Transformation of Medical Service Ontology to Relational Data Models

Osamu Takaki[a], Izumi Takeuti[b], Koichi Takahashi[b], Noriaki Izumi[b],
Koichiro Murata[c], Mitsuru Ikeda[a] and Kôiti Hasida[b]
[a]School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST),
Nomi, Japan, {takaki, ikeda}@jaist.ac.jp
[b]National Institute of Advanced Industrial Science and Technology (AIST),
Ibaraki, Japan, {takeuti.i, k.takahashi, n.izumi, hasida.k}@aist.go.jp
[c]School of Medicine, Kitasato University, Sagamihara, Japan, murata-k@kitasato-u.ac.jp

***Abstract*—For assessment of medical service quality in hospitals, it is important to define quality indicators for evaluating medical services and to calculate their values based on data in hospitals databases. Thus, one needs a proper method to correctly calculate the value of a given quality indicator based on data in medical databases. This paper introduces a method to transform Medical Service Ontology (MSO) to relational data models in medical databases, where MSO is an ontology that provides vocabulary words for developing quality indicators. To this end, this paper defines a virtual data model called the Global Data Model (GDM) and a transformation of MSO to GDM by grouping concepts and properties in MSO. Based on the transformation, a quality indicator defined with MSO can be transformed into queries on GDM automatically. Moreover, by developing mappings from GDM to relational data models in medical databases, a quality indicator on MSO can be transformed into queries on the data models, using which the value of the quality indicator is calculated based on data in medical databases.**

Keywords-component; quality indicator; ontology; data model; medical database

I. INTRODUCTION

A quality indicator is a measure of medical service quality that is represented numerically. For example, “in-hospital mortality after stomach cancer surgery” is a quality indicator that indicates the quality of surgeries of stomach cancers in a hospital [1, 2]. For assessment of medical service quality in hospitals, it is important to define quality indicators for evaluating medical services and to calculate their values based on data in hospital databases. Many hospitals and institutes in medical science have developed and published quality indicators and their data [1, 2, 3, 4, 5].

However, despite these efforts, an established framework to develop or maintain quality indicators has not yet been developed. In fact, even though certain groups of hospitals have begun to give quality indicators to collaborating hospitals, to gather their data, and to compare them, it is not easy to compare them fairly. For example, mortality, re-hospitalization rate, incidence of bedsores, and incidence of complications are difficult to compare, since there are no suitable standard definitions of related concepts. Moreover, there often exist various types of gaps in the development of quality indicators and in the calculation of their values. Such gaps arise from a difference in the vocabulary of medical services and/or its interpretations, and they often make the results of evaluations of the quality of medical services unsuitable.

To realize sharing of quality indicators among multiple hospitals and to compare data fairly, a proper framework that helps develop and/or improve quality indicators according to their own environment is needed. We call such a framework QI-framework and describe it by QI-FW. To realize QI-FW, it is desirable to develop a system that creates quality indicators with self-checking vocabulary words and constructs quality indicators and their interpretations. To this end, we developed a representation system of quality indicators, which we describe by QI-RS [6, 7, 8]. QI-RS gives a graph-based representation of a quality indicator that is rigorous and easily understandable, and the system is based on an ontology called “Medical Service Ontology (MSO).” QI-FW helps to develop quality indicators in QI-RS and to calculate their values based on medical databases.

The main purpose of this paper is to present a method to calculate the value of a given quality indicator in QI-RS based on data in medical databases. To this end, this paper introduces a method to transform MSO to data models for relational databases, where MSO is an ontology that provides vocabulary words to define quality indicators. We first define a virtual data model called the Global Data Model (GDM) as a standard model for calculating the values of quality indicators, and introduce transformation of MSO to GDM by grouping concepts and properties in MSO. Based on the transformation, a quality indicator defined using MSO is transformed to queries in the GDM. Moreover, by developing mappings from GDM to data models on given medical databases, a quality indicator in MSO is transformed to queries on the data models, and the value of the quality indicator can be calculated based on the data in the given medical databases.

