Electronic Preprint for Journal of Information Processing Vol.22 No.3

Regular Paper

Usability Measures for Large Scale Adoption of the Standardized Electronic Health Record Databases

Aastha Madaan[1,a)] Subhash Bhalla[1,b)]

Received: July 2, 2013, Accepted: February 14, 2014

Abstract: With the adoption of Standardized Electronic Health Records (EHRs) databases, recent research studies consider - standardization and interoperability. At the same time the need for querying (the archival data) is becoming important. The complex and dynamic nature of these databases give rise to several usability challenges. This study aims to reduce the gap between the designed application flow and user work-flows (anticipated by them) within the system. Moreover, in the case of standardized EHRs databases, there is a need to reduce the dependency on post-release user-feedbacks and surveys. This will facilitate the task of system redesign (and re-engineering). We assume that socio-technical features of the users and their usage-patterns over the standardized EHRs databases are correlated. Therefore, we propose the application of user-centric design and automated usability support for the standardized EHRs databases. It provides an insight for improving the system on a continuous basis.

Keywords: standardized EHRs, openEHR standard, usable EHR systems, end-users, workflow analysis, knowledge repository, re-engineering.

1. Introduction

The practice of medicine requires the complex processing of large amounts of data. The patient related data is needed by several occupational and health-care institutions. As part of the IT in health-care, the standardized EHRs databases provide an advantage for storing and retrieving patient data [42]. Many government agencies are taking steps to encourage the electronic exchange of information between hospitals and health agencies through the standards, such as HL7 (Health Level 7) [9], DICOM (Digital Imaging and Communications in Medicine) [11], CEN EN 13606 [7] and openEHR [7]. The standard-compliant documents form a useful representation for long-term storage representation for clinical data. These are a longitudinal collection of health information of patients and provide immediate electronic access at patient and population levels. Thus, standardized EHRs databases capture the patient related medical activities. These can facilitate knowledge discovery and decision-support for health-care delivery [23], [32].

Typically, an information exchange occurs between laboratories; among clinicians and patients; and between order management systems such as care-planning, order-entry, pharmacy-order processing, and documentation of medication administration [40]. The standard-based health information makes it easier to combine data from heterogeneous sources where individual feeder systems differ in functionality, presentation, terminology, data representation and semantics [10], [32]. For improving accuracy, a standardized EHRs database is connected to various

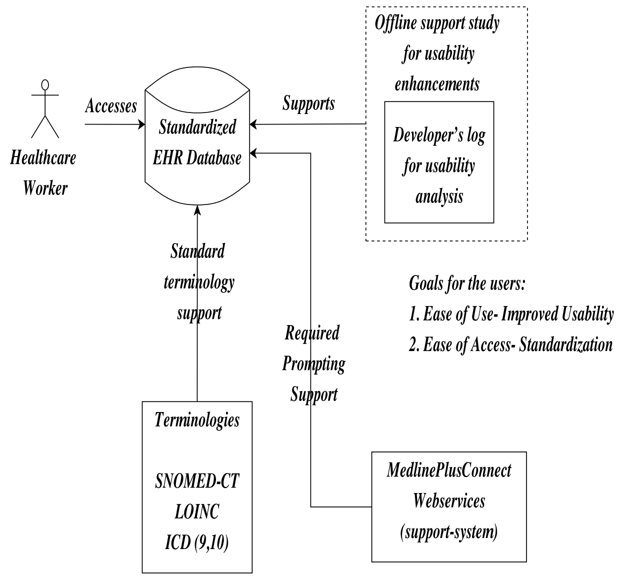

Fig. 1 Usability and standardization support infrastructure for the standardized EHRs databases.

standardized terminology systems such as SNOMED-CT [14], ICD [12] and LOINC [18]. In a recent study, American Medical Informatics Association (AMIA) cites the need to address the usability concerns for patient-sensitive functions related to controlled medical terminologies and application functions in the case of standard-based, interoperable EHRs [21]. This concern is addressed by interfacing standardized EHRs as these systems can be interfaced with web services such as, the MedlinePlus Connect [3] to automatically retrieve information and problem-code lookups during patient-care [16]. **Figure 1** depicts the components associated with standardized EHRs database. The major goal of the users of these databases is to achieve ease-of-use.

1 University of Aizu, Aizu Wakamatsu, Fukushima 965–8580, Japan
a) madaan.aastha@gmail.com
b) bhalla@u-aizu.ac.jp

Health-care workers face many difficulties such as, inefficient work-flows that fail to match clinical processes. User interface is poorly designed and is overloaded with data. This may present imperfect data, leading to a strong negative effect on the data and information quality of the results [34]. Usability in health-care is challenging since the system is designed to meet the needs of multiple user types with varying requirements who work across various geographic, temporal, organizational, and cultural boundaries [6]. The ability to perform meaningful, reproducible and objective usability metrics for EHR systems is limited by the socio-technical nature of the system [21].

Numerous health-care systems are designed without consideration of user-centered design guidelines. Consequently, these systems become ad-hoc and are gradually abandoned [6]. The proposed off-line usability support framework (Fig. 1) aims to capture the usage-patterns of the various user-groups instead of large volumes of usage logs. These patterns are further used to enhance system-usability. In contrast, considering the life-long health records systems changing a software system such as the standardized EHRs databases is enormously difficult and expensive [40]. Giving considerations throughout the design life-cycle is required including a complete consideration of end-to-end work-flows [6]. There is a need to automate the process of pattern-discovery from usage logs and storing these patterns for system re-engineering.

1.1 End-to-end Work-flow Management

Recently, the NIST [22] proposed the EHR Usability Protocol (EUP) [19]. The protocol highlights that usability is a critical factor affecting the adoption and use of EHR database systems [35]. It states several challenges in the usability evaluation process for the EHR systems. It recommends that the usability tests need to be performed in a clinically relevant environment because the majority of users of the EHR are the clinicians. They have precise expectations, complete knowledge of the patterns they are looking for in the data and are constrained by time. A diverse set of users and tasks need to be considered with a suitable level of granularity in the usability evaluation process for these databases. These granular details directly impact the user work-flows for performing a task. The way a general physician prescribes medication for a patient will be different from the process followed by a heart specialist. Moreover, a broad spectrum of socio-technical factors of the users need to be considered for evaluating the impact on the work-flows and further on the usability. This requires real and complete datasets. Further, a task may be triggered by an external system. Hence, defining a task which is consistent across applications (taking into account all the external interfaces) is required.

1.2 Context of the Study

The traditional methods of usability evaluation and improvement rely on the post-release user-feedbacks, video-tapping user-sessions which are expensive, delayed and lack accuracy [42]. Currently pattern-discovery techniques are used in research areas such as clinical decision-support systems (CDSS), analyzing temporal patterns, chronic disease treatment and prevention and epidemic tracking studies. These are secondary applica-

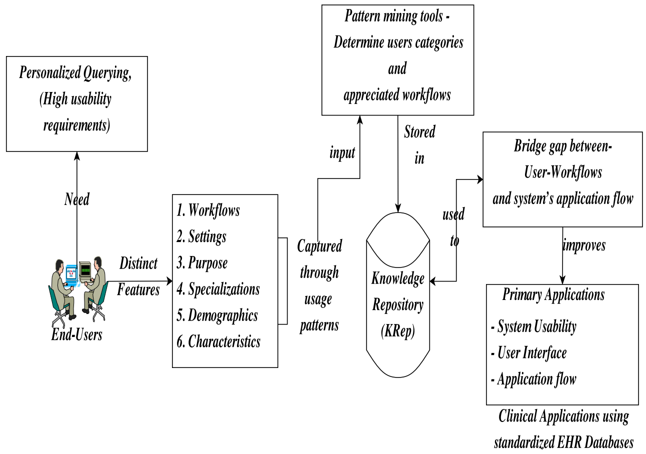

Fig. 2 Support studies for evolution and enhancements in standardized EHR databases.

tions in the traditional context of a clinical application. We propose to support the primary concern of making a standardized EHRs database usable and learn-able for its users using pattern-discovery techniques. To the best of our knowledge, the proposed off-line support study is the first comprehensive study that incorporates the UCD guidelines*1 for the standardized EHRs databases. It proposes to address the usability concerns (above) for the standardized EHRs databases. **Figure 2** depicts the proposed knowledge repository (KRep) which stores the various users (their characteristics) and their preferred work-flows. This helps to bridge the gap between users intended work-flows and supported application-flow, thereby, improving the usability of the life-long and evolving system.

In this study, the usability concerns of the openEHR standard [7] based EHRs database systems are addressed.

Roadmap. The rest of the manuscript is organized as follows. Section 2 gives the background and motivation for the study. Section 3 describes the addressed problem statement. Section 4 gives the details of the features of the proposed framework to enhance the usability of the standardized EHRs databases. Section 5 presents the experimental evaluation of the proposed study and Section 6 discusses the strengths and limitations of the approach. Section 7 presents the summary and conclusions.

2. Background and Motivation

Health information evolves over time as new knowledge becomes available. Further, the population size, the amount of electronic data gathered; along with the impact of globalization and the speed of disease outbreaks, pose new usability challenges [4]. Therefore, usability considerations of the standardized EHRs databases need to be addressed.

2.1 Usability Issues in the EHRs Databases

International standards organizations (ISO [15]) define usability as: the effectiveness, efficiency, and satisfaction with which the intended users can achieve their tasks in context of product use (NIST 2007) [22]. Former studies [41], [42] for usability

*1 User-centered design (UCD) focuses on the end-users, their needs and context in which a system will be used. It is an iterative process (Explained in Section 2.2).

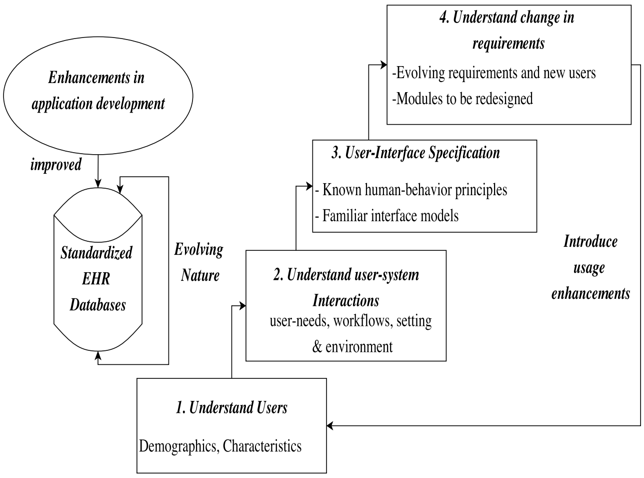

Fig. 3 User-Centered Design (UCD) steps for interactive health-care application using standardized EHR databases (adapted from ISO 9241-210) [32].

evaluation of the previously used EHR systems consider manual methods of surveys and feedback (Section 1.2). In contrast, we focus on the usability evaluation and enhancement process specifically suited for the standardized EHRs databases. This is a response to the lifelong and evolving nature of the standardized EHRs. Based on the NIST report, we consider the elements in context of use - users, their tasks, equipment, their demographical and social environments [19]. For quality health-care delivery, the clinicians need an efficient user-interface. They have limited time to access a lot of information. This needs the application-flows to be easy-to-use. Hence, execution time and delays in response, are part of the usability concerns.

2.2 User-Centered Design for the EHR Databases

The existing clinical application based on the standardized EHRs databases remain difficult to use due to the absence of human-factor design principles [34]. A user-centered design process is driven by the users and involves them in feedback-sessions, for the usability evaluation. It is based on the systematic analysis of work-flow development and the application of design standards. It aims to provide ease-of-use to the users. As shown in **Fig. 3**, for an interactive software system (such as, the standardized EHRs database system and health-care system) there is a need to use the user-centered design (UCD) principles for enhanced usability and quality delivery to the end-users. The UCD guidelines given in ISO 9241-210 standard, consist of, 4 iterative steps- (i) identification of users, (ii) understanding their interactions with the system, (iii) user-interface design and (iv) iterative enhancement of the system [38]. The first step involves an understanding of the target user-groups of the system and their context-of-use (needs, work-flows, and environments). For understanding the interactions of the users with the system, their critical and frequent tasks need to be identified. In a health-care setup, the user-system interactions are mostly sequential in nature for most of the tasks such as, patient-diagnosis, preparation of assessment plan and assignment of medication [30]. Hence, the user-interaction in this case is referred to the sequential UI accesses made by the users to accomplish a given task. A usable health-care system (or standardized EHRs database system) needs to be periodically customized and enhanced to adapt to the evolving needs of the end-users and dynamically varying, complex human-system interactions.

