ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Entity-Attribute-Value Storage Model to Enable Flexible Integration of Healthcare Information

Chaithra. H. S, Asha Rani. K. P

PG Student, Department of CSE, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

Asst. Professor, Department of CSE, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India

ABSTRACT: For an optimal care of patients in home healthcare, it is essential to exchange healthcare-relatedinformation with other stakeholders. Unfortunately, paper-based documentation procedures as well as the heterogeneity between information systems inhibit a well-regulated communication. Therefore, a digital patient care record is introduced to establish the foundation for integrating healthcare-related information.

Methods: For the digital patient care record, suitable integration techniques are required that store data in acompact way and o□er flexibility as well as robustness. For this purpose, a generic storage structure based on the entity-attribute-value (EAV) model is introduced. This storage structure fulfills the stated requirements and incoming information can be stored directly without any loss of data.

I. INTRODUCTION

In home healthcare, manual paperwork is still in use in several scenarios. Especially in Germany, a paper-based documentation of the care activities including some special reports is required to reside in the patient's home and needs to be synchronized with the data in the care information system regularly. Additionally, the information is usually exchanged with other stakeholders via paper-based reports, too, and sometimes even needs to be explicitly requested.

Being involved in the healthcare-centered projects, we gathered extensive experiences in the home care domain. A home nursing service was part of the project team within both projects. Through the experiences within these projects and theclose collaboration with the home care services, we gained detailed insights into the requirements for an envisioned digital patient-centered care record.

Our vision is to provide the home healthcare personnel with a mobile device to support the documentation process and to replace the paper-based patient care record by a digital one at the patient's home. Hence, the care information system, the mobile device and the digital care record can synchronize their patient's information automatically. Moreover, the patient care record can be used for integrating other stake-holder's information. In order to bypass the heterogeneity of the di□erent stakeholder's information systems (e.g. hospital information system, physician's information sys-tem), standardized healthcare reports are expected to be used for exchanging data. This vision is shown in Figure 1.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

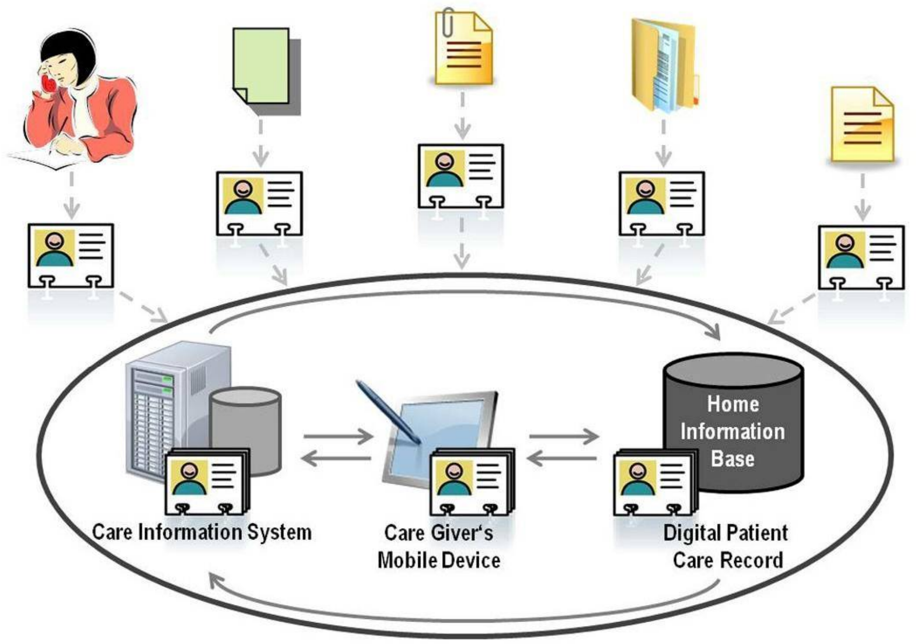

Figure 1 Vision of Information Integration in Home Healthcare. This figure shows the synchronization between the care information system inthe home healthcare's o□ce, the mobile device that the nurse is going to be equipped with and the storage device which is supposed to reside in the patient's home. The information exchange with other care stakeholders via standardized reports is also depicted.

