© 2017 IJSRSET | Volume 3 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Preserving Security and Privacy in Big Healthcare Data

Dr. V. Goutham

Professor and HOD of CSE in Teegla Krishna Reddy Engineering College, Telangana, India

ABSTRACT

Big knowledge has essentially modified the manner organizations manage, analyze and leverage knowledge in any trade. one amongst the foremost promising fields wherever huge knowledge will be applied to create a modification is attention. huge attention knowledge has extended potential to boost patient outcomes, predict outbreaks of epidemics, gain valuable insights, avoid preventable diseases, cut back the value of supplying and improve the standard of life generally. However, preferring the allowable uses of knowledge whereas conserving security and patient's right to privacy may be a tough task. Big data, in spite of however helpful for the advancement of life science and important to the success of all attention organizations, will solely be used if security and privacy problems square measure self-addressed. to confirm a secure and trustworthy huge knowledge atmosphere, it's essential to spot the constraints of existing solutions and envision directions for future analysis. during this paper, we've got surveyed the progressive security and privacy challenges in huge knowledge as applied to attention trade, assessed however security and privacy problems occur just in case of huge attention knowledge and mentioned ways that within which they'll be self-addressed. we tend to primarily targeted on the recently planned ways supported anonymization and coding, compared their strengths and limitations, and unreal future analysis directions.

Keywords: Security and privacy, Big healthcare data, Security lifecycle, Anonymization, Encryption

I. INTRODUCTION

Change is that the new norm for the worldwide attention sector. In fact, conversion of health and patient knowledge is undergoing a dramatic and elementary shift within the clinical, operative and business models and usually within the world of economy for the predictable future. This shift is being spurred by aging populations and manner changes; the proliferation of software system applications and mobile devices; innovative treatments; heightened specialise in care quality and value; and evidence-based drugs as against subjective clinical calls—all of that square measure resulting in provide vital opportunities for supporting clinical decision, up supplying, management and political affairs, surveilling malady, observation adverse events, and optimizing treatment for diseases touching multiple organ systems.

As noted higher than, huge knowledge analytics in attention carries several advantages, guarantees and presents nice potential for remodeling attention, however it raises manifold barriers and challenges.

Indeed, the issues over the large attention knowledge security and privacy square measure enhanced year-by-year. in addition, attention organizations found that a reactive, bottom-up, technology-centric approach to determinative security and privacy needs isn't capable defend the organization and its patients.

Motivated therefore, new data systems and approaches square measure required to stop breaches of sensitive data and alternative styles of security incidents therefore on create effective use of the large attention knowledge.

In this paper, we tend to discuss some attention-grabbing connected works and gift risks to the large health knowledge security similarly as some newer technologies to redress these risks. Then, we tend to specialise in the large knowledge privacy issue in attention, by mentioning numerous laws and laws established by completely different restrictive bodies and citing some possible techniques wont to make sure the patient's privacy. Thereafter, we offer some planned techniques and approaches that were reportable within the literature to upset security and privacy risks in attention whereas distinguishing their limitations. Lastly, we provide conclusions and highlight the long run directions.

Big data has fundamentally changed the way organizations manage, analyze and leverage data in any industry. One of the most promising fields where big data can be applied to make a change is healthcare. Big healthcare data has considerable potential to improve patient outcomes, predict outbreaks of epidemics, gain valuable insights, avoid preventable diseases, reduce the cost of healthcare delivery and improve the quality of life in general. However, deciding on the allowable uses of data while preserving security and patient's right to privacy is a difficult task. Big data, no matter how useful for the advancement of medical science and vital to the success of all healthcare organizations, can only be used if security and privacy issues are addressed. To ensure a secure and trustworthy big data environment, it is essential to identify the limitations of existing solutions and envision directions for future research. In this paper, we have surveyed the state-of-the-art security and privacy challenges in big data as applied to healthcare industry, assessed how security and privacy issues occur in case of big healthcare data and discussed ways in which they may be addressed. We mainly focused on the recently proposed methods based on anonymization and encryption, compared their strengths and limitations, and envisioned future research directions.

