Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 9, September 2014, pg.564 – 572

RESEARCH ARTICLE

IEHRC - An Approach towards Interoperable E-Health Records over Cloud using ECHISTAR

Ayesha Kounin[1], Mohammed Ali Shaik[2], K.Pranav Kumar[3]

[1]Pursuing M.Tech in CSE & JNTU Hyderabad, India

[2]Department of CSE ARTI, Warangal, Telangana, India

[3]Department of CSE ARTI, Warangal, Telangana, India

[1] ayeshakounin2015@gmail.com, [2] niharali@gmail.com, [3] pranavkumar47@gmail.com

Abstract-- We propose and present a cloud-based approach for designing an interoperable electronic health record (EHR) system implemented in a cloud computing environment that provides several benefits to all the stakeholders or end users among healthcare ecosystem such as patients or providers or payers, etc. Due to lack of data interoperability standards in society proposals of solutions are been considered to be major obstacle in the sharing of healthcare data between different stakeholders. Here in this paper we propose a system that is ECHISTAR: Electronic cloud health information systems technology architecture that achieves electronic semantic interoperability through the use of a generic design methodology that uses a reference model which defines a general purpose set of data structures and architecture type oriented models that define the clinical data attributes. ECHISTAR application components are designed by using the cloud component model approach that consists of loosely coupled components that communicate asynchronously with one another in a brokered fashion by using an approach for electronic semantic interoperability, data integration, and essential security.

Keywords— EHR, data integration, electronic health records, healthcare

I. INTRODUCTION

Electronic ecosystem consists of distinct healthcare providers such as doctors, physicians, consultants, specialists, etc. and distinct areas of end-users or payers such as health insurance companies, pharmaceutical companies, IT recruiter firms, and of course patients. The process includes provision of massive healthcare data that is available in the form of structured data and unstructured data on dissimilar data sources such as relational databases, file servers, flat files, etc. and in various formats. When a patient is admitted to a hospital, his or her information is entered as input to electronic health record (EHR) systems after which the physicians diagnose the patient and the diagnostic information from medical devices such as CT scanners or MRI scanners, etc. are stored in EHR systems. In this process of diagnosis doctors will retrieve the health information of patient and analyze it to diagnose and cure the illness in this process a doctor can take expert advice by sharing the case information with consulting specialists.

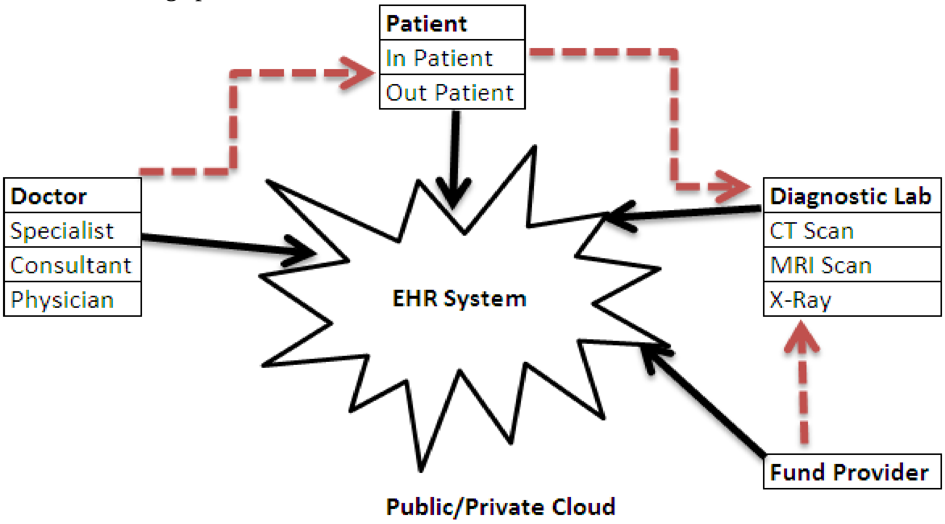

Fig. 1 Cloud computing environment applied for IEHRC.

In the present nurturing trend of cloud computing the environment provides several benefits to all the stakeholders or end users in the healthcare ecosystem by providing electronic health information management system, electronic laboratory information system, electronic radiology information system, electronic pharmacy information system, etc.