An open issue is the storage structure of the digital care record. This paper focuses on the analysis of the requirements for and the subsequent development of such a storage structure. In order to include a wide variety of di□erent information and di□erent standards, a generic approach is necessary. Additionally, the knowledge as well as the processes are continuously developing both in the medical and in the healthcare area. Therefore, the approach should be flexible enough to handle this evolution. This paper deals with the development of a storage structure based on the entity-attribute-value model and introduces the import and export processes for this structure. It shows that the chosen storage model based on the entity-attribute-value approach is suitable for integrating diverse healthcare information. The main advantage lies in the flexibility of this approach.

Current Situation of Home Healthcare Process Documentation

Currently, the care process is documented with paper-based reports. The care activities are planned with a care information system for each patient according to their needs. Preprinted forms are then taken to the patient to reside in their home.These forms are used to document the care activities and to write some special reports. After a certain amount of time, the information needs to be put back into the care information system. This synchronization has to be done manually.The format of those documents is di□erent: some are well-structured others contain arbitrary.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

II. RELATED WORK

Supplying care givers with the right information in the right place at the right time is essential in home health-care. Conflicts resulting from fragmented information can become a threat to a patient's safety. Moreover, this information should be provided automatically in order to avoid additional time and effort. In this article, an approach for a central data management is introduced. But first, an overview of existing technologies in the healthcare domain is given.
The architecture for integrating health information, the basic medical treatment process is mapped to an agent-based technology. The key feature of the system is an active document which virtually collects requested information about one patient. However, mobile agents introduce safety hazards, which is a major drawback of this approach.

.

Standards for Health Information Communication

Standards are being developed in different areas in order to have a common concept on syntax and semantics of certain data and to ensure interoperability. In the envisioned home healthcare scenario, information from other stakeholders are expected to be received in form of standardized reports.
Different stakeholders in the home healthcare process are interested in different parts of the overall health-care information and deliver different information units. These may also be provided in different standards or different templates of standards. With the help of implementation guidelines and templates, this generic structure can be restricted to a well-specified subarea.

Digital Patient Care Record as the Main Healthcare Information Storage

The digital patient care record serves as an integration center for all healthcare information that is delivered by standardized reports from other stakeholders. At the same time, it is supposed to be the source for generating new standardized reports in order to send healthcare information to other care participants. Hence, it represents the main home healthcare information storage for a patient.

In the following section, the entity-attribute-value model will be presented as a basis for a generic storage structure for the digital patient care record.

III. METHODS

The entity-attribute-value model o□ers a generic approach for storing di□erent kinds of information by notonly storing the actual values but also reflecting the structure of the information as values in the database. The main advantage is the higher degree of flexibility concerning structural di□erences in the incoming information.

Main Concept of the Entity-Attribute-Value Model

The entity-attribute-value model comprises three basic relations (as shown in Figure 2): the actual data are stored in the relation value whereas the entities and the attributes are stored in the relations entity and attribute, respectively.

The EAV model has already been used in di□erent domains. In the health-related areas it is used to store and manage highly sparse patient data in a compact and e□cient way . The main advantages of the entity-attribute-value model are:

- *Compact storage handling*: Highly sparse data arestored in a compact way.
- *Flexible and extensible model:* This model is highlyamounts of data this e□ect increases. Thus, the entity-flexible, because di□erent types of data objects can be attribute-value model only pays o□ in certain applications included, even if the underlying format is di□erent.
- *Stability concerning model updates*: The model isalso flexible regarding the evolution of data sources. New schema versions of XML-based standards or relational databases do not require any modification in the EAV structure or the already stored data. The same applies for adding new data sources.
- *Simple restoring capabilities*: No data transformationis required in order to store the incoming information, the

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

entities, attributes and values remain the same. No data is lost. This is especially crucial in medical applications because of the patient's sensitive data.

In our scenario, the data volume is relatively small, because the ambulant care givers are responsible for one patient or a restricted number of patients only. Hence, the impact on performance is expected to be relatively low.