Successful related works:

Seamless integration of greatly various huge attention knowledge technologies cannot solely change North American country to realize deeper insights into the clinical and structure processes however additionally facilitate quicker and safer output of patients and make bigger efficiencies and facilitate improve patient flow, safety, quality of care and also the overall patient expertise notwithstanding however pricey it's.

Such was the case with South Tyneside NHS Foundation Trust, a supplier of acute and community health services in northeast European nation that understands the importance of providing top quality, safe and compassionate look after the patients in any respect times, however desires a stronger understanding of however its hospitals operate to boost resource allocation and wait times and to confirm that any problems area unit known early and acted upon [4].

Another example is that the UNC Health Care (UNCHC), that could be a non-profit integrated attention system in North geographic region that has enforced a replacement system permitting clinicians to quickly access and analyze unstructured patient knowledge victimisation natural-language process. In fact, UNCHC has accessed and analyzed huge quantities of unstructured content contained in patient medical records to extract insights and predictors of admittance risk for timely intervention, providing safer look after speculative patients and reducing re-admissions [5].

Moreover within the us, the Indiana Health info Exchange, that could be a non-profit organization, provides a secure and strong technology network of

health info linking over ninety hospitals, community health clinics, rehabilitation centres and alternative attention suppliers in Indiana. It permits medical info to follow the patient hosted in one doctor workplace or solely in an exceedingly hospital system [6].

One more example is Kaiser Permanente medical network based mostly in Calif.. it's over nine million members, calculable to manage massive volumes of knowledge starting from twenty six.5 Petabytes to forty four Petabytes. [7].

Big knowledge analytics is employed additionally in Canada, e.g. the baby hospital of Toronto. This hospital succeeded to boost the outcomes for newborns at risk of serious hospital infections. Another example is that the Greek deity project, that could be a newborns observation platform designed mercy to a collaboration between IBM and also the Institute of Technology of Ontario. It supported the acquisition and also the storage of patients' physiological knowledge and clinical data system knowledge for the target of on-line and real time analysis, retrospective analysis, and data processing [8].

In Europe and specifically in Italia, the Italian medicines agency collects and analyzes an oversized quantity of clinical knowledge regarding overpriced new medicines as a part of a national gain program. supported the results, it should evaluate the medicines costs and market access terms [9].

In the domain of mHealth, the planet Health Organization has launched the project "Be Healthy Be mobile" in African nation and underneath the mDiabetes initiative it supports countries to line up large-scale comes that use mobile technology, above all text electronic communication and apps, to control, forestall and manage non-communicable diseases like polygenic disorder, cancer and cardiovascular disease [10].

Privacy and security concerns in big data:

Security and privacy in huge knowledge area unit vital problems. Privacy is commonly outlined as having the flexibility to shield sensitive info regarding in person classifiable health care info. It focuses on the utilization and governance of individual's personal knowledge like creating policies and establishing authorization needs to confirm that patients' personal info is being collected, shared and utilised in right ways in which. whereas security is usually outlined because the protection against unauthorized access, with some as well as express mention of integrity and accessibility. It focuses on protective knowledge from pernicious attacks and stealing knowledge for profit. though security is significant for shielding knowledge however it's scant for addressing privacy. Table one focuses on extra distinction between security and privacy.

Table 1. Differentiation between security and privacy

Security	Privacy
Security is the "confidentiality, integrity and availability" of data	Privacy is the appropriate use of user's information
Various techniques like Encryption, Firewall, etc. are used in order to prevent data compromise from technology or vulnerabilities in the network of an organization	The organization can't sell its patient/user's information to a third party without prior consent of the user
It may provide for confidentiality or protect an enterprise or agency	It concerns with patient's right to safeguard their information from any other parties
Security offers the ability to be confident that	Privacy is the ability to decide what information of

Security	Privacy
decisions are respected	an individual goes and where to

Security of big healthcare data:

While attention organizations store, maintain and transmit large amounts of information to support the delivery of economical and correct care, the downsides area unit the shortage of technical support and borderline security. Complicating matters, the attention trade continues to be one among the foremost prone to publically disclosed knowledge breaches. In fact, attackers will use data processing strategies and procedures to search out out sensitive knowledge and unleash it to the general public and so knowledge breach happens. Whereas implementing security measures remains a posh method, the stakes area unit regularly raised because the ways in which to defeat security controls become a lot of refined.

