT world only benefit from this development? Big Data is a service that has immense potential in application to various businesses. In fact, Big Data is already changing the way businesses work. The author repeatedly stresses the need for organizations worldwide to understand what Big Data is and how it can be used to full advantage. The benefits and implications are far too vast for any business owner to ignore.

The service does have a few considerations to be taken into account before one embarks on a strategy to incorporate it. Big Data requires a different culture, it is omnipresent and it needs to be protected, hence airtight security measures are necessary.

The books further go beyond just informing the reader about Big Data. It draws on extensive research and numerous practical case studies, in a way providing a path and useful takeaways to implement a successful big data strategy that best serves an individual's business. It also offers advice on putting Big Data on the balance sheet of an organization and determining its return on investment, along with emphasizing the important role Big Data plays in customer relations,

public relations as well as human resources.

Furthermore, a whole chapter is dedicated to how Big Data can change 16 different industries, namely: agriculture, automotive, consumer goods, education, energy, financial services, gaming, healthcare, non-profit, manufacturing, oil and gas, telecom, transportation, retail, travel, and government.

Readers also get an insight into successful practices of companies such as Amazon, Disney, Nike, and Wal-Mart, among others, who have implemented Big Data services.

After reading this book, reader can expect to have the knowledge of following things:

- 1) History of the Big Data
- 2) Meaning of Big Data
- 3) 7Vs
- 4) Some important realities of using and incorporating Big Data in organization
- 5) Effect of Big Data on business
- 6) Trends of Big Data
- 7) Gamification
- 8) Privacy, Security and Ethics of Big Data
- 9) Presentation of Big Data on the company's Balance Sheet
- 10) Future prospective of Big Data

Analysis and Critique

This book written by Mark Van Rijmenam, talks about how the days of gigabytes and megabytes are forgone and how the Big Data which has been collected through various digital sources in the world is making us introduce new terms for describing it. This book offers clear and easy to understand explanations and insights for managers and top management to get a feel of Big Data and its management. It does not focus on a particular industry or size but talks, in general, about the development of Big Data strategies. Mark also stressed on the fact that the organizations cannot simply continue to use their old processes and ignores the uses and benefits of Big Data. Since, Big Data is vital for company's performance, they have to learn the tools and techniques sooner or later.

It also explored the concepts that every manager should consider while developing a strategy for Big Data.

This book provides the reader with the guidelines, tenets and also the lessons which he might use while implementing this technology. It also suggests treating Big Data as an asset for an organization and how to ascertain return on investment for this. It covers the trends of Big Data which are affecting the organizations, about some other important technologies and also about the various set of analyzes.

Moreover, this books also provides guidelines on how not to sacrifice security and to give due respect to the rights of privacy of the buyers.

However, there were some aspects which this book fails to adequately address:

- 1) This book does not provide the meanings of various Big Data tools and if you are a beginner in this field, you may find it little difficult to relate to each and every concept mentioned in this book.
- 2) Also, most of the examples and cases were from developed countries. Being an Indian reader, I would have appreciated if there were cases on How to develop Big Data strategy in developing nations.

To summarize, this book talks about how the emergence of Big Data changing the environment and working styles of business and also the government. It attracts all those readers who have some interest in the innovation and the development of an organization. It can serve as a guide to all those who have bid something for the growth and future perspective of a particular organization and can also serve as an important resource for Mangers/Leaders who want to be sure that their company is not left behind by the competitors.

Shweta Grover joined IIM Indore as a participant of Fellow Programme in Management in Economics area in 2015. She is a graduate in B.A. in Economics from Delhi University and Masters in Financial Economics from Gokhale Institute of Politics and Economics, Pune. During her two years of association with Aon Hewitt, Gurgaon, she gathered experience in HR Analytics, Engagement, Labour economics, etc.

Semi-Organic Growth - Tactics and Strategies behind Google's success

Kakul Modani

[George T. Geis (2015). *Semi-Organic Growth - Tactics and Strategies behind Google's success*, Wiley, Pages 240, ISBN: 978-1118933220.]

We live in a world where traditionally a company can grow in two ways - either organically i.e. by growing internally through increase in production or customer base, etc. or inorganically by way of mergers and acquisitions. Now, the Author explains to us in the book an alternate growth path - the Semi-Organic Growth way.

