International Journal of Management, IT & Engineering
Vol. 7 Issue 10, October 2017,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

SOCIAL MEDIA ANALYTICS IN HEALTH CARE

Subash Thota [*]

Abstract

Recent years have seen expanding enthusiasm for the patient-focused care and call to focus on enhancing the patient experience. In the meantime, a substantial number of patients are utilizing the web to describe and share their experiences. We believe the growing availability of patients' accounts of their care on blogs, social networks, Twitter and hospital review sites presents an intriguing opportunity to advance the patient-centered care agenda and provide a unique quality of health data. Social media has progressed beyond being a tool for sharing private lives like pictures, videos, and messages especially by young individuals to fostering serious and useful discussion on technology and business. On the other hand, the need for 'Get it right the first time' and minimize costs remain a major concern in the healthcare industry. Customer feedback thus plays a vital role in growing business. In this article, we review the importance of social media analytics in improving business in the healthcare industry.

With the advent of social media into healthcare discussions, it has emerged as a vital source of information which if analyzed could unleash new insights to enhance Health Care. With the emergence of Big Data analytics, trend and predictive analysis have garnered useful business insights to a wide range of industries. One of the significant challenges, researchers trying to address is the interoperability among patient and their records. Here, in this article, we attempt to illustrate the hurdles and various possibilities of Social Media and how streaming data from APIs can be subjected to Big Data analytics in the field of Health Care.

Key Words and Phrases

Data Analytics, Health Analytics, Social Analytics, Data Management, Information Quality, Data Mitigation, Metadata, Data Profiling

[*] **Data Architect**

Introduction

In general, consumers obtain information on any product through advertisements in media (could be print, electronic and social), salespersons, family, friends, neighbors, and acquaintances. They perceive little difference between the similar products from different brands. In the traditional contract-centered world, consumers are routed to agents based on their perceived business value, purchase history, and status. Thanks to the advancement of social media, most of the buyers today have access to a more trusted source of information in the form of social networking sites, where customers share their authentic experience which is then accessible to a larger audience.

A recent study on US Hospitals shows that larger hospitals (having > 200 beds) are early adopters of social media. The study also highlights both Facebook and Twitter are equally popular with 30% and 28% followed by Youtube and LinkedIn with 19% and 18%, blogs are distant from these channels with 5% popularity only. These percentages are computed among the hospitals who adopted social media as a key factor in their success and growth.

Social media strategy

It is an approach that helps enterprises identify the key factors to be monitored and measured to its success. Social listening and Business Analysis are the key areas of this process.

Social Listening

Social listening is the key essence in the social media analytics as there is a dedicated system (like Google analytics) available to monitor the social media.

Social listening gives key insights about

- Loyalcustomers
- Customercomplaints
- Product feedback andanalysis
- Competitoranalysis
- Positive or negative sentiments about thebrand

Health Care Data is a systemized collection of the comprehensive history of a patient's health information in a digitized format, which can be updated, shared and accessed widely among all healthcare providers and organizations. US Government mandates the use of Health Care Data and strives to foster consistent measures taken among all stakeholders, physicians and healthcare providers in improving the quality of the data captured. A term ‘meaningful use' was coined to state improving overall quality, safety, the efficiency of treating healthcare ailments with extensive coverage of wide patient's health information.

Major Challenges in Health Care Data - Data Capture and Interoperability

Improvements to health care data have come a long way since its inception, and it requires huge improvisations concerning the quality of data captured, sourcing the data from various sources viz., patient's health record, medication history, details of lab tests that have been recorded in each of his hospital visits/stay. Interoperability is a stumbling block to sharing healthcare data among different healthcare entities. Providers use different healthcare data software, and the problems arise because of the variation among the healthcare data vendors in the underlying architecture. There is no centralized mechanism that can be used to consolidate and query healthcare data records among different vendors. There are radical changes concerning health care data's extensibility, i.e., providing support for healthcare data software to mobile applications and measures to formulate a means of communication among different healthcare data systems.