The remainder of this paper is organized as follows. Section II explains quality indicators and an overview of the QI-FW. Section III explains QI-RS based on our previous studies. Section IV explains a method to transform MSO to GDM and to generate queries on GDM from quality indicators in QI-RS. This section contains the main contribution of the paper. Section V extends the method to transform quality indicators in QI-RS to queries on data models of medical databases. The last section explains related works and conclusions.

II. OUTLINE OF QI-FRAMEWORK

A. *Quality Indicator*

A quality indicator consists of a *name* (or a *label*) and a *calculating formula*. For example, "five-year survival rate of stomach cancer patients" is the name of a quality indicator, and its calculating formula is given as follows [2].

Calculating Formula: Data for the above quality indicator is obtained by calculating the rate of the following numerical values.

Numerator: Number of inpatients that satisfy the condition defined in the denominator and that survived more than five years since they were diagnosed with stomach cancer

Denominator: Number of patients that were diagnosed with stomach cancer

The value that is obtained from a quality indicator by using the calculating formula and data in a hospital (or hospitals) is called "the value of a quality indicator (in a hospital (or hospitals))" or "the data of a quality indicator (in a hospital (or hospitals))." We assume that the values of quality indicators are essentially calculated from data in medical databases.

Even though herein we distinguish between a quality indicator and its calculating formula, we will often refer to the calculating formula of a quality indicator simply as "a quality indicator," unless stated otherwise.

B. *Overview of the QI-Framework*

Here, we briefly explain an overview of the framework QI-FW to develop quality indicators and to calculate their values based on medical databases.

As mentioned in Section I, proper sharing of the definition of a quality indicator is not straightforward. To address the above problems, it is significant to establish a way to unify the vocabulary of quality indicators and their interpretation including their whole structures. Thus, toward solving this problem, we are in the process of developing QI-FW. (At this time, QI-FW has not yet been completed implementation, and it is currently under development.)

QI-FW consists of (1) a representation system QI-RS of quality indicators (Section III), (2) medical databases in hospitals, and (3) mapping systems (Section V). Moreover, QI-RS has MSO as its main component (Section III.A). Medical staff and system engineers who administrate medical databases (and knowledge engineers, if necessary) collaborate in developing and improving MSO.

QI-FW users are assessors of the medical service quality of a hospital (or hospitals) based on data in medical databases, who are patients, medical staff, and so forth (Fig. 1). They can develop quality indicators in QI-RS via some interface of QI-FW. A quality indicator *Q* in QI-RS is expressed as a graph (Section III.B). Some nodes in *Q* are concepts in MSO, while edges in *Q* are properties in MSO.

On the other hand, system engineers who manage medical databases are responsible for developing and improving mapping systems between GDM and data models in medical databases. Concepts and properties in MSO are translated to tables in the Global Data-Model (GDM), which is a virtual relational data model (Section IV). According to the translation and mapping systems described above, *Q* is translated to queries on the data models and an algorithm on the data retrieved by the queries. Through the queries and the algorithm, the user can calculate the value of *Q* based on data in the medical databases.

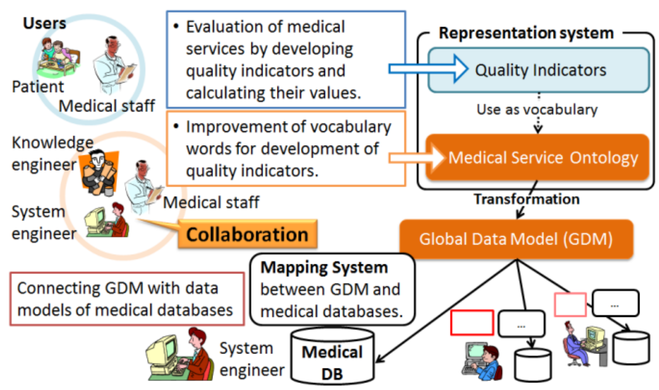


Figure 1. Overview of QI-framework.

III. REPRESENTATION BASED ON MSO

In this section, we explain a graph-based representation system of quality indicators, which we call QI-RS. QI-RS is developed based on the concept of considering a quality indicator as a combination of the quantification target and the quantification and development method of the target and the method independently. We represent the target of the calculation or quantification as a graph that we call an objective graph, and we represent the way to calculate or quantify it by a concept that we call a quantifying concept. Furthermore, objective graphs are constructed based on ontology that we call Medical Service Ontology (MSO).