With public cloud based EHR systems hospitals do not need to spend a significant portion of their budgets on purchasing or maintenance of IT infrastructure since public cloud service providers provide on-demand provisioning of hardware and software resources with pay-per-use pricing models by which hospitals can save on upfront capital investments in hardware and data center infrastructure by using public cloud-based EHR systems and pay only for the operational expenses of the cloud resources used by them when required.

Hospitals can access patient data stored in public or private cloud and share the data with other hospitals when a doctor wants to view patient history or biological relatives history only after the patient provides access to his or here information stored in the cloud using SaaS applications to hospitals so that the process of admissions, care, and discharge processes can be streamlined.

Physicians can upload diagnostic reports, prescription to the cloud with an intention of accessing such information by doctors remotely for diagnosing the illness in the proposed system we have provided a provision where a patient can manage his or her prescriptions and associated information such as dosage, quantity, and frequency, and provide this information to their healthcare provider due to which health payers can increase the effectiveness of their care management programs by providing value added services and giving access to health information to members.

II. RELATED WORK

The Rajiv Aarogyasri Health Insurance Scheme (RAHIS), Quality Medicare For All[1] is a information systems that is the largest single medical system in Andhra Pradesh and Telengana states of India that caters to a quarter of the nation's below poverty line population as shown in the below Fig.. This scheme helps people belongs to BPL families means the group of individuals as indicated in ,Health Card or ,White ration card (BPL card) and the maximum sum insured per family is Rs.1,50,000.00, where this benefit can be availed by an individual or family and a buffer amount of Rs. 50000.00 is also provided for the needy as a buffer amount.

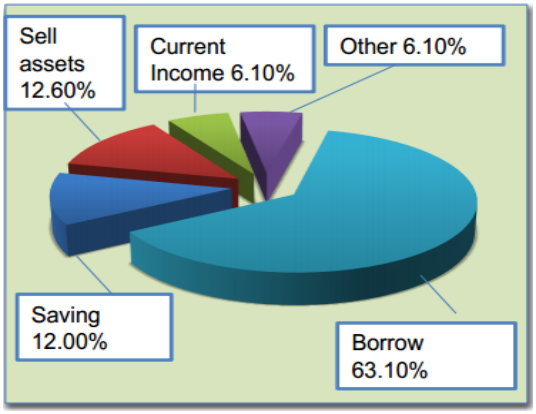

Fig. 2 Illustration of below Poverty line.

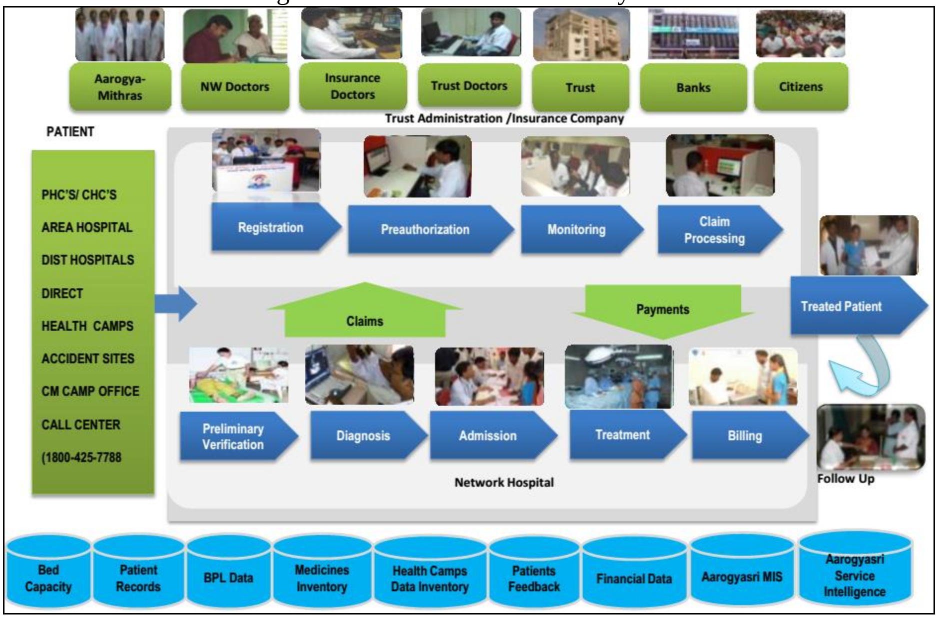

Fig. 3 ECHISTAR implementation on RAHIS.