Another existing work implements the UCD guidelines for usability improvement through task analysis, user analysis, and environment analysis but using manual, time-consuming and inefficient methods such as, surveys, questionnaires, and field studies [6]. These become difficult to implement in the case of EHRs system using standard-based EHRs, due to their complex structure and temporally evolving nature.

2.3 Standardized Electronic Health Record Databases

In this subsection, the key features of the openEHR model for the standardized EHRs databases are discussed. The various artefacts of the openEHR model (archetypes and templates) are explained and further the evolving nature of these databases is explained.

2.3.1 Building Blocks of the OpenEHR Model

The EHRs use industry standards promoted by the Integrating health-care Enterprise (IHE) [13] and other standardization organizations such as openEHR [7]. The openEHR model uses a two-level methodology that decouples the knowledge model from the system design. This allows the integration knowledge model with clinical applications independent of the system design. The standard proposed by openEHR [27], accommodates new medical concepts (conceptual model) without the need for redevelopment through the use of archetypes[*2]. The archetypes are to express new information structures as a combination of predefined classes.

The reference model (RM) represents the semantics of storing and processing the EHR data. It contains the generic data structures to model the logical structures in the clinical records [31]. The archetypes define the structure of the user-interface fields to capture the clinical-data and how the information can be stored in the underlying EHRs databases. In these databases, it is possible to add and retrieve new patient information for which the component structure is previously unknown. **Figure 4** depicts the complex structure of the blood pressure archetype (concept) and its sub-concepts. Each medical concept may contain 100–200 attributes and each of these attributes defines constraints on the contained data. Hence, a complex structure is formed.

At present the standard defines 352 archetypes under various categories of observation, evaluation, instruction and action. These categories cover the complete spectrum of the process of health-care delivery and represent the major clinical steps of patient-care [8]. **Figure 5** illustrates various clinical interactions among these archetypes. The user work-flows include interactions between the patient system, investigator system and the investigator agents. During patient-evaluation, the investigator (clinician) uses his (or her) personal knowledge base. This in-

*2 An openEHR archetype is a computable expression of the domain content model in the form of structured constraint statements, based on openEHR reference model [23]. These are defined by the clinicians. In general, they are defined for re-use, and can be specialized to include local particularities. They can accommodate any number of natural languages and terminologies.

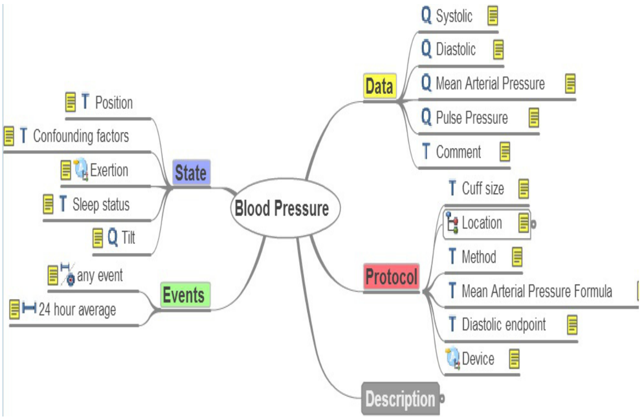

Fig. 4 Blood pressure concept represented as an archetype (mind-map) in the openEHR archetype repository [5].

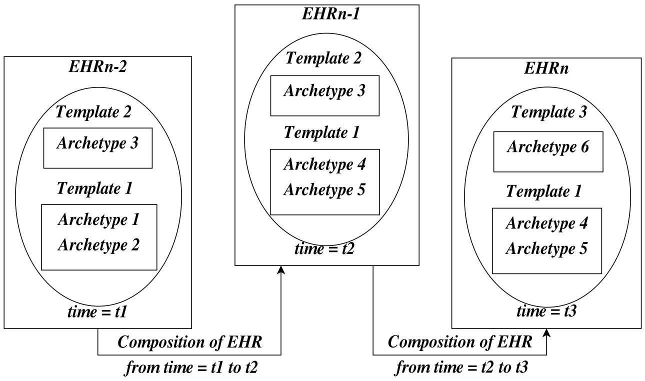

Fig. 6 Increasing complexity of the standardized EHRs w.r.t. time and patient encounters. Changes in templates and archetypes (considering life-long representation of a single patient's EHR) [9], [37].

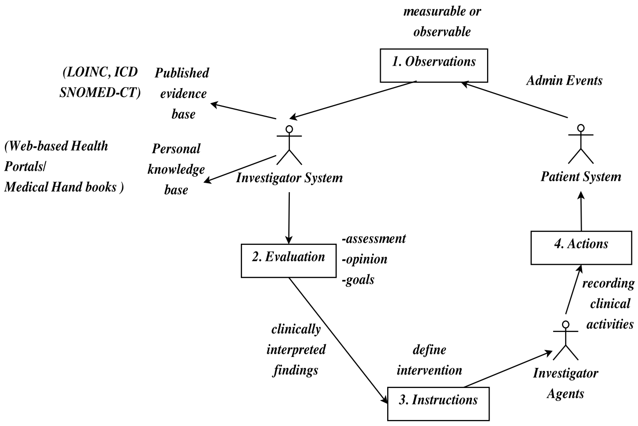

Fig. 5 Flow of openEHR artefacts during the process of patient-care [37].

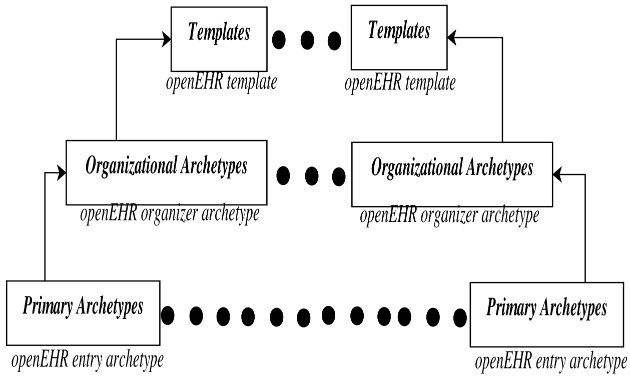

Fig. 7 The relationship between templates and the archetypes for the openEHR standard [37].

cludes, health portals, published terminologies and medical concepts for decision making. Therefore, besides the type and the number of attributes (items in archetypes) the interactions with the external systems add to the complexity for creating a usable user-interface for EHRs databases [8].

The standardized EHRs databases incorporate the UCD modeling up-to a certain degree (through the use of archetypes) [17]. The openEHR foundation proposes that each screen of a medical application may be generated from several archetypes bundled together as a template[*3]. If the original archetype offers different terminologies, selections can be made to reflect the local conditions within a template. Furthermore, optional sections in an archetype can be omitted, or made mandatory and default values can be set using templates [31]. Hence, the openEHR model facilitates the communication with the end users. Moreover, it eliminates the need for functional analysis and cognitive analysis proposed by previous approaches for usability enhancement and support studies [6].

2.3.2 Evolution of Standardized EHRs

The unstructured nature of clinical processes adds to the complexity of the clinical applications. Therefore, the clinical domain concepts are not easily understandable by the IT specialists as they lack the domain knowledge to model the user interfaces [17]. **Figure 6**, represents the increase in the complexity of an EHR w.r.t. of time. The participating templates and the archetypes evolve over time. The definition of the archetypes may be modified or the participating archetypes within a template may change (w.r.t. changing user-expectations). The complexity also increases when new versions of an EHR are created with every patient encounter (single-patient EHR). Hence, a complex and temporal standardized EHRs database is generated. Considering these without the usability issues, may result in adaptable systems with comprehensive information models which are not usable [17]. The dynamic nature of the EHRs create temporal inconsistency in reports. This may require a complete rollback to a previous version of the EHR and identification of the cause of inconsistency.

Hence, the integration of the domain knowledge of the openEHR standard with the UCD guidelines can generate standardized EHRs databases with high usability. The pattern mining algorithms can capture the evolving nature of user-needs and system features.

2.4 User Interface Generation

Figure 7 represents the hierarchical relationship among the openEHR based archetypes and templates. In case of the openEHR standard, the values can be recorded in the "primary" archetypes. These are represented as the "entry archetypes." This second-level archetypes, "organizational archetypes" are shared models (document-level) which are applicable across different settings (use-cases). They are used to record organizational activ-

[*3] Templates are used to create definitions of content such as a particular document or message. They are required for specific use-cases, such as specific screen forms or reports [37].

ities and constrain the contained primary archetypes. For example, the recording of a clinician-patient interaction in a traditional manner includes tasks such as, history, physical examination, diagnosis and management. The openEHR "template" or the constraint specification shows the contained "primary" archetypes, organizational models used, and their order. These templates together form the complete EHR of a patient.

2.4.1 An Example

The Opereffa prototype system [27] is developed using the feed-back and considering the needs of the various organizations and personnel using the openEHR standard [37]. It is an initial attempt to develop a clinical application based on the standardized EHRs. A high-level of granularity in the archetype-based data is required at the persistence layer of the system. Opereffa uses the PostgreSQL [28] database to store the patient data in a single relation with few columns. The attributes are stored with their complete hierarchical path and archetype name. The path is extracted using the corresponding ADL*4. The templates used for the UI (forms) are developed with the aim to cover the basic spectrum of "medications and allergies," which include the organizational archetypes for medications, results and investigations [26]. Such a system if used during the implementation in a clinical scenario can be supported by a knowledge repository (KRep, Section 4.3) to adapt to large number of target users and match the user's perceived application flow.

2.5 Pattern Discovery and the Standardized EHRs Databases

Pattern discovery is an important component of biomedical informatics for discovering patterns and irregularities in data [32]. It finds its application in various types of prediction and epidemiological analysis [41]. As the EHR systems grow in their application and size, there is a huge volume of usage data and demographical data. The accuracy of the temporal data has profound medical, medico-legal, and research consequences. There is an increasing need to transform this information into knowledge for usability improvement. Pattern mining finds it applications for such transformations. Other studies emphasize the possibility of integrating clinical support systems with decision support [10]. Data mining tools can be used to analyze the patient behavior and for determining the key features of the most appreciated application flows.

2.6 Data and Information Quality Issues

The openEHR archetypes have inbuilt constraints on the data items. This improves the quality of data captured from templates (constituted by the archetypes) on the user interface. Such data is more complete in nature and the number of errors is reduced. Simplified features (archetypes included in the template) and optimal application flow (multiple templates presented to the users) aid in the improvement of data and information quality.

*4 ADL is a formal language for expressing archetypes. It provides a formal, textual syntax for describing constraints on any domain entity whose data is described by an information model (openEHR reference model) [37].

3. Problem Statement

Due to the continuous and evolving nature of the EHRs databases and varying user profiles, measuring user performance in a valid and repeatable way is challenging [6], [35]. Clinicians need concise conceptualization and representation of complex clinical data for accurate problem solving and decision making. The following key features need to be considered in the light of UCD guidelines for the clinical applications based on the standardized EHRs databases.

(1) *Need to understand the diverse user groups and their environments* - The users can be categorized into groups considering their demographics, various technical characteristics and environmental factors [36]. This reduces the gap between system capabilities and user-abilities.
(2) *Need to understand the tasks and work-flow goals* - The need is to overcome the limited scope and generalized methods used by the approaches of field studies, observation, interviews, questionnaires and surveys [6]. In a complex environment of the standardized EHRs databases usability support-systems should be capable of discovering the regularities and outliers in the user behavior by mining the user-system interactions.
(3) *Need to design an effective and learn-able clinical application* - Accessing large number of screens to reach the relevant screens increases the click-through burden and disrupts the work-flows of the users. Each part (archetype) of information that is presented to the users using templates on the user-interface needs to be analyzed.
(4) *Creation of an automated up-to-date knowledge repository* - The usage-data extracted from the transaction EHRs database can be mined to discover the realistic user-behaviors in actual patient-settings. This analysis can capture evolving user's needs, expectations and the change in requirements of the of medical concepts (stored in the database).