Role of Big Data Analytics in Health Care Data

There is a plethora of room for data analytics and trend analysis with these huge chunks of data. The challenge exists in the structured storage medium that is being used to capture healthcare data. Healthcare industry is purely agile, and huge chunks of heterogeneous data need to be captured across all stages of a patient's treatment and instances that may be extremely beneficial. There is a huge potential of identifying useful, tangible information through these uncovered sources of information. It is essential that these unstructured data chunks need to be captured, queried and mined across different hospitals, clinics from all geographical areas and a new set of useful data can be traced and incorporated as part of healthcare data. Healthcare data records of every individual are to be validated against the agile, dynamic raw set of data, whose source can be personalized health record of individuals, data captured as part of patient hospitalization. i.e., Discharge summary, analysis, medications, symptoms, post-treatment recovery, etc.

The forerunner of the problems for sluggish growth of data analytics is the limitations concerning information exchange, data interoperability and an automated mechanism to do Big data analytics on varied sources of data. Due to the need to secure patient health information, various healthcare entities viz., physicians, drug analysts, data scientists hesitate to share information amongst them. Currently, Big data analytics is limited to querying data from external sources, healthcare data records of the patients within a hospital or a small network of hospitals and research are going on to adopt machine learning streaming. Practices and improvements in health care data efficiency are discussed and shared only during periodic meetings between different healthcare participants.

Role of Key Health Care Data Vendors In US

Although there are many providers that facilitates capturing of hospitalization, medication, patient history, symptoms,andobservationsdigitally,thereareessentiallyfewmajorplayersplayingakeyroleinHealth CareData software. Most of the hospitals either use one of these top vendors. Proven track record based on experience, easy user interface and structured workflow to capture details are the reasons for popularity. Storage and retrieval of patient information, compatibility with the Health Care Data format produced by other hospitals in the same geographical region are major factors, for a healthcare provider to choose a Health Care Data vendor. With anever-burgeoningnumberofpatientswiththeplethoraofcomplaints,atvariousstagesoftreatment,hospitals tend to capture critical data in Health Care Data as customary. Health care providers strive to uncover the huge hidden insights that could be tapped out of these huge chunks of raw data is converted into data of ‘meaningful use' as described by the Government. Top healthcare data service providers are AdvancedMD, Allscripts, Practice Fusion, eClinicalWorks, AthenaHealth, etc., Most of these providers save form data in XML/ JSON format and provide REST APIs that helps developers to customize unique solutions. These vendors are analyzing the possibility of sharing Health Care Data over hospitals, researchers, drug dealers, healthcare authorities, etc., for various purposes by opening APIs to the public.

Improvisations in Health Care Data Software

The benefits of Health Care Data, both at the macro and micro level is huge and it ranges from research, patient safety, large-scale statistical analysis over a geographical area, faster diagnosis, easy sharing across different healthcare practitioners, etc., Each Health Care Data vendor have its research group that liaises with hospitals, patients, gathers market trends and feedback forum from the end users to continually improvise the software overall. This improvisation usually includes correction of software bugs, more customized workflow for easy capturing of data and pre-populated values for different fields that facilitate an end user to choose from. All vendors adopt these improvisations continually to stay agile and to improve customer care. The area of improvement that needs immediate consideration is the means of sharing vital information across these vendors that benefits all of them overall.

Challenges in Practical Implementation of Big Data in Health Care Data

The nature of healthcare data itself amounts to huge quantities of data that is unstructured, heterogeneous and not refined. Big data analysis would yield phenomenal useful, actionable information that can be put to patient use. The most root level information that can be extracted from the huge patient recordset could be regression analysis of disease or symptom, identification of repetitive record set, most common keywords, indexing of the parameters to name a few. More deep data analytics could be employed to get deeper, penetrative insights into hidden trends behind huge data that could be of vital value to the business, research, drug dealers, physicians, patients alike. There are many tools, algorithms, and analytics that are in nascent stages of development and are being tested with a minimal set of data. On an initial phase, big data analytics could be used across different patient records that use same Health Care Data software. This has proved effective in many hospitals and physicians in arriving at some micro-level decisions. This is also easier, as the structure of the data is similar because of the same Health Care Data vendor. On a larger scale, the patient records across different hospitals and a geographical area could be queried against to identify needed insights.