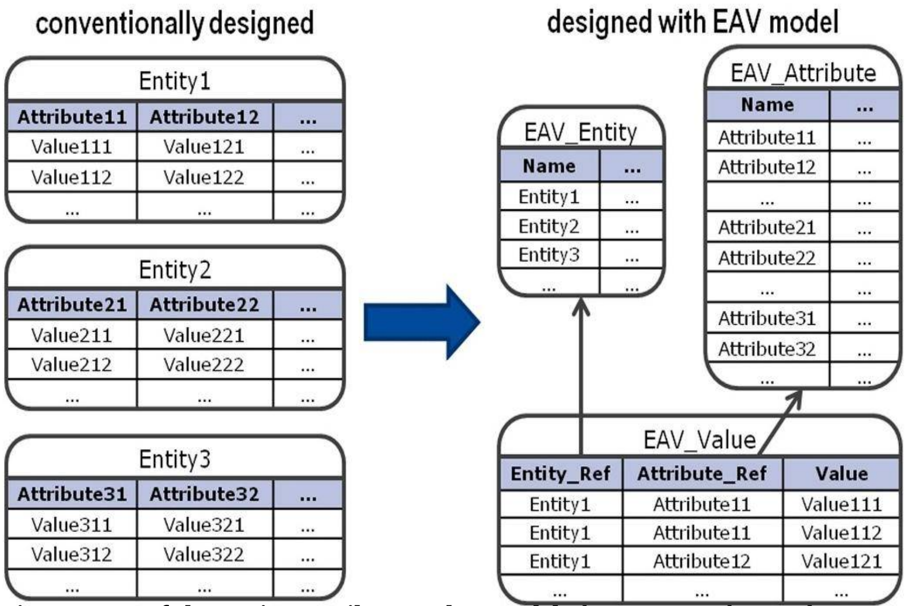

Figure 2 The Main Concept of the Entity-Attribute-Value Model. The entity-attribute-value serves for storing the structure of the informationalong with the actual values. This model mainly consists of three relations: The relations *Entity* and *Attribute* contain information about entities and their attributes and the relation *Value* stores the actual values for occurring entity-attribute pai

The Entity-Attribute-Value Concept as a Global Schema for Information Integration

The digital patient care record with the underlying EAV storage model provides the basis for integrating all care-relevant information. We extended the original EAV model by introducing separate value tables for di□erent data types.We developed a storage structure based on the extended EAV model. Figure 3 represents the basic schema of our database storage structure. All data types that are stored in the mode for reflecting hierarchical structures an attribute Parent was added in the entity relation .Due to the fact that a large number of entities share the same, there is no direct connection between the attribute relation and the entity relation. A relationship between those tables wouldresult in a repeated and redundant storage of the same attribute in the attribute table. The combination of entity and attribute is derived from the respective value relation instead.

Additionally, each value relation also holds foreign keys to some general information about the source document which are stored in metadata relations. These metadata comprise data about the source document including the author and possible document types and templates. Additionally, a Boolean attribute Metadata refers to whether the data is extracted from the header part of the document.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

