Development of an Enhance Medical Intelligence Process using Ontology Based and Virtual Data Integration Technique

Agbasonu, Valerian Chinedum., Amanze, Bethran Chibuike., Eleanya, Mirian C
Department of Computer Science, Imo State University,
Owerri, Imo state
Email: amanzebethran@yahoo.com
DOI:10.56201/ijemt.v10.no8.2024.pg79.96

Abstract

A Hybrid model for enhance medical intelligence process using ontology based and virtual data integration technique is a work that aims to build a medical system that has the ability to detect and suggest cure for an ailment with minimal effort. Healthcare system has been a major source of worry across the globe in recent time due to emergence of different types of diseases and epidemics. The number of health care personnel in various hospitals falls short of the number required most especially in terms of specialists. The medical system in Nigeria today suffers from lack of fast, accurate, reliable and intelligent software solutions that can help healthcare practitioners make decisions that would solve urgent, and in some cases, complex medical problems in real-time. Also, the cost of processing and analyzing large volumes of data in a medical environment is high most especially in terms of time consumption. So, this work proposes a patient-oriented design for integration of large volumes of data in order to improve database validity compared to procedure-oriented design that multiplies the redundancy of data. The research will address the problems, by having a hybrid of ontologies and virtual data integration in order to enhance medical intelligence process. A hybrid model for enhance medical intelligence process using ontology based and virtual data integration technique will be develop. The design provide for a database system for storing medical records, software for enhanced Medical Intelligence Process that will be more user-friendly, flexible, adaptive, intelligent, agile and automatic in integrating and analyzing medical data thereby helping medical practitioners at various levels to make realistic intelligent and real-time decision on critical health issues. Object Oriented Analysis and Design Methodology (OOADM) will adopt in the design of the system. The system will achieve integration of various patient's medical records from different hospitals using ontology based and virtual data integration technique that will allow clinic data of one patient collected together to form a Combinational resource, and could be accessed by physician if authority is assigned to the physician. The Hybrid technique using both Ontology-based data integration and virtual data integration technique for disease control procedure will achieve a good accuracy in predicting the disease control procedure.

Keywords: Hospitals, Patients, VDI and Ontology

Introduction

Healthcare system has been a major source of worry across the globe in recent time due to emergence of different types of diseases and epidemics. The number of health care personnel in various hospitals falls short of the number required most especially in terms of specialists. Having effective healthcare system will help in delivering the information to the doctors, patients at the point of emergency and it will improve the quality of healthcare in emergency medical service and medical intelligence. Medical intelligence is the ability to detect and cure an ailment on time with minimal effort. It requires vast knowledge on disease symptoms, cure, and this can only be achieved by having a data warehouse build from the knowledge of medical experts. To apply medical intelligence effectively, the healthcare condition of the patients must be ascertained. The health-care condition of a patient is defined as all the past and current medical and social information about the patient that may affect the professional immediate and short-term management of that patient. In this thesis, this information corresponds to all the diseases, syndromes and social issues that are diagnosed for the patient, the signs and symptoms (including family medical history), the problem assessments performed (i.e., medical, social, cognitive, and mobility tests), and the current interventions, either pharmacological, rehabilitative, nurse care, social care, counseling, and special medical services. In order to describe patient conditions, several national and international encoding systems have been proposed for diseases and procedures (Shortliffe, 2015). Also, the ability to describe patient's condition on time will result in a speedy healthcare delivery system. This requires techniques for data integration, warehousing and representing knowledge in biomedicine. At some time, health records of the patients and detail of the doctors are stored in different hospitals or stored in different location of the database, it is difficult to collect these records (Boyi, 2014). For that, ontology can be constructed to resolve these problems and to make correct decision at emergency period. Ontologies are one of the most successful ways of representing actionable knowledge in biomedicine (Rosse, 2013). Two of the reasons for this success are their ability to capture biomedical knowledge in a formal but simple, powerful and incremental manner, and their easy application in the reasoning processes performed by medical decision support systems (Ouwens, 2015). In health care, the most common, complex and resource-consuming clinical cases to deal with correspond to chronically ill patients, who are a kind of patients that deserve long term and simultaneous assistance provided by several sorts of professionals, as for example family doctors, specialists, nurses, or social workers. In order to deal with this highly variable kind of patients, we need mechanisms to personalize the knowledge describing both the condition of these patients (each individual patient is a potential different case with specific diseases, syndromes, social needs, signs and symptoms), and the intervention plan for these patients (the actions to be followed for different patients are eventually very varied). But we also need mechanisms to assess whether the decisions and recommendations on these patients are correct or not in part because the possibilities of over- and under-treat these kinds of patients can be very high. Also, integration becomes a crucial challenge as heterogeneous data is generated by various healthcare systems. This has to do with how to integrate various types of data including patient demographics and environmental data, clinical monitoring systems, pathology and radiology imaging data, textual data from clinical reports etc. Bianchi et al. (2019) describe the integration of clinical, environmental and genetic data and point out with examples how ontologies