A Typical MapReduce Implementation

MapReduce a typical distributed programming model, which acts as a vital tool in Big data analytics, can process large chunks of patient record of a hospital. The records could be in a document format or can be converted to JSON or XML for more efficient processing. Typically, a mapper parses the raw input data and retrieves useful key- value pairs that could be refined. The reducer further refines the key-value pairs by eliminating duplicate values and redundant data, generating a more refined result set. To quote a more practical example, assume a hospital attempting to find useful trends of Diabetes medication. To arrive at a meaningful insight, a typical MapReduce framework would be to search for the occurrence of keyword 'Diabetes' and compute the number of occurrences.

Mapper

Parses the input data and looks for 'diabetes' related field by a key value based searching or through indexing. The

Health Care Data entries are converted into useful key-value pairs and passed onto reducer.

Reducer

A reducer process the structured data from the mapper and count occurrences of each field, eliminate repeated entries and stores unique set of parameters related to 'Diabetes' that are more refined. Adopting Big Data analytics on a larger scale that transcends hospitals, geographical area, and different Health Care Data software is a very ambitious, futuristic challenge which is continually evolving. Some typical analytics over a large hospital network base would involve parallel, distributed processing of Health Care Data records and map data as structured,usefulkey-valuepairsineachofthehospital.Thisistobeconsolidatedwithotherrecordsetsobtained

from various other hospital networks. The structured record set to be further processed for the elimination of redundancy and remove data that do not add much value.

This comprises huge technical and process level impediments,

- Hospitalsmayhavestrictrulesandethicstosecurepatient detailsthatpreventthemfromsharingdata across other hospitalnetworks.

- Authorization- Enforcingsecuredtransactionofdataandensuringthattheauthorizedpersonisaccessing data is a huge challenge.
- Hospitalsandhealthcareprofessionalsmayfearsharingofvital,refinedinformationmayprove beneficial toothercompetitorsregardingbusinessinsightsandlossofitsvitalbusinessprospects.
- Technically, input data could be of any data type/format depending on the Health Care Data software used.Itcouldvarylargelyregardingvolumeandcomplexity.Differenthospitalswouldneedresultsofthe analysis in a different format to process itfurther.

Big data analytics on a large scale could be practically viable only if these impediments are overcome and only if legal and policy barriers of different healthcare organization are addressed.

Role of Social Media in Health Care Data Improvisation

There are a plethora of sources for capturing of Health Care Data and to make it inclusive; social media acts as the last mile connectivity to end users viz., patients, healthcare professionals, and drug researchers alike. There is a different approach of viewing social media. It could be viewed as a centralized platform for people across different areas and walks of life express, share their concerns, insights, and views. In a scenario where hospitals, physicians, researchers are striving to find more data on ‘meaningful use,' social media is a boon as it contains useful insights and consolidated views. It can be viewed as one large repository of consolidated views, experiences, feedback, and advice. With intelligent data mining, Intelligence Business insights, vital Health Care Data parameter could be obtained. The potential sources of copious, online information are physician's blog, patient support groups, and health discussion forums, internal closed group of professionals sharing their insights and findings, drug researchers and dealers sharing their insights among authenticated list of closed groups, etc., A major impediment is not to violate Health Information Privacy and Accountability Act (HIPAA). The key players in social media must take due care in divulging critical patient details. The details shared is to be generic and to be posted to a closed network of authorized people. For instance, a closed group of Facebook members. There must be a centralized system or strategy that prevents exploitation of this vital information other than research.