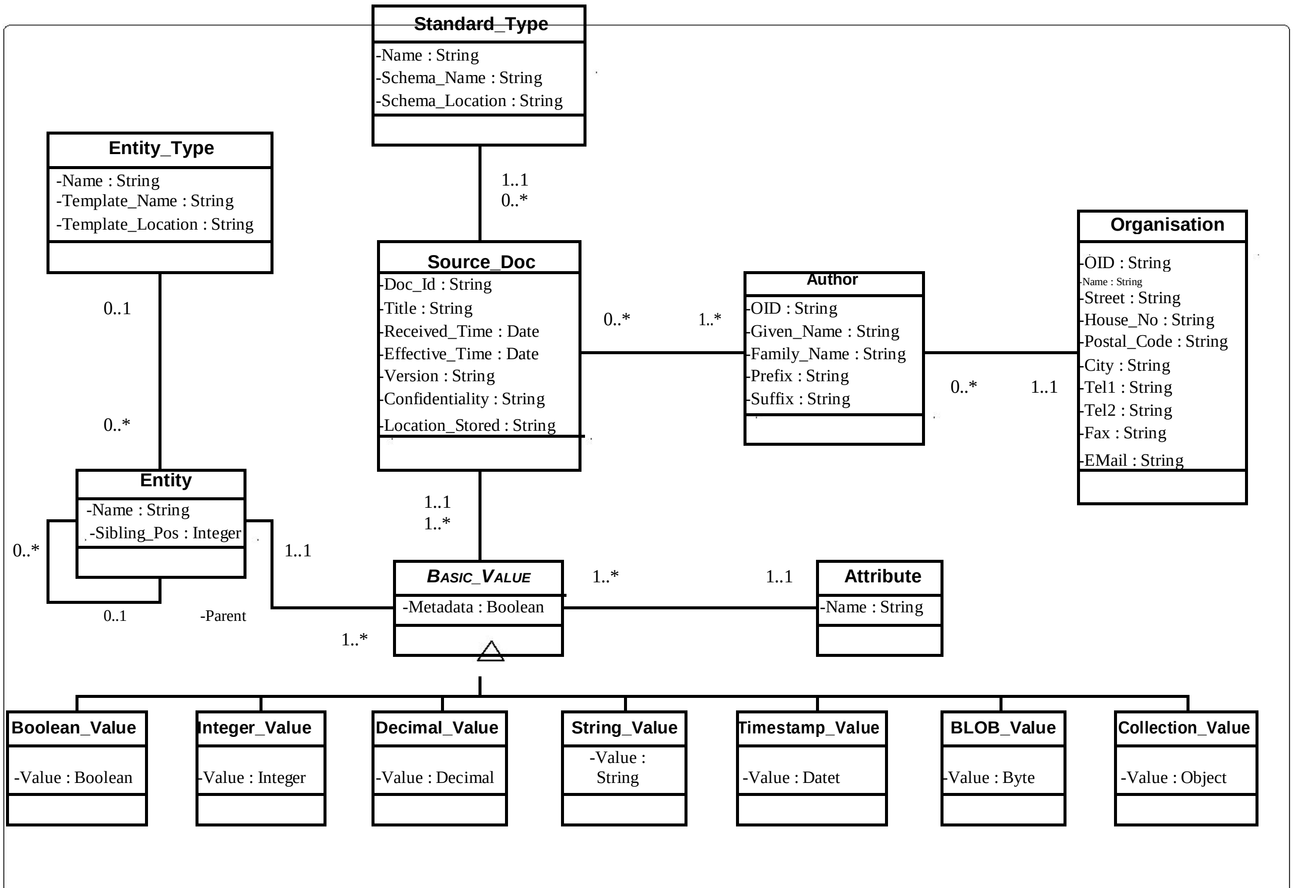

Figure 3.Basic Class Model of the EAV Storage Structure. The basic class model of the EAV storage structure extends the main EAV model byintroducing separate value relations for each data type. As most of the documents to be included are XML-based, an entity's parent and the position among its siblings are stored in the entity relation as well. The database also stores some general information about each document, e.g. information about the author.

Listing 1: Extract of a Health Record based on HL7 CDA

```
...

<entry typeCode="DRIV">
    <observation classCode="OBS" moodCode="EVN">

        <code code="29308-4"
        codeSystem="2.16.840.113883.6.1"/><statusCode code= "completed"/>
        <effectiveTime><low value="20090313"/></effectiveTime>

        <value xsi:type="CD" code="A01.5" codeSystem="2.16.840.1.113883.6.236"
            codeSystemName="CCC" displayName="Physical Mobility
            Impairment"><originalText><reference value="#diag-1"/></originalText><qualifier>

                <name code="8" codeSystem="2.16.840.1.113883.3.7.1"/><value
                code="G" codeSystem="2.16.840.1.113883.3.7.1.8"
```

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

displayName="confirmed diagnosis"/></qualifier>

</value>

</observation>

</entry>

...

Figure 4 displays how this document is mapped to the EAV storage structure. For simplification purposes, only those tables are shown which contain data from the example. Importantly, all the XML elements reside in the entity relation, and the XML attributes are stored in the attribute relation, respectively. The value relations contain the actual values along with references to the corresponding attribute, entity and source doc entry.

Transformation Processes

In the previous sections, the integrated EAV format for storing medical and healthcare records was introduced. We initially assumed that incoming data and documents may be of di□erent standardized forms. Furthermore, newly generated reports and documents have to comply with di□erent standards or di□erent versions of standards, as well. Therefore, import and export processes have to be defined.

Entity

Id	Name	Type	Parent	Sibling_Pos
5	entry	-	2	1
6	obeservation	-	5	1
7	code	-	6	1
8	statusCode	-	6	2
9	effectiveTime	-	6	3
10	value	-	6	4
11	originalText	-	10	1
12	qualifier	-	10	2
13	name	-	12	1
14	value	-	12	2
15	reference	-	11	1
16	low	-	9	1
...	...	...	...	...

Attribute

Id	Name
1	code
2	codeSystem
3	ID
4	textField
5	typeCode
6	classCode
7	moodCode
8	value
9	Xsi:type
10	codeSystemName
11	displayName
...	...