are used to normalize data from various disparate systems. In order to support the Ontology content accessing, this thesis presents a hybrid model to locate and get hospital data which are stored in heterogeneous hospital information systems. In the proposed method, clinic data of patient is defined as resource with unique URL address. Related clinic data of one patient is collected together to form a combinational resource, and could be accessed by physician if authority is assigned to the physician. A wrapper system integrating this two personalization and the decision support tool was also implemented in the new system. With this system, the ontology can be directly maintained (and extended) by health-care professionals without any intervention of information technology specialists. As the ontology grows up with new medical concepts and properties, the diversity of patients that our system is able to deal with increases, the possible intervention plans can be more detailed and accurate, and the decision support system automatically becomes more powerful. Therefore, the health-care utility of the system is exclusively dependent on the incorporation of new knowledge in the case ontology. The expected result from the thesis will shows that the clinic data of patients could be accessed more conveniently from any hospital. In healthcare service, doctors, patients, physicians play a major role and they also involved in an entire servicing. Doctors need to access the patient record from anywhere by storing it in a distributed manner. Patients also need to know about the doctor's availability in a given hospital. With this, an integrated healthcare system that is more intelligent and uses intelligent process will be achieved.

Statements of the Problem

In healthcare system all over the world, the amount of patient-oriented data is constantly growing. More hospitals are opening up with various departments / units. For example, the Intensive Care Unit (ICU) is an extremely data intensive environment where large volumes of data from patient monitoring and observations are recorded continuously. Such patient-oriented data could be generated from medical devices, laboratory results, electronic prescriptions, therapeutic decisions, clinical observed values by physicians and nurses. These data are disintegrated and access to the data has to been done by requesting for it through the various hospital units. This is not only time consuming but obsolete. The following are the existing challenges of the medical system;

1. Lack of fast, accurate, reliable and intelligent software solutions that can help healthcare practitioners make decisions that would solve urgent, and in some cases, complex medical problems in real-time.
2. Cost of processing and analyzing large volumes of data in a medical environment is high most especially in terms of time consumption.

So, this work propose a patient-oriented design for integration of large volumes of data in order to improve database validity compared to procedure-oriented design that multiplies the redundancy of data. The research is to address the problems, by having a hybrid of ontologies and virtual data integration in order to enhance medical intelligence process.

Aim and Objectives of the Study

The aim of this work is to develop a hybrid model for enhance medical intelligence process using ontology based and virtual data integration technique. The specific objectives include:

i. To design a database system for storing medical records.
ii. To develop software for enhanced Medical Intelligence Process that will be more user-friendly, flexible, adaptive, intelligent, agile and automatic in integrating and analyzing medical data thereby helping medical practitioners at various levels to make realistic intelligent and real-time decision on critical health issues.
iii. To develop a system that will integrate the various patients' medical records from different hospitals using ontology based and virtual data integration technique that will allow clinic data of one patient collected together to form a Combinational resource, and could be accessed by physician if authority is assigned to the physician

Literature Review

Marut (2016) presented an article title "Ontology-based Clinical Reminder System to Support Chronic Disease Healthcare". In his presentation, he said that improving quality of healthcare for people with chronic conditions requires informed and knowledgeable healthcare providers and patients. Decision support and clinical information system are two of the main components to support improving chronic care. In his paper, he describes an ontology-based information and knowledge management framework that is important for chronic disease care management. Ontology-based knowledge acquisition and modeling based on knowledge engineering approach provides an effective mechanism in capturing expert opinion in form of clinical practice guidelines. The framework focuses on building of healthcare ontology and clinical reminder system that link clinical guideline knowledge with patient registries to support evidenced-based healthcare. He describes implementation and approaches in integrating clinical reminder services to existing healthcare provider environment by focusing on augmenting decision making and improving quality of patient care services. The paper was focused on clinical reminder service and didn't integrate electronic health record (EHR) standards and this is the research gap established in his work. In a research work (Madhura, 2020) titled "A data integration platform for patient-centered e-healthcare and clinical decision support", the researchers proposed an open data integration platform for patient, clinical, medical and historical data across multiple health information systems. As an open platform, it can accommodate and integrate further heterogeneous data sources such as data streams generated by wearable Internet of Things (IoT) devices. As an integration platform, it facilitates centralization of data assets. This centralization empowers every stakeholder in a patient-centered care setting to actively participate in decision-making. A range of analytics and reporting solutions, such as data warehouse, interactive dashboards, and predictive analytics tools, can be deployed upon this open data integration platform. The proposed platform is currently being adapted and implemented to address patient-centered healthcare and clinical decision support requirements in a sports injury clinic at a not-for-profit private hospital in Melbourne, Australia. Use Case based demonstration of the platform's suitability for holistic information management, decision-support, and predictive analytics justify its role in the advancement of e-healthcare. The work suggested than advanced analytics, data