Accordingly, it's important that organizations implement attention knowledge security solutions that may shield necessary assets whereas conjointly satisfying attention compliance mandates.

II. BIG DATA SECURITY LIFECYCLE

Data collection phase:

This is the apparent initiative. It involves collection knowledge from completely different sources in varied formats. From a security perspective, securing huge health knowledge technology may be a necessary demand from the primary part of the lifecycle. Therefore, it's necessary to collect knowledge from sure sources, preserve patient privacy (there should be no arrange to determine the individual patients within the database) and check that that this part is secured and guarded. Indeed, some mature security measures should be wont to make sure that all knowledge and knowledge systems area unit protected against unauthorized access, disclosure, modification, duplication, diversion, destruction, loss, misuse or felony.

Data transformation phase

Once the info is out there, the primary step is to filter and classify the info supported their structure and do any necessary transformations so as to perform important analysis. a lot of generally, knowledge filtering, enrichment and transformation area unit required to enhance the standard of the info previous analytics or modeling part and take away or befittingly modify noise, outliers, missing values, duplicate knowledge instances, etc. On the opposite facet, the collected knowledge could contain sensitive data, that makes very necessary to require spare precautions throughout knowledge transformation and storing. so as to ensure the protection of the collected knowledge, the info ought to stay isolated and guarded by maintaining access-level security and access management (utilizing an intensive list of directories and databases as a central repository for user credentials, application logon templates, countersign policies and shopper settings), and process some security measures like knowledge anonymization approach, permutation, and knowledge partitioning.

Data modeling phase:

Once the info has been collected, remodeled and hold on in secured storage solutions, the info process analysis is performed to come up with helpful information. during this part, supervised data processing techniques like agglomeration, classification, and association may be used for feature choice and prophetical modeling. Further, there conjointly exist many ensembles of learning techniques that improve accuracy and hardiness of the ultimate model. On the opposite facet, it's crucial to produce secure process atmosphere. In fact, the main target {of knowledge|of knowledge|of information} miners during this part is to use

powerful data processing algorithms that may extract sensitive data. Therefore, the method of information mining and therefore the network elements normally, should be designed and guarded against data processing based mostly attacks and any security breach which will happen, furthermore as check that that solely approved employees add this part. This method helps eliminate some vulnerabilities and mitigates others to a lower risk level.

Knowledge creation phase:

Finally, the modeling part comes up with new data and valued knowledges to be utilized by call manufacturers. These created knowledges area unit thought-about sensitive knowledge, particularly in an exceedingly competitive atmosphere. Indeed, attention organizations tuned in to their sensitive knowledge (e.g. patient personal data) to not be publically discharged. consequently, security compliance and verification area unit a primary objective during this part.

At all stages of huge knowledge lifecycle, it needs knowledge storage, knowledge integrity and knowledge access management.

Technologies in use

Various technologies are in use to make sure security and privacy of massive attention knowledge. most generally used technologies are:

1) Authentication:

Authentication is that the act of building or confirming claims created by or regarding the topic are true and authentic. It serves important functions inside any organization: securing access to company networks, protective the identities of users, and guaranteeing that the user is actually UN agency he's pretence to be.

The information authentication will cause special issues, particularly man-in-the-middle (MITM) attacks. Most cryptologic protocols embody some type of end authentication specifically to forestall MITM attacks.

2) Encryption:

Data encryption is associate economical means that of preventing unauthorized access of sensitive knowledge. Its solutions defend and maintain possession of information throughout its lifecycle—from the information centre to the end (including mobile devices utilized by physicians, clinicians, and administrators) and into the cloud. coding is helpful to avoid exposure to breaches like packet sniffing and stealing of storage devices.