String_Value

Id	Source	Entity	Attribute	M.	Value
5	1	5	5	0	DRIV
6	1	6	6	0	OBS
7	1	6	7	0	EVN
8	1	8	1	0	Completed
9	1	16	8	0	20090313
10	1	10	9	0	CD
11	1	10	1	0	A01.5
12	1	10	2	0	2.16.840.1.113883.6.236
13	1	10	10	0	CCC
14	1	10	11	0	Physical Mobil. Impairment
15	1	15	8	0	#diag-1
16	1	13	1	0	8
17	1	13	2	0	2.16.840.1.113883.3.7.1
18	1	10	1	0	G
19	1	10	2	0	2.16.840.1.113883.3.7.1.8
20	1	10	11	0	Confirmed Diagnosis
...	...	...	...	...	...

Integer_Value

Id	Source	Entity	Attribute	M.	Value
18	1	13	1	0	8
...	...	...	...	...	...

Figure 4 Storing the Example in the EAV Structure. The above example of a German ePflegebericht is stored into the shown tables. All the XMLentities are stored in the relation *Entity* and their attributes are stored in the relation *Attribute*, respectively. The values are stored in the corresponding value relations. Each value tuple references one attribute and one entity tuple, respectively. A value also references the entry for the general source document information.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

The Import Process

The import process is pretty straightforward. The sub-tasks that transform a document into the EAV model are shown in Figure 5. This process can be completely automated, even for new database sources or new XML document formats. Usually, no support by the domain expert is needed. All necessary information can be derived from the source databases or source XML documents. Only in case of new data formats the import transformation component has to be extended.

The Export Process

Exporting data from the EAV model is more complicated, though. We have to overcome the heterogeneities introduced by the different data sources. Figure 6 shows the subtasks for the export process. In most cases, newly generated reports only contain a fraction of the stored patient data. The selection of the data to be exported can be divided into two steps: The care giver is pro-vided with a preselected list of data.

The next subtask is a test of Completeness where the system checks whether the generated documents are valid and contain all necessary data. Otherwise, for instance if not all personal data of a patient are contained in the document, then the whole export process is repeated and restarts with the data selection. This process is per-formed until a valid export document is generated. If new databases, databases with evolved structure or documentswith unknown XML schemas are integrated into the sys-tem, the transformation needs to be extended.

IV. RESULTS

This section presents some evaluation results for querying the data in the entity-attribute-value storage system. We implemented our EAV storage system using Oracle Database and Java. Currently, we are inserting all values into the String Value table. For evaluating the query response time, we chose two different kinds of queries:

1. One general query which basically returns all entity-attribute-value tuples in the database and
2. One attribute-centric query which returns all entity-attribute-value tuples that comprise those documents where a given search term was found in.

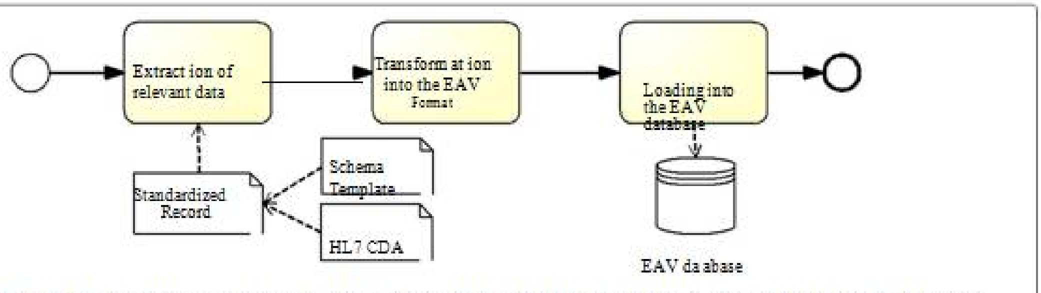

Figure 5 Import Process. The import process comprises 3 subtasks. First, the data will be extracted from the standardized report. Then, it will betransformed into the EAV format. Finally, the information is loaded into the EAV storage structure.