The Author has defined Semi-organic growth as under, *"Semi-organic growth is generated when revenue results from products or services that emerge when acquired technology-related assets attach to a company's existing capabilities in a complementary manner."*

The questions that immediately pops up in our minds are - How do you understand the semi-organic growth way? How is it different from the traditional paths? Why is it even necessary to have a third way?

The Author has answered all the pertaining questions through judiciously choosing a company that has been a serial acquirer in a patterned manner in the last decade and has followed a semi-organic growth model. He picked Google, a Fortune 500 company, with a history of more than 200 acquisitions out of which two-third have been successful and synergic. Not only the Author analyzes the mergers and acquisitions of Google in the past decade in order to explain the readers the semi-organic growth pattern but also answers questions in mind deftly.

The Author explains how Google, unlike its competitors, have had a series of acquisitions since its inception, even before it went public on August 19th, 2004. This company boasts about 200 odd purchases in various related as well unrelated fields while being more than successful in it as per the industry standards, commands a study on it from Mergers & Acquisitions (M&A)

perspective. The author uses infographics so to make the readers understand and appreciate the deals in a better way and attempts to answer any question that the readers may have on acquisitions done by Google.

At the beginning of the book, he provides a framework to measure the impact of an M&A and attempts to explain it to the readers with respect to Google. It is essential to follow a pattern, if any, by which Google makes these acquisitions and does it have anything to do with its first acquisition and the way it was done? The book throws light on this aspect in the initial chapters. The author explains the following main aspects of M&A program for it to be sound:-

1. Strategy - This point revolves around thoroughly understanding the rationale behind the M&A.
2. Deal Economics - It mentions that the valuation of the target using a constellation approach (i.e. using more than one valuation methodology) instead of going for North Star valuation metric. Synergy analysis is also a part of this aspect.
3. Organizational Design - This part of the program tries to measure how well the integration will happen. Also, the reader gets an idea that organizational design's process of integrating varies from one acquisition to the other.
4. Deal Dynamics - Finally, we come to a position where the company talks about the structure of the M&A deal, like whether they deal would be cash or stock or a mix both.

Then, the performance evaluation of the M&A deal is done. The M&A effectiveness of a deal can be measured by value-creation, strategic relatedness or organizational behavior. Google's performance evaluation in its M&A activities is done by the value-creation school of thought as well as strategic relatedness. However, the Author

explains how the acquisition of NEST labs (a home automation producer, acquired for USD 3.2 billion) is an example of strategic un-relatedness as the acquisition doesn't pertain to Google's core ad-tech business. The third school of thought for M&A effectiveness involves organizational behavior.

The book introduces the reader to new terms like Acqui-hire, which has been practiced by Google while making smaller acquisitions. Acqui-hire means acquiring a company for its talent, where technology, product, etc. may or may not be the reason for the acquisition. The perfect example of acqui-hire for Google is the acquisition of Applied Semantics (ASI) which was a pre-IPO acquisition done in April 2003. This acquisition also imprinted Google and charted a way for its future acquisitions.

The Author states the fact about an acquisition for which Google paid USD 41.5 million in cash plus stock and USD 60.9 million worth of stock options in a cash plus stock deal for ASI. This acquisition proved to be extremely fruitful for Google as it resulted in a massive contribution to its advertising revenues for the years to come. AdSense (an ASI product) contributed around 25% to Google's advertising revenues in the next decade. This acquisition provided Google with an inclination towards M&A activities as well as performing the M&A activity in a certain way i.e. the semi-organic way. Hence, the book contains an entire chapter on this acquisition and its details in which even the problems that Google had to go through for the semi-organic integration of the ASI to Google are mentioned.

Going forward, the book forces the reader to think whether Google's acquisition spree is a result of merely using its ideal cash pile or part of experimentation by the inquisitive founders or a well-thought of a plan. Once the reader starts asking these questions, the book tries to answer it in its upcoming chapters.

Generally, in a business book, competition is talked about in a cohesive manner. However, this book does it differently. Google's competition with Apple is discussed in an initial chapter in which its competition with Facebook and Amazon is addressed in the later chapters. It looks like a well thought of move. Once you

get the basic understanding of Google's M&A activity, the reader expects that the next chapter would talk about the M&A deals in a detailed manner. The book doesn't, in fact it takes a turn and brings Apple in the picture.

