Structuring Clinical Workflows for Diabetes Care
An Overview of the OntoHealth Approach

M. Schweitzer¹; N. Lasierra²; S. Oberbichler¹; I. Toma²; A. Fensel²; A. Hoerbst¹

¹UMIT – University for Health Sciences, Medical Informatics and Technology, Research Division for eHealth and Telemedicine, Hall in Tirol, Austria; ²University of Innsbruck, STI – Semantic Technology Institute, Innsbruck, Austria

Keywords
Electronic health records, medical records, diabetes mellitus, workflow, semantics

Summary
Background: Electronic health records (EHRs) play an important role in the treatment of chronic diseases such as diabetes mellitus. Although the interoperability and selected functionality of EHRs are already addressed by a number of standards and best practices, such as IHE or HL7, the majority of these systems are still monolithic from a user-functionality perspective. The purpose of the OntoHealth project is to foster a functionally flexible, standards-based use of EHRs to support clinical routine task execution by means of workflow patterns and to shift the present EHR usage to a more comprehensive integration concerning complete clinical workflows.

Objectives: The goal of this paper is, first, to introduce the basic architecture of the proposed OntoHealth project and, second, to present selected functional needs and a functional categorization regarding workflow-based interactions with EHRs in the domain of diabetes.

Methods: A systematic literature review regarding attributes of workflows in the domain of diabetes was conducted. Eligible references were gathered and analyzed using a qualitative content analysis. Subsequently, a functional workflow categorization was derived from diabetes-specific raw data together with existing general workflow patterns.

Results: This paper presents the design of the architecture as well as a categorization model which makes it possible to describe the components or building blocks within clinical workflows. The results of our study lead us to identify basic building blocks, named as actions, decisions, and data elements, which allow the composition of clinical workflows within five identified contexts.

Conclusions: The categorization model allows for a description of the components or building blocks of clinical workflows from a functional view.
1. Introduction

1.1 Scientific Background

Health care is facing an ever-increasing number of multimorbid patients worldwide. This is due to various reasons, such as the aging population or the increasing number of patients with chronic non-communicable diseases in Western countries [1]. Higher patient volumes, complex cases, and the proliferation of technology in health care are leading to a serious increase in health care expenses. Diabetes mellitus, for example, results in economic costs that rose from $174 billion (2007) to $245 billion (2012) in the United States [2]. Hence, there is an urgent need to keep costs low while at the same time guaranteeing or even improving the quality of patient care.

Information systems and, specifically, the electronic health record (EHR) play an important role in achieving such an ambitious goal by e.g. fostering information exchange between various health care providers or by increasing the transparency of information and patient flows within the health care system [3]. The concept of an EHR is to share institution-specific patient data between various health care providers in order to establish a comprehensive and longitudinal collection of the patient’s health and health care data. This intention requires the active participation of the patient in his/her treatment by adding and providing health-related self-documented data.

An obvious requirement of and major challenge to EHRs is interoperability at all layers of communication between the different EHR systems as well as their involved local information systems. With regard to technical and semantic interoperability and information semantics, important standards and best practices have already been developed over the last few years. One of these is the Integrating the Healthcare Enterprise (IHE) initiative [4]. IHE describes best practices based on available standards or de-facto standards, such as CEN13606 [5], openEHR [6], and Health Level 7 (HL7) Clinical Document Architecture (CDA) [7], to gain maximum interoperability of health care IT. IHE is used by many projects worldwide, including the Austrian electronic health record (ELGA) [8], epSOS [9], and EHR4CR [10].

However, one problem that has not been tackled sufficiently is the domain of functionality perceived by users of trans-institutional information systems such as EHRs. Similar to local information systems, EHR systems are still regarded as monolithic systems from the user-functionality perspective. This means that interaction with the EHR system is based on single and individual queries performed manually and for specific purposes by the medical personnel. These queries are usually executed within a clinical workflow conducted by the medical personnel in order to obtain a result that is relevant for a given medical activity. Hence, supporting complete clinical workflows and executing all needed queries in a semi-automatic way within such workflows would lead to an improvement in efficiency and would support health care providers in their daily tasks.