Standard Gateway of Information-API

An API is typically a software-to-software interface. Interoperability is a major stumbling block across Health Care Data software vendors. Many Health Care Data vendors have not extended their API to be used by other vendors or extended by other mobile apps. Policy and concern for security are a major reason for this. Developers should design software and API in such a way that it is foolproof and immune to security threats. Opening Health Care Data vendor's API to a public or a set of trusted external sources would mean different vendors who cater to physicians, researchers, and patients can develop applications that communicate with the Health Care Data software and accesses vital information that can be used. Opening API would also mean Health Care Data software could benefit by communicating with other APIs.

A Business Model that Nurtures API Sharing

We propose a business model that coordinates and share Health Care Data insights among its peers by open APIs. Physicians, drug researchers, patients use one or other software module on their end. We propose an application that is centralized and monitors API communication between different entities, thereby making real-time streaming API content as the essential source of Data analytics. Monitoring API usage and guarding it against any security threats is vital in ensuring safety and criticality of data are protected against exploitation. Unlike social networking sites, healthcare information is critical, and it is the sole accountability of the healthcare provider against any misuse. Opening APIs to the public has a huge potential risk of exposing this critical data and may lead

to a violation of HIPAA (Health Insurance Portability and Accountability Act). Health Care Data vendors, healthcare providers, physicians, patients, drug researchers, social networking portals can have an agreement of exposing their APIs through an authenticated application that is centralized and acts as a bridge to social media APIs.

It is essentially a paradigm shift, as it nurtures an API to API communication which is streamlined, continually integrated, mined along with data streams of other vendors. Thus, the insight of formulating an inclusive, more detailed capture of 'meaningful data' as enforced by HITECH act, can happen in a coordinated, inclusive approach rather than disjointed analytics that happen within an organization and which are oblivious to other improvements, enhancements that happen in other organizations. We propose to create a secure, sustainable framework where all Health Care Data vendors design their APIs in a modular, extensible way that is secure and immune to virus, network or other attacks. These APIs can communicate with one another through secured, encrypted API key. Leveraging social media in this context is more beneficial and vital. Health Care Data could potentially benefit from Facebook, Twitter APIs in identifying new parameters of ‘meaningful use' to be captured. A centralized application acts as a Bridge to social media APIs and Health Care Data vendors who have to open their APIs.

Processing of Health Care Data API’s

A typical Health Care Data vendor exhibits its data either in JSON/XML. The Health Care Data widely covers complex data capturing all vital health information of a patient. A Health Care Data vendor, after careful analysis of the requirement, develops tailor-made software to cater to the needs of end users. Thus, we have heterogeneous vendors, exposing XML or JSON when queried against its API. We must perform various analytics and mine the data across all Health Care Data records of each API. Typically, a fault-tolerant, distributed, parallel computation is apt to cater to this need. More precisely, we must adopt a map-reduce framework that categorizes and breaks complexly related data into smaller units, viz., key-value pairs. The key-value pair is the ideal format to capture unstructured data that can be subjected to various Dataanalytics.

A mapper typically breaks data chunks into smaller key-value pairs. A combiner combines the result set of differently distributed mappers across all data nodes based on a user requirement. Reducer module refines the dataset, eliminates redundancy and exhibits data that are more statistical. i.e., it gives a high-level topology of the records present, number of instances of each entry, number of different key-value pair variations, etc., Data exhibited from the reducer, can be directly subjected to stream analytics to uncover various underlying key trends. Health Care Data API analytics by nature is more of real-time, streaming analytics that requires immediate insights and results rather than static batch processing over huge chunks of static data queried over a period. The relevant, structured key-value result sets are obtained by querying Health Care Data API’s in parallel. To be more precise, we are required to parse multiple Health Care Data record set that is either in JSON or XML over multiple APIs.