The book explains that Apple and Google have displayed an extreme behavior when it comes to M&As. While one is a serial acquirer, the other has been organically growing by choice barring a few small acquisitions, possibly believing, as mentioned in the book "real men don't do M&A." This was the case till 2009. Enter 2010, Apple and Google have got into dyadic cascading. In November 2009, Google acquired Admob for USD 750 million in stock and in reply, Apple bought Quattro Wireless from the same industry. Then in March 2010, Apple acquired Intrinsity to enhance battery life in mobiles while Google in April 2010 acquired Aguilux. This dyadic cascading has continued from that point forward. Apple started making huge acquisitions (eg. Beats for USD 3.2 billion), unlike its earlier attitude of growing organically. The book provides infographics of acquisitions of Google and Apple for 2010 and 2013 to let the reader appreciate how Apple changed its attitude towards acquisition.

Keeping aside the information this book provides on Google's M&A deals, the key takeaway of the book is learning how to build an M&A market model. As per the author, "M&A market modeling fundamentally involves chunk building- accumulating meaningful pieces of information about market segments/ subsegments, companies, and deals. A robust market model can only be built with persistent effort." The book attempts to develop a robust M&A market model for Google and its acquisitions where the categories and the basis are described vividly. To begin with, the Google M&A model is divided into three main categories - 1) Media deals, 2) Internet software/services deals, 3) Technology platform deals, and 4) Others. Further, the model subdivides these categories. For instance Media deals are divided into Advertising that is further subdivided into Internet Advertising, Mobile Advertising, Publishing Advertising, Radio Advertising, Television Advertising etc. The author also tries to make the reader understand the difference between

M&A modeling and the traditional classifications such as SIC (Standard Industrial Classification) or NAICS (North American Industry Classification System). The M&A model distinguish each and every acquisition on the basis of its connection (direct or indirect) with the main category and provide it with an icon. A click on the icon gives all the available information on that particular deal in a prescribed format. The model also accepts specific queries that relate to the problem. A comparison between two principal companies' specific acquisitions can also be made. The book provides step-to-step instructions on building the model and also speaks about its scope.

The forthcoming chapters are dedicated to the three categories and descriptively talks about the various acquisitions that Google performed in the field of Media, Internet and Technology Platform. The book provides an exhaustive summary of the deals related to that particular area. For instance, the technology platform mentions the acquisition of Nest Labs by Google in USD 3.2 billion. This is an entirely unrelated to core business acquisition for Google, but Nest Labs are in the company of producer of home automation products. Hence, this deal comes under technology and a different subcategory.

While going through the first half of the book, a question arises in the reader's mind as to know certain other aspects of the M&A deals like the failed deals and the repercussions of those or may be the extent to which Google and its competitors made the deal disclosures. As the reader moves further, these questions are answered in the book. The book has put Motorola Mobility acquisition by Google under the chapter that talks about the failed or shuttered deals. As the reader goes through it, the book has inputs that says that Motorola Mobility deal hasn't been a failed deal for Google per say. These instances and the numbers supporting the discussion ignite curiosity in the mind of the reader. Also, considering the wide variety of acquisitions that Google has made, the book suggests that it should have different parameters for evaluation of a deal with respect to target's position in the life cycle, product expectations, etc.

The book also touches upon how secretive American companies like Apple, Google etc. are when it comes to deal disclosures. SEC guidelines for disclosure give the leeway in the materiality concept as it says, "The omission or misstatement of an item in a financial report is material if, in the light of surrounding circumstances, the magnitude of the item is such that it is probable that the judgment of a reasonable person relying upon the report would have been changed or influenced by the inclusion or correction of the item." It is because of this reason that Google mentions its acquisition of Waze Limited for a cash consideration of USD 969 million separately whereas it clubs the rest of the acquisitions of the same period together and hence the public doesn't come to know the target's name at times. Here, the M&A model comes in handy if the user wants to run a query asking for the M&A details of the deals whose values are disclosed by Google. The last portion of the book talks about the deal dynamics of the M&A program. Like, Google has used stock consideration in four deals till now (as per the disclosures) that includes ASI and YouTube.