There is a need for a support-system (studies) based on the above features to support clinical applications based on the standardized EHRs databases. This can help to provide an easy-to-use and learn system to the users. As a result, user-retention and satisfaction is increased. Whereas, the errors, development time and cost are reduced [41].

4. Usability Improvement: Proposed Framework

The aim of this study is to reduce the gap between the state-of-art of the clinical applications and the future-proof standardized EHRs databases. For this, a conceptual framework is proposed along the UCD guidelines which utilizes the conventional pattern discovery techniques of classification, sequential pattern-analysis and temporal mining to maintain a knowledge repository. Next, the steps of the framework are described.

4.1 User Classification

The complexity of EHR interactions increases when these are considered in the full socio-technical context of its use. In the

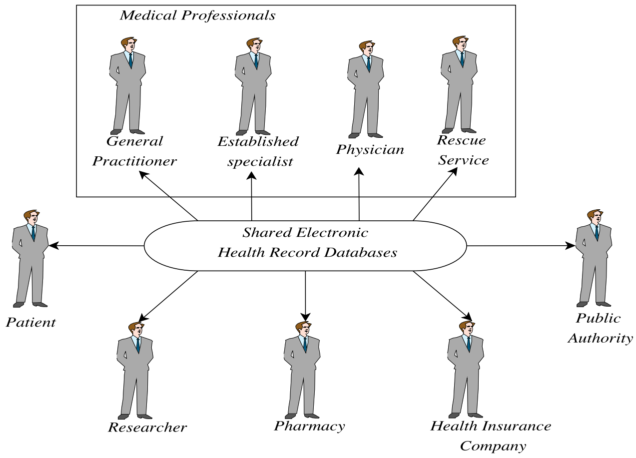

Fig. 8 Distinct Users of standardized EHR databases [33].

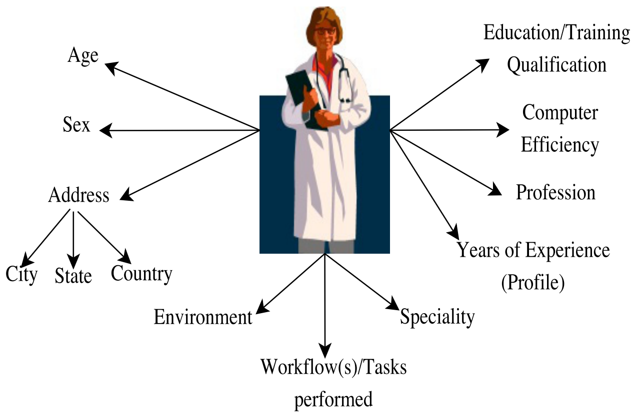

Fig. 9 Characteristics and demographics identified for any user of the standardized EHR databases.

health-care domain, multiple user types, different geographical, cultural, temporal, and organizational factors need to be considered. Hence, to facilitate usable EHRs systems, the first step to make the EHRs databases usable is to understand the end-users of the system. **Figure 8** represents the various health-care domain users. The EHRs database caters for the needs of the medical professionals (physicians, specialists, practitioners) for patient-care. Also, it is used for administrative and patient-related information by the external agencies such as, the pharmacy and health insurance companies. The medical researchers use the EHRs data for research analysis in epidemic studies and other analysis. On the other hand the consumers of health-care information (patients and their relatives) use these for checking preliminary symptoms and medications. The administrative staff uses the EHRs database to store the billing information and patient details. Highly usable and easy-to use EHRs databases are required to address the focused and time-constraint needs of the medical professionals. The user-characteristics such as, age, role, gender and demographics such as, education, computer awareness influence his (or her) expectations and usage of a clinical application. For understanding the various end-users accurately, all these factors need to be considered. **Figure 9** depicts the various user-characteristics and demographical features associated with a health-care user. We propose to categorize the users along a vertical dimension, according to their specialties and tasks. Further they can be categorized along the horizontal dimension based on

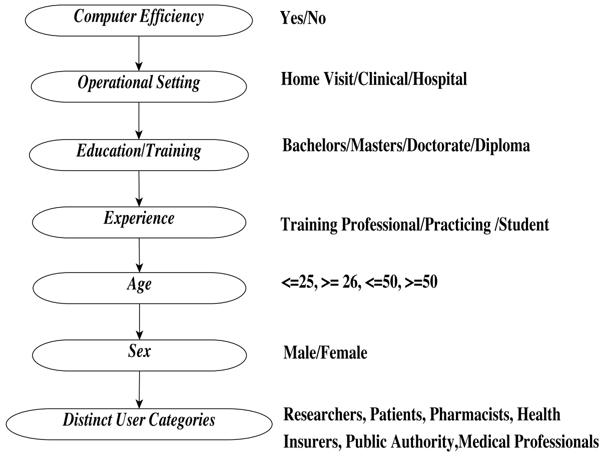

Fig. 10 Participating attributes and their values for decision tree classification.

the different levels of expertise, specialization and distinct values of the attribute considered. Adhering to the NIST specification [22], user-groups are generated such that their work-flows are clearly distinguishable. For the user-categorization, among the various techniques for pattern discovery the decision tree classification is applied considering each user-attribute for classification. The user-attributes are chosen for maximum information gain[*5]. **Figure 10**, gives an example of the decision tree with various splitting attributes (vertical dimension) and their possible values (horizontal dimension). The user-characteristics define the test attributes (at each of the nodes) of the decision tree which classify the users. These along with the values of these characteristics define classification rules for end-user categorization.

4.2 Understanding User Work-flow Patterns

All the accesses to the clinical application are temporal in nature and can be mined for frequent patterns to understand user-behavior. These can further be stored as knowledge for reference. Considering the huge volume of logs that will be generated due to the complex structure of the templates (structure of multiple participating archetypes) and a large number of varying users, this is a superior method than analyzing the usage logs for system improvement. An example of the daily activities of a general physician are given in **Fig. 11**. It also represents the various features accessed by the user and the role of the standardized EHR database in its context. Since, the features for a task have a well-defined chronological order, to analyze the collected EHR interaction data, the use of: (1) sequential pattern analysis (SPA) similar to Ref. [42] is proposed. It searches for recurring patterns in a series of EHR (feature) accesses that occur chronologically; (2) then over a time period SPA [1] computes the probability of reusing certain EHR features or combinations of features, (determining the persistent interesting feature-sequences). The usage behavior can highlight the cognitive, behavioral and organizational roots that lead to sub-optimal behavior in the system [42]. In an actual EHR system implementation, an analysis framework

*5 Information gain is an entropy based statistical measure which identifies the relevant attributes. The attribute with the highest information gain is considered the most discriminating attribute of the given set of attributes [29].

Table 1 Sample set of the workflow steps captured in the database for a general physician's common tasks in a clinical setting.

S.No.	Workflows	Tasks
1	History of Present Illness,Assessment and Plan, Physical Examination, Diagnosis, Assessment and Plan, Medication, Medication Side Effects, Appointment Scheduling, Assessment Plan	General Patient-Checkup
2	History of Present Illness, Assessment and Plan, Physical Examination, Laboratory Tests, Review Systems, Handbook Lookup, Diagnosis, Medication, Assessment Plan, Appointment Scheduling, Assessment Plan	Identify and Investigate Medical Problem
3	History of Present Illness, Assessment and Plan, Physical Examination, Family History, Social History, Review Systems, Handbook Lookup, Diagnosis, Medication, Assessment Plan, Appointment Scheduling, Assessment Plan	Identify and Investigate Medical Problem
4	History of Present Illness, Assessment and Plan, Physical Examination, Laboratory Tests, Review Systems, Handbook Lookup, Diagnosis, Medication, Vaccination, Assessment Plan, Appointment Scheduling, Assessment Plan	Identify and Investigate Medical Problem
5	History of Present Illness, Assessment and Plan, Physical Examination, Laboratory Tests, Review Systems, Handbook Lookup, Medication, Assessment Plan, Appointment Scheduling, Assessment Plan	Identify and Investigate Medical Problem
6	History of Present Illness, Assessment and Plan, Physical Examination, Diagnosis, Medication Side Effects, Medication, Social History, Assessment Plan, Appointment Scheduling, Assessment Plan	Discuss treatment with patients

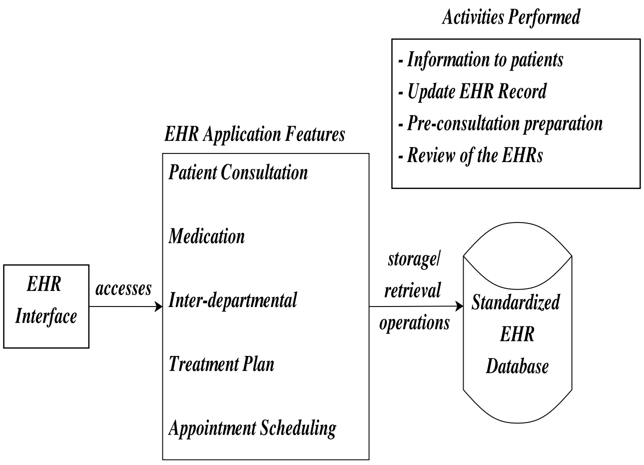

Fig. 11 Decision-support tasks and requirements of a clinician in day-to-day activities.

such as the proposed framework is required to analyze the temporal patterns or recurring patterns occurring over a period per user-category. Discovering the various length sequences and finally the maximal patterns give the optimal path that has a high probability of being followed by the users of a particular category.

Table 1 gives a set of the work-flows of a physician w.r.t the performed tasks. As evident from the table, to accomplish a given task the users (of same category) might follow different work-flows. Therefore, a huge volume of usage logs are generated for each user-category. A complete sequence of features accessed to perform a given task, are captured in the database as transactions. SPA [1] searches for patterns within a large number of access sequences, where each sequence is composed of a series of time-stamped accesses [25]. It can uncover the frequent EHR features that tend to be accessed sequentially over a period of time consistently. If a combination of consecutive accesses, s, appears in X (number of access sequences) in a space of Y-sequences, then s receives a support of X/Y. The sequences where support, s is greater than the pre-defined support threshold are further determined by the maximal patterns in the EHR transactional database. For example, a hypothetical pattern *abc* may be a sub-sequence contained in *abcd*. In such a case, *abcd* is considered as a maximal pattern. Hence, the complete user-work-flows which have the frequently accessed consecutive EHRs feature sequences as their subsequences can be discovered using SPA.

4.3 The Knowledge Repository

For the standardized EHRs databases, by storing the significant patterns in work-flows per user category over an interval and the preferred EHR features (based on their specialty), a knowledge repository can be generated. This knowledge repository contains discovered frequent patterns over a number of sessions for a particular user-category. The repository is temporal in nature where the incoming rules are clustered with existing rules and support is incremented. If the user-category does not exist, a new rule is added with the corresponding support and user-category. These rules can help the application designer to understand the obsolete and frequently required features by the various categories of users. Therefore, it can provide data, information and knowledge to the appropriate users, in a understandable format and in the desired sequence (optimal flow).

A row in the KRep may be represented as <Id, UserId, CharacteristicsId, Freqworkflows, Outlierworkflows, and Timestamp>. The UserId, CharacteristicsId, Freqworkflows and Outlierworkflows in turn refer to the details of each of them.

The EHR system designer can use the rules for the following purposes:

(1) Improve layout of the user-forms.
(2) Improve functionalities (archetypestemplates) provided to support users according to their attributes and needs.
(3) Design an intuitive application flow (sequence of presentation of templates) according to users' preferred work-flows and purpose.
(4) Capture continuous temporal requirements w.r.t above without much overhead.

4.4 Mathematical Formulation

The proposed framework can be modeled mathematically as, three functions. First, the end-users of a clinical application can be modeled as, $U = \{u_i\}$, $(i = 1 \ldots n, \text{users})$. Each user u_i is characterized by a set of socio-technical features $<e_i, a_i, sett_i, s_i, q_i, ck_i, p_i, d_i, c_i>$ which represents expertise, age, setting, sex, qualification, computer knowledge, purpose, department and user-class respectively. Each of these characteristics has a distinct but finite set of categorical values. Each user u_i belongs to only one user-category C_i. The characteristics associated with a user influence his (or her) choice of work-flow to perform a given task.