Healthcare organizations or suppliers should make sure that coding theme is economical, straightforward to use by each patients and care professionals, and simply protractible to incorporate new electronic health records. what is more, the amount of keys hold by every party ought to be decreased .

Although varied coding algorithms are developed and deployed comparatively well (RSA, Rijndael, AES and RC6, DES, 3DES, RC4, IDEA, Blowfish …), the right choice of appropriate coding algorithms to enforce secure storage remains a troublesome downside.

3) Data masking:

Masking replaces sensitive knowledge components with associate elusive price. it's not really associate coding technique that the original price cannot be came from the covert price. It uses a technique of de-identifying knowledge sets or masking personal identifiers like name, social insurance range and suppressing or generalizing quasi-identifiers like date-of-birth and zip-codes.

A significant advantage of this method is that the value of securing a giant knowledge preparation is reduced. As secure knowledge is migrated from a secure supply into the platform, masking reduces the

necessity for applying extra security controls thereon knowledge whereas it resides within the platform.

4) Access control:

Once genuine , the users will enter associate data system however their access can still be ruled by associate access management policy that is often supported privileges and rights of every professional licensed by patient or a trustworthy third party. It is then, a robust and versatile mechanism to grant permissions for users. It provides refined authorization controls to confirm that users will perform solely the activities that they need permissions, like knowledge access, job submission, cluster administration, etc.

A number of solutions are planned to deal with the safety and access management issues. Role-based access management (RBAC) and attribute-based access management (ABAC) area unit the foremost well-liked models for EHR. RBAC and ABAC have shown some limitations after they area unit used alone in medical system. To satisfy necessities of fine-grained access management however security and privacy conserving, we recommend adopting technologies in conjunction with different security techniques, e.g. encryption, and access management strategies.

5) Monitoring and auditing:

Security watching is gathering and work network events to catch the intrusions. Audit means that recording user activities of the care system in written account order, like maintaining a log of each access to and modification of information. These area unit 2 facultative security metrics to live and make sure the safety of a care system.

Intrusion detection and hindrance procedures on the total network traffic is kind of difficult. to deal with this downside, a security watching design has been developed via analyzing DNS traffic, information processing flow records, hypertext transfer protocol traffic and honeypot data. The steered resolution includes storing and process knowledge in distributed sources through knowledge correlation schemes. At this stage, 3 probability metrics are calculated to spot whether or not name, packet or flow is malicious. counting on the score obtained through this calculation, associate alert happens in detection system or method terminate by hindrance system. in keeping with performance analysis with open supply huge knowledge platforms on electronic payment activities of an organization knowledge, Spark and Shark manufacture quick and steady results than Hadoop, Hive and Pig.

Privacy of massive health care data:

The invasion of patient privacy is taken into account as a growing concern within the domain of massive knowledge analytics as a result of the emergence of advanced persistent threats and targeted attacks against data systems. As a result, organizations area unit in challenge to deal with these completely different complementary and significant problems. a happening reportable within the Forbes magazine raises associate alarm over patient privacy [42]. within the report, it mentioned that focus on Corporation sent baby care coupons to a teen-age lady unbeknown to her oldsters. This incident impels analytics and developers to think about privacy in huge knowledge. they must be able to verify that their applications adapt to privacy agreements which sensitive data is unbroken non-public no matter changes in applications and/or privacy rules.

Privacy of medical knowledge is then a crucial issue that should be seriously thought-about. we tend to cite within the next paragraph a number of laws on the privacy protection worldwide.

Data protection laws:

More than ever it's crucial that care organizations manage and safeguard personal data and address their risks and legal responsibilities in regard to process personal knowledge, to deal with the growing brush of applicable knowledge protection legislation.