Semi-organic growth's most essential aspect is the talent the target brings. It is the most important trait of this type of growth. The author emphasizes on the fact that how important it is for the founders to be a part of the team post the acquisition and that comes under contingent consideration. To make the founders stay for a while, there is an Earn-out provision where in some milestones are set post-acquisition based on which the target would receive a remaining deal amount. The book talks about other such mechanisms to motivate founders as well as employees and how well semi-organic growth has become Google's way of acquiring the firm. The last part of the book emphasizes on the M&A integration and describes four major types of acqui-hires and how each and every type has enhanced Google's offering. An interesting part of the book is that it also tells the reader where the founders of the target company went after they left Google. It is a piece of information that the reader would don't mind knowing. Like, Twitter co-founder Evan Williams previous firm Pyra Labs (Blogger) was acquired by Google. Twitter was founded by him after he left Google.

Once the reader finishes *Semi-Organic Growth - Tactics and strategies behind Google's success*, it is clear that the book has lived up to its title as now the reader has valuable information on mostly all the aspects of the M&A deals undertaken by Google. Having said so, the book lacks with respect to letting the reader know about the company's perspective over any deal post an acquisition. The content would have been richer if there were quotes directly from the Horse's mouth which in this case is Google. However, one can't shy away from appreciating the references that are provided at the end of each chapter that helps in substantiating the author's writing. There are aspects of the book that are explained in great detail like the deals under the three categories namely Media, Internet and Technology Platform but there are also elements in the book that have been touched at the periphery like Google's competition with Amazon. As they say, too many cooks spoils the broth. Luckily for this book, despite concentrating on many subtopics, it doesn't look fragmented. A reader from the finance/strategy/Information Technology background might feel the necessity of a deeper understanding in building the M&A market model after going through the book but the book suffices in the construction of the foundation of the model. The provision of 50 infographics at various places in the book provides visualization to the user who would result in enhancing the reader recall when there will a discussion on such topics.

A curious thought may run into the reader's mind after looking at the title of the book as Google is now Alphabet. Alphabet has become the holding company and has Google Ventures, Google Capital, Fiber, Calico, Nest as the heads along with Google. Google comprises of

YouTube, Android, Search, Apps, Maps and Ads. The change in organizational structure questions a lot of inferences drawn in the book as when the book was written it was Google, not Alphabet. The silver lining being, the book can be used to draw a comparison between Google and Alphabet's semi-organic growth path if need be. If Google had become Alphabet in 2014 instead of 2015, the book might have been written differently.

About the Author - George T. Geis teaches at UCLA Anderson in the areas of mergers & acquisitions, financial modeling, entrepreneurship, and accounting. He has been voted Outstanding Teacher of the Year at UCLA Anderson five times, the most recent honor being given by the MBA class of 2012. Geis has also served as Associate Dean and Faculty Director of UCLA Anderson's Executive MBA program. He is currently Faculty Director of Anderson's Mergers and Acquisitions Executive program. Dr. Geis is an expert on M&A activity in technology, communications and media markets. Geis received a B.S. "summa cum laude" and with "honors in mathematics" from Purdue University, an M.B.A. from University of California, Los Angeles, and a Ph.D. from the University of Southern California. (Source:<http://www.anderson.ucla.edu/programs-and-outreach/information-systems-research-program/faculty-and-staff/geis>)

Kakul Modani currently pursuing FPM in finance and accounting from IIM Indore. She holds a Masters degree in Financial Engineering from NYU and have worked as an Equity research analyst for the FMCG and Media sector. Monitoring the capital market is my favorite hobby. Apart from this, she has a penchant for urdu poetry, good vocals, kids and clouds.

Code Halos: How the Digital Lives of People, Things, and Organizations are Changing the Rules of Business

Suganya Balakumar

[Malcolm Frank, Paul Roehrig and Ben Pring (2014). *Code Halos: How the Digital Lives of People, Things, and Organizations are Changing the Rules of Business*, Wiley, Pages 256, ISBN: 978-8126548606.]

In the evolution of the human species if Industrial revolution were to be described as a major step; the technology revolution can be considered as a giant leap. This revolution has been made possible by the generation of innumerable amount of data around us. Imagine the vast oceans of data around every being - text messages, every website visited, songs listened to, every click made, Facebook likes etc.- Humongous. The arduous task for the Code Halo (inclusive of big data) companies is to make meaning from the colossal amounts of data, and to apply the understanding for business strategies and practices.