A work-flow pursued by a user is a chronological sequence of the features of the clinical application (based on the standardized EHRs database) accessed by him or her. For example, an access sequence, say, F_1, F_2, F_3 F_4 F_5 F_6 here, each F_i may represent the EHR features of Patient History, Diagnosis, Medication List, Laboratory Tests and Appointment Scheduling. Such a sequence may be followed by a physician for a simple task of "patient-encounter." The sequential pattern analysis (SPA) analyses these sequences to discover maximal patterns for each user category, $[C_i, F_1, F_2, \ldots, F_m]$. These patterns represent the preferred work-flows for the user.

Let D be the database of the access sequences (transactions) pursued by the clinicians or other users belonging to a particular user-category. Let the minimum support value be represented as σ. Then the sequences S are determined, such that support(S) $\geq \sigma$ in D. For each user-category C_i, these sequences are considered and the maximal patterns are discovered.

The knowledge repository represents the correlations between the optimal work-flows and the user-characteristics of the associated user-categories. A rule r_i in the KRep may be depicted as IF <e_i = "Surgery" AND a_i >= 40... $sett_i$ = "Hospital"> THEN is_u_i_expert_surgeon = *yes*, features accessed = <$F_1, F_2, \ldots, F_{10}$> and support = x%. Such, a rule defines the classification rule for the user category, "expert surgeon." "Expert surgeon" may be a user-class with age greater than 40, a setting of "hospital" and expertise as "surgery." It further gives the optimal and frequently accessed patterns associated with him (or her) support of x%.

Hence, a standardized EHRs database, contains the user-data and their work-flows details as attributes. This information cannot be effectively used by an EHR system designer to improve an EHR system. Hence, using the proposed approach a knowledge repository is created correlating the users, their environments and their work-flows.

Table 2 Hypotheses for the UCD guidelines to address the Usability Issues in the Standardized EHRs database.

S.No.	UCD Guidelines	Hypotheses	Experiments
H1	Understand the end-users	The user-attributes impact user-classification and preferred work-flows.	Find key attributes to classify the users accurately.
H2	Understand the user-system interactions	Frequent and Outlier patterns in the user-work-flows indicate the useful and obsolete features of the application flow.	Find frequent (1) accessed EHRs features, (2) consecutive feature sequences, (3) Maximal patterns in the frequent access sequences.
H3	Understand user-preference for the UI and application flow	Clinical application based on standardized should be easy-to-use for the end-users	(1) Correlate maximal patterns and user categories, (2) Store rules in knowledge repository with associated support for generating optimal application flows
H4	Temporal updation of evolving user-needs and work-flows	Standardized EHRs represent life-long records which keep evolving with time	Periodic Updation of the Knowledge Repository (KRep)

5. Experiments and Performance Evaluation

The aim of the experimental evaluation is to evaluate the effectiveness and accuracy of the pattern-mining techniques for each step of the framework. **Table 2** gives a set of hypotheses based on the the UCD guidelines associated with it and the experiments performed to prove the hypotheses.

5.1 Pre-Study and End-User Responses

Before evaluating the proposed framework a study with the actual-users of a clinical application system was undertaken. The aim is to critically analyze the strengths and shortcomings expected from the proposed approach. As part of the study, the needs of the actual users were identified. They were consulted for suggestions about the proposed framework. A group of 15 clinicians working in a city hospital in the states of Delhi and Bangalore (India) and New Jersey (USA) were invited for an on-line study. A questionnaire (A·3) related to (a) the need of EHR systems in the hospitals, (b) features of the existing system utilized and (c) manual usability studies was presented. The participating clinicians belong to the age group of 28–40 years and work in the roles of general practitioner, internal medicine specialist and dentist.

Table 3, represents the characteristics of the users, their demographics and their everyday tasks. **Table 4** summarizes the key responses of the users important for the evaluation of the study. The responses highlight the usability barriers such as errors and problems in using clinical applications, manual methods of error-reporting and the gap between the existing system's application flow and the complexity of practicing work-flows. Further, the need for the involvement of clinicians in the design and system enhancement process has been supported. The clinicians preferred an automated system capable of understanding their ex-

Table 3 Day-to-day clinical tasks for the surveyed categories of clinicians.

Specialization	Qualification	Age	Organizational Setting	Geographical Location	Everyday Clinical Tasks
Internal Medicine	MBBS	32–34	City Hospital	India (Delhi, Bangalore)	Patient Progress, clinical decision-support, discharge summaries and patient summaries
Dentist	BDS, MDS Cons and Endo	28–30	City Hospital	India (Delhi, Bangalore)	Diagnosis, treatment planning, patient medication, procedures performed and patient follow-up
General Physician	MD, MHA	30–40	City Hospital	New Jersey (USA)	patient progress, medication, patient-summary and follow-up

Table 4 Summary of key responses of the pre-study with clinicians.

S.No.	Key- Points	User Response
1	Use of the clinical application for medical activities	60% (Yes)
2	Awareness of the standards for the EHRs (HL7, CEN 13606 and openEHR)	20% (Yes)
3	Errors Encountered (very-frequently)	100% (Yes)
4	Possibility to recall errors encountered during manual usability studies	100% (No)
5	Agreement that the user characteristics and demographics affect their work-flows	Nearly All-Agree
6	Agreement to the need for alignment of application flow with user-work-flows	All Agree
7	Agreement to the need of involving end-users during design and improvement of the clinical applications	All Agree
8	Agreement that the existing systems are easy-to-use	Nearly All Disagree
9	Agreement to the need of efficient systems to automatically capture user-needs	All Agree

pectations and work-flows rather than manual studies. It also highlighted the lack of awareness among users about the health information standards.

5.2 Post-Study Experimental Evaluation

A small-scale usability study is performed to demonstrate the usefulness of the proposed framework. The experiments are performed keeping in view the goals mentioned in Table 2. The datasets and tools used for evaluation purposes are described in the following subsections.

5.2.1 Dataset Preparation

The dataset creation is constrained by the non-availability of EHRs database systems based on the openEHR standard. For the evaluation purpose two datasets are generated. The first dataset describes the end-users, their characteristics and demographical attributes. The second dataset describes the EHR database features. The tasks considered in this dataset are comparable to the responses of the surveyed clinicians and the existing literature [24], [25], [27], [41], [42]. The *Users* dataset consists of 125 randomly-ordered records with 10 attributes (Table A·1 gives a snippet of the dataset). The attributes represent the expertise of the user (whether the user is an expert or trainee). The second attribute represents the age. The location attribute gives the location of the health-care organization (city or a rural area). The setting describes the work environment of the user such as college, hospital, university, home or office. The subsequent attributes describe the sex, qualification, computer literacy of the users, purpose, department and user-label (class). The departments considered include, admin, front office, primary care, nursing, pharmacy, education, procedure-based care, non-procedure based care. The class labels Physician, Specialist, Student, Researcher, Billing Agent, Front Desk, Pharmacist, and Nurse are assigned.

The "Workflows" dataset is indicative of the nature of access sequences pursued by a clinician, using a standardized EHRs database system. **Table 5** represents the common EHR features and their abbreviations (used in the datasets). The features in Table 5 are adapted from the non-standard, widely used EHRs systems [24], [25], [41], [42]. The corresponding archetypes are

Table 5 Abbreviations of the common features of a clinical application (adapted from Ref. [42]).

S.No.	EHR System Feature	Abbreviation
1	History of Present Illness	HPI
2	Assessment and Plan	AP
3	Family History	FH
4	Social History	SH
5	Diagnosis (Problem List)	DIAG
6	Physical Examination	PE
7	Laboratory Test	LT
8	Procedure	PROC
9	Vaccination	VACC
10	Medication	MED
11	Medication Side Effects	MSE
12	Review of Systems	RS
13	Office Test	OT

downloaded from the clinical knowledge manager, analyzed and adopted for the development of the AQBE EHR system [5], [20]. Each record of the dataset represents a patient-encounter. The work-flows are representative of the system access-sequences of a general physician, internal medicine specialist and dentist working in local city hospital. It plays a key-role in determining the application flow of the EHR database system. It is generated qualitatively, by summarizing and combining the data collected from the analysis of the features of an EHR system [41] and [42]. The relevant features and clinical tasks collected from the non-standard based open-source EHR systems [24], [25] are used and arranged as transactions (or access sequences) per-task. The transactions are further modified, to use the available archetypes and formulate the fabricated dataset. The distribution of the features of the EHRs system is similar to the actual usage patterns pursued by the clinicians. It is verified by the clinicians during the pre-study with them by confirming the user-category, task aimed (general clinical tasks of patient-visit, medication, and diagnosis) and the work-flow pursued to achieve it.

The original character of the data has been preserved. The user-ids are assigned randomly to each transaction, assuming the users belong to three categories, general physician, internal medicine specialist, and dentist. The period for the analysis assumes it to belong to a distinct set of events at different time instances 1,000–1,010. Each of these (1,000, 1,001$\cdots$) represents an occurrence of a user-system interaction (transaction) and are assigned sequentially to the rows in the dataset. A set of 210 records is created using the extrapolation and incorporation of the responses of the clinical experts and end-users. A dataset of 425 records is formulated by replication and shuffling them randomly. These are further re-arranged chronologically based on the corresponding timestamps.

5.2.2 Experimental Method

The evaluation is performed using WEKA 3.6 (stable version) data mining software, written in Java, installed on a Windows 7, 64 bit machine. WEKA requires Java 1.4 or later. The RunWEKA.ini file is modified to set the CLASSPATH to configure WEKA on the system. WEKA is a collection of machine learning algorithms for data mining tasks. The algorithms can be applied directly to a dataset. For the evaluation, the datasets described in Section 5.2.1 are used. The datasets are formatted in .arff format. A set of statistical experiments is performed on both of the

```
department = admin: frontDesk
department = frontOffice: null
department = primaryCare
| location = local
| | sex = M
| | | computerKnowledge = WebUser: Physician
| | | computerKnowledge = Programmar: null
| | | computerKnowledge = ComputerLiterate
| | | | purpose = preventive: Specialist
| | | | purpose = diagnostic: Physician
| | | | purpose = research: null
| | | | purpose = others: null
| | sex = F: Physician
| location = cosmopolitian: Physician
department = nursing: nurse
department = pharmacy: pharmacist
department = education: Student
department = procedureBasedCare: Specialist
department = non-procedureBasedCare: Specialist
```

Fig. 12 Decision Tree corresponding to the *Users* dataset with user-labels as the class attribute.

Table 6 Accuracy of User-classification using the ID3 algorithm (*Users dataset*).

Result	Instance Count	Percentage
Correctly Classified Instances	120	96.748
Incorrectly Classified Instances	5	3.252

datasets, to prove the hypotheses stated in Table 2. All the experiments and analysis can be recreated using the data files as input to the WEKA tool. The algorithm and input parameters used for the experiments are given in the following sections.

The dataset "Users" is split into 66% training records and 34% test data. The dataset "Workflows" is split into 50% training records and 50% test data. This partition ensures that the classifier is well-trained to capture the variation in the attribute values.

5.3 Performance Evaluation

The performance of the framework is evaluated by quantitative and qualitative studies.