visualization, monitoring and reporting functionalities for clinical decision support will should be added and customized in future work and this is identified as the research gap. Medical decision support systems based on machine learning was presented by (Chih-Lin, 2019). The central idea of the dissertation is aimed at facilitating personal health care, reducing costs of health care, and improving outcomes. They proposed a new machine-learning algorithm for three disjointed health care problems: hospital referral, cost-effective diagnosis, and lifestyle recommendation. In the hospital referral and lifestyle recommendation projects, individualized recommendation is generated based on the input of personal characteristics and preferences. The systems can then return the best individual solution (hospital selection or the plan of lifestyle changes) that fits one's preference and personal considerations. In the cost-effective diagnosis project, the recommendation of a test is provided based on individual information (including symptoms and previous test results). The recommended test has the highest potential to cross (or get close to) the treatment (or non-treatment) threshold. In other words, they optimize diagnosis in terms of the number of tests and the amount of cost without sacrificing accuracy (sometimes improving accuracy). For the lifestyle project, the dataset is the description of a specific population. First, the recommended lifestyle is thus the best lifestyle observed from that population. Some known healthy behavior may not be recommended when most people don't have the behavior. For example, when most people consume saturated fat more than 7% energy (this cut-point is suggested by American Heart Association), the system may not recognize the benefit of satisfying this criterion. In addition, due to bias, the recommended lifestyle may not be the best for patients not in that population. Sozan (2019) presented a thesis titled "Intelligent System for Identification Heart Diseases" to the graduate school of applied sciences of near east university. I her submission, she said that most of existing traditional medical systems are based on the knowledge of experts-doctors. In her thesis, the application of soft computing elements is considered to automate the process of diagnosing diseases, in particularly diagnosing of a heart attack. The research work offers probable help to the medical practitioners and healthcare sector in making instantaneous resolution during the diagnosis of the diseases. The intelligent system predicts heart attacks from the patient dataset utilizing algorithms and help doctors in making diagnose of these illnesses. In the study, three techniques such as a neural network (back propagation), Fuzzy Inference System (FIS) and Adaptative Neuro-Fuzzy System (ANFIS) are considered for the design of the prediction system. The systems are designed using data sets. The data sets contain 1319 samples that include 8 input attributes and one output. The output refers presence of a heart attack in the patient. For comparative analysis, the simulation results of the ANFIS model is compared with the simulation results of the neural network-based prediction model. The ANFIS model has shown better performance and outperformed NN based model. The obtained simulation results demonstrate the efficiency of using ANFIS model in the identification of heart attacks. Serdar (2014) presented an "Intelligent Systems in Patient Monitoring and Therapy Management". The research pointed out that although today's advanced biomedical technology provides unsurpassed power in diagnosis, monitoring, and treatment, interpretation of vast streams of information generated by this technology often poses excessive demands on the cognitive skills of health-care personnel. In addition, storage, reduction, retrieval, processing, and presentation of information are significant challenges. These problems are most severe in critical care environments such as intensive care units (ICUs) and operating room (ORs) where many events are life-threatening and thus require

immediate attention and the implementation of definitive corrective actions. The article focuses on intelligent monitoring and control (IMC), or the use of artificial intelligence (AI) techniques to alleviate some of the common information management problems encountered in health-care environments. The article presents the findings of a survey of over 30 IMC projects. A major finding of the survey is that although significant advances have been made in introducing AI technology in critical care, successful examples of fielded systems are still few and far between. Widespread acceptance of these systems in critical care environments depends on a number of factors, including fruitful collaborations between clinicians and computer scientists, emphasis on evaluation studies, and easy access to clinical information. In a research work by (Vishesh, 2010) titled "Personal health record system (PHRS) and integration techniques with various electronic medical record systems", the researcher proposed a Web-based PHRS that can store data in a cloud-based architecture. The new system is web based and is built on J2EE technology which will enable users to access and share the medical health information from any place at any time with desired care provider. Integration of healthcare information systems is a complicated task and full of challenges. To effectively handle the large volume of medical imaging data system uses the open-source system. The cloud computing based architecture was used which allows consumers to address the challenge of sharing medical data. PHRS provides a complete and accurate summary of the health and medical history of an individual by gathering data from many sources. This makes information accessible online to anyone who has the necessary electronic credentials to view the information. Dipti (2010) in a paper titled "An Expert System for Diagnosis of Human Diseases" said that detecting diseases at early stage can enable to overcome and treat them appropriately. Identifying the treatment accurately depends on the method that is used in diagnosing the diseases. A Diagnosis expert system (DExS) can help a great deal in identifying those diseases and describing methods of treatment to be carried out taking into account the user capability in order to deal and interact with expert system easily and clearly. Present expert system uses inference rules and plays an important role that will provide certain methods of diagnosis for treatment. In research, (Ighoyota, 2017) proposed a web-based expert multi-fever diagnosis system using a novel fuzzy symptom classifier with human self-observed physiological symptoms; Considering malaria, Lassa, dengue, typhoid and yellow fever. The fuzzy-symptom classifier has two stages. Fist stage is fever type confirmation using common fever symptoms, leading to five major fuzzy rules and the second phase is determining the level of infection (severe or mild) of the confirmed type of fever using unique fever symptoms. Furthermore, Case studies during the system implementation yielded data collected from 50 patients of having different types of fever. The analysis clearly shows the effectiveness and accuracy in the system performance through false result elimination. In addition, acceptability of the system was investigated through structured questionnaire administered to same 50 patients. This result clearly indicates that the system is well accepted, by users and considered fairly easy to use, time and cost saving. Alexander (2011) developed a business intelligent system and said that Business Intelligence Systems (BI) describe a form of data driven Decision Support Systems (DSS) that integrate a variety of concepts and technologies to gather, store and analyses data. Traditionally the focus of BI is on strategic and tactical decision support by providing decision makers a centralized and holistic view on organizational data. Today businesses are generating increasingly larger amounts of data due to regulatory requirements, business needs and new technologies. Managing and using this data in