"As the authors define it: A Code Halo is the field of digital information that surrounds any noun-any person, place, or thing. More often than not, that virtual self can provide more insight into-and thus generate more value from-the physical entity alone"

Through a rich narrative the authors attempt to answer the questions whether individuals/ businesses were able to see the technology-based sea change coming? Were they able to capture the commercial opportunity? And, they direct the reader that using technology is at the discretion of the individual/business, but by ignoring the significance of these transition businesses may become extinct. The book also encourages the readers to think and relate to their everyday activities.

The Digital Economy

The leaders, winners and outliers of today's world are playing a different game; fighting a different fight; and winning with a new set of rules. With the immense amounts of data generated by us, using our widgets, there is an unseen digital economy that is quickly emerging. This started with the commercialization of the internet, and according to W. Brian Arthur of the Santa Fe Institute and the Palo Alto Research Center (PARC):

"This Second Economy will surpass the world's first (or industrial) economy in scale and scope by 2025. The virtual, digital economy will exceed the industrial economy-the cumulative output of mankind's development over several millennia- in terms of transactions, revenues, and value creation, in a mere 30 years"

What made, or makes, the digital economy successful, rightly mentioned by the authors, is their business model - the creation and management of Code Halos. According to the authors, "Amazon eliminating Borders was not a book thing, nor was Apple outrunning Nokia and BlackBerry a mobile

phone thing, every one of these industries' disruptions was a Code Halo thing." The emergence of these and much other, exuberant growth in digital economy reveals two major patterns - the Code Halo formed the basement for the commercialization of consumer technology and there was outsized growth in a very consistent manner. What is surprising is that the skills for creating, implementing and managing these are not taught in any curriculum in the education system.

What has made such business models successful is our interwoven personal lives with the internet. In the past two decades, we have highly personalized and blended our virtual world into our real world. According to Mark Parker, the CEO of Nike, "The digital and physical worlds are starting to come together - it's only the tip of the iceberg in terms of what's coming."

Tyes and Elements of Business Code Halos

The authors enunciate that such interweaving is possible through the five Code Halo solutions - Customer, Product, Employee, Partner and Enterprise - which the authors call "Holocracy". The authors have provided with an extraordinary emphasis to the term "Holocracy", grabbing the reader's attention as the book unfolds.

You and I, as customers have personal Code Halos. This, in the hands of the company, is its lock, and the key to it is providing us with personalized customer experience by applying individual code. It provides targeted commercial interactions. For example, Amazon provides personalized shopping experience; it provides the right items in the right place at the right time. The products are the software and hardware that collect our information. Our Code Halos connect with the Internet of Things, so that value and meaning is generated over time. For example, our smart phone and the installed games and apps, which are increasingly network aware, are products that can gather information about us. As personal Code Halos enrich customer experience, Employee Code Halos enrich work experience. Sharing of our work history helps in knowledge transfer, individualizing and transforming work experience. It facilitates getting the right work, to the right person, at the right time. It also helps in improving productivity, employee retention and hiring. For example, LinkedIn helps in employee-employee and employee-employer communication. The Partner Code Halo integrates the product and employee code halos, it brings clarity and insight to real time decision making. And finally, the Enterprise Code Halo is an aggregation of all four Code Halos. It gives the organization an identity based on what it does and how it does it.

"As the number of interconnected things expand, almost every device, product, and object - from simple consumer goods to complex industrial equipment - is a candidate for a Code Halo"

To collect and assess the required data the Code Halo elements are employed - amplifier, application interface, algorithm, data and business model.

Amplifiers are the devices that enable interaction (Computer devices, industrial machines, wearables). Application interface refers to the inbuilt softwares. Algorithm, the most important of the five elements, processes the data to make meaning. Data is then analysed and used in Business models.

The authors enhance the fact that the organization's Code Halo explains more about the organization than its campaigns or annual reports. This forms the basis

for understanding the growth in the digital economy.

The SMAC Stack

When I purchase a book in Amazon.in I get further recommendations of books to buy, this is based on the preferences of other individuals who have similar tastes like mine. When I search for flights between Indore and Bangalore in MakeMyTrip.com I see advertisements of MakeMyTrip.com in the other websites I visit (for the same route). This is because MakeMyTrip.com records my searches, and follows me to the other websites it has partnered with. This is how Google and Facebook earn its revenues. This exchange of information by companies is possible because codes connect with codes on four technologies- Social, Mobile, Analytic and Cloud (SMAC) - and they work together to know the minutest of details about our personal likes and dislikes. The authors have made a significant contribution by providing us with this integrated view.