A. *Medical Service Ontology*

In this sub-section, we define MSO as a vocabulary for calculating formulas of quality indicators. For example, the formal calculation in Section II.A uses the words "patients," "hospitalize (hospitalization)," and "aged." In fact, to define quality indicators, we need words for describing characteristics of patients, events (medical services) in hospitals, predicates about patients, and so forth. MSO is developed in the ontology developing tool Semantic Editor [9].

1) *Patients*

First, we describe the basic concepts related to patients and their attributes in Fig. 2. Yellow rounded rectangles denote concepts, and pink rounded rectangles denote attributes. In general, pink rounded rectangles in diagrams in Semantic Editor denote properties.

 ISBN: 978-1-61208-230-1

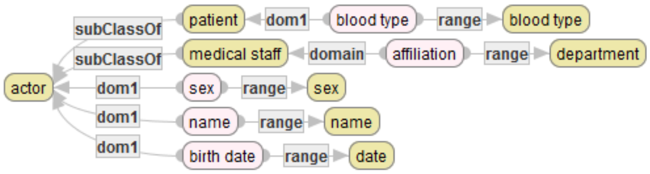


Figure 2. Basic concepts and attributes related to patients.

In this paper, we classify properties between concepts into attributes of concepts and relations between concepts. The concept [patient] has attributes ⟨blood type⟩, ⟨sex⟩, ⟨name⟩, and ⟨birth⟩, where we describe a concept by brackets and labels and an attribute by angle brackets and labels. The values of these attributes are supposed to be immutable for a patient.

2) Events

Next, we explain concepts related to events in a hospital. An event is defined to be what a medical staff or a hospital executes for a patient or what happens to him/her (Fig. 3).

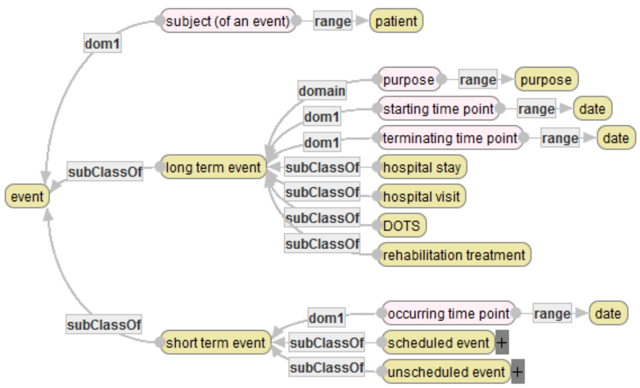


Figure 3. Basis concepts and attributes related to events (partial).

Events are classified into long-term and short-term events. While a long-term event such as a hospital stay (hospitalization) usually takes multiple days, a short-term event does not usually take more than two days. Moreover, short-term events are further classified into scheduled and unscheduled events. Usual medical services are regarded as scheduled events, while accidents such as deaths are regarded as unscheduled events. For example, “admission,” “discharge,” “diagnosis,” “examination,” and “operation (surgery)” are typical scheduled short-term events, while “death,” “falling,” and “bone fracture” are typical unscheduled short-term events. Each typical event is further classified into detailed classes. For example, examination events are classified into about thirty types of examinations.

Each long-term event has attributes including the subject (target patient), purposes, the starting date, and the ending date, while each short-term event has attributes including the subject and occurring time (Fig. 3). We omit a detailed explanation of them due to space limitations.

3) States of Patients

The state of a patient denotes the health state or a condition of a patient at a time point. The diagram in Fig. 4 defines the following main states: age, state of life or death, state of disease that a patient suffers from, and basic body properties. These states are used to describe a feature of a patient as a target of a medical service or an outcome of a patient that cannot be represented by any event.