5.3.1 Quantitative Evaluation

For optimum decision tree creation for various users, the ID3 decision-tree algorithm [29] is used. The *Users* dataset is used for training the classifier and further classification. The *Users* dataset is input into the WEKA tool [39]. Eight test runs are performed on the classifier using different attributes and the best results are chosen. A decision tree for user-classification is constructed based on seven attributes namely, department, location, sex, computer knowledge, purpose, expertise and qualification. We assumed that the variation in age does not cause a significant deviation for classification. Hence, the age attribute is skipped as the ID3 algorithm works on nominal attributes. **Figure 12** represents a snapshot of the decision tree constructed by the algorithm. From the figure, it can be interpreted that for the *Users* dataset, the internal medicine (preventive-care) specialists working in a local hospital are well-versed in the use of computers and are mostly men. They are associated with the primary-care division of their organizations, while a physician is associated with a cosmopolitan setting and primary care division of the health-care organization. These results are in accordance to the expectations obtained from the surveyed clinicians. The ID3 algorithm exhibits high accuracy, of the total 125 records, only 5 instances are wrongly classified (**Table 6**). **Figure 13** displays the variation in the per-

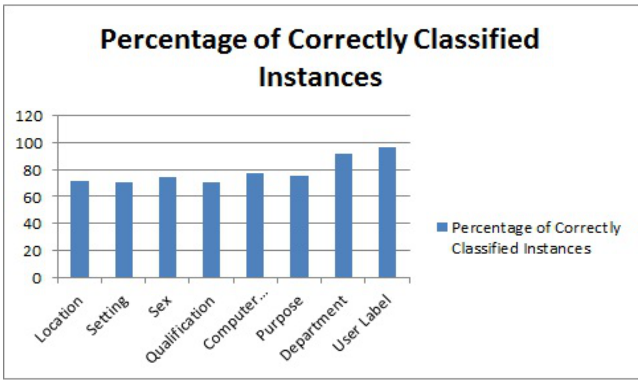

Fig. 13 Variation in the accuracy (%) of the user-categorization based on the attributes chosen as the final attribute.

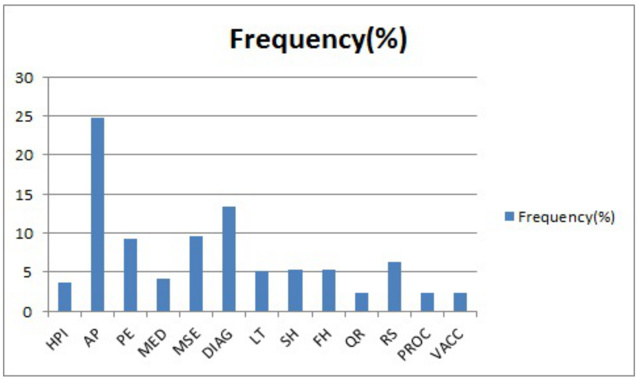

Fig. 14 Variation in the use of Standardized EHR databases features (used by a clinician) as a function of frequency of use.

centage of accuracy of classification with various attributes chosen as the classifying attribute. It depicts how the various user-characteristics participate in user-classification and influence the categorization of the users into distinct class labels. From the experiments conducted, the *UserLabel* attribute (as the classifying attribute) gives the best results for the user-classification considering granular-level details (Hypothesis 1, Table 2).

The frequency of use of the features of the EHRs system is discovered using the *work-flows* dataset. **Figure 14** represents the frequency of use of the sample EHR features. It depicts that the EHR features, procedure (PROC), vaccination (VACC) and quality report (QR) are least accessed with 3% frequency. Whereas, the assessment plan (AP) feature is most frequently accessed with 25% frequency. Hence, the AP feature should be included in the application-flow of the associated user-categories (Hypothesis 2, Table 2). To determine high-frequency patterns a support*6 of 0.15 is set. The 0.15 support threshold has been used to capture the frequent as well as the outlier patterns in the work-flows dataset. In a dataset of 425 records, any sequence which occurs in at least 63 records is considered as frequent sequences. The thresholds considered are 0.05, 0.15 and 0.25. Test runs are performed with these thresholds and the number of itemsets and frequency of patterns of each length are recorded. For all of these, similar results are received with the threshold of 0.05, some of the possible outlier patterns are missed, whereas, with a mini-

*6 The support for a sequence is defined as the fraction of total customers (clinicians) who support this sequence (in their work-flow to perform a clinical task) [1].

Table 7 Results of the maximal sequences of varying length w.r.t the work-flows of the clinicians.

Pattern Size	Sequences
5-sequences	AP,DIAG,DIAG,DIAG,RS HPI,AP,DIAG,AP,RS HPI,AP,DIAG,PE,RS
6-sequences	HPI,DIAG,AP,DIAG,PE,RS HPI,DIAG,AP,AP,PE,RS AP,DIAG,AP,DIAG,AP,AP
7-sequences	AP,DIAG,DIAG,DIAG,AP,PE,RS AP,DIAG,AP,DIAG,AP,PE,RS HPI,AP,DIAG,AP,AP,PE,RS
8-sequences	AP,DIAG,AP,DIAG,DIAG,AP,PE,RS AP,DIAG,AP,AP,DIAG,AP,PE,RS DIAG,AP,DIAG,AP,DIAG,AP,PE,RS
9-sequences	HPI,AP,DIAG,AP,AP,DIAG,AP,PE,RS AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS
10-sequences	HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS

Table 8 From the *Workflows* dataset- Most frequent consecutive feature accesses (forming the higher order maximal patterns for knowledge discovery) and their level-of-support.

Pattern Length	Maximal Patterns	# of occurrences	Support Proportion
2-sequences	DIAG,AP	91	0.21
	AP,DIAG	84	0.20
	AP, PE	34	0.08
	DIAG, DIAG	23	0.05
	PE, RS	15	0.04
	HPI, AP	19	0.04
3-sequences	AP,DIAG,AP	47	0.11
	DIAG,AP,DIAG	38	0.08
	DIAG,AP,PE	36	0.08
	HPI,AP,DIAG	22	0.05
	AP,DIAG,DIAG	21	0.05
	AP,PE,RS	20	0.05
	DIAG,DIAG,AP	12	0.03

mum threshold of 0.25, some of the reasonably even frequent sequences are considered as outliers, hence the choice of 0.15 was made for the experiments of the study.

Clinicians have different work-flows for tasks (represented as access-sequences in the Workflows dataset). The system providers on the other hand, need to align the application flow to the most optimal pattern (maximal pattern discovered) for making the EHR system usable. **Table 7** displays the frequent patterns in user-work-flows grouped by their length (patterns with length 5 to 10 are given). The maximum length of the sequential patterns is 10 for the work-flows dataset. The table demonstrates that the pattern <HPI, AP, DIAG> (<History of Patient Illness, Assessment Plan, Diagnosis>) is the most frequent pattern across the varying lengths of the patterns. This patterns occurs with 80% frequency. The work-flows containing this subsequence are further analyzed to obtain the maximal patterns[*7]. The work-flows with features (QR, PROC, VACC) may be considered as obsolete and diminished with the application flow presented to the users (Hypothesis 2, Table 2).

Table 8 displays the frequent consecutively accessed features of an EHRs database along with their recurrence rate and support. These are the subsequences considered to obtain the maximal (optimal) patterns. It is evident from the table that a user accessing diagnosis (DIAG) features uses the assessment plan (AP) with 21.4% frequency. The components <HPI, AP> and <AP, DIAG> of the frequent subsequence <HPI, AP, DIAG> occur with 19% and 4.5% frequency respectively. Hence, they form a maximal sequence <HPI, AP, DIAG> of length 3 (Table 8). The frequent maximal and obsolete patterns of the work-flows are determined at this step by the above analysis (Hypothesis 2, Table 2). **Figure 15** displays the comparison of the lengths of the interesting patterns discovered. This explains the number subsequences that are considered for each length as a candidate set for the higher order subsequence. The length 2 and 3 patterns represented in Table 8 represent the consecutive patterns with their respective support thresholds. These are further analyzed, to obtain frequent patterns of each length until a pattern of length 10 is discovered. A full-fledged implementation of an EHR system, will support a large number of tasks, as a result a large number of longer access sequences might be pursued by the users as compared to the considered "work-flows" dataset. **Figure A·5** (Appendix) shows an example list of archetypes required by a single EHR feature, "heart failure summary" in the European Union Semantic Health Net Project (EU-SHN Project) [5]. As shown in the figure, for a single UI form there are around 15 participating archetypes (features). Hence, the number of features in the complete system will be very large.

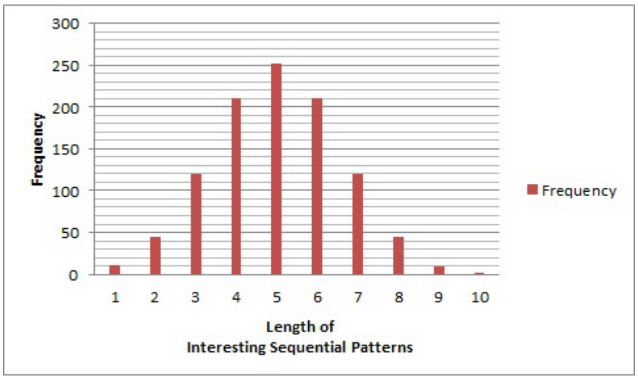

Fig. 15 A comparison between the length of the interesting patterns discovered and the frequency of their access (*Workflows* dataset).

Table 9 gives a qualitative overview of the possible variation in the resulting maximal frequent patterns due to the variation in the EHRs features accessed w.r.t to tasks. The table shows that by altering the minimum support threshold, the proposed framework can handle various distributions of access sequences (EHRs features) to obtain maximal patterns. The expected number of rules and frequent patterns are given based on the execution of the sequential pattern mining on the snippets of "work-flows" dataset (extracting transactions based on tasks and sequences) using the WEKA tool [39].

The key feature to reduce overhead is that the knowledge repository (KRep) of the usability support infrastructure contains only maximal sequences (rather than huge volume of usage logs). These maximal patterns and the associated users form the rules in the knowledge repository.

Considering, the consecutive feature accesses of *<DIAG, AP>* (Table 8), with frequency of 91 occurrences, the support of the pattern is approximately 0.21 (considering a total of 425 sequence of feature-accesses) which is greater than the support

*7 An itemset X is a max-pattern if X is frequent and there exists no frequent super-pattern Y ⊃ X [2].

Table 9 Qualitative overview of the influence of distribution of EHRs features in the user-work-flows in the analytical studies.

S.No.	Sequential EHR Features	Tasks	Support Threshold	Expected Frequent Patterns	Number of rules in Knowledge Repository
1	Variable	Single	Low	Few Patterns, High Frequency	Few
2	Similar	Single	High	More Patterns, Low Frequency	More
3	Similar	Distinct	Low	More Patterns, Low Frequency	More
4	Variable	Distinct	High	More Patterns, Low Frequency	More

threshold of 0.15. which means that 21% of times the feature DIAG leads the Assessment plan (AP) feature in the given sample dataset. Whereas, the assessment plan (AP) feature leads the DIAG feature in approximately 19% of the user-system interactions. These features with support greater than the predefined support-threshold represent the preferred features by the physicians whole performing various tasks. On the other hand, the features *<HPI, AP>* has a frequency of 19 which represents a support of 0.04 which is below the predefined support-threshold. Hence, these consecutive features are not preferred frequently by the users (physician). Using this analysis the system designer can customize the application flow to match the perceived (preferred) application flow by the users (Hypothesis 3, Table 2).

A sample task, General Patient Checkup, its corresponding work-flows and maximal patterns of distinct lengths are given in **Table 10**. A subset of 10 work-flows is considered for the task as an example to show the effectiveness of step two of the framework. The table shows the run information from the WEKA tool. A support threshold of 0.15 was used for the analysis. Since, a large number of itemsets of different lengths are generated varying from length 1 to 10 (considering a 10 feature EHR system), few of the frequent patterns of length 2, 3, 9, 10 are given. Based on the sample, optimal work-flows (of length 10) are obtained. These are stored in the knowledge repository (KRep). An EHR system designer now can refer to this knowledge (optimal paths) for understanding the desired work-flows to realign the application flow accordingly rather than analyzing the raw usage logs which may vary according to users' even for a single task. The analysis considers all the combinations of (consecutive) features to obtain the maximal patterns. Hence, in a real world setting, a larger number of logs needs to be analyzed by a system designer. The proposed framework can reduce this overhead significantly and provide an effective result.