The SMAC Stack provides the raw materials for Code Halos. SMAC technologies are the infrastructure of the new Code Halo economy. Explaining the SMAC Stack:

Social technologies: All software that helps in interpersonal communication, like blogs, email, social networking etc.

Mobile technologies: All portable hardware devices and the software, network and applications that support those devices. For example smart phones, tablets, wearables etc.

Analytics technologies: A software supply chain that helps in collecting, organizing, managing and analyzing myriad amounts of heterogeneous data.

Cloud computing: The network of remote servers on the internet that helps store and process data, instead of owning and hosting them internally. This allows for greater flexibility of resources and lower costs.

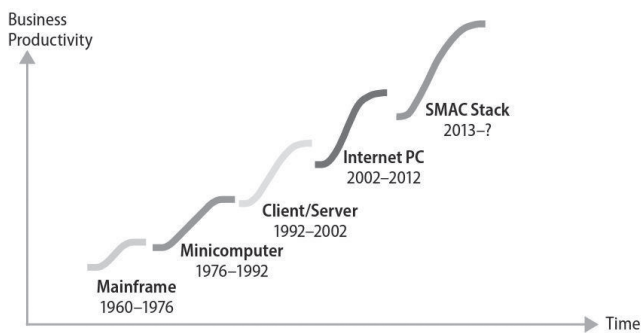
Initially the SMAC technologies were disconnected, but now there is a high degree of integration, mainly fostered by the information exchange between the participating companies. With the development of the SMAC technology the internet age has evolved to the digital economy age. The authors explain this as the fifth wave

of Corporate IT - from Mainframe, Mini Computer, Client/server, and Internet PC to SMAC Stack.

The Fifth IT Wave

The first four IT Waves have enhanced business productivity. It has helped in reducing costs, generating revenue and providing productivity gains. The same is expected from the present SMAC technology model. No technology is an island now. The combination of the SMAC technologies has resulted in the multiplier effect, interlinking the technologies to provide meaningful data for further business implications.

"For a Code Halo to have an impact on performance, it must be integrated into well-codified and well-understood business processes- such as sales, customer service, research and development, or supply chain management"



In the past decade Apple, Google, Facebook, Amazon, Pandora, and Netflix have generated market value of more than \$1 trillion, and this has been made possible because of the creation and management of Code Halos. It depends on companies to leverage this new technology. For example, Zune was a portable media player marketed by Microsoft, launched as a competition to Apple's iPod. It started in 2008 but was discontinued in 2012, mainly because it failed to build proper partnerships. It did not build apps, it did not build games, and it did not generate information. Thus, it was unable to provide any meaningful user experience and remained as a fancy music player.

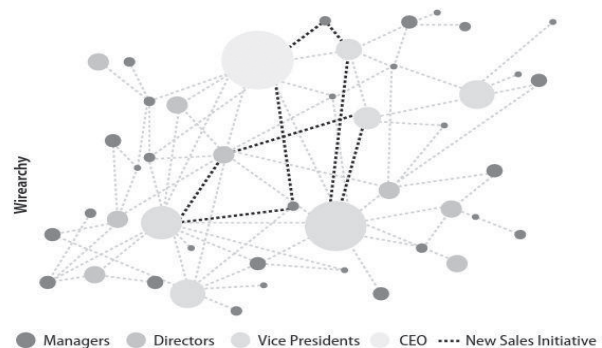
"The Code Halo opportunity is vast, and the downside for missing this trend will be harsh and swift."

Principles of Code Halo Strategy

The authors provide a "new rule of code" that will help business survive and thrive in the Code Halo world. The authors recommend businesses to adopt the four principles of Code Halo strategy: (i) To deliver beautiful products and experiences (ii) Avoid evil (iii) Wirearchy principle and (iv) To nurture Code Halo heroes.

Firstly, the Code Halo solutions must provide users with beautiful products and solutions, like apple iPod and unlike iPhone 6. The users must experience the "Wow! How did they read my mind?" effect. Secondly, the Code Halo companies must learn to ignore the voices of the dark side. Issues of privacy and ethics will erupt, people will be hacked, government interventions will grow and advertisements and media will continue to trouble. But, the winning organizations must recognize, understand and actively manage such negative issues.