A. De-identification

De-identification may be a ancient technique to ban the revealing of confidential data by rejecting any information that may establish the patient, either by the primary technique that needs the removal of specific identifiers of the patient or by the second method wherever the patient verifies himself that enough identifiers area unit deleted. notwithstanding, associate offender will probably get additional external data help for de-identification in huge knowledge. As a result, de-identification isn't ample for safeguarding huge knowledge privacy. It can be additional possible through developing economical privacy-preserving algorithms to assist mitigate the chance of re-identification. The ideas of k-anonymity, l-diversity and t-closeness are introduced to reinforce this ancient technique.

- k-anonymity during this technique, the upper the worth of k, the lower are the likelihood of re-identification. However, it should result in distortions of data and therefore bigger information loss as a result of k-anonymization. what is more, excessive anonymization will create the disclosed knowledge less helpful to the recipients as a result of a number of the analysis becomes not possible or could manufacture biased and incorrect results. In k-anonymization, if the quasi-identifiers containing knowledge area unit wont to link with different in public obtainable knowledge to spot people, then the sensitive attribute (like disease) in concert of the symbol are unconcealed.

  Various measures are planned to quantify data loss caused by anonymization, however they are doing not replicate the particular quality of information. Therefore, we tend to move towards L-diversity strategy of information anonymization.

- L-diversity it's a style of cluster primarily based anonymization that's utilised to safeguard privacy in knowledge sets by decreasing the roughness of information illustration. This model (Distinct, Entropy, Recursive) is associate extension of the k-anonymity that utilizes strategies as well as generalization and suppression to scale back the roughness of knowledge illustration in an exceedingly manner that any given record maps onto a minimum of k completely different records within the data.

- T-closeness could be a more improvement of l-diversity cluster based mostly anonymization.

The t-closeness model (equal/hierarchical distance) [46, 50] extends the l-diversity model by treating the values of AN attribute clearly, taking into consideration the distribution of knowledge values for that attribute. the most advantage of this method is that it intercepts attribute speech act, and its drawback is that as size and style of information increase, the chances of re-identification increase too.

HybrEx

Hybrid execution model [55] could be a model for confidentiality and privacy in cloud computing. It utilizes public clouds just for AN organization's non-sensitive information and computation classified as public, i.e., once the organization declares that there's no privacy and confidentiality risk in mercantilism the information and activity computation thereon victimization public clouds, whereas for AN organization's sensitive, personal information and computation, the model executes their personal cloud. Moreover, once AN application needs access to each the personal and public information, the appliance itself conjointly gets divided and runs in each the personal and public clouds. It considers information sensitivity before a job's execution and provides integration with safety.

The four classes during which HybrEx MapReduce permits new forms of applications that utilize each public and personal clouds square measure as shown in Figure 2.

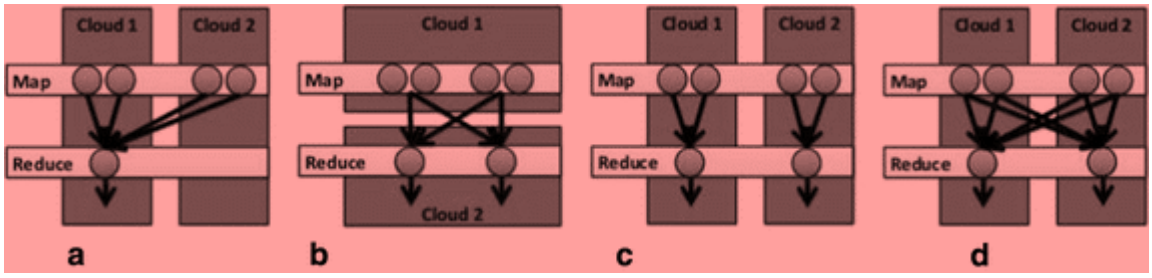

Figure 1.

The four Execution categories for HybrEx MapReduce.

a. Map hybrid.
b. Horizontal partitioning.
c. Vertical partitioning.
d. Hybrid

Map hybrid (1a) The map phase is executed in both the public and the private clouds while the reduce phase is executed in only one of the clouds.