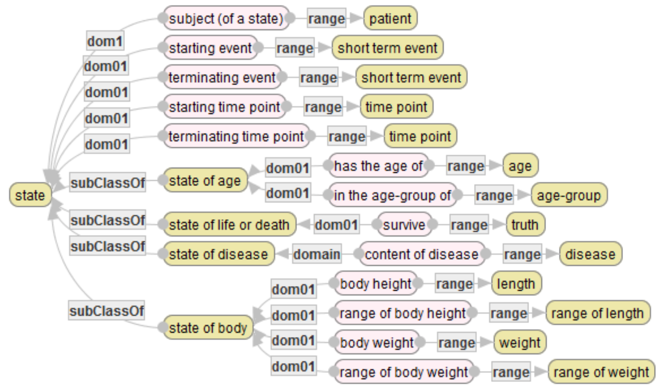


Figure 4. Basis concepts and attributes related to states of patients.

Remark. We often identify a concept with its extension, which is the set of instances of the concept. For example, in a hospital *H*, the concept “patient” is identified with the set of patients of *H*.

4) Main Relations in MSO

In this sub-section, we define properties in MSO that are not attributes. We call them *relations*. We define the primary relations between concepts as follows.

Relations of patients and events: These relations are defined between the concept [patient] and concepts of events. For example, the following relation denotes the relations between patients and their diagnosis (we describe a relation by angle brackets and a label).

⟨subject (of an event)⟩ ⊆ [diagnosis] × [patient].

Note that these relations share the same name, “subject (of an event).” We omit the explanation of the relations between patients and other events due to limitations of space.

Relations of patients and states: These relations are defined between [patient] and concepts of patients’ states. For example, the following relation denotes the relationship between patients and their state of disease.

⟨subject (of a state)⟩ ⊆ [patient] × [state of disease].

These relations also share the same name “subject (of a state)” and all concepts of patients’ states have the common attributes of starting time points and terminating time points. We omit the explanation of the relations between patients and other states.

Relations of time ordering: These relations are defined between the concepts of events and patients’ states. For example, the following relations denote the relationships between operations.

⟨before more than <p>⟩ ⊆ [operation] × [operation],
⟨before less than <p>⟩ ⊆ [operation] × [operation],
⟨after less than <p>⟩ ⊆ [operation] × [operation], and
⟨after more than <p>⟩ ⊆ [operation] × [operation].

 ISBN: 978-1-61208-230-1

Here, "<p>" denotes a parameter. For example, the relation ⟨before more than <2 weeks>⟩ consists of a pair <op_1, op_2> if op_1 and op_2 are performed and if op_1 is performed more than two weeks before op_2.

Belonging relations of events: These relations are defined between concepts of events with no term and events with terms. For example, the following relation denotes the relations between operations and hospital stays that have operations.

⟨belonging⟩ ⊆ [operation] × [hospital stay].

The relation contains a pair (op, sty) of an event of an operation op and that of a hospital stay sty if op is performed in the duration of sty.

5) Special property OR

MSO has a special property with the label "OR." The domain and range of the OR property is the most general concept "class," which is a superclass of all concepts in MSO. Thus, we can give this property to all pairs of concepts in MSO. The property is used in the definition of objective graphs in the next section (Case 3 for the definition).

B. Objective Graphs

In the first part of Section III, we explained that QI-RS represents a quality indicator as a combination of a target concept of quantification and a quantification method. In this sub-section, we define the representation of a quantification target as a graph.

To define a quality indicator such as the example in Section II.A, one needs to define special concepts such as "inpatients that were diagnosed with stomach cancer and that survived more than five years since they received the diagnosis" as a quantification target. In most cases, such a special concept can be defined with a basic concept [patient] or [event] and some conditions on it that specializes the basic concept. Furthermore, such conditions can be defined with properties and (other) basic and/or special concepts. MSO defines the basic concepts and properties, while objective graphs represent special concepts and conditions to define special concepts based on MSO.

1) Syntax of Objective Graphs

Definition 1. An objective graph $\mathbb{G}$ consists of five components ($N(\mathbb{G})$, $R(\mathbb{G})$, $E(\mathbb{G})$, $L(\mathbb{G})$, $C(\mathbb{G})$), where

(i) $N(\mathbb{G})$ is a set of nodes,
(ii) $R(\mathbb{G})$ is the root node,
(iii) $E(\mathbb{G})$ is a set of edges,
(iv) $L(\mathbb{G})$ is a label function on $N(\mathbb{G}) \cup E(\mathbb{G})$, and
(v) $C(\mathbb{G})$ is a concept.

We define $\mathbb{G}$ by the induction on the structure of the node labels, as follows.