Based on a service-orientated approach, the OntoHealth project [11] establishes a functionally flexible layer that connects on top of standards-based EHR systems and makes it possible to answer and solve clinical problems by means of customizable workflow patterns which are generally not limited to a certain organization or EHR system. Although the basic framework and components of the system will be generic, clinical functionalities of the system will be specifically designed for the diabetes domain. Diabetes mellitus (DM) is a chronic illness whose prevalence has increased in recent years. Some estimates indicate that 7.7% of the world’s adult population will suffer from diabetes by the year 2030 [12]. Although diagnostics and treatment of diabetes is well-established, diabetes care requires complex clinical information from various medical domains.

1.2 Objectives

The objectives of this paper are to: 1) introduce the underlying concept of the OntoHealth project by presenting the architecture of the intended system; 2) present a functional categorization of workflows and building blocks regarding workflow-based interactions with EHRs in the domain of diabetes.

The remainder of the document is organized as follows: Section 2 presents the state of the art regarding the implementation of clinical workflows and main benefits of service composition for this aim. Methods used for the systematic literature review as well as considerations regarding the Onto-
Health architecture are presented in Section 3. A schematic overview of the modules that compose the OntoHealth architecture together with the results from the systematic literature review is presented in Section 4. Finally, methods and results as well as their clinical relevance are discussed.

2. Electronic Health Records, Clinical Workflows and Service Composition

Structuring of the functional level of electronic health records is still a problem that has not been sufficiently tackled by the majority of current solutions in the health care sector. Electronic health records are still very often regarded solely as records allowing basic access to data, such as viewing and retrieving existing documents, and not as integrated, complex systems of data and functionality. Besides digital formats and interoperability, “meaningful use” of EHRs requires improving their functionality, supporting the implementation of customizable workflows, and fostering their integration within daily routine care in the clinical setting [13, 14]. Also, the integration of EHRs within common clinical workflows has been pointed out as a major prerequisite for their adoption and implementation [14]. A clinical workflow can be defined as a set of activities conducted in the hospital setting that lead to the achievement of a clinical result related with the assessment, diagnosis, or general supervision of the patient. Hence, EHR usage can be integrated as part of these activities and can also be used to execute implementation in an efficient and flexible manner.

Experiences in other domains, such as the business domain, have demonstrated that service orientation is a suitable means for structuring complex distributed software systems, reducing cost, and easing the composition of new or adapted applications by using existing services [15]. The latter, in particular, commonly known as service composition, is one key feature of service orientation. Service composition typically requires combining multiple existing services, which in turn can be combined even further to result in higher-level functionalities. In this way, a particular goal which cannot be fulfilled by each of the services alone can be jointly fulfilled as part of the composition. A service composition contains execution constraints on how services are to be combined in order to fulfill the particular goal. Such constraints can be created manually, semi-automatically, or automatically [16]. In the manual approach, a human expert creates the composition; in a semi-automatic approach, she/he is supported by a tool; while in automatic composition, the composition is created without any human interaction. The semi-automatic and automatic compositions bring more flexibility and enable easy adaptation to changes. Semantics have been recognized as an enabler for (semi-)automation of various service-related tasks, including service composition, leading to an entire field of research known as Semantic Web Services [17]. By semantically specifying service interfaces and how they can be consumed, intelligent service composition algorithms can create new services with minimal user support. Furthermore, if some of the services in the composition are failing, they can be dynamically replaced by other similar services.

Semantically modeled and operated workflows, as well as (semi-)automatic composition based on semantics, which brings flexibility and scalable adaptation, could help to solve a common problem for standard eHealth infrastructures, i.e. the lack of proper solutions for on-the-fly, painless integration of services in order to provide new functionalities.

3. Methods

3.1 OntoHealth Architecture

The OntoHealth architecture aims at supporting the dynamic composition and execution of clinical workflows related to EHR interaction focused on the domain of diabetes mellitus. Hence, in order to implement this functionality, we based our design on tools and methods for composition and integration of services in a clinical environment.