Thirdly, businesses should establish a "Wirearchy" organizational structure. The term Wirearchy is defined for the Millennial as the organization structure where members earn their status through knowledge and willingness to share. These are dynamic networks with interconnected nodes and free of ranks. With such a structure the employees are obliged to be an active part of the community in the Wirearchy. The Facebook organizational structure is one such, which defies, but not completely eradicates, the age old command and control hierarchy. The authors list the rules to follow for success in the new organization model. Though the authors created an enduring effect throughout the book with the term "Holocracy", they fail to do the same with "Wirearchy".



Lastly, it conciliates the reader to find that the authors have prioritized the role of "Halo Heroes", who would make use of the one-time opportunity to step forward and lead the organization. Halo Heroes would progress in such a way to tear down the walls between IT and business, and lead the new IT organization into place. The businesses should endeavour to accommodate virtual work habits with the physical work habits; this will enable the business and IT to work together.

Cross Roads Model

The most interesting part of this book is the description of the Cross road model. The authors have elaborated this to a considerable extent that a business leader would implement such a process without second thoughts.

Imagine the days when Mark Zuckerberg had the idea of building a "hot or not" game for Harvard students, also called "Facemash". And then he went on to create "thefacebook.com" with the intention of creating a universal website that can connect people in the university. After expulsion and lawsuits he created the social networking site "Facebook" in 2004. Facebook has now become a global internet phenomenon from a Harvard social-networking website. This phenomenon, the authors' detail, is the Cross roads models.

Metaphorically the authors explain how nature feels the exuberance when lightening is about to strike as "ionization", where the Code Halos doesn't exist and the situation in the industry is ripe for its introduction. They continue to describe how a "spark" is like lighting the match, a small spark can ignite into a fire for the good or the bad, and likewise a small idea leads to the evolution of the Code Halo process. "Enrichment" is when the spark turns into a blaze, the Code Halo solutions scale up increasing the number of users and generating more data, building meaningful services and solutions. At "cross roads" is when the markets flip and the companies experience massive transitions. And "after cross roads" the companies set on an accelerated journey which is difficult to stop, be it negative or positive.

"There are decades where nothing happens; and then

there are weeks where decades happen" - Vladimir Lenin

The authors provide guidelines for the organizations to follow based on the Cross roads model, but each organization might have different requirements and also may prefer not to follow these.

According to the authors the business begins to realize that there is something strange, there is a "future of work", which has to begin with an "Action plan" which is the Ionization phase. The main reason that businesses do not "Ionize" is because there are voices in the organization pulling them back and there is no "will to innovate". The authors provide a set of actions that the businesses could undertake - beginning from preparing the organization for innovation, mapping its value chain, scanning the market for signs of Ionization, listening for new voices within the organization, and finally picking the Code Halo targets. Though there might be several ideas worth exploring in the organization the Code Halo target must be the one that delivers real impact.

When the innovations come to life a Spark is created. In this stage the authors advocate businesses to embrace mass personalization, build Code Halos around the main character of the progress and pilot and fine tune the business model. The challenge for the organizations here would be to pick the right Spark from the countless ideas available to them.

Once the Code Halo solutions are ready it must augment products and services to provide with meaningful information and data, this is the Enrichment stage, where the Spark turns into a blaze. Organizations learn during this stage, and algorithms develop with business analytics. At this stage the authors suggest a set of actions for the businesses - to provide a balanced focus of all the five elements of the Code Halo anatomy, to get the right data and user experience, to use the right metrics, to capture code and make meaning, and to expect the unexpected.

Code Halos: How the Digital Lives of People, Things, and Organizations are Changing the Rules of Business presents itself as a practical guide for business to participate in the ongoing digital revolution. It provides

the businesses with guidelines to change process and attitudes of the company.

With the blending of the virtual and physical world, every individual and businesses have lost their personal space. A new economy is being created. Similar to the previous economic models in the world, this economy will also develop in a scattered manner, bringing with it the opportunities to develop. This book helps this transition easier. It helps businesses see what is available for them and what awaits them in the future. It is no Bible to follow, but basic guidelines to adapt. The authors do not provide a path for success but show businesses that they could do much better, than they are doing right now, if they participate in the Code Halo economy. It is a series of recommendations that would help people and business successfully navigate through these new waters.