The knowledge repository is an on-line support base for the standardized EHRs database, it can be updated periodically (weekly or fortnightly) depending upon the usage of the system and system-designer. The rules associated with a user-class and their support is updated depending on the user-categorization and the work-flow analysis (example given in Section 4.3). For the considered datasets, the work-flow subsequence *<DIAG, AP>* is stored with a support of 0.21 for the user-class "physician" and the values (set) of socio-technical features used for user-classification is also stored. The periodic analysis to maintain an upto-date KRep may be performed fortnightly on it. This requires the updating of the support for the given work-flow subsequence, with the addition of a previously not existing subsequence. These rules are easily comprehensible by the system designer for system redesign and enhancement rather than performing usage-log analysis (Hypothesis 4, Table 2).

5.3.2 Qualitative Evaluation

Table 11 presents a qualitative study of off-line usability support studies using measures of effectiveness, efficiency and satisfaction. For each task a set of expert and trainee users are considered (assuming that the expertise of the user directly impacts on his or her task performance). For the trainee users, the effectiveness with which a task is performed is defined as a goal to accomplish a task irrespective of the sequence chosen or the number of clicks required. On the other hand, for the expert users it is defined as whether the task is accomplished using the optimal path and fewer clicks. For the tasks in the procedural and non-procedural settings, it is defined as whether the task is accomplished or not. The efficiency of the system is compared w.r.t the tasks, as the successful completion of the task in minimal amount of time. The satisfaction is defined as the rating which is given by the user to the system while performing the task ("Easy", "Very Easy" or "Difficult"). A heuristic comparison is done based on the responses of the clinical users invited for the pre-study. It is evident from the comparison (Table 11) that these qualitative measures are significantly improved by the application of the proposed framework.

Note. The requirements and the assumptions considered in the proposed framework are in agreement with the responses of the surveyed clinicians. Hence, the framework is expected to show minimal or no deviation in the actual environments (larger and broader application scenarios).

6. Discussions

The given experimental results highlight the achievements of the proposed framework according to the hypotheses stated in Table 2. These evaluate the application of the pattern-discovery techniques for the primary purpose of improving the usability of the EHRs databases. **Figure 16** depicts the process for improved interactions between a user and the standardized EHRs database through the use of the on-line feedback for usability enhancement (available through the knowledge repository). The rules captured in the knowledge repository represent the correlation between the user-categories and their work-flows. The EHR designer can use these rules to anticipate and design the user preferred application-flow. The templates can be added or modified and removed accordingly. Hence, the interface can be redesigned as per the end-user expectations. For example, some new archetypes or EHR extracts may be included in the template of a user-interface form for different specializations.

Table 10 From the *Workflows* dataset- Most frequent consecutive feature accesses (forming the higher order maximal patterns for knowledge discovery) and their level-of-support.

Task	General Patient Checkup
Workflows (*sample set = 11 rows*)	1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS,1000 1,HPI,AP,DIAG,LT,AP,MSE,DIAG,AP,PE,AP,1001 1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS,1002 1,PE,PROC,VACC, LT,DIAG,AP,MSE,QR,AP, MSE,1003 1,PE,PROC,VACC, LT,DIAG,AP,MSE,QR,AP, MSE,1004 1,AP,MED,SH, FH,LT,DIAG,AP,MSE,RS, MED,1005 1,AP,MSE,SH, FH,LT,DIAG,AP,MSE,RS, MED,1006 1,AP,MED,SH, FH,LT,DIAG,AP,MSE,RS, MED,1007 1,AP,MSE,SH, FH,LT,DIAG,AP,MSE,RS, MED,1008 1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS,1009 1,HPI,AP,DIAG,LT,AP,MSE,DIAG,AP,PE,AP,1010
Run Information (*WEKA Tool*)	Run information Scheme - weka.associations.GeneralizedSequentialPatterns -S 0.15 -I 0 -F -1 Relation - sequential_test_set Instances - 11 Attributes - 12 user Feature1 Feature2 Feature3 Feature4 Feature5 Feature6 Feature7 Feature8 Feature9 Feature10 Timestamp Associator model (full training set) GeneralizedSequentialPatterns Number of cycles performed- 10 Total number of frequent sequences- 2046 Frequent Sequences Details (filtered)- (given below)
Length 2, Frequent Patterns (*Total number of patterns = 90*)	HPI,AP AP,MED AP,SH MED,SH HPI,DIAG AP,DIAG HPI,AP AP,AP DIAG,AP AP,FH
Length 3, Frequent Patterns (*Total number of frequent patterns = 240*)	HPI,AP,DIAG HPI,AP,AP HPI,AP,DIAG HPI,AP,AP HPI,AP,DIAG HPI,AP,AP HPI,AP,PE HPI,AP,RS AP,MED,SH AP,MED,FH
Length 9, Frequent Patterns (*Total number of patterns = 20*)	HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,RS HPI,AP,DIAG,AP,DIAG,AP,DIAG,PE,RS HPI,AP,DIAG,AP,DIAG,AP,AP,PE,RS HPI,AP,DIAG,AP,DIAG,DIAG,AP,PE,RS HPI,AP,DIAG,AP,AP,DIAG,AP,PE,RS HPI,AP,DIAG,DIAG,AP,DIAG,AP,PE,RS HPI,AP,AP,DIAG,AP,DIAG,AP,PE,RS AP,MED,SH,FH,LT,DIAG,AP,MSE,RS AP,MED,SH,FH,LT,DIAG,AP,MSE,MED
Length 10, Optimal Path	HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS AP,MED,SH,FH,LT,DIAG,AP,MSE,RS,MED

6.1 Applicability of the Framework

The framework is generic in nature and is used for automation of usability analysis in large EHRs systems. It is applicable to any standard or non-standard based EHRs system. It can automate the usability studies. Major standards for the EHRs, HL7, CEN 13606, and openEHR follow a dual-level modeling approach. Therefore, the framework can readily be applied to the EHRs systems based on any of the mentioned standards. In the

Table 11 Relative comparison of the improvement in task performance using qualitative measures.

S.No.	Tasks	Effectiveness	Efficiency	Satisfaction
1 (a)	Record Patient Demo-graphics (Trainee)	90–100%	✓	Easy-to-Very-Easy
1 (b)	Record Patient Demo-graphics (Expert)	90–100%	✓	Very-Easy
2 (a)	General Patient Checkup (Trainee)	90–100%	✓	Easy-to-Very-Easy
2 (b)	General Patient Checkup (Expert)	90–100%	✓	Very-Easy
3 (a)	Identify and Investigate medical problem (Procedural Setting)	90–100%	✓	Easy-to-Very-Easy
3 (b)	Identify and Investigate medical problem (Non-procedural Setting)	90–100%	✓	Easy-to-Very-Easy
4 (a)	Discuss treatment with patient (Trainee)	90–100%	✓	Easy-to-Very-Easy
4 (b)	Discuss treatment with patient (Expert)	90–100%	✓	Very-Easy
5 (a)	Provide Clinical Summary of visit (Procedural Setting)	90–100%	✓	Easy-to-Very-Easy
5 (b)	Provide Clinical Summary of visit (Non-procedural Setting)	90–100%	✓	Easy-to-Very-Easy
6 (a)	Prescribe Medication (Procedural Setting)	90–100%	✓	Easy-to-Very-Easy
6 (b)	Prescribe Medication (Non-procedural Setting)	90–100%	✓	Easy-to-Very-Easy

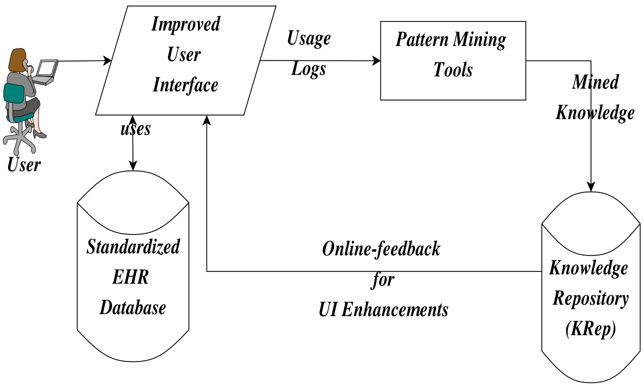

Fig. 16 Enhanced Decision making with the support of Usability Support Study (Framework).

case of openEHR standard the archetypes can be designed from scratch, or adapted from preexisting ones. However, the micro-details of the application such as, the features provided by the system, the usability concerns require to be understood and implemented. Primarily, it needs complete details of the end-users, their demographical data, other characteristic data, their usage logs and the application flow provided by the system. It performs analysis on the user-data and the usage logs, w.r.t the features provided by the EHR system. The rules in the knowledge repository (KRep) are coupled with the results of the pattern mining and depend on the system to which the framework is applied. The EHR system designer does not need to query the end-users or analyze system-logs; rather he only needs to refer to the rules stored in the knowledge repository.

Different archetypes are aggregated into one by means of archetypes templates, which also support semi-automatic derivation of user-interfaces. As explained in Sections 1 and 2 the archetypes and templates have a complex structure and store a large amount of data. The complexity is further intensified as they evolve over time w.r.t changing or new needs of the users. Hence, due to the volume of data and its complex structure, traditional usability methods cannot address the large-scale usability concerns in a temporally evolving environment sufficiently. Hence, to address these concerns, the proposed automated framework facilitates EHRs systems enhancements.

Limitations of the Study. The study aims to improve the usability of the EHRs database system by automating the iterative process of understanding the end-users and their tasks, work-flows using the analytical tools, thus, eliminating the need to use surveys, questionnaires, field studies for improving system usability. However, the actual use of the system's features can only be analyzed in a real-world setting. The participating archetypes and templates in an EHR database system vary according to the environments in which they are deployed. The framework may show some deviations in a real-world setting but these are expected not to deflect the purpose of automation of the usability improvement.

The standardized EHRs create difficult problems as a result of the continuity of template revisions and the evolution of archetypes, and these are done at the cost of user ignorance. For example, when an archetype or a template is revised the end-users are not aware of it. To make them available for use by the end-users is a challenge for the usability studies. The new ways of capturing work-flows and usage are required in such complex scenarios.

A large portion of the usability concerns has been addressed in this study, and further investigation is ongoing. For the case of deviating work-flows, the size of the problem may be as large as the one which the proposed framework is trying to cope with. The understanding of the deviating cases may involve algorithms to analyze whether the deviation in the flow is temporary and long-term, further whether the features have become obsolete or not and whether it should be removed or if modifying the feature will be sufficient. It may also need to find the reasons for deviation and impact on other features of the EHRs system.

An Example. Considering the results of the experiments performed (Table 7, Table 8 and Fig. 12), the framework successfully identifies the class of users "internal medicine specialists" with characteristics <Male, local-hospital, preventive-care, high computer-literacy, primary-care>. This category of users is associated with the frequent work-flow pattern <HPI, AP, DIAG>. This implies that the features "History of Patient Illness" (HPI), "Assessment Plan" (AP) and "Diagnosis" (DIAG) are preferred consecutively by the users. The pattern, user-information and the associated support are stored as a rule in the knowledge repository (KRep). Each of these features can be customized by updating/adding or deleting the archetypes from the corresponding templates to improve data quality. The least accessed (low-frequency of use) features such as Quality Reports (QR) (Fig. 14) can be removed from the application-flow.

7. Summary and Conclusions

The standardized EHRs databases store life-long EHRs of the patients which continuously evolve over time. The dynamic nature of these EHRs and the complex structure of the participating archetypes give rise to critical usability barriers. The former usability studies using manual methods of post-release surveys and user-feedback sessions are not applicable in the complex environment of these databases. In this study, we propose an automated support framework to address these usability barriers. The proposed framework is evaluated by a pre-study with the actual end-users of the EHRs databases. The datasets for the evaluation study has been prepared by using the responses of actual end-users. The results successfully correlate the users and the interesting work-flows for application customization (and modification). It considers the granular-level factors that impact on the usability.

Pattern mining techniques of decision-tree classification, SPA and temporal association mining give results with high accuracy on the considered data. In addition, the framework helps the system designer to infer the corrective actions such as focused application-flow per user-class (and given task). In view of the continuous evolution of the standardized EHRs databases, the proposed enhancements will meet the challenges for improved usability.

References

[1] Agrawal, R. and Srikant, R.: Mining Sequential patterns, *Proc. 11th International Conference on Data Engineering* (*ICDE 1995*), pp.3–14, IEEE Computer Society (Mar. 1995).