Storm- Spout and Bolts

Apache Storm is suitable to process large volumes of streaming data that are continuous. Data flow is spontaneous,anditistreatedasdatatuples.AdaptingStormabstractionsandmodelto ourrequirementisthekey to arrive at useful, refined data. The storm has three abstractions, viz., Spout- Source of streams in computation that reads data streams from Health Care Data API's through a queuing agent/Event Hub. A Queuing agent abstracts the complex interoperability involved in obtaining data from Health Care Data API takes care of buffering and the amount of data streams/tuples to be emitted to the Spout depending on the number of active storm clusters and the workload on each of them. Customizations could vary based on a variety of factors viz., the capacity of the data node, number of nodes available to process the streams and the need for low latency. Data queried against Health Care Data API's can be considered a stream of tuples that needs to be parsed and broken into useful key-value pairs as per ourneeds.

Essentially this would be the underlying functionality of the Spout module. Depending on the number of data nodes, multiple tuples queried from multiple APIs can be processed simultaneously. A bolt module is responsible for carrying out the business functionality through filtering, aggregation, querying to arrive at a refined output stream. Topology defines the network of spouts and bolts that could vary based on the business logic. Complex business requirements would need spout emitting tuples to multiple bolts in parallel, and each bolt emits its output stream to another bolt sequentially. The spouts and bolts may bore a many to many relationships to one another.

Adopting YARN

Latest enhancements in YARN data model would prove effective in parallel streaming and processing of iterative, real-time streaming of existing Health Care Data entries. YARN facilitates multiple access to Storm clusters and can perform batch, iterative and real-time processing in parallel. This means we can do a streaming analytics to query a Health Care Data API, all the while running a parallel iterative comparison of the Health Care Data. The Storm environment with a queuing system, spout and bolts essentially perform map-reduce routine and arrive at refined, result set without redundancy. We can adopt a comparative study of records with the same key but different values and choose a useful value that is more descriptive.

Social Media API as a Potential Source of Health Care Data Enhancements

In real-time practical application, the requirement to do predictive analysis/trend analysis will typically be agile and the requirement would be dynamic. We propose to havea fault tolerant, robustsystemthatqueriesagainstHealth Care Data APIs and process huge streams of data to arrive at a more refined output stream that could be subjected to further data mining analytics. The useful, identified patterns can be shared with close group of authenticated professionals, patients, researchers scattered over various social media portals viz., Facebook, LinkedIn, Twitter, subscribed blogs etc., this will trigger an informative, healthy debate over the findings and some valuable insights that could benefit Health Care Data enhancements in return. Health Care Data is gradually enhanced to capture more detailed, in-depth coverage of patient medical record that is inclusive and efficient in arriving at a faster diagnosis/solution. It's an improvement regarding quality as well as quantity. This enhancement happens over periodic meetups and reviews of potential parameters that could be included in the Health Care Data among physicians, drug manufacturers, pharma giants, government agents and SMEs. This periodic meeting is transformedtoonlinediscussionsduetothepenetrativelastmilereachofdigitaltechnology.

Leveraging Social Media APIs with Storm: Facebook's graph API provides interactive, in-depth coverage that aides in accessing all information viz., status, comments, likes, user details, etc., after a foolproof authentication process. It is effective in accessing and posting data over a REST channel. We propose a framework that acts as an application that can be categorized into different modules viz., a storm network with a spout-bolt topology and a queuing system to provide communication with Health Care Data and social media APIs. Facebook could authenticate this application, share an encrypted access key and allow it to access, POST content through its REST APIs to the closed group or discussion forum whose members are patients, physicians, drug researchers alike.

Querying Facebook Group Page for Useful Insights

There are great hidden insights in each of the Facebook pages, which if carefully analyzed, can be transformed into useful business insights. Likes, statuses, comments, posts shared all could be transcended into critical trend analysis. For instance, a patient posting a status on the symptoms he experienced during a medication could trigger some likes and comments. The number of likes and the comments would mean the number of people who have faced similar experiences. The number of shares of a post would mean it's a true experience and has valuable content which can be queried for insights. Any insightful post by a physician or SME could spark a series of comments that support and contradict with the views of the post. These could be transpired into a plethora of trend analysis. An intelligent system that behaves in lines of machine learning should be designed that can capture and persist these unstructured data and mine. In the wake of implementing 2-factor authentication, authorizing application is more secure, and any possible leakage of potentially vital information can be avoided. Along with authenticating the application, Facebook also generates Time-based One Time use Password (TOTP). This means anyvital,healthcareinformationisonlywithinaverifiedFacebookpageanditcanalsobemadetorestrictusersto share content outside of thegroup.