With innumerable examples the authors bring to the readers the importance of the growing digital economy, and the opportunities available. Though the book provides a model, and provides guidelines it needn't be accepted by the readers at its face value. The authors fail to explain the economic situations that the businesses must consider. Also, there is very limited mention about the expected competitions. When industries grow, it attracts new entrants, competitions grow. For this, the businesses must create strategic architecture for themselves, with core competencies, so that they could sustain and grow irrespective of the competitive environments.

The guidelines provided by the authors are too lengthy. The authors have also failed to realize that individual

decisions depend on perceived value - be it the firm or the customer. For example, the authors recommend "to deliver beautiful products and experiences", but "beauty" depends upon the individuals' perspectives' and 'environment'. I doubt if such generalized statements can be provided as guidelines.

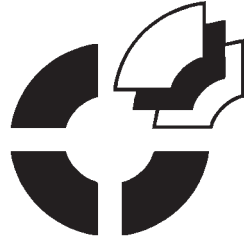
These criticisms aside, the authors provide an insightful thought into the future. Their ideas are systematic, far-fetched and exploratory in nature. As the authors mention businesses should aim at identifying the Code Halo opportunity, adopt the principles of the Code Halo strategy and implement the crossroads model. This book is a practical guide for businesses in the new digital economy.

This book would enable the managers to answer the questions posed by the authors in the beginning - "Did you see the technical change coming? And, if so, what did you do about it?"

But the managers must look beyond the recommendations that the book provides, to suit their requirements.

Suganya Balakumar is a doctorate participant in the Economics area at the Indian Institute of Management Indore. She holds a Master's degree in Economics from Madras School of Economics, and a Bachelor's degree in Economics from Madras Christian College. Her areas of interests include Development economics and Institutional Economics.

Her personal interests include traveling, writing and blogging.



सिद्धिमूलं प्रबन्धनम्
भा. प्र. सं. इन्दौर
IIM INDORE

INDIAN INSTITUTE OF MANAGEMENT INDORE
Prabandh Shikhar, Rau-Pithampur Road, Indore 453 556 (M.P.) India

Guidelines for Author(s)

The manuscript should be accompanied by an abstract of 150-200 words and up to six keywords. All the manuscripts submitted for this section will undergo double blind review process after being vetted by the Editor.

Research Articles should:

- Have a clear statement of purpose
- Have a clear statement of importance; why the paper was written and what it contributes to the body of knowledge
- Have clearly stated implications for further research
- Have clearly stated implications for practice
- Be well-written and readable
- Present reliable and valid conclusions appropriate to the methodology employed
- Be completely free of spelling and grammatical errors

For the complete author(s) guidelines, please visit:

<http://www.iimidr.ac.in/iimi/index.php/publications-consulting/indore-management-journal>

Manuscript Submission:

The manuscript may be submitted to the following link:

<http://www.iimidr.ac.in/iimi/index.php/online-submission>.

The authors will receive an acknowledgement receipt. In case of any difficulty in online submission, manuscript may be sent directly to imj@iimidr.ac.in.

Each submission should include author information (names, affiliation, address, phone number, fax number and email, author(s) profile within 100 words).

Address

The Editor

Indore Management Journal

Publication Office

Indian Institute of Management Indore

Prabandh Shikhar, Rau-Pithampur Road, Indore-453 556, Madhya Pradesh, India

Phone- (0731)-2439 670

Email: imj@iimidr.ac.in

CONTENTS

EDITORIAL

- 1 Big Data: Road Ahead for India
Madhukar Dayal, Sachin Garg and Rubaina Shrivastava
- 15 Big Data - Intuitive Lessons for HR Managers
Debolina Dutta
- 17 A review of Evolution of Theories and Models of Technology Adoption
Rajesh Sharma and Rajhans Mishra
- 30 Participating in Social Networking Sites (SNS):
Mediating Role of Self-disclosure and the Effects of Well-being
Manoj Das
- 39 Government Process Re-engineering for an E-governance Implementation for
Motor Vehicle Registration in India
Suresh Subramoniam and Dev Twinky
- 47 Internet of Things (IoT) and Smart Technologies:
Framework for Temporal Data Mining Concerning Smart Meters
Shashi Kant Srivastava

Book Review

- 56 Think Bigger: Developing a Successful Big Data Strategy for your Business
Shweta Grover
- 59 Semi-Organic Growth - Tactics and Strategies behind Google's success
Kakul Modani
- 63 Code Halos: How the Digital Lives of People, Things, and
Organizations are Changing the Rules of Business
Suganya Balakumar