Case 1. Assume that the following data are given:
(a) concept C,
(b) attributes $A_1,\ldots, A_n$ of C, and
(c) values $a_1,\ldots, a_n$ of $A_1,\ldots, A_n$, respectively.
Then, we define an objective graph $\mathbb{G}$, as follows.
(i) $N(\mathbb{G}) := \{*_0, \ldots, *_n\}$,
(ii) $R(\mathbb{G}) := *_0$,
(iii) $E(\mathbb{G}) := \{f_1,\ldots, f_n\}$, where each f_i is an edge from $*_0$ to $*_i$.
(iv) $L(\mathbb{G})(*_0) := C$,
$L(\mathbb{G})(*_i) := a_i$ for $i = 1,\ldots, n$, and,
$L(\mathbb{G})(f_i) := A_i$ for $i = 1,\ldots, n$,
(v) $C(\mathbb{G}) := C$.
Note that if n = 0, then $N(\mathbb{G})$ is the singleton set $\{*_0\}$ and $E(\mathbb{G})$ is an empty set.

Case 2. Assume that the following data are given:
(a) an integer n with $n \geqq 1$,
(b) a set of objective graphs $\{\mathbb{G}_0, \ldots, \mathbb{G}_n\}$,
(c) a set of relations $\{R_1,\ldots, R_n\}$, where each R_i is a relation between $C(\mathbb{G}_i)$ and $C(\mathbb{G}_0)$,
(d) a function $n(i,j)$: $\{1,\ldots,n\}^2 \rightarrow \mathbb{N}$, where $\mathbb{N}$ is the set of integers.
(e) a set of relations $\{R^{i,j}_1,\ldots, R^{i,j}_{n(i,j)}\}$, where $0 \leqq i \leqq n$ and j with $0 \leqq j \leqq n$, and each $R^{i,j}_k$ is a relation between $C(\mathbb{G}_i)$ and $C(\mathbb{G}_j)$. (Note: if $n(i,j) = 0$, then $\{R^{i,j}_1,\ldots, R^{i,j}_{n(i,j)}\}$ is an empty set).
Then, we define an objective graph $\mathbb{G}$, as follows.
(i) $N(\mathbb{G}) := \{*_0, \ldots, *_n\}$,
(ii) $R(\mathbb{G}) := *_0$,
(iii) $E(\mathbb{G}) := \{f_1,\ldots, f_n\} \cup (\cup_{0\leqq i\leqq n,\ 0\leqq j\leqq n}\{f^{i,j}_1,\ldots, f^{i,j}_{n(i,j)}\})$, where each f_i is an edge from $*_i$ to $*_0$ and each $f^{i,j}_k$ is an edge from $*_i$ to $*_j$.
(iv) $L(\mathbb{G})(*_i) := \mathbb{G}_i$ ($i = 0,\ldots, n$),
$L(\mathbb{G})(f_i) := R_i$ ($i = 1,\ldots, n$) and,
$L(\mathbb{G})(f^{i,j}_k) := R^{i,j}_k$ ($i,j = 0,\ldots, n$ and $k = 1,\ldots, n(i, j)$).
(v) $C(\mathbb{G}) := C(\mathbb{G}_0)$.

Case 3. Assume that the following data are given:
(a) an integer n with $n\geqq 1$,
(b) a set of objective graphs $\{\mathbb{G}_0, \ldots, \mathbb{G}_n\}$, where each $C(\mathbb{G}_i)$ ($i = 1,\ldots, n$) is a subclass of $C(\mathbb{G}_0)$.
Then, we define an objective graph $\mathbb{G}$, as follows.
(i) $N(\mathbb{G}) := \{*_0, \ldots, *_n\}$,
(ii) $R(\mathbb{G}) := *_0$,
(iii) $E(\mathbb{G}) := \{f_1,\ldots, f_n\}$, where each f_i is an edge from $*_0$ to $*_i$.
(iv) $L(\mathbb{G})(*_i) := \mathbb{G}_i$ ($i=0,\ldots, n$) and
$L(\mathbb{G})(f_i) :=$ OR ($i=1,\ldots, n$), where OR is a special property defined in Section IV.
(v) $C(\mathbb{G}) := C(\mathbb{G}_0)$.