RAHIS program is designed such that it brings awareness among the rural masses about the diseases they are possessing and provides free medical advice and medicines to the rural people, around 47.45 Lakh people have been screened at their door step by specialist doctors of various network hospitals at 28,008 Health Camps and around 2,97,547 Patients referred from Health Camps to network hospitals 340 network hospitals which are conducting around 984 camps in a month across Telengana and Andhra Pradesh states. The main emphasis is on the Tribal and Backward areas. The working process or RAHIS is depicted in Fig 3.

RAHIS is not a single application, but a collection over 180 application packages/modules.

Traditional EHR systems are based on client– server architectures. The RAHIS front end comprises of applications such as ADT, pharmacy, radiology, etc. The applications communicate with the server through a remote procedure calls (RPC) broker. RAHIS server comprises of RPC Broker, Kernel/Tools (such as TaskMan, package manager, FileMan-provides APIs and utility functions and an object based database management system where each application module generates at least one global data file and contains clinical, administrative, and computer infrastructure related information.

The main underlying technology of most of RAHIS applications is both a procedural programming language and a hierarchical or multidimensional key value database. RAHIS is implemented in two phases phase-I and phase-II.

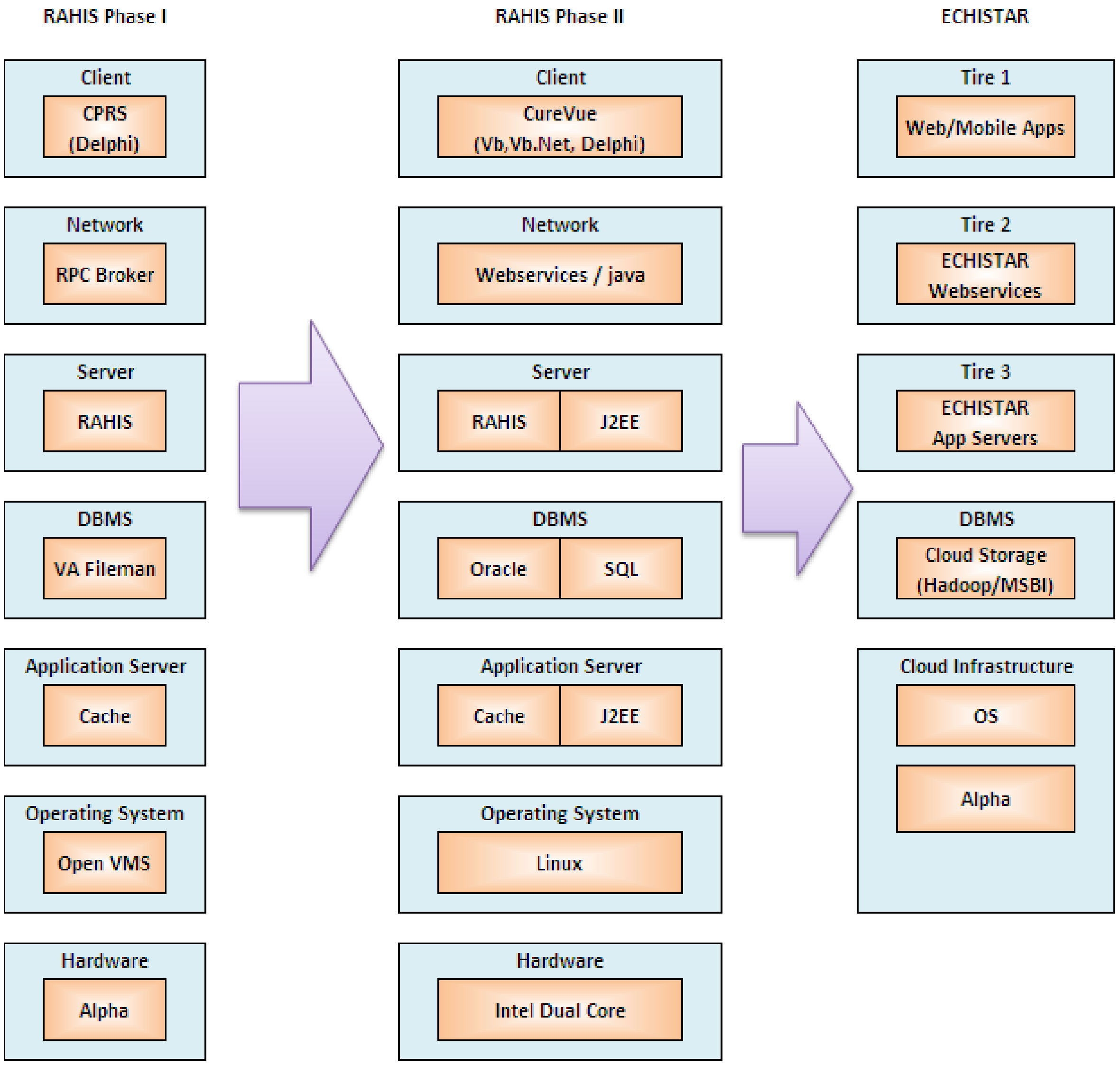

Fig. 4 Technology stacks of RAHIS Phase-I, RAHIS Phase-II, and the proposed ECHISTAR system. Both RAHIS are client–server systems, whereas ECHISTAR is a cloud-based system.

ECHISTAR uses the cloud component model approach for application design described in our previous work for the design of mobile applications that can utilize the capabilities of the next generation of cellular networks.

III. MOTIVATION

In this section, we describe the motivation for EHR based on cloud computing.

A. *Design Methodologies*

Traditional EHR systems have been designed using the methodologies described as follows [3].

1) Unstructured approach: This approach consists of unstructured data.

2) Big model approach: This approach consists of structured data. A separate table is maintained for each clinical concept leading to a large number of tables.

3) Generic approach: This approach allows a wide variety of data to be stored in generic data structures. A constraint mechanism is used to ensure that the stored data are valid in terms of the clinical domain.

RAHIS Phase-II follows the big model approach where each application module generates at least one global data file that is stored in the MUMPS database. ECHISTAR follows the generic approach where a reference model defines the general purpose set of data structures and the archetype model defines the clinical data attributes.