business decisions can be difficult because of the volume of the data, time pressure and general complexity of today's business problems. The traditional BI concept does not fully reflect these operational and local requirements and should adapt to this new environment and these requirements to better support businesses in their decision-making activities. Agent and Multi Agent technology is often mentioned as an approach to design and develop flexible and distributed software systems. The technology is used in this research to design the Multi Agent Enhanced Business Intelligence (MAEBI) framework that focuses on distributing decision making capabilities throughout an organization. Core to the MAEBI framework is the so-called Decision Unit (DU) that encapsulates BI functionality with the extension of a Decision Execution (DE) module that allows implementing (changing business process) a decision without human interaction. The agent-based design allows embedding a DU in the problem domain to make decisions with a local perspective. A prototype, pMAEBI (p=pricing), was implemented in the context of multi store retail pricing. Pricing is an important and complex problem for retailers and it allows demonstration of some of the capabilities of a MAEBI based system. To evaluate the pMAEBI system a simulation testbed was implemented to analyses the prototype in comparison to a traditional "centralized" system. Simulation results indicate that the pMAEBI managed stores performed better (in terms of profit) than the comparison stores. An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients was proposed by (David, 2020). He stated that chronically ill patients are complex health care cases that require the coordinated interaction of multiple professionals. A correct intervention of these sort of patients entails the accurate analysis of the conditions of each concrete patient and the adaptation of evidence-based standard intervention plans to these conditions. There are some other clinical circumstances such as wrong diagnoses, missing information, unobserved related diseases or prevention, whose detection depends on the capacities of deduction of the professionals involved. An article published by (Agustina, 2018) reviewed and mad a comparison of Ontology-Based Data Integration Methods. The researcher stated that data integration system provides a uniform interface to distributed and heterogeneous sources. These sources can be databases as well as unstructured information such as files, HTML pages, etc. One of the most important problems within data integration is the semantic heterogeneity, which analyzes the meaning of terms included in the different information sources. This survey describes seven systems and three proposals for ontology-based data integration. An important feature is that all of them use, in some way, ontologies as the way to solve problems about semantic heterogeneity. In the paper, they show similarities and differences among the systems by providing a framework for comparison and classification. Bostjan (2014) proposed Automating ontology-based information integration using service orientation. He argued that with the rise of the Internet, globalization and the increasing number of applications used inside organizations, there is an emerging need to integrate information across heterogeneous information systems. Service oriented architecture (SOA) is seen as a general answer to intra-organizational as well as inter-organizational integration problems. While service-oriented systems have been well studied, there are still some challenges remaining unanswered. One of them is automation of service execution. The paper proposes a method for automated execution of Web Services. Based on Web Service execution automation, the proposed approach is bridging the gap between ontology-based integration and service-oriented architecture by enabling dynamic and transparent integration of information which is