Each f_i and each $f^{i,j}_k$ in Case 2.(iii) are called a *main edge* of $\mathbb{G}$ and an *optional edge* of $\mathbb{G}$, respectively.

Let $\mathbb{G}$ be an objective graph. If $\mathbb{G}$ is defined in Case 1 of the above definition, then it is called an *atomic* objective graph. If $\mathbb{G}$ is defined in Case 2, then it is called a *complex* objective graph. Finally, if $\mathbb{G}$ is defined in Case 3, then it is called an *OR-type* objective graph.

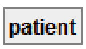
2) Example of an Objective Graph

Here, we give an example of an objective graph. Let us consider the quality indicator "five-year survival rate of stomach cancer patients". The definition of the quality indicator is the ratio of the number of patients surviving five years to all stomach cancer patients, where a "stomach cancer patient" is a patient who was diagnosed with stomach cancer, and a "five-year surviving patient" is a patient diagnosed with stomach cancer but who is alive five years after the diagnosis. Thus, we will first express the set of five-year surviving patients in Fig.5. To this end, we construct three objective graphs $\mathbb{G}_0$, $\mathbb{G}_1$, and $\mathbb{G}_2$, as follows.

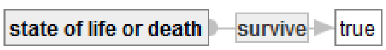
(1) $\mathbb{G}_0 = (\{*\}, *, \emptyset$ (empty set), L_0, [patient]), where $L_0(*) =$ [patient].

(2) $\mathbb{G}_1 = (\{*_0, *_1\}, *_1, \{f_1{:}*_0\rightarrow *_1\}, L_1,$ [diagnosis]), where $L_1(*_0) =$ [diagnosis], $L_1(*_1) =$ ⟪stomach cancer⟫, $L_1(f_1) =$ ⟨result⟩ and [diagnosis] denotes an event concept, ⟪stomach cancer⟫ denotes an instance of the concept of diseases, and ⟨result⟩ denotes an attribute of the concept [diagnosis]. Note that the range of ⟨result⟩ is the concept of diseases.

(3) $\mathbb{G}_2 = (\{*_0, *_1\}, *_1, \{f_1{:}*_0\rightarrow *_1\}, L_2,$ [state of life or death]), where $L_2(*_0) =$ [state of life or death], $L_2(*_1) =$ ⟪true⟫, $L_2(f_1)$ = ⟨survive⟩, [state of life or death] denotes the viability status of a patient, ⟪stomach cancer⟫ denotes an instance of the concept of diseases, and ⟨result⟩ denotes an attribute of the concept [diagnosis]. Note that the range of ⟨result⟩ is the concept of diseases.

(1) $\mathbb{G}_0$= 

(2) $\mathbb{G}_1$= 

(3) $\mathbb{G}_2$= 

We next construct an objective graph $\mathbb{G}$ of "five-year surviving stomach cancer patients", as follows.

(i) $N(\mathbb{G}) = \{*_0, *_1, *_2\}$,

(ii) $R(\mathbb{G}) = *_0$,

(iii) $E(\mathbb{G}) = \{f^1{:}*_1\rightarrow *_0, f^2{:}*_2\rightarrow *_0, f^{21}{:}*_2\rightarrow *_1\}$,

(iv) $L(\mathbb{G})(*_i) = \mathbb{G}_i$ $(i = 0, 1, 2)$,

$L(\mathbb{G})(f^1)$ = ⟨subject (of the event)⟩,

$L(\mathbb{G})(f^2)$ = ⟨subject (of the state) ⟩,

$L(\mathbb{G})(f^{21})$ = ⟨after more than <5 years>⟩,

(v) $C(\mathbb{G}) = C(\mathbb{G}_0)$ = [patient].

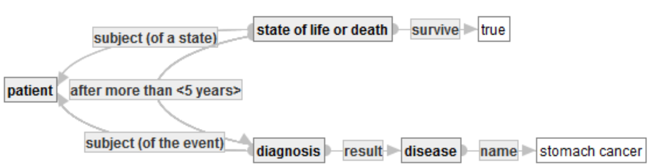


Figure 5. Objective graph $\mathbb{G}$ describing five-year surviving