Posting periodic useful insights derived from Health Care Data API to Facebook

Patients, physicians and other people linked to healthcare are always on the lookout for potential information through online portals and blogs. Inconsistencies and discrepancies are hugely unavoidable. From an information seeker's perspective, our approach would prove more effective and the information provided is actual statistical

facts more than assumptions and points of view. As we're attempting to design a model in which Health Care Data enhancements and information provided to users are mutually benefitted, the subjects in which the users interested are identified and can be queried from Health Care Data APIs.

Analyzing Twitter Feeds for Useful Insights

Adopting analytics in twitter feeds pose fewer challenges comparatively to Facebook as the feeds are streamlined and can be easily identified with the hashtag. The application must be authenticated with the encrypted twitter key or through a 2-step authentication process. Critical statistics obtained from Health Care Data APIs can be tweeted to a twitter feed with a hashtag. The followers who are subscribed to the application can re-tweet, comment, mark favorite, etc., Insights can be harnessed by following the comments, number of re-tweets of the hashtag. For enhanced security, only users holding verified account can be allowed to subscribe to the feed.

Future Enhancements

We propose to adopt a model in which Health Care Data applications and social media communications through a third-party API that integrates with the insights garnered from streaming data tuples. Machine Learning can be adapted to trigger a real-time vital data capture on the onset of critical events, i.e. when a patient experiences critical symptoms which if captured could produce valuable insights. Machine Learning could prove more efficient as it triggers analytics only when a critical event occurs. Public cloud clusters are a major security concern for critical data and adopting Storm clusters is more efficient in a cloud environment. Healthcare providers can opt for private internal cloud setup or a hybrid cloud setup in which critical data reside at private cloud, and less critical operations can happen over public cloud. Cloud ecosystem comparatively efficient than on-premise setups as new clusters can be formed and dropped dynamically as the need arises. Adoption of Apache Spark model would further prove efficient for faster data processing. Data structuring and refinement can happen in memory, and processed data can be directly streamed without the need for Storm clusters.

Business Analysis

Customer behavior like their likes, dislikes, concerns, complaints, problems, and positives can be gathered using social media. Analyzing these factors can help enterprises improving their communication and targeting along with their profits. Thus, analyzing and collating social media gives an overview of consumer sentiments, the buzz around the products, etc. To use this data effectively for actual business recommendations there should be certain KPIs and frameworks defined. KPIs those can be measured using social media analytics.

Brand awareness and reputation

KPIs under this category depict how deep the brand has penetrated public. Below are some useful KPIs measured under this category

❖ Shareofvoice=No.ofConversationsmentioningabrand/TotalNo.ofindustryconversations

Sentiment analysis = Positive, Negative or Neutral

- ❖ Social engagement = No. of People who are talking about the product / Total No. of people talking about similarproducts

Marketing program effectiveness

KPIs those depict how effectively the marketing strategy working for the product/brand falls under this category.

- ❖ Social reach = Total No. of followers across allplatforms

Growth = Weekly/Monthly/Quarterly

- ❖ Engagement = (No. of Likes + No. of comments + No. of blogs) / Total No. of published posts (Engagement per socialplatform)

Customer behavior

This category depicts how the customer feels about a brand or a product.

- ❖ Social sentiment = No. of comments for the product / No. of comments for all similarproducts

- ❖ Socialinterestgenerated =No.oflikesandcommentsfortheproduct/No.oflikesandcommentsforall similarproducts

Sales: KPIs that define the impact of social media on sales fall under this category.