OntoHealth builds on standards and solutions from composition and workflows that are available in the eHealth domain extended with semantics. In OntoHealth, we take a semi-automatic ap-
proach in which generic predefined composition templates created by a user, assisted by a machine, can be selected and customized to realize service compositions useful for specific domains and situations. The project delivers flexible support for the modification of the workflow processes by utilizing rules and decision logics. Flexibility of the composition process can be reached through incremental composition of services based on the reaction of the user at each step of the composition.

3.2 Systematic Literature Review

As a precondition to the development of the final architecture, as well as specific services for the domain of diabetes, it is necessary to identify and analyze health care professionals’ clinical workflows as well as their informational or functional needs regarding the treatment of patients. Although fostering the quality of care by using EHRs involves various stakeholders, including the patient him/herself, the OntoHealth project focuses on the meaningful interaction between EHRs and physicians in their daily practice of diabetes care.

To achieve this goal, a systematic literature review was conducted based on the PRISMA Statement [18] combining scientific literature and other pertinent sources, such as e.g. clinical guidelines, best practices, and legislation. We decided to analyze both sources independently, beginning with the scientific literature described in the present paper.

According to PRISMA, the initial search in the Identification phase was conducted to retrieve papers from the selected databases given in ▶Table 1. The search terms were built using the defined keywords or derived MeSH terms supplied by the National Center for Biotechnology Information (NCBI). Each search considered title and abstract for selection. The databases ACM Digital, EMBASE, IEEEExplore, CiteSeer, and DARE did not allow using an advanced search for MeSH terms or abstract lookup. These were searched using the predefined keywords. Only papers which were published between the years 1995 and 2013 were considered for further analysis. All duplicates were removed. In order to manage the complete review process, a shared Excel sheet was used containing at least title and author(s) of all papers.

The Screening phase excluded all documents that did not fulfill a certain document type (▶Table 1) or were not written in English.

During the third step (Eligibility), a group discussion of three of the authors was held to define a set of inclusion and exclusion criteria common to our objectives. These criteria were used to select meaningful documents by reviewing title and abstract. The inclusion and exclusion criteria which were applied are depicted in Table 1. This step was collaboratively performed by three of the authors. The quantity of included and excluded papers were collected in the aforementioned Excel sheet as well. All papers were divided into two groups, which were then reviewed by the authors separately. The authors revised title and abstract and assigned the appropriate exclusion or inclusion criteria. A control sample was shared between the authors to validate the assignment of criteria to each paper within the control sample.

In the last phase (Included), the remaining set of papers was full-text analyzed using a quantitative content analysis as described in Mayring [19] and Bortz [20]. All eligible papers were separated and reviewed by two authors finding pertinent statements concerning the inclusion criteria from Table 1. The utilization of MAXQDA software allowed for the coding of all found statements on-the-fly and building a first categorization system. The software describes a created categorization as “code” and an assigned statement to a certain category as “coding”. After reviewing 25% of all included papers, the categorization system was redefined and generalized by further group discussion.

4. Results

4.1 OntoHealth Architecture

Built upon the principles of service composition and semantic integration, ▶Figure 1 depicts a schematic overview of the OntoHealth architecture and its consisting modules. The modules of the basic architecture can be described according to its three main functionalities: 1) Workflow definition:
The interaction between the health care professional and the architecture is established by means of a graphical user interface (GUI). This GUI allows the physician to execute different clinical workflows that may require several interactions with the EHR in an easy and fast manner. For example, a physician could obtain information regarding side effects of a medication given the particular conditions of the evaluated patient and then come up with a better alternative in a one-step interaction. In addition, the GUI allows the definition and composition of new clinical workflows using predefined building blocks associated to common tasks in clinical daily routines. Hence, the GUI enables the communication between health care professionals and the workflow generator module. According to a predefined ontology, this module provides a workflow instance that is compliant with the ontology and filled according to the health care professional's requirements. The specification of this compound of sequence actions for a specific patient's case is denoted as clinical workflow specification, the generalization of these specifications has been denoted as clinical workflow pattern. As shown in Figure 1, this information serves as an input for the orchestrator module.