Vertical partitioning (1b) Map and cut back tasks square measure dead within the public cloud exploitation public knowledge because the input, shuffle intermediate knowledge amongst them, and store the lead to the general public cloud. an equivalent work is completed within the personal cloud with personal knowledge. the roles square measure processed in isolation.

Horizontal partitioning (1c) The map part is dead solely publically clouds, whereas the cut back part is dead during a personal cloud.

Hybrid (1d) The map part and also the cut back part square measure dead on each public and personal clouds. knowledge transmission among the clouds is additionally potential.

The problem with HybridEx is that it doesn't cope with the key that's generated at public and personal clouds within the map part which it deals solely with cloud as AN someone.

C. Identity based mostly anonymization

It is a kind of data sanitisation whose intent is privacy protection. it's the method of either encrypting or removing in person diagnosable data from knowledge sets, in order that the folks whom the information describe stay anonymous. the most problem with this method involves combining anonymization, privacy protection, and massive knowledge techniques to research usage knowledge whereas protective the identities.

Intel Human Factors Engineering team required to safeguard Intel employees' privacy exploitation website access logs and massive knowledge tools to reinforce convenience of Intel's heavily used internal internet portal. They were needed to get rid of in person distinguishing data (PII) from the portal's usage log repository however during a means that didn't influence the employment of massive knowledge tools to try to to analysis or the flexibility to re-identify a log entry so as to research uncommon behavior.

To meet the numerous edges of Cloud storage, Intel created AN open design for anonymization that allowed a spread of tools to be utilised for each de-

identifying and re-identifying journal records. within the implementing design method, enterprise knowledge has properties completely different from the quality examples in anonymization literature [58]. Intel conjointly found that in spite of masking obvious Personal Identification data like usernames and IP addresses, the anonymized knowledge was defenseless against correlation attacks. once exploring the tradeoffs of correcting these vulnerabilities, they found that User Agent data powerfully correlates to individual users. this is often a case study of ANonymization implementation in an enterprise, describing needs, implementation, and experiences encountered once utilizing anonymization to safeguard privacy in enterprise knowledge analyzed exploitation massive knowledge techniques. This investigation of the standard of anonymization used k-anonymity based mostly metrics. Intel used Hadoop to research the anonymized knowledge and acquire valuable results for the Human Factors analyst. At an equivalent time, it learned that anonymization must be quite merely masking or generalizing bound fields—anonymized datasets have to be compelled to be fastidiously analyzed to see whether or not they square measure prone to attack.

Summary on recent approaches used in big data privacy

In this paper, we've investigated the protection and privacy challenges in massive knowledge, by discussing some existing approaches and techniques for achieving security and privacy during which tending organizations square measure probably to be extremely useful. during this section, we tend to targeted on citing some approaches and techniques conferred in several papers with stress on their focus and

Summary on recent approaches used in big data privacy

Focus	Limitations
Discusses experiences and issues encountered when successfully combined anonymization, privacy protection, and Big data techniques to analyze usage data while protecting the identities of users	It still uses K-anonymity technique which is vulnerable to correlation attack
Proposed the privacy preserving data mining techniques in Hadoop, i.e. solve privacy violation without utility degradation	Its execution time is affected by noise size
Introduced an efficient and privacy-preserving cosine similarity computing protocol	Need significant research efforts for addressing unique privacy issues in some specific big data analytics
Discussed and suggested how an existing approach “differential privacy” is suitable for big data	This method depends totally on calculation of the amount of noise by the curator. So, if curator is compromised the whole system fails
Proposed a scalable two-phase top-down specialization (TDS) approach to anonymize large-scale data sets using the MapReduce framework on cloud	It uses anonymization technique which is vulnerable to correlation attack
Proposed various privacy issues dealing with big data applications	Customer segmentation and profiling can easily lead to discrimination based on age

Focus	Limitations
	gender, ethnic background, health condition, social, background, and so on
Proposed an anonymization algorithm (FAST) to speed up anonymization of big data streams	Further research required to design and implement FAST in a distributed cloud-based framework in order to gain cloud computation power and achieve high scalability
The novel framework proposed into achieve privacy-preserving machine learning	The training data are distributed and each shared data portion of large volume, is not able to achieve distributed feature selection
Proposed methodology provides data confidentiality, secure data sharing without Re-encryption and access control for malicious insiders and forward and backward access control	Limiting the trust level in the cryptographic server

These accumulated complexness and limits create the new models harder to interpret and their dependability less straightforward to assess, compared to previous models.