[2] Bayardo, R.J.: Efficiently mining long patterns from databases, *SIGMOD Rec.*, Vol.27, No.2, pp.85–93 (online), DOI: 10.1145/276305.276313 (1998).

[3] Burgess, S., Dennis, S., Lanka, S., Miller, N. and Potvin, J.: MedlinePlus Connect: Linking Health IT Systems to Consumer Health Information, *IEEE IT Pro*, pp.22–28 (May-June 2012) (online), available from ⟨http://www.nlm.nih.gov/medlineplus/connect/medlineplusconnecthandout.pdf⟩.

[4] Canlas, R.D.: Data Mining in Healthcare: Current Applications and Issues, Carnegie Mellon University, Master Thesis (2009).

[5] CKM: Clinical Knowledge Manager, available from ⟨http://openehr.org/knowledge/⟩ (accessed 2012-11).

[6] Constance, M.J., Todd, R.J. and Zhang, J.: A user-centered framework for redesigning health care interfaces, *Journal of Biomedical Informatics*, Vol.38, No.1, pp.75–87 (2005).

[7] Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A. and Laleci, G.: A survey and analysis of Electronic Healthcare Record standards, *ACM Comput. Surv.*, Vol.37, No.4, pp.277–315 (2005).

[8] Helma van der, L., Austin, T. and Talmon, J.: Generic screen representations for future-proof systems, is it possible? There is more to a GUI than meets the eye, *Journal of computer methods and programs in biomedicine*, Vol.95, No.3, pp.213–226 (2005).

[9] HL7: Health Level Seven, available from ⟨http://www.hl7.org⟩ (accessed 2011-08).

[10] Hui, J., Knoop, S. and Schwarz, P.: HIWAS: Enabling Technology for Analysis of Clinical Data in XML Documents, *PVLDB*, Vol.4, No.12, pp.1260–1271 (2011).

[11] Hussein, R., Engelmann, U., Schroeter, A. and Meinzer, H.P.: DICOM structured reporting: Part 1. Overview and characteristics, *RadioGraphics*, Vol.24, No.3, pp.891–896 (2004).

[12] ICD: International Classification of Diseases, available from ⟨http://apps.who.int/classifications/icd10/browse/2010/en⟩ (accessed 2013-01).

[13] IHE: Integrating the Healthcare Enterprise, available from ⟨http://www.ihe.net/⟩ (accessed 2013-01).

[14] IHTSDO: SNOMED-CT® User Guide (online), available from ⟨http://ihtsdo.org/fileadmin/user_upload/doc/download/doc_UserGuide_Current-en-US_INT_20130131.pdf⟩.

[15] ISO: International Standards Organization standard, available from ⟨www.iso.org/⟩ (accessed 2012-10).

[16] Jin, J., Ahn, G.-J., Hub, H., Michael, J.C. and Zhang, X.: Patient-centric authorization framework for electronic healthcare services, *Journal of Computers and Security*, Vol.30, No.2-3, pp.116–127 (Mar.-May, 2011).

[17] Kashfi, H.: Applying a User Centered Design Methodology in a Clinical Context, *Studies in health technology and informatics*, Vol.160, No.2, pp.923–931 (2010).

[18] LOINC: Logical Observation Identifiers Names and Codes, available from ⟨http://loinc.org/⟩ (accessed 2013-01).

[19] Lowry, S.Z., Quinn, M.T., Ramaiah, M., Schumacher, R.M., Patterson, E.S., North, R., Zhang, J., Gibbons, M.C. and Abbott, P.: The EHR Usability Evaluation Protocol (EUP), Technical Evaluation, Testing and Validation of the Usability of Electronic Health Records, *NIST Interagency Report 7804* (2012).

[20] Madaan, A., Chu, W., Yaginuma, D. and Bhalla, S.: Quasi-Relational Query Language Interface for Persistent Standardized EHRs: Using NoSQL Databases, *DNIS*, pp.182–196 (2013).

[21] Middleton, B., Bloomrosen, M., Dente, M.A., Hashmat, B., Koppel, R., Overhage, J.M., Payne, T.H., Rosenbloom, S.T., Weaver, C. and Zhang, J.: Enhancing patient safety and quality of care by improving the usability of electronic health record systems: Recommendations from AMIA, *Journal of the American Medical Informatics Association*, DOI: 10.1136/amiajnl-2012-001458 (2013).

[22] NIST: National Institute of Standards and Technology, available from ⟨http://www.nist.gov/index.html⟩ (accessed 2012-10).

[23] Committee on Data Standards for Patient Safety: Key Capabilities of an Electronic Health Record System, The National Academies Press (2003).

[24] OpenEMR: OPENEMR, available from ⟨http://www.oemr.org/⟩ (accessed 2012-10).

[25] OpenVista: OpenVista EHR Solution, available from ⟨http://www.medsphere.com/solutions/openvista-for-the-enterprise⟩ (accessed 2012-07).

[26] Opereffa: Opereffa Issues and Other details, available from ⟨http://listarchives.openmrs.org/FW-openEHR-announce-Opereffa-an-openEHR-reference-application-td3059093.html⟩ (accessed 2013-01).

[27] Opereffa: UCL, The openEHR Reference Framework and Application, available from ⟨http://opereffa.chime.ucl.ac.uk:80/introduction.jsf⟩ (accessed 2011-08).

[28] PostgreSQL: PostgreSQL database, available from ⟨http://www.postgresql.org/⟩ (accessed 2011-11).

[29] Quinlan, J.R.: Induction of Decision Trees, *Mach. Learn.*, Vol.1, pp.81–106 (1986).

[30] Ramakrishnan, N., Hanauer, D.A. and Keller, B.J.: Mining Electronic Health Records, *Computer*, Vol.43, No.10, pp.77–81 (2010).

[31] Sachdeva, S. and Bhalla, S.: Semantic interoperability in standardized electronic health record databases, *Journal of Data and Information Quality* (*JDIQ*), Vol.3, No.1 (2012).

[32] Sachdeva, S., Madaan, A. and Bhalla, S.: Discovery of Patterns to Improve Usability of Electronic Health Record Systems, *IPSJ Journal*, Vol.52, No.12, pp.327–339 (Jan. 2012).

[33] Schabetsberger, T., Ammenwerth, E., Gobel, G., Lechleitner, G., Penz, R., Vogl, R. and Wozak, F.: What are Functional Requirements of Future Shared Electronic Health Records, *Proc. MIE2005* (*Studies in Health Technology and Informatics*), ENMI (2005).

[34] Schumacher, R.M.: Commentary: Electronic Health Records and Human Performance, *Journal of Oncology Practice*, Vol.6, No.3, pp.125–126 (2010).

[35] Schumacher, R.M. and Jerch, K.E.: Measuring Usability in Healthcare IT: It's a practice, Not a competition, *Interactions Magazine* (July-Aug. 2012).

[36] Swanson, A. and Lind, S.: Usability Testing EHRs: Examples from Front Lines, *Interactions Magazine*, Vol.18, No.6, pp.54–58 (2011).

[37] Beale, T., Heard, S., Kalra, D. and Llyod, D.: The OpenEHR Reference Model: EHR Information Model, *The OpenEHR release 1.0.2* (2008).

[38] TIGER: Designing Usable Clinical Information Systems, available from ⟨http://www.tigersummit.com/uploads/Tiger_Usability_Report.pdf⟩ (accessed 2012-11).

[39] The University of Wakaito: WEKA, available from ⟨http://www.cs.waikato.ac.nz/ml/weka/⟩ (accessed 2011-07).

[40] WALKER: EHR Safety: The Way Forward to Safe and Effective Systems, *J. Am. Med. Inform. Assoc.*, Vol.15, No.3, pp.272–277 (2008).

[41] Zhang, J. and Muhammad, F.W.: TURF: Toward aunifiedframework of EHR usability, *Journal of Biomedical Informatics*, Vol.44, No.6, pp.1056–1067 (Dec. 2011).

[42] Zheng, K., Padman, R., Johnson, M.P. and Diamond, H.S.: An Interface-Driven Analysis of User Behavior of an Electronic Health Records System, *J. Am. Med. Inform. Assoc.*, Vol.16, No.2, pp.228–237 (2009).

Appendix

A.1 Dataset Snippets

Tables A·1 and **A·2** give the snippets of the datasets used for the evaluation of the proposed framework.

A.2 Pre-Study with the Clinicians

Table A·3 gives the details of the questions used to interview the actual clinicians. It also explains the kind of response expected from them.

A.3 Prototype EHR system based on openEHR standard and Usability Concerns

The AQBE system is an EHR system, based on the openEHR standard in the early stages of development [20]. The prototype system makes use of a small set of archetypes from clinical knowledge manager (CKM). In the following figures, the usability concerns with respect to the features of this EHR system are highlighted. In the **Fig. A·1**, the "data inserter" user-interface is shown. A list of archetypes, a form for patient-details and a query formulation table are given. A dynamic template (form) corresponding to the selected archetype (in the drop down menu) is generated. There is a need to cater to the usability challenges that are foreseen. Most of these challenges can be addressed, if the end-users and their preferred work-flows are available as rules (guidelines) for the system designer.

The classification of users and understanding their relationship with various attributes helps to select the features of the system that are mandatory and those which are optional. The customization of archetypes can be performed on the basis of user attributes and expected work-flows.

Figure A·3 and **Fig. A·4** represent a template for the blood pressure archetype generated dynamically after a user selects the concept from the list (**Fig. A·2**). These depict the long list of attributes for a single (blood pressure archetype). The complexity increases when a user selects multiple archetypes and needs forms (templates) with multiple archetypes.

Table A·1 A snippet of *Users* dataset.

@ relation Users
@attribute expertise {expert, trainee} @attribute age real @attribute location {local, cosmopolitian} @attribute setting {clinical, hospital, university, home, office} @attribute sex {M, F} @attribute qualification {Bachelors, Masters, More} @attribute computerKnowledge {WebUser, Programmar, ComputerLiterate} @attribute purpose {preventive, diagnostic, research, others} @attribute department {admin, frontOffice, primaryCare, nursing, pharmacy, education, procedureBasedCare, non-procedureBasedCare} @attribute UserLabel {Physician, Specialist, Student, Researcher, BillingAgent, frontDesk, pharmacist, nurse}
@data expert,65,local,clinical,M,Masters,ComputerLiterate,diagnostic, primaryCare, Physician expert,60,local,hospital,F,Masters,ComputerLiterate,diagnostic, primaryCare, Physician expert,63,cosmopolitian,hospital,M,Masters,ComputerLiterate ,diagnostic, primaryCare,Physician expert,50,cosmopolitian,hospital,M,More,ComputerLiterate,diagnostic, primaryCare,Physician expert,48,cosmopolitian,hospital,M,More,ComputerLiterate,preventive, pharmacy,pharmacist

Table A·2 A snippet of *Work-flows* dataset.

@ relation Workflows
@attribute user {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15} @attribute 'Feature1'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature2'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature3'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF, PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature4'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature5'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature6'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature7'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature8'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature9'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @attribute 'Feature10'{HPI,AP,FH,SH,DIAG,PE,LT,PROC,VACC,MED,MSE,RS,OT,AA,AF,EP,CI,QR,FB,HR,LTR,DF,PA,CSO,DO,SF,MHR,BF,CT} @timestamp{1000, 1001,1002,1003,1004,1005,1006,1007,1008,1009,1010}
@data 1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS, 1000 2,HPI,AP,DIAG,LT,AP,MSE,DIAG,AP,PE,AP, 1000 1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS, 1001 1,HPI,AP,DIAG,AP,DIAG,AP,DIAG,AP,PE,RS, 1002 1,PE,PROC,VACC, LT,DIAG,AP,MSE,QR,AP, MSE, 1002 1,PE,PROC,VACC, LT,DIAG,AP,MSE,QR,AP, MSE, 1003

Table A·3 Questionnaire for the pre-study performed with the end-users (clinicians).

S.No.	Questions
1.	Do you use an electronic health record (EHR) system in your day to day activities (such as, patient assessment/ medication/ patient revisit and so-on)? [YES/NO]
2.	If YES, please tell us which software or software provider?
3.	If NO, please tell us the reason (not available/ time consuming/ difficult to use/ complex screens)?
4.	Have you heard of the standards for EHRs such as HL7, CEN 13606 or openEHR? [YES/NO]
5.	Have you ever used a clinical application based on any of these standards [YES/ NO]?
6.	What are the main tasks or activities you think are important in everyday clinical activities that can be performed by using a clinical application? (e.g., patient diagnosis/referrals/revisit of patient/assignment of assessment plan to the patient/clinical task such as, surgery etc.)
7.	Does the application flow (the order in which the screens appear) of the system serve your purpose easily? [YES/NO]
8.	How frequently do you face difficulties or errors in accessing a clinical application [Very frequently/Frequently/Rarely/Not at all]?
9.	How do you report the errors encountered (report to admin staff of the clinic or hospital/ fill up a form/ Others) Please specify?
10.	Were you consulted before the clinical application was deployed at your organization? [YES/NO]
11.	Did you receive any training session before using the system?
12.	Do you agree if your needs were considered during system design rather than during post-use surveys/ interviews would result in a user-friendly system? [YES/NO]
13.	Do you agree a clinician's qualification, specialization and work-experience affect the use of clinical-application? [YES/NO]
14.	Do you agree the geographical location of the clinician and the setting (clinic or hospital) affects the use of clinical application? [YES/NO]
15.	Do you agree age of the health-care users affect their use of the clinical application? [YES/NO]
16.	Do you agree clinical application's flow must be customized to suit your common daily tasks? [YES/NO]
17.	Do you agree that the application flow should match the end-user work-flows within a clinical application? [YES/NO]
18.	Do you recall all the issues/errors encountered by you during system use during the post-release feedbacks and surveys or interviews? [YES/NO]
19.	Do you agree the IT providers should capture your needs based on your usage of the system automatically or ask for manual feedbacks or interview sessions? [YES/NO]
20.	Please give us some reason for the above question.
Name: Qualification: Age: Department: Organization: Address:	

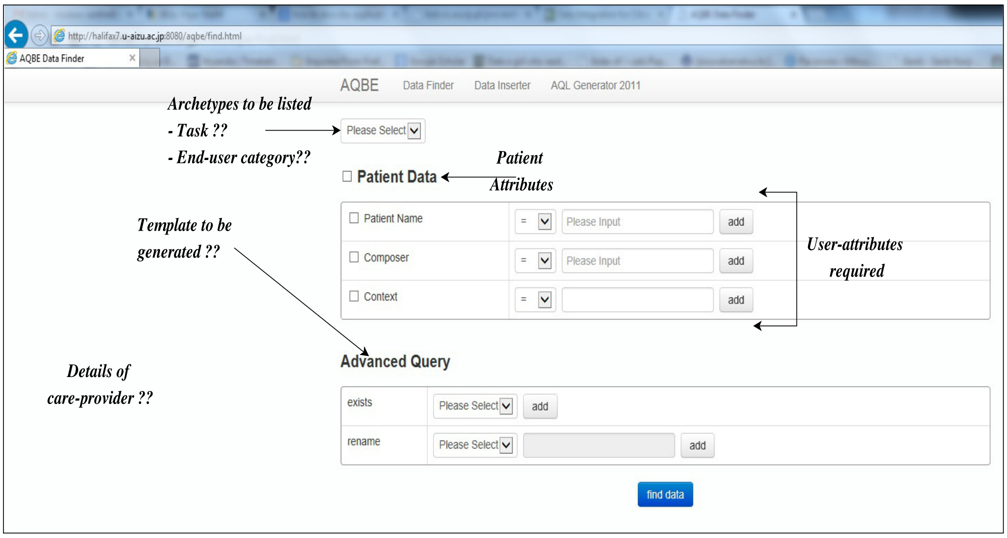

Fig. A·1 AQBE prototype system representing the usability concerns — (i) Flow of contents on the user-interface, (ii) mandatory and optional user-attributes, (iii) archetypes required based on the needs of the (specialized) health-care environment, (iv) when and where the template needs to be generated, (v) Distinguishing the mandatory and optional user data.

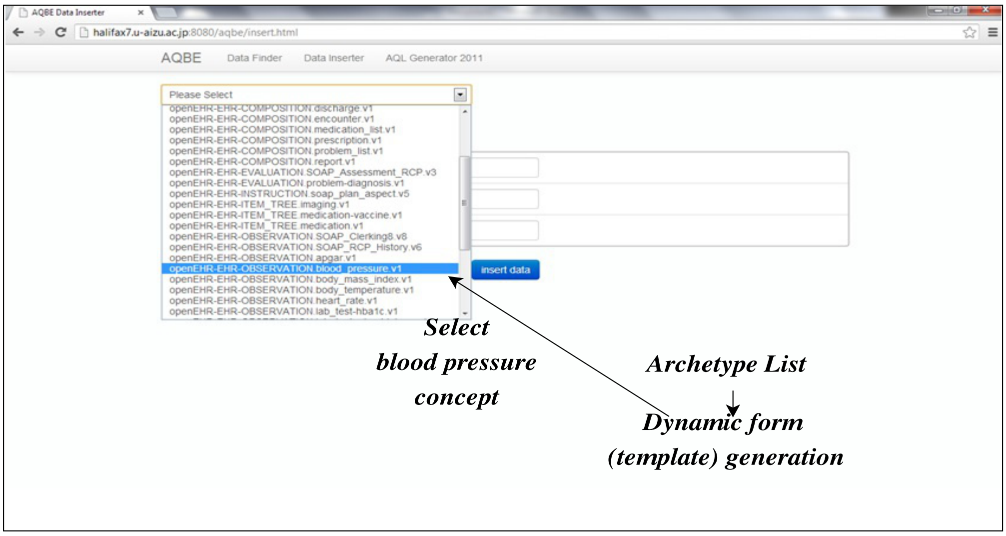

Fig. A·2 Archetype list in the AQBE system, used for selection of archetype(s) for dynamic form generation.

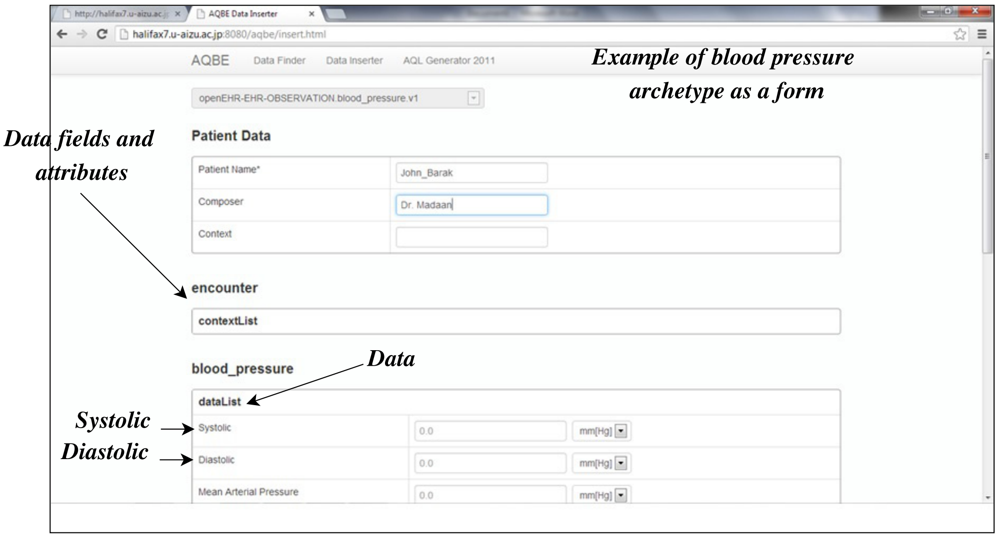

Fig. A·3 Example form corresponding to the blood pressure concept in the AQBE system representing the various attributes in the concept.

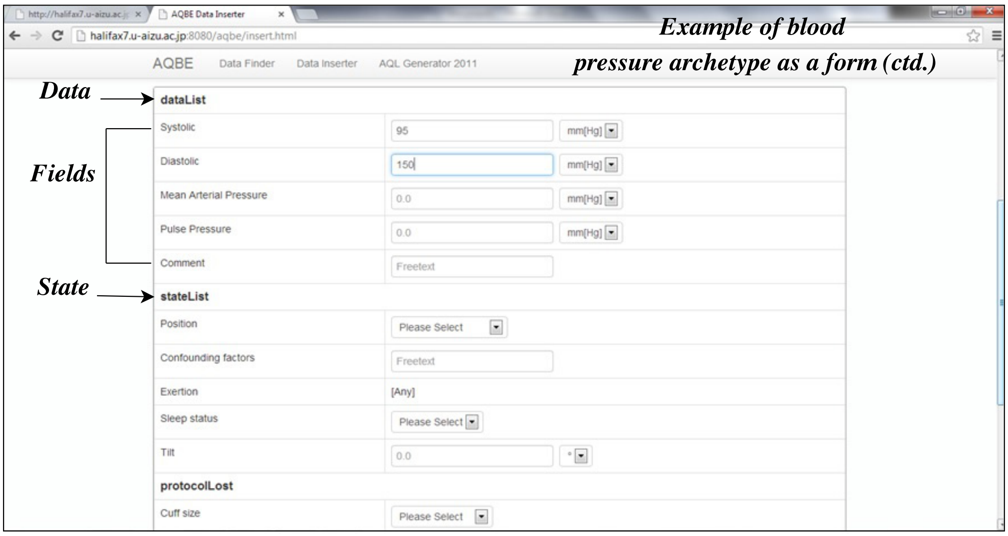

Fig. A·4 Example form corresponding to the blood pressure concept (Fig. A·3) continued.

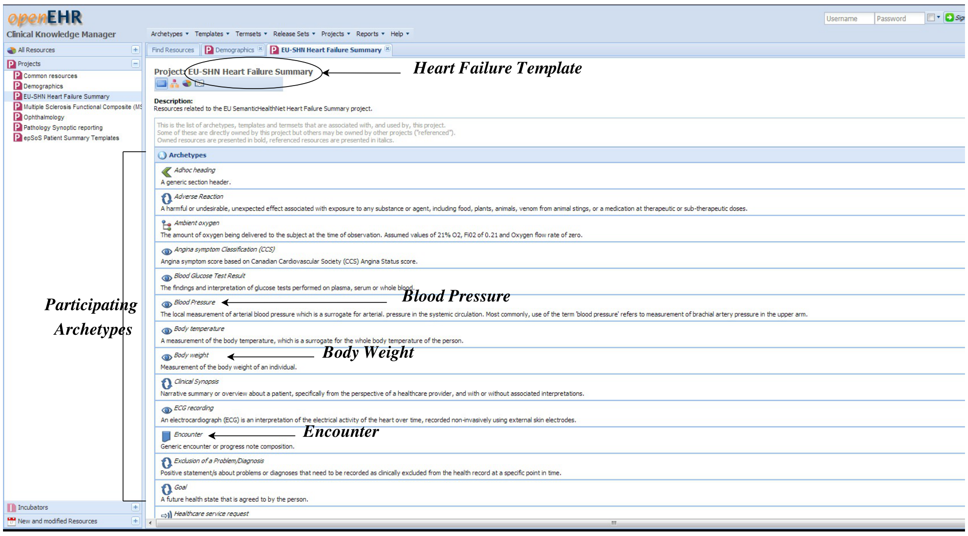

Fig. A·5 The participating archetypes of "heart summary failure" of the EU-SHN Project.

Aastha Madaan received her B.Sc. (Hons) in Computer Science from the University of Delhi in 2006. She received her M.Sc. in Computer Science from the University of Delhi in 2008. She has worked as a software developer with Aricent Group, India for nearly 3 years. She has been pursuing a Ph.D. in Computer Science from University of Aizu, Japan, since April 2011.

Subhash Bhalla received his B.Tech in Computer Science in 1978. He received his Ph.D. degree in Computer Science in 1984 from the Indian Institute of Technology, Delhi, India. His research interests include performance characterization of real-time dependable databases and the design of new databases to support multimedia information systems, data modeling and transactions and distributed algorithms. He is currently working as professor at University of Aizu, Japan. He has 28 years of experience as faculty scientist and researcher.