The orchestrator module is responsible for selecting the most suitable service available in the semantic web service bus for the execution of each of the tasks that comprises the clinical workflow given as an input. For this task, the module examines each part of the workflow, communicates to the available services, retrieves semantic information, and matches both descriptions for the best selection.

The orchestrator module relies on an underlying linked data infrastructure which enables the integration of different data sources [21]. It exploits the integration of various data types, instances, and links between the data sets of the linked data in order to figure out relations between data produced and consumed by services. In this way, output data of a service can be related to input data of another service and thus a composition of the two can be identified as possible. Composition and orchestration is thus based on the integration of linked data and services via linked data services [22] using lightweight semantic descriptions based on standards (e.g., RDF [23], SPARQL [24]).

The semantic web service bus to be developed within the project facilitates the task of composing and orchestrating services and will enable a flexible interaction with EHRs. It relies on a linked data infrastructure which enables the integration of different data sources. The semantic web service bus provides a semantic abstraction over the EHR's functionalities. It includes formal descriptions of the EHR's functionalities based on semantics enabling agents to orchestrate and compose these functionalities based on their semantic description (see the Orchestrator module).

Finally, the EHR module allows for the retrieval of required information and storage of information when appropriate for all involved services within a workflow from EHRs. The EHR module will be based on relevant IHE profiles comprising Cross-Enterprise Document Sharing (XDS), Audit Trail and Node Authentication (ATNA), and Cross-Enterprise Document Workflow (XDW), which are all part of the IHE IT Infrastructure Technical Framework [25]. Hence, the OntoHealth architecture sets up an IHE-based information system that allows native, dynamic workflow support regardless of the type of clinical workflow.

4.2 Categorization Model

Figure 2 summarizes the literature review with respect to the papers selected for each step of the review. As shown in the figure, the 543 initially found papers from the database search led to 87 papers that met the predefined inclusion criteria based on the analysis of abstract and title. The following content analysis was then conducted to find statements related to tasks and complete clinical workflows in order to create a categorization model for defining and structuring building blocks and characteristics of physicians' workflows in the domain of diabetes care. According to Niazkhani et al. [26], a clinical workflow is defined as "the allocation of multiple tasks of a provider or of co-working providers in the processes of care and the way they collaborate." Using this definition in combination with the inclusion criteria, diverse findings concerning tasks and complete clinical workflows could be obtained from the raw data.

In total, 802 single statements were identified from the data, which were generalized and categorized in a three-part identification model. Using this model, a workflow and its containing compo-
nents can be classified by 1) Actions, 2) Data Elements and 3) Contexts. An overview of the categorization is depicted in Figure 3, which also includes the main levels of subcategories. The full categorization can be obtained from the authors.

Every clinical workflow consists of several tasks and decisions which define a closed action or activity executed by a health care professional or a system. The tasks do not need to be atomic, but they are also able to represent complex actions by building another sub-workflow. This hierarchical structure establishes a dynamic use of the proposed categorization model. Each task consists of a generic Action that belongs to a certain Data Element which describes the specific application content of the action. For example, the action “Retrieve data” can be applied over the data element “Laboratory tests”. The data element “HbA1c”, as a sub-item of “Laboratory tests”, can then form a task “Retrieve HbA1c”. Finally, the Context defines the environment a certain task is executed in. For example, “Retrieve HbA1c” in the context of “Continuing & Monitoring” shows the purpose of the task. The decisions are part of the process flow within a workflow. They will be determined in the next step of the literature review using other literature sources as clinical guidelines. We found different diabetes-related guidelines in the analyzed literature. All 91 guideline codings build examples for tasks, consisting of action and data element, decisions, or complete workflows, but they were not considered during this study.

A complete sample workflow could be “Retrieve the summary of laboratory measurements of a certain diabetes patient”. Using the proposed categorization model, this workflow comprises one task for each laboratory value (e.g. “HbA1c”, “foot examination results”, “eye examination results”, “physical exercise adherence”) using a “retrieve” action and is followed by the task “Display summary” to visualize the requested results.

Statements found within the groups Action, Data Element and Context were further classified according to different subcategories.