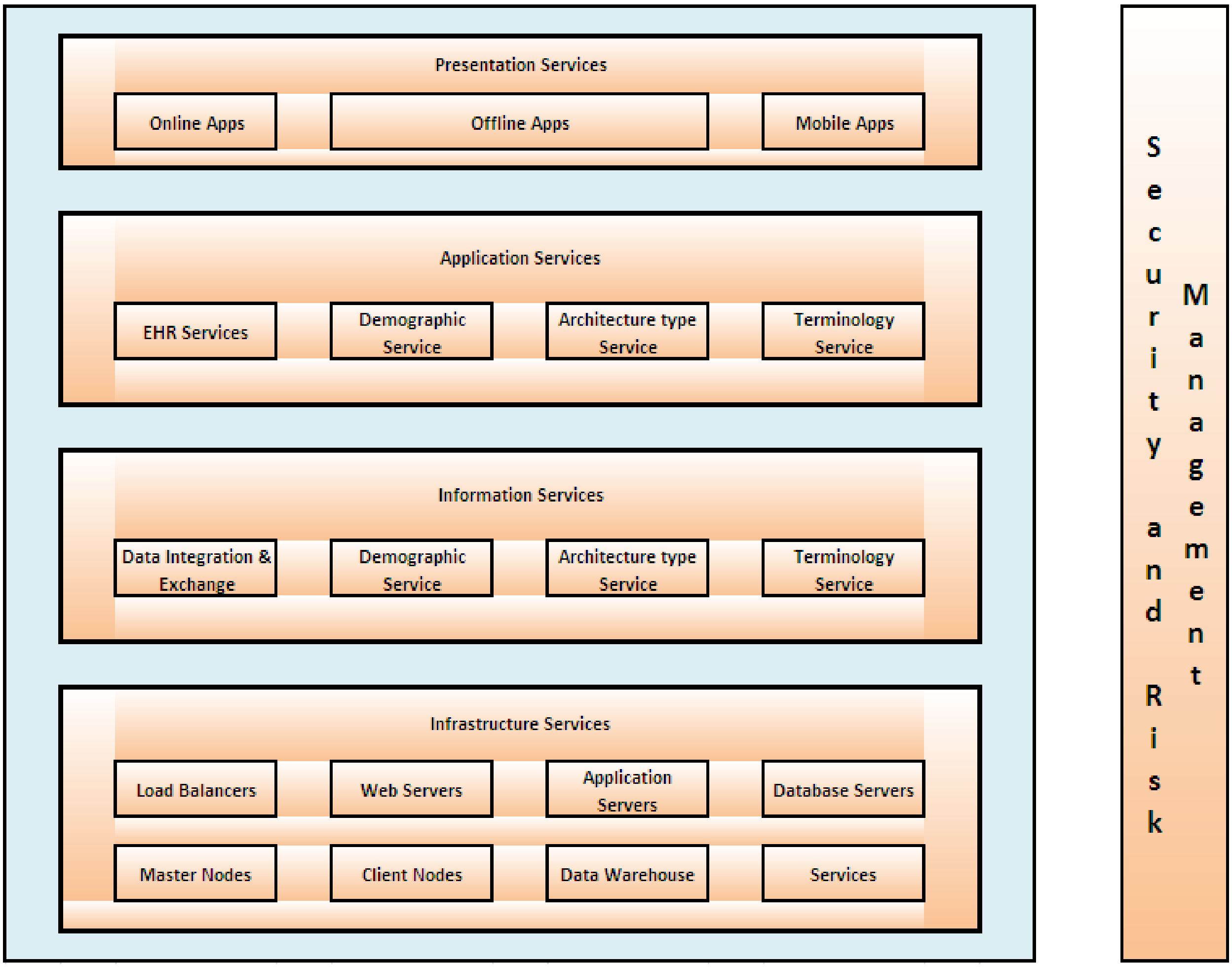

Fig. 5 Architecture of proposed ECHISTAR

B. Data Interoperability

The main two challenges faced by traditional HER system is data integration and interoperability which use different and often conflicting technical and semantic standards that leads to data integration and interoperability problems by using different languages, and different technology generations by finding the consequence is that EHR systems that are fragmented and unable to exchange data.

Acquiring medical data from different sources requires a higher level of data interoperability where most of the medical information systems store clinical information about patients in proprietary formats where as in interoperable EHR system will contribute to more effective and efficient patient care by facilitating the retrieval and processing of clinical information about a patient from different sites.

C. Loose Coupling

ECHISTAR adopts the cloud component model approach along with the software engineering best practices such as loose coupling between various application components in which the cloud component model approach is adopted instead of hard-wired links or the components interface that clearly specifies functional and service boundaries by relying on use of API interfaces and web services interfaces.

IV. PROPOSED SYSTEM

The Fig. 5 shows the layered architecture of the proposed ECHISTAR system where the infrastructure services layer consists of various cloud instances such as load balancers, application servers, Hadoop master, and slave nodes on which ECHISTAR is deployed.

The information services layer also consists of a data integration engine which allows integration of data from multiple disparate data sources into the cloud model for data storage with clinical concepts and the data governance module where the application services layer provides various services such as EHR service, demographic service, archetype service, and terminology service.

ECHISTAR achieves semantic interoperability by using a two-level modeling approach that makes the system more robust by separating data from clinical knowledge by using data storage model and an architectural model where data storage model defines entities for data storage and represents the semantics of storing data and architectural model defines the clinical concepts by representing domain-level structures and constraints on the generic data structures defined by the data storage model which contains a header section that includes the architecture type metadata, definition section contains the modeled clinical concept, and ontology section describes the entities defined in the definition section, here architecture types are separate from the data and are stored in a separate repository and are deployed at runtime using templates that specify groups of records of all vital signs of a patient such as blood pressure, pulse, etc.