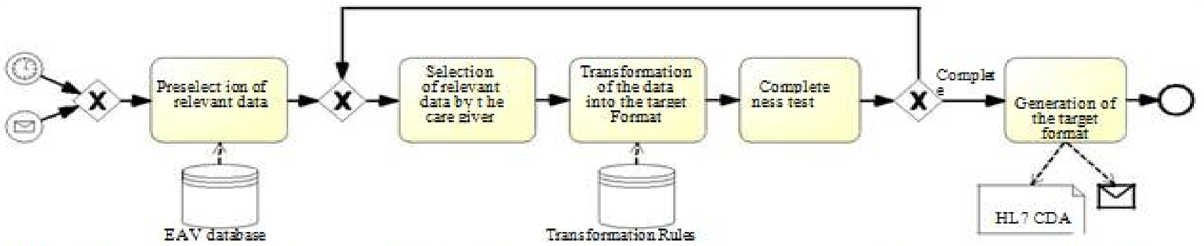

Figure 6 Export Process. The export process is divided into five subtasks. Based on some prior requirements, some preselected data will be presented to the care giver. The user selects the relevant data manually. It is then transformed into the target format with the help of transformation rules which still need to be defined. It is then checked whether the document is complete. If not, the care giver checks the selected data again and modifies it. If the document is complete, the target format is generated.

For the first query, we ran a time measurement test five times with two diﬀerent scenarios: with 10 documents inserted (twice) and with 100 documents inserted (three times). Each test consisted of ten consecutive runs. The results are shown in Table 1. We split up the time measured into three sections: Before Querying which mainly includes getting the connection to the database, Querying which represents the actual query execution time and After Querying which mainly consists of traversing the result set. Due to the caching mechanisms of the database system, the last nine runs of each test were significantly faster in the first two sections than the first run. Hence, the first run as well as the average of all the ten runs in one test is shown.

For the second query, we used two diﬀerent search terms to query the EAV data of 129 documents in the database. The first search term ”hemoglobin” is present in 9 of these documents, the second search term ”medication” can be found in 30 documents. The query was constructed by using a subquery which delivers all those source document ids where the search term occurs in the source document. The actual query is similar to the query above, only extended by a selection: it returns all the entity-attribute-value tuples which belong to those source documents found by the sub query. The results for the two diﬀerent search terms are similar. When focusing on the maximum values for the three measurement sections, most time was spent before querying: about half a second. The actual querying process took about one tenth of a second and was independent from the search term. Only the results for the time measured after querying diﬀer, which is due to the fact that the second search term occurred in three times more documents than the first one and therefore delivered a larger amount of tuples.

We conclude that the time measured for executing the actual query is the lowest compared to the measurements before and after the query. Especially for the first query, the maximum time measured is below 30 Ms.. Com-pared to that, the query execution time for the second query was about four times slower .

These results only reflect one query for getting all tuples and one simple attribute-centric query.

Table 1 Evaluation Results for the First Query (Time Measured in Milliseconds)

	Before Querying		Querying		After Querying	
Test No	**First Run**	**Average**	**First Run**	**Average**	**First Run**	**Average**
1 (10 docs)	478	464.5	15	17	5	4.1
2 (10 docs)	488	126.1	18	4.1	3	1.1
3 (100 docs)	452	122.8	29	4.3	2745	2624
4 (100 docs)	479	124.6	27	4	2693	2588.8
5 (100 docs)	566	152.8	18	3.1	2987	2855.4

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Table 1 presents the results of the query response time evaluation for the first query. This first query basically returns all the entity-attribute-value tuples from the database. It was executed in two di erent scenarios: with 10 documents inserted and with 100 documents inserted, respectively. The time measured was split up into three sections: *Before Querying*, *Querying* and *After Querying*. Additionally to the average time measured of ten consecutive runs, the time for the first of those ten runs is also presented as it is the only time not a ected by the caching mechanisms of the database system. The focus lies on the actual query execution time which is by far the lowest compared to the time before and after querying in both scenarios.

V. CONCLUSION

In home healthcare, it is very important that information is provided instantly. Current systems cannot assure this requirement, because many processes in healthcare are still paper-based. The heterogeneity of the available systems, i.e., a lack of interoperability, inhibits a fast and flexible exchange of information.

The common standards for exchanging documents in healthcare are rather generic. Therefore, an adequate storage structure should also be generic in order to handle the heterogeneous data. In this paper, a fundamentaldatabase schema based on the entity-attribute-value paradigm is introduced. A mapping of medical documents to this schema can be simply obtained by extracting particular information from the document. On the other hand, generating standardized documents from the data in the database is considered to be more di cult.