Actions
In total, 183 statements were classified as Actions. These statements were again classified according to different activities of medical routine. Most of them are related to “Retrieve data” (n = 33), which describes all actions of obtaining any kind of data. Especially for the diabetes domain, these actions could be subdivided into “Retrieve patient’s self-documentation” and “Collect relevant clinical data”. The latter comprises “Retrieve single data element” as it could be e.g. for a laboratory result and “multiple data elements” which describes e.g. a summary or report of multiple laboratory results. Depending on the point of view of a certain action, “multiple data elements” will rather relate to complete further workflows.

Data Elements
For the group Data Element, 437 codes were identified in the content analysis. The most prominent sub-items are “Medical patient data” (n = 291), “Treatment” (n = 63), and “Administrative data” (n = 43). Some of the Data Elements’ subcategories relate to one object (e.g. a patient) or more objects (e.g. group of patients) depending on the situation. For an overview of all categories and subcategories, see Figure 3. Further subdivisions were available in all framed categories but only “Medical patient data” displays them in the figure. Holding subdivisions is not restricted to assigning elements to the category itself, which means that “Administrative data” could also be a valid data item within a certain task. Table 2 shows examples of selected categories with found statements and relevant references.

Contexts
Extending the classification proposed by Veinot et al. [27] for classifying patient-centered EHR use in clinical consultations, context of use can be categorized by: “Priming” (n = 8), “Structuring” (n = 5), “Assessing & Diagnosis” (n = 14), “Informing and Awareness” (n = 138), and “Continuing and Monitoring” (n = 10). Examples of selected statements concerning a certain Context are shown in Table 2.
5. Discussion

5.1 Methods

The categorization builds on acquired statements from a qualitative content analysis. Literature sources are related to well-known online literature databases. The search terms were carefully determined in group discussion, but it cannot be assured that all pertinent papers concerning the requested content were retrieved. Also, a search for gray literature was not included during this study. Coding differences from authors during the content analysis were solved by discussion. However, a qualitative approach always involves some subjectivity.

5.2 Results

The integration of the flexible use of electronic health records is a crucial factor in successfully supporting a physician's daily practice [14]. The developed categorization makes it possible to describe all components or building blocks within clinical workflows. The results presented in this paper are essential for both the design of the services included in the semantic service bus and the clinical workflow models. The latter govern the adequate execution of workflows over the services included in the bus. In order to establish helpful constraints in the vast medical domain, the study is limited to the domain of diabetes. It might be possible to generalize the proposed categorization model, but further investigations need to be done to prove this. The dynamic model using tasks as single actions connected with a certain data element or as further workflows enables a categorization in a hierarchical and flexible manner. Furthermore, tasks could either be specific or generic depending on the assignment of a certain data element or not.

Note that the given categorization does not describe any workflow process rules. The model only provides a tool describing the components or building blocks of a decomposed workflow from a functional view. Using another way, Hsieh [66] e.g. focused more on describing the detailed process related to clinical workflows using Petri nets. This offers workflow-related details that are important for process execution (detailed identification, time of execution, etc.) but does not classify the particular components according to their functional meaning.

Other studies, such as Brindisi et al. [67], try to adjust a certain clinical application system to the workflows a physician is used to. In contrast, OntoHealth tries to offer a context-aware functionality independent of a specific system by building on EHR standards and best practices. Therefore, the categorization model provides a possibility to describe building blocks of clinical workflows from a user perspective without relating to any specific application system.

Outlook and Clinical Relevance

The conducted scientific review is followed by involving other pertinent literature sources to gather further requirements in order to extend or modify the present categorization model. As a next step, direct observations in the hospital setting will be held in order to identify executed workflows by shadowing diabetes specialists. The combined result of literature review and observations will be used for the following expert interviews.

The project OntoHealth is utilizing the gathered categorization model in further steps for implementing the semantic service grid and service prototypes for diabetes care. The detailed work and progress of the OntoHealth project can be followed online [11]. The establishment of a standards-based flexible EHR usage to a more comprehensive integration within complete clinical workflows by (semi-)automatic service orchestration allows the physician to be the main actor in defining and executing information requests within an EHR system. This could be compared to “apomediation” as referred to in [68], which describes that there is no need for an intermediary when an actor is accessing or retrieving information or executing any other health care workflow. The OntoHealth project creates a functionality access to EHR systems which establishes this direct access for health care professionals to patient-related data.