The proposed data integration engine is designed based on technologies such as Hadoop and MSBI (Microsoft Business Intelligence) frameworks where the data integration engine converts EHR data from different sources to flat files which are stored in HDFS distributed storage. A MapReduce-based bulk loader loads data from the flat files into MSBI and into Hive which is a data warehouse system for Hadoop that facilitates easy data queries and the analysis of large data sets stored in Hadoop compatible file systems by using Java APIs or the SQL-like query language provided by Hive called HQL.

ECHISTAR supports distributed querying of healthcare data using hQuery open source framework since the EHR system can handle massive amount of data while meeting the performance requirements as it maintains separate MSBI tables for EHR data and demographic or identity data which provides additional security as the EHR data alone provides no information on the identity of the patient it belongs to.

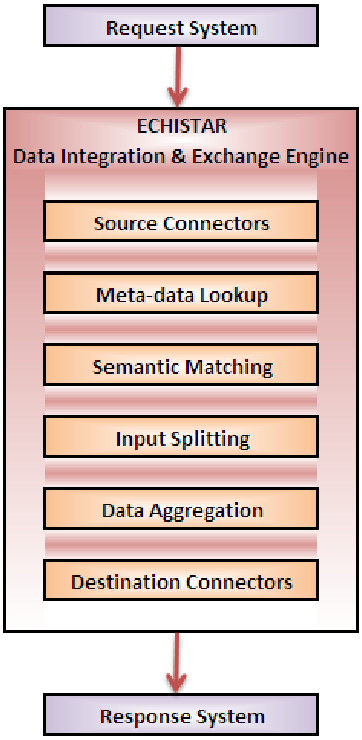

Fig. 6 Architecture of ECHISTAR's data integration

ECHISTAR's data integration engine is built on the Mirth Connect open source integration engine and the FM Projection toolkit which integrates EHR data with RAHIS phase I and RAHIS phase II by viewing standard database queries and reporting tools that allows hooking up any tool that knows JDBC or ODBC to tie up with the mysql database. In the process of integration a source connector connects to an external system is conFig.d against data sources then a metadata lookup is performed to discover the semantics of data elements in the source file then ECHISTAR maintains metadata repositories for all the data types supported and type casted by the data integration engine and generates an intermediate XML file which has all the data elements related to source file along with the annotations.

Metadata lookup is performed to discover the semantics of data elements in the source file a ECHISTAR maintains metadata repositories for all the data types supported by the data integration engine since the lookup process is data driven and produces an intermediate XML file which has all the data elements in the source file along with the annotations and bugs which are obtained by a metadata repository of the source format.

The biggest obstacle in the area of cloud computing is security and privacy issues of healthcare data stored in the cloud is outsourcing in nature where a third party will do the job for us without knowing what it is by protecting once individual health information. ECHISTAR adopts single sign on (SSO) for the process of authentication by enabling users to access multiple applications after signing in only once by which the identity is recognized for the rest of time, when a user tries to access ECHISTAR a request is generated and user is redirected to the identity provider where the request is parsed and authenticated by providing a SSL encryption when transmitting assertions and messages by using digital signature mechanism to prevent replay attacks.

ECHISTER provides an open standard for authorization that allows resource owners to share their private data stored on one site with another site without handing out the credentials. Here an application requests access to resources which are further granted permission to access the resources in the form of token and matching shared-secrets, these tokens are created and issued with a restricted scope and limited lifetime by a cloud owner and later revoked independently, here a system administrator grants permissions and access control policies that are stored in the user roles and data access policies databases. A Role-based access control framework will grant access permissions to healthcare data for all types of users based on the assigned roles and data access policies.

Identity management services consistent of various methods for identifying users and maintaining associated identity attributes for the users across multiple organizations in the proposed application ECHISTAR implements on 256-bit Advanced Encryption Standard (AES-256) which is a data encryption standard for message encryption algorithms. ECHISTAR data is stored in HBase and first encrypted with AES-256 encryption and then inserted into HBase.

Data integrity ensures that the data is not altered in an unauthorized manner after it is created or transmitted or stored, ECHISTAR uses message authentication codes (MAC) to detect both accidental and deliberate modifications in the data by providing cryptographic checksum over data for provision of assurance that the data is not altered or changed.