III. CONCLUSION

Whereas the potential opportunities offered for giant knowledge within the care square measurena are unlimited (e.g. drive health analysis, data discovery, clinical care, and private health management), there square measure many obstacles that impede its true potential, together with technical challenges, privacy and security problems and adept talent. huge knowledge security and privacy square measure thought-about large obstacles for researchers during this field.

In this paper, we've in short mentioned some made connected work across the planet. we've additionally bestowed privacy and security problems in every part of huge knowledge lifecycle at the side of the benefits and flaws of existing technologies within the context of huge care knowledge privacy and security.

We in the main reviewed the privacy preservation ways that are used recently in care and mentioned however coding and anonymization ways are used for health care knowledge protection likewise as bestowed their limitations. to boot, there square measure a lot of varied techniques embrace concealment a needle in a very rick, Attribute primarily based coding Access management, Homomorphic coding, Storage path coding and then on. However, the matter is often obligatory.

In this context, as our future direction, views consist in achieving effective solutions in privacy and security within the era of huge care knowledge. As well, privacy ways ought to be increased.

Also with the speedy development of IoT, the bigger the amount, the lower the standard. Consequently, quality of information shouldn't be affected a lot of by privacy protective algorithms to urge the acceptable result by researchers. And to travel additional, we are going to try and solve the matter of accommodative security and privacy models by simulating various approaches to ultimately support higher cognitive process and designing methods.

IV. AUTHORS

Dr V. GOUTHAM is a Professor and Head of the Department of Computer Science and Engineering at Teegala Krishna Reddy Engineering College affiliated to J.N.T.U Hyderabad. He received Ph.D. from Acharya Nagarjuna University M.Tech from Andhra University.His research interests are Software Reliability Engineering, software testing, software Metrics, and cloud computing.

V. REFERENCES

[1]. Burghard C. Big data and analytics key to accountable care success. Framingham: IDC Health Insights; 2012.

[2]. Fernandes L, O'Connor M, Weaver V. Big data, bigger outcomes. J AHIMA. 2012;83:38–42.

[3]. David Houlding, MSc, CISSP. Health Information at Risk: Successful Strategies for Healthcare Security and Privacy. Healthcare IT Program Of ce Intel Corporation, white paper. 2011.

[4]. South Tyneside NHS Foundation Trust. Harnessing analytics for strategic planning, operational decision making and end-to-end improvements in patient care. IBM Smarter Planet brief. 2013.

[5]. UNC Health Care relies on analytics to better manage medical data and improve patient care. IBM Press release. 2013.

[6]. Indiana Health Information Exchange. http://www.ihie.org/. Accessed 24 Mar 2016.

[7]. Transforming healthcare through big data, strategies for leveraging big data in the healthcare industry. Institute for Health. 2013.

[8]. Artemis. http://hir.uoit.ca/cms/?q=node/24. Accessed 21 May 2016.

[9]. Groves P, Kayyali B, Knott D, Kuiken SV. The big data revolution in healthcare, accelerating value and innovation. 2013.

[10]. WHO. Mobile phones help people with diabetes to manage fasting and feasting during Ramadan. Features. 2014.

[11]. Sophia Genetics. Product & Technology Overview 2014.

[12]. CynergisTek, Redspin. BREACH REPORT 2016: Protected Health Information (PHI) 2017.

[13]. Podesta J, et al. Big data: seizing opportunities, preserving values. Executive Office of the President. 2014;1:2013.