- ❖ Socialsales=No.ofSalescomingfromsocialchannel/Totalnumberofsales

Influence and Amplification are the two-advanced metrics that can be measured using social media analytics. Influence

Consumers with large followers/friends in the social media like Twitter and Facebook are the key factors in this category.Aninfluencerwithblogswithlargeaudiencesandfrequentcommentsabouttheproductsonsocial networking sites is more likely to influence others to buy a product. Following performance indicators can be measured using this influencerdata.

- ❖ Share of the influencer’s voice = No. of Brand mentions by the influencer / Total No. of relevantmentions

- ❖ Influencerengagement=Influencer’sawarenessofthebrand/Totalnumberofinfluencers

Social Graph

A social graph depicts each user as a node by connecting by arrows (Online relationships) with other nodes. As more and more users are connected over social media, its social graph can become wider and wider and can potentially be a rich source of information for enterprises. Any node can be selected, and the online relationships between the people can be analyzed to read the information useful for the enterprises. This can be used to identify the super influencers the frequency with which these influencers can comment on the products and services. High density implies closely packed network and the potential to spread the word quickly which sets an ideal target for marketing.

Below concerns can be analyzed using this analysis.

- Key interests, issues in thenetwork
- How strong is thenetwork?
- Isitastronginfluentialnetworkoraretherespecificindividualswhoinfluencetherestofthenetwork?

Conclusion

As mentioned earlier social media is no more meant only for sharing private lives it has become a forum to discuss many things including wellness programs and products. Most of the hospitals, wellness providers are adopting social media as the platform to endorse their brands and products. As the enterprise becomes more dependent on external signals to influence decisions and actions, the ability to ‘speak data' will become a critical skill for anyone who works within the social and adaptive business. The above KPIs are powerful for the enterprise to measure, analyze and interpret data from social media to cater business decisions, to develop the product, to understand the market and to influence consumers.

Healthcare industry has critical data related to patients, physicians, Drug manufacturers. Owing to criticality and legal barriers in exposing data between different providers, analytics on a large scale is a challenge, and it is at the onset of being addressed with increased security mechanisms. Thus, major improvisations concerning Health Care Data, periodic meetings among individual Medicaid and Medical claim providers to discuss the problems/trends will witness a paradigm shift towards automated, intelligent systems that consolidates all provider-specific patient records. On the advent of this change, tapping the potential of APIs and social media will prove largely beneficial.

References

- Cappiello, C., Francalanci, C. and Pernici, B. (2003). Time-related factors of data quality in multi-channel information systems. *Journal of Management Information Systems,* 20 no. 3, 71-91.

- Greengard, S (1998). Don't let dirty data derail you. *Workforce*, 77, no. 11, 107-8.

- Lederman, R., Shanks, G. and Gibbs, M.R. (2003). Meeting privacy obligations: the implications for information systems development.

- Liebenau, J. & Backhouse, J. (1990). *Understanding in Formation: an Introduction*. Palgrave Macmillan.

- Marsh, R. (2005). Drowning in dirty data? It’s time to sink or swim: A four-stage methodology for total data quality management. *Database Marketing & Customer Strategy Management*, 12, no. 2, 105-112. http://link.springer.com/article/10.1057/palgrave.dbm.3240247

- Sellar, S. (1999). Dust off that data. *Sales and Marketing Management*,151, no. 5, 71-3.

- Thota, S., 2017. Big Data Quality. *Encyclopedia of Big Data*, pp.1-5. https://link.springer.com/referenceworkentry/10.1007/978-3-319-32001-4_240-1

- Vayghan, J. A., Garfinkle, S. M., Walenta, C., Healy, D.C., & Valentin, Z. (2007). The internal information transformation of IBM. *IBM Systems Journal*, 46 no. 4, 669-684.

ABOUT THE AUTHOR

- Subash Thota works as Data Architect and specializes in Big Data, Cloud, Data Integration and Data Analytics with significantexperienceinProjectManagement,Agile,andDataGovernance.Subashhaswrittenseveralpapersinthe field of Big Data, the Cloud, andAnalytics.