HIPAA/HITECH are used as regulations that requires log data on the accesses to PHI for maintenance of accountability purposes it also facilitates logs which maintains all read and write accesses to patient health records which maintains all information such as when a token is accessed or actions performed on token and whether a exception is raised, if raised what is the exception description.

V. EXECUTION AND RESULTS

We deployed ECHISTAR on the PHP cloud infrastructure where the tier-1 consists of web servers and android app for load balancing then tier-2 consists of application server Apachi and tier-3 consists of a cloud-based distributed batch processing infrastructure such as Hadoop and MSBI, the main data base implementation is done on HBase which is used as a database layer and is a distributed non-relational column oriented database that runs on top of HDFS by providing fault-tolerant way of storing large quantities of sparse Aurogyasri data in the form of flat files and images.

For measuring the scalability of ECHISTAR many series of experiments have been conducted by us with very large data sets with almost 10,00,000 patient health records and the data sets for the experiments were generated synthetically by the algorithm and patient record data is used for experiments consisted of diagnosed problems, medications, vital signs, etc.

We have used Weka 3.6 Data mining tool to generate results and they are:

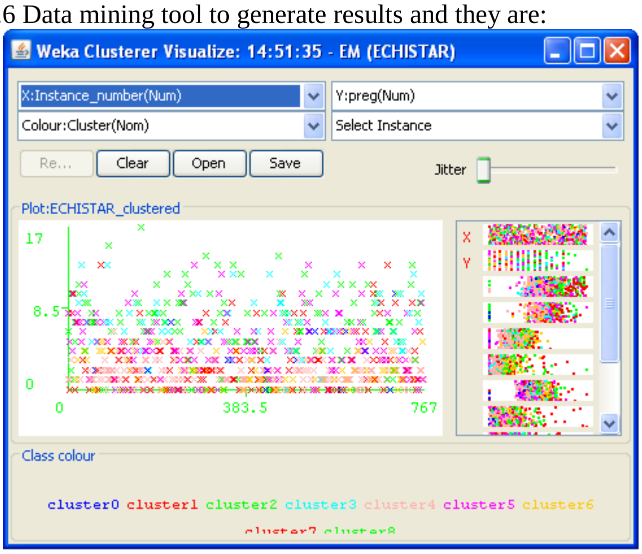

Fig. 7 Cluster analysis of ECHISTAR data.

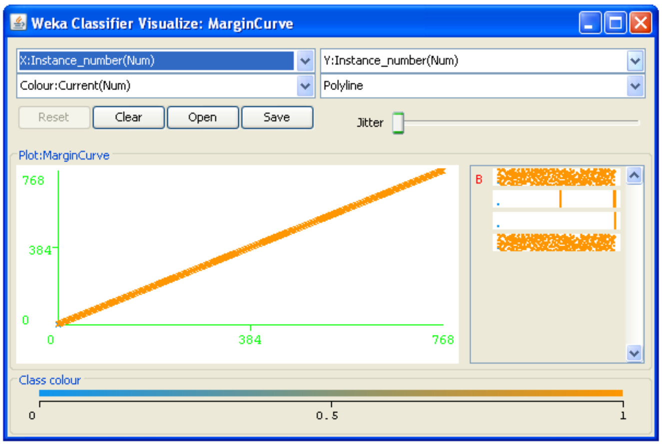

Fig. 8 Classification of ECHISTAR data.

And below is the output chunk generated on the data provided.

=== Run information ===
Scheme: weka.clusterers.EM -I 100 -N -1 -M 1.0E-6 -S 100
Relation: ECHISTAR
Instances:7680
Attributes:9
Preg plas pres skin insu mass pedi age class
Test mode: evaluate on training data
=== Model and evaluation on training set ===
EM
==
Number of clusters selected by cross validation: 9

	Cluster								
Attribute	0	1	2	3	4	5	6	7	8
	(0.03)	(0.11)	(0.13)	(0.1)	(0.25)	(0.23)	(0.12)	(0.02)	(0.02)
== preg									
mean	3.9034	1.5715	7.5167	6.5193	1.6279	4.7407	2.3972	5.9978	3.2096
std. dev.	3.1713	1.4802	3.2402	2.9664	1.2627	3.3012	1.9538	3.1802	4.1346
plas									
mean	103.0753	147.3413	141.0528	110.5914	99.9371	121.1021	122.4256	153.4527	140.4861
std. dev.	16.6232	30.9517	35.7507	28.8575	21.1453	27.8133	21.8205	31.2459	24.373
pres									
mean	0.0049	73.1341	77.0567	76.8168	65.7426	74.1094	71.6251	83.0705	0.1514
std. dev.	0.6221	14.421	10.5618	10.1902	10.2404	12.5058	12.08	9.7853	2.888
skin									
mean	0	34.2458	31.9466	30.5496	23.3558	0	33.1474	0.7417	3.9111
std. dev.	15.9522	9.906	7.7691	12.0826	8.0542	15.9522	10.6131	2.2115	9.4715
insu									
mean	0	182.3554	157.3839	48.7443	62.9186	0	162.3848	23.5146	0.1365
std. dev.	115.244	158.2586	156.2192	57.6517	52.8202	115.244	116.7673	81.8094	4.3004
mass									
mean	18.561	37.3254	34.9624	31.7416	28.9775	31.0445	35.368	29.4171	36.6186
std. dev.	13.8774	8.7732	5.5031	5.3704	5.9071	7.0182	6.2361	10.677	7.3488
pedi									
mean	0.3203	0.7262	0.613	0.3705	0.4119	0.3692	0.487	0.6999	0.4946
std. dev.	0.2175	0.5354	0.3298	0.2068	0.2436	0.2321	0.2919	0.4798	0.2621


```
age
  mean          30.1742  28.3448  43.9476  43.7547  23.6227  36.6204  27.7527  56.2464  31.5495
  std. dev.     11.2186   5.3519   8.1344  11.0816   2.5065   11.944   4.8958  11.1521   6.6209

class
  tested_negative    18.6122  11.1036  11.1079  68.9642 184.4546 114.3313   88.074  10.1124   2.2398
  tested_positive     4.4756  72.5373  88.0818   6.8737  10.0472  65.0047   6.0009  10.5016  13.4773
  [total]            23.0878  83.6409  99.1896  75.8379 194.5018  179.336  94.0749   20.614  15.7171
Clustered Instances

0      170 (  2%)                  1      760 ( 10%)
2     1040 ( 14%)          3      730 ( 10%)
4     2080 ( 27%)          5     1470 ( 19%)
6      830 ( 11%)                  7      420 (  5%)
8      180 (  2%)
Log likelihood: -28.54483
```

VI. CONCLUSION AND FUTUREWORK

In this paper, we proposed a design of a cloud-based HER system by name ECHISTAR which addresses the problems faced by traditional client–server EHR systems by adopting a two level modeling approach for achieving semantic interoperability and data integration engine by which a ECHISTAR allows aggregating healthcare data from disparate data sources mainly taken from Aarogyasri and supports advanced security features and addresses the key requirements of HIPAA and HITECH. ECHISTAR has better interoperability, scalability, maintainability, portability, accessibility, and reduced costs as compared to traditional client–server EHR systems. Future work will focus on the development of a cloud based information integration and informatics framework for healthcare applications apart from Aarogyasri where the framework will allow development of smart and connected healthcare applications backed by massive scale healthcare data integrated from heterogeneous parallel and distributed healthcare systems within a scalable cloud infrastructure.

REFERENCES

[1] Rajiv Aarogyasri [Online]. Available: https://www.aarogyasri.gov.in/
[2] (2012). KPMG, The Cloud Changing the Business Ecosystem[Online]. Available: www.kpmg.in
[3] J. Patrick, R. Ly, and D. Truran, “Evaluation of a persistent store for openEHR,” inProc. HIC HINZ, 2006, pp. 83–89.
[4] Weka 3: Data Mining Software in Java [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/
[5] Medsphere Systems Corporation. (2010).From MUMPS to Java: OVID Unleashes Power of Open-source Health IT[Online]. Available: http: //www.medsphere.com/ovid-white-paper.