provided by services. The problem of splitting the query into static and dynamic query was not addressed fully and that is the research gap identified in the work. Ali et al. (2018) proposed a Virtual Data Integration Framework for information sharing. In the research work, he said that Data Integration is the process of combining data residing at homogeneous, autonomous, and heterogeneous data sources, and providing users with a unified global schema (GS). A Mapping Approach for fully Virtual Data Integration System Processes was proposed by (Ali, 2018). He said that multiple data sources across the organization fuels the need for data integration. Data integration system's users pose queries in terms of an integrated schema and expect accurate, unambiguous, and complete answers. Vinoth (2019) presents a paper titled "Ontology based Public Healthcare System in Internet of Things". He is of the opinion that Internet of Things is a growing technology that is predicted to discover new drugs and medical treatments. The efficiency and quality of healthcare have high potential features as flexibility, adaptability, affinity, cost shrinkage, and high speed. This technology helps us to understand the specific risks related to security and privacy. This paper targets on a Healthcare information system based on ontology method. In particular, security and privacy challenges are analyzed in the proposed Ontology-based healthcare information system. Taqdir, et al (2017) proposed that technologically integrated healthcare environments can be realized if physicians are encouraged to use smart systems for the creation and sharing of knowledge used in clinical decision support systems (CDSS). While CDSSs are heading toward smart environments, they lack support for abstraction of technology-oriented knowledge from physicians. Therefore, abstraction in the form of a user-friendly and flexible authoring environment is required in order for physicians to create shareable and interoperable knowledge for CDSS workflows. The proposed system provides a user-friendly authoring environment to create Arden Syntax MLM (Medical Logic Module) as shareable knowledge rules for intelligent decision-making by CDSS. They pointed out that existing systems are not physician friendly and lack interoperability and share-ability of knowledge. In this paper, they proposed Intelligent-Knowledge Authoring Tool (I-KAT), a knowledge authoring environment that overcomes the above-mentioned limitations. Shareability is achieved by creating a knowledge base from MLMs using Arden Syntax. Interoperability is enhanced using standard data models and terminologies. However, creation of shareable and interoperable knowledge using Arden Syntax without abstraction increases complexity, which ultimately makes it difficult for physicians to use the authoring environment. Therefore, physician friendliness is provided by abstraction at the application layer to reduce complexity. This abstraction is regulated by mappings created between legacy system concepts, which are modeled as domain clinical model (DCM) and decision support standards such as virtual medical record (vMR) and Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT). They represent these mappings with a semantic reconciliation model (SRM). The objective of the study is the creation of shareable and interoperable knowledge using a user- friendly and flexible I-KAT. Therefore, they evaluated the system using completeness and user satisfaction criteria, which was assessed through the system- and user-centric evaluation processes. For system-centric evaluation, they compared the implementation of clinical information modelling system requirements in the proposed system and in existing systems. The results suggested that 82.05% of the requirements were fully supported, 7.69% were partially supported, and 10.25% were not supported by their system. In the existing systems, 35.89% of requirements were fully supported, 28.20% were partially supported, and

35.89% were not supported. For user-centric evaluation, the assessment criterion was 'ease of use'. The proposed system showed 15 times better results with respect to MLM creation time than the existing (Taqdir, 2017). The research done by Richter and Weber (2016) illustrated that to have high-quality content of case base, they should be prepared using effective and dependable sources. There must be a regular or a uniform distribution of cases of the problems (Richter, et al 2016). When cases are not appropriately distributed or fairly among problems, this results in the existence of any problems without solutions while others may have redundant and useless cases. The CBR systems' accuracy will be enhanced when the medical dataset of EHR is pre-processed. The step of data pre-processing is about CBR as well as applying the many techniques of artificial intelligence, such as genetic algorithm, k-nearest neighbor (KNN), Bayesian network, and fuzzy approach (Gu, 2010). The steps of data preprocessing steps involve handling missing data, feature selection and weighting, data integration, discretization of data, normalization, and outliers' detection and removal (Gu, 2010). These data preprocessing steps are applied on EHR for converting database structure of EHR to case base structure and transforming EHR generic data to specified case base. Among the studies conducted, which focused on the missing data problem, was by Jagannathan and Petrovic (2009). The authors of this study illustrated that case base, which contains missing data, is a big problem. It affects negatively on the CBR system's performance. So, these missing values must be treated using imputation approaches like mean/mode method, KNN imputation, etc. However, the missing data values of some attributes have been handled while others are not treated. The performance of CBR applications is enhanced when applying preparation algorithms on the case base, such as feature selection. Every CDSS system, which is based on knowledge, needs a step of preprocessing to produce a high-quality dataset. For obtaining superior results during the retrieval process, cleaning, and normalizing data steps are done on retrieval algorithms like KNN algorithm (Jagannathan, 2009). A clinical expert system has been proposed in (Asma, 2011) which adopts decision tree approach for prediction of presence and absence of diabetes among the instances. The author used PIMA Indians Diabetes Data Set, which collects the information of patients with and without developing diabetes. Their research went through two phases. The first phase is data pre-processing including attribute identification and selection, handling missing values, and numerical discretization. The second phase is a diabetes prediction model construction using the decision tree method and they used Weka software throughout all the phases of their study. Nassim (2014) investigated and evaluated a model framework, for diagnostic decisions, based on a cognitive process and a Semantic Web approach. Fuzzy cognitive maps (FCM) are a cognitive process applying the main features of fuzzy logic and neural processors to situations involving imprecision and uncertain descriptions, in a similar way to intuitive human reasoning. The research explored the use of this method for modeling clinical practice guidelines, using Semantic Web tools to implement these guidelines and for the formalization process. Twenty-five clinical and 13 diagnosis concepts were identified, to represent the problem of urinary tract infection diagnosis. With the proposed methodology, each cause-effect relationship between an observation and diagnosis is described by one or more fuzzy rules, thus producing the rule-based FCMs. From each fuzzy rule, an inference is generated, depicting the degree of confidence the research have in the influence concerned. Moreover, based on the available fuzzy rules, some of the relationships (weights) may change in value or degree of confidence before the final diagnosis is reached, taking into account cases and observable states

that contribute to a different diagnosis. This is an important issue in FCM construction and diagnosis, as specific initial states of observables can be assessed for patients. The result of the research work is summarized in the development of a platform able to interact with heterogeneous data and formalize knowledge from the model. The results presented here follow the steps in the methodology described above and the inference mechanism for FCM (Eulersharp/plugin). 92% (65/70 patients: 37 males and 28 females) diagnosis proposed by the system was in fully agreement with the guidelines.

Methodology Adopted

The new system will bring together the benefits of ontology-based data integration technique as it relates to seamless transition, solving inconsistencies in semantics and accuracy issues in syntactic and that of virtual data integration technique as it relates to hiding of technical jargons from users and unifying integrated and reconciling views of data residing at different sources as well as at different location for users. These would be implemented in a Business Intelligence environment using health sector, thereby enhancing the existing models that adopted exclusively either of the two techniques. The interoperable regional healthcare system lies in the maximum usage of medical and health resources by integrating important medical technologies and sharing the medical resource and information. That includes three specific objectives as follows:

1. The new digital medical service pattern and service process standard with the complete modern medical information system
2. The integration of national healthcare sharing platform to achieve the unified procedure, and service sharing among medical institutes in certain area will combined with the monitoring and evaluating system for the sharing services.
3. Provision of services, such as two-way referral, telemedicine imaging consultation, online medical consultation, remote medical record and test result query, online medication consultation and patient ‘s following-visit, in the community and hospitals of intermediate and advanced level.

The above-mentioned objectives demand a sophisticated information structure that features the following four core requirements:

1. The integration of existing information systems of the medical institutions in the country for the capability of medical resource sharing.
2. The integration platform of medical institutions for public information service to government, medical organization, community and citizen on disease control.
3. The data center of medical information in the hospitals as the main data source for the integration platform realizing the medical information sharing.
4. The process and rules for the operation of hospitals and interoperable healthcare service platform to keep the system efficient and stable.

Since the information of healthcare domain is diversified and dynamic because of the heterogeneous and distributed information resources and the large amount of daily updated data produced by many information systems from medical institution, the key problem for constructing integrated healthcare system is the efficient management of medical workflow and the effective integration of mass data resources. Therefore, the modularization design and extensibility with the hierarchy structure are essential to this system. In the proposed system, the healthcare centers are responsible for the medical data from each regional data center to monitor disease. Each regional information integration platform is composed of hospital, clinics and community healthcare center by which different levels of medical institutions are able to share the data from data centers and realize two-way referral and medical record lending. The information sharing among different medical systems not only needs to provide full accessibility to the data but also requires the interoperability among these systems. The main problems caused by bringing together heterogeneous and distributed computer systems are summarized as semantic heterogeneity and structural heterogeneity. Semantic heterogeneity refers to the variation of semantic meaning in medical information resources which will lead to the semantic conflicts and complication for data integration. Structural heterogeneity means that the same data will be described in different structures by different systems because of various application systems, DBMS and operating systems. In this proposed business intelligence system, we will take advantage of ontologies for the explication of medical knowledge as a possible approach to overcome the problem of semantic heterogeneity and the problem of structure heterogeneity will be solved by the use of XML based data integration together with data warehouse. In the realm of knowledge management, ontology provides both the theoretical basis and the applied methods to morph the different knowledge modalities to form a unified knowledge object since they can be used for the identification and association of semantically corresponding information concepts. In the healthcare sharing platform, ontologies are used for describing semantic meaning of information source explicitly in order to solve semantic heterogeneity. We will adopt hybrid ontologies in our research. In effect, the proposed ontology-based and virtual data integration architectural process would be based on the use of ontology which explicitly captures knowledge about different types of data sources and virtual aspect uses mediators to bring about the real-time and agility aspect of the system. Generally, database schemas are regarded as static, but ontology schemas are typically assumed to be highly dynamic and are an evolving object(s).

Justification of the Study

The key concept of the propose medical intelligence hybrid model adopting ontology-based (OBDI) and virtual data integration (VDI) techniques would have the ability to ensure abstraction of data that comes from multiple sources in varying schemas, syntactic accuracy and to have a seamless transition from data into information then into action.

The advantages of the propose system include the following;

1. The system ensures seamless transition from a practical workspace into a virtual business-oriented analysis world expected by business users.

2. The system is based on hybrid architecture and also relies on elements such as system vocabulary and local ontology per each heterogeneous data source.
3. The system reduces syntax errors, structural and semantic heterogeneity and redundancy which leads to increased availability and degree of completeness.
4. With the layers of ontology-based and virtual data integration both in the system seamless transition and hiding of technical jargons from users is feasible.
5. With both approach to data integration in the system, there would be reduce cost of processing, maintenance and risk in the project as well as increased availability will be feasible.
6. The system will ensure real-time processing, analyzing and accessing of data.
7. The system will be intelligent, reliable, adaptive, flexible and agile.
8. The ontology aspect would see to the reduction in lateness in the decision process which would allow users to take and make faster or fastest business decisions accessing current business data in its proper level of abstraction, thereby enhancing the ability of organization to adapt it as new necessities or in business changes.

These justify the implementation of the system to help health sector manage and control diseases and patients record.

REQUIRED TOOLS AND AVAILABILITY

The computer system is divided into software and hardware. Both works together to achieve the desired goal in any application developed. In the web-based security system developed, the following are required.

Hardware Requirements

For the implementation of the hybrid model for enhanced business intelligence process, the following hardware is required:

- 2.4 GHZ of processor speed
- 4GB RAM
- 180 GB Hard disks
- Internet Modem
- Colored Monitor
- Printer

Software Requirements

The following software has to be installed on the computer system to run the new system developed.

- window 7, Window 8 or windows 10

- Microsoft Dream Weaver 8
- Wamp Server
- JQuery
- Fireworks
- Php-Mysql
- Java Virtual Machine

Result

Performance Evaluation

A performance metrics can be derived from the confusion matrix as show in table 1 and equation 1, which show the accuracy (AC) of the hybrid model for enhanced business intelligence process.

Table 1: **Confusion Matrix**

		Observed	
		True	**False**
Predicted	**True**	TP	FP
	False	FN	TN

$$AC = \frac{a+d}{a+b+c+d} \quad (1)$$

a = True Positive

b = False Positive

c = False Negative

d = True Negative

During the testing, 20 different disease control procedure were performed on the system to see how it can accurately identify the disease based on the symptoms and at the same time prescribe the treatment process. Three stags were followed in the performance testing; the stages are:

1. Ontology-based data integration technique for disease control procedure
2. Virtual data integration technique for disease control procedure

3. Hybrid technique using both Ontology-based data integration and virtual data integration technique for disease control procedure

Table 2: **Confusion matrix applied to test dataset of Disease Control using Ontology-based data integration technique**

		Observed	
		True	**False**
Predicted	**True**	13	2
	False	3	2

Table 2 shows that out of 20 disease control test carried out using ontology-based data integration technique for disease control procedure; 13 diagnoses are True Positive and was predicted correctly, 2 diagnosis were true while they are not, 3 diagnosis was detected to be False while it is not thereby allowing the disease to go undetected. Finally, 2 False was detected where no disease was found and it was correct. A model of performance metrics can be derived from the confusion matrix as show in equation 3, which show the level of accuracy of the ontology-based data integration technique for disease control procedure.

Substituting the values we have

AC = (13+2) / (13+2+3+2)

AC = 0.75 i.e. 75% accuracy in predicting the disease control procedure using ontology-based data integration technique.

Table 3: **Confusion matrix applied to test dataset of Disease Control using Virtual data integration technique**

		Observed	
		True	**False**
Predicted	**True**	12	3
	False	4	1

Table 3 shows that out of 20 disease control test carried out using Virtual data integration technique for disease control procedure; 12 diagnoses are True Positive and was predicted correctly, 3 diagnosis were true while they are not, 4 diagnosis was detected to be False while it is not thereby allowing the disease to go undetected. Finally, 1 False was detected where no disease was found and it was correct. A model of performance metrics can be derived from the confusion matrix as show in equation 3, which show the level of accuracy of the Virtual data integration technique for disease control procedure.

Substituting the values we have

AC = (12+1) / (15+3+4+1)

AC = 0.65 i.e. 65% accuracy in predicting the disease control procedure using Virtual data integration technique

Table 4: **Confusion matrix applied to test dataset of Disease Control**

		Observed	
		True	**False**
Predicted	**True**	18	0
	False	1	1

Table 4 shows that out of 20 transactions, 18 diagnoses are True Positive and was predicted correctly. One diagnosis was detected to be False Negative while it is not. Finally, 1 False Positive was detected. A model of performance metrics can be derived from the confusion matrix as show in equation 3, which show the accuracy of the system.

Substituting the values we have

AC = (18+1) / (18+0+1+1)

AC = 0.95 i.e. 95% accuracy in predicting the disease control procedure

Conclusion

Utilizing ontology-based data integration and virtual data integration is an attractive avenue as it is also a key factor for enabling interoperability. However, integrating vast amount of information

from different sources is a difficult, complex and demanding task. The use of ontology-based data integration systems and virtual data integration tools to automate partly the data integration task and reduce this effort has been achieved in this work. The establishment of local ontologies must represent the vocabulary used in the domain in order to recognize the synonyms relation and the hierarchical relations between the concepts. On this basis, the ontology matching becomes less time-consuming than the global schema matching as the method aims to reduce the amount of integration decisions and the number of rules.Advances in intelligent systems, e.g., "Intelligent Information Agents" for the Internet, will help doctors in accurately carrying out disease control procedures. Emerging and more mature standards such as "Extensible Markup Language" (XML), "Ontology Web Language" (OWL) and Web Services based on "Simple Object Access Protocol" (SOAP), "Universal, Description, Discovery, and Integration" (UDDI) and "Web Service Description Language" (WSDL), will also help to resolve many software-level interoperability problems. The application developed in this thesis relies on these Web interoperability standards in order to integrate information dynamically.

Acknowledgments

We would like to express our gratitude to the management of Imo State University, Owerri who has supported our research through TETFUND.

References

Shortliffe, E.H. , Perreault, L.E. , Wiederhold, G. , Fagan (Eds.), L.M. (2015) Medical Informatics: Computer Applications in health-care and Biomedicine, Springer, New York (2015), pp. 212-256

Boyi, X., Li, D., Hongming, C., Cheng, X., Jingyuan, H., and Fenglin, B. (2014) "Ubiquitous Data Accessing Method in IoT-Based Information System for Emergency Medical Services, " IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, VOL. 10, NO. 2, MAY 2014.

Rosse, C. , Mejino, J.L.V. (2013) A reference ontology for biomedical informatics: the Foundational Model of Anatomy J Biomed Inform, 36 (2003), pp. 478-500 Article Download PDF View Record in Scopus Google Scholar

Ouwens, M. , Wollersheim, H. , Hermens, R. , Hulscher, M. , Grol, R. (2015) Integrated care programmes for chronically ill patients: a review of systematic reviews. Int J Qual Health-Care, 17 (2) (2005), pp. 141-146

Bianchi, S. et al., (2019) "Biomedical data integration - capturing similarities while preserving disparities," Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., vol. 2019, pp. 4654–4657, 2019.

Marut, B. (2016) Ontology-based Clinical Reminder System to Support Chronic Disease Healthcare. Article in IEICE Transactions on Information and Systems · March 2016 DOI: 10.1587/transinf.E94.D.432 · Source: DBLP

Madhura, J., Dinithi, N., Daswin, S., Damminda, A., Brian, D., Kate, E .W. (2020) A data integration platform for patient-centered e-healthcare and clinical decision support. Research Center for Data Analytics and Cognition, La Trobe University, Victoria, Australia b School of Allied Health, La Trobe University. Victoria, Australia

Chih-Lin, C. (2019) Medical decision support systems based on machine learning. PhD (Doctor of Philosophy) thesis, University of Iowa, https://doi.org/10.17077/etd.o5gmwvxk

Sozan, S. M. (2019) Intelligent System for Identification Heart Diseases. A thesis submitted to the graduate school of applied sciences of near east university

Serdar, U. (2014) Intelligent Systems in Patient Monitoring and Therapy Management. Knowledge Systems Laboratory, Stanford University, 701 Welch Road Bldg. C Palo

Vishesh, V. (2010) Personal health record system and integration techniques With various electronic medical record systems. A Thesis Submitted to the Faculty of The College of Computer Science and Engineering in Partial Fulfillment of the requirements for the Degree of Master of Science Florida Atlantic University , Boca Raton, Florida

Dipti, P. S., Santosh, K. P. (2010) An Expert System for Diagnosis of Human Diseases. International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 13

Ighoyota, B. A. and Sujatha, P. (2017) Fuzzy Based Multi-Fever Symptom Classifier Diagnosis Model. I.J. Information Technology and Computer Science, 2017, 10, 13-28 Published Online October 2017 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijitcs.2017.10.02

Alexander, P., Joseph, L.(2011) "Multi Agent Enhanced Business Intelligence for Localized Automatic Pricing in Grocery Chains" A dissertation submitted in fulfillment of the requirements of the degree of Doctor of Philosophy for the School of Information Technology, Bond University. December 2011. www.fulltext.pdf/ Retrieved on August 13, 2017

David, R., FrancisReal, J. (2020) An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients. Journal of Biomedical Informatics Volume 45, Issue 3, June 2020, Pages 429-446

Augustina, B., Alejandra, C. and Nieves, R. B. (2014) "Ontology-Based Data Integration Methods: A Framework for Comparison", retrieved on July, 2016

Bostjan, G., Vili, P. (2014) Automating ontology based information integration using service orientation. Faculty of electrical engineering and computer science University of Maribor Smetanova ulica 17, 2014 Maribor SLOVENIA

Ali, Z. (2018) A Mapping Approach for Fully Virtual Data Integration System Processes. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 12, 2018

Alto, CA 94304, USA

Taqdir, A., Maqbool, H., Wajahat. A., Muhammad, A., Jamil, H., Rahman, A., Waseem, H. , Arif, J., Byeong, H., Sungyoung, L. (2017) Multi-model- based interactive authoring environment for creating shareable medical knowledge, *Department of Computer Science and Engineering, Kyung Hee University, Seocheon-dong, Giheung-gu, Yongin-si 446-701, Gyeonggi-do, Republic of Korea*

Richter, M.M., Weber, R.O. (2016) Case-based reasoning. Springer-Verlag Berlin

Jagannathan, R., Petrovic, S. (2009) Dealing with missing values in a clinical case- based reasoning system. In international conference on computer science and information technology (pp. 120-4). IEEE.

Asma, A., A. (2011) Decision Tree Discovery for Diagnosis of Type II Diabetes.", IEEE International Conference on Innovations in Information Technology (IIT), pp.303

Nassim, D., Elpiniki, I. P., Jos, D., Hans, C. and Marie-Christine, J., (2014) Clinical Decision Support System based on Fuzzy Cognitive Maps, INSERM UMR_**S** 872, Eq 20, Medicine Faculty, Pierre and Marie Curie University , Paris 6, France