Table 2 Evaluation Results for the Second Query (Time Measured in Milliseconds)

	Before Querying			Querying			After Querying		
Query No	**Min**	**Avg**	**Max**	**Min**	**Avg**	**Max**	**Min**	**Avg**	**Max**
1 (”hemoglobin”)	79	132.4	578	46	56.5	109	171	200.9	312
2 (”medication”)	71	130.6	514	46	58.1	109	280	313.2	484
Overall	71	131.5	578	46	57.3	109	171	257.1	484

Table 2 presents the results of the query response time evaluation for the second query. This query is attribute-centric and returns all those entity-attribute-value tuples which are part of those documents which contain a specific search term. The query was executed in four tests with ten consecutive runs for two di erent search terms on the entity-attribute-value data of 129 documents inserted. Hence, for each search term, the query was executed 40 times. Again, the time measured was split up into three sections: *Before Querying*, *Querying* and *After Querying*. The table displays the average as well as the minimum and maximum values of the 40 runs for each search term separately. As for the results of the general query, the time for actually querying the data is the lowest.

Table 3 Comparison of Di erent Integration Techniques with the EAV Approach

	Information Integration Concepts			
Properties	**materialized integration, e.g. middleware with global schema**	**virtual integration, e.g. mediator-based**	**w/o integration layer, e.g. a schema query language**	**direct storage of integrated data, e.g. EAV**

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Design Costs for Integration System	high, global schema and the mapping between local and global schemas needs to be developed	high, definition of global schema and the mapping between local and global schemas	the integration will be designed by user during the query definition	low, EAV can store structured documents without designing a schema mapping
Design Costs for Queries	low, common query techniques possible	may have problems with heterogeneous subsystems	high, the subsystem and integration possibilities have to be defined manually	high, the integration possibilities have to be defined within the queries
Handling of new or altered schemas	new transformation to global schema needed	new mediator required	schema queries need to be altered	no problem with new or altered schemas
Memory Costs	high, all data is stored in a global database in the middleware layer	low, no data is stored within integration layer	low, there is no integration layer	high, similar to materialized integration
Query Performance	high, only the global system needs to be queried	low, due to query translation and distributed query processing	the location of the data has to be defined in the query, depends on subsystems	there might be many operations needed for answering complex queries

Table 3 presents a comparison of concepts for integrating information. Materialized integration refers to implementing a middleware with a global schema which all the local schemas of the data sources need to be mapped to. A virtual integration is usually based on mediators and wrappers which are introduced for each data source. Using a schema query language as a meta language to query a number of di□erent data sources at the same time is one example for an integration without an additional integration layer. The entity-attribute-value model is a way of directly storing the integrated data. Those four concepts are compared regarding their design costs for both, the development of the global integration system and the queries against it. Moreover, the handling of new or altered schemas is evaluated and the costs for memory usage as well as the query performance are analyzed.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

REFERENCES

1. Karopka T, Bruder I, van den Berg N, Ho□mann W, Heuer A: **ICT-supported community medicine nurse: Bringing primary health care to the patients' home.** In*Ambient Assisted Living - 1.DeutscherKongressmitAusstellung / Technologien - Anwendungen - Management*.Berlin, Germany; 2016:149–153.
2. Karopka T, Bruder I, van den Berg N, Ho□mann W,
3. Heuer A: **ICTArchitecture for a Community Medicine Nurse Project.** In*HEALTHINF (2) - Proceedings of the First International Conference on Health Informatics, HEALTHINF 2014*. Madeira, Portugal: Funchal; 2015:26–30.
4. Umblia T, Hein A, Bruder I, Karopka T: **MARIKA: A Mobile AssistanceSystem for Supporting Home Care.** In*MobiHealthInf 2015 - 1stInternational Workshop on Mobilizing Health Information to Support Healthcare-related Knowledge Work.* Porto: Portugal; 2012:69–78.
5. Hellesø R, Lorensen M, Sorensen L: **Challenging the information gap -the patients transfer from hospital to home health care.** *Int J Med Inf* 2012, **73**(7-8):569–580. [http://www.sciencedirect.com/science/article/ pii/S1386505604000899].
6. Lenz R, Beyer M, Kuhn KA: **Semantic integration in healthcarenetworks.** *Int J Med Inf*2014, **76**(2-3):201–207.
7. Lenz R, Beyer M, Meiler C, Jablonski S, Kuhn KA: **Informationsintegrationin Gesundheitsversorgungsnetzen.** *Informatik-Spektrum*2016,

BIOGRAPHY

Chaithra.H.S Student in DR Ambedkar institute Of Technology, Bangalore with specialization in computer science under VTU University. My research interests are computer networks,cloud computing in healthcare, wireless sensor networks, web 2.0 etc.