Conflict of Interest

The authors declare that they have no conflicts of interest in the research.
Human Subjects Protections
Human and/or animal subjects were not included in the project.

Acknowledgments
This work was supported by the Austrian Science Foundation (FWF), project number P 25895-B24.
Fig. 1 Overview of the general OntoHealth architecture

Fig. 2 Results of each step in the literature review.
Fig. 3  Description and categorization model of clinical workflows including cardinalities between the components. Actions, Data Elements and Context show contained items at which framed items form subcategories holding further elements. Numbers show assigned codings to the categories.
Table 1  Setup for the systematic review according to PRISMA, comprising inclusion and exclusion criteria for every phase.

<table>
<thead>
<tr>
<th>Identification</th>
<th>MEDLINE/PubMed, Cochrane, ACM Digital, IEEExplore, EMBASE, Taylor &amp; Francis, Citeseer, DARE, BioMed Central, ScienceDirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords</td>
<td>Diabetes; electronic health record; EHR; clinical workflow; health information technology; information system</td>
</tr>
<tr>
<td>MeSH Terms</td>
<td>“Diabetes Mellitus”[Mesh], „Workflow”[Mesh], „Guidelines as Topic”[Mesh], „Guideline”[Publication Type], „Practice Guideline”[Publication Type], „Health Planning Guidelines”[Mesh], „Medical Informatics”[Mesh], „Information Systems”[Mesh], „Medical Records Systems, Computerized”[Mesh], „Medical Records, Problem-Oriented”[Mesh], „Health Records, Personal”[Mesh], „Electronic Health Records”[Mesh], „Medical Record Administrators”[Mesh], „Medical Record Linkage”[Mesh], „Medical Records”[Mesh]</td>
</tr>
</tbody>
</table>

| Screening      | • Papers not written in English  
|                | • Symposium reports  
|                | • Opinion papers  
|                | • Letters, interviews  
|                | • Commentaries and discussion reports  
|                | • Documents that used abbreviations that stood for other terms than our purpose (e.g. author’s name) |

| Eligibility     | • General statements or tasks regarding functional or non-functional requirements for workflows regarding the interaction with EHRs or other electronic information systems in the context of diabetes or tasks that are included in a clinical workflow.  
|                | • Organizational requirements |

| Exclusion Criteria | • Pure medical studies about diabetes e.g. concerning medical errors or new findings  
|                    | • Studies in emergency departments  
|                    | • Outpatient cases  
|                    | • Workflows out of the physician’s office or medical consultation |

| Included        | Same inclusion criteria as for eligibility |
Table 1  Selected sample categorizations for Data Element and Context including statements found and references.

<table>
<thead>
<tr>
<th>Component</th>
<th>Category</th>
<th>Example Source Statement</th>
<th>Assigned References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Element</td>
<td>HbA1c</td>
<td>“The records of Type 2 diabetics were searched for … glycosylated haemoglobin (HbA1c).”</td>
<td>[28–54]</td>
</tr>
<tr>
<td></td>
<td>Patient-reported data</td>
<td>“… interpret the results of blood glucose self-monitoring …” [43]</td>
<td>[27, 43, 49]</td>
</tr>
<tr>
<td>Context</td>
<td>Informing and Awareness</td>
<td>“… some providers combined patient data sets in order to help patients understand the health consequences” [27]</td>
<td>[27, 29, 30, 32, 33, 35, 39, 41, 43, 50, 51, 55–62]</td>
</tr>
<tr>
<td></td>
<td>Continuing and Monitoring</td>
<td>“… doctors monitor patients’ levels of adherence to care processes (e.g. annual assessment, review visits, education sessions, laboratory tests) and self management as well as their status of attainment of treatment targets.” [55]</td>
<td>[27, 28, 55, 56, 62–65]</td>
</tr>
</tbody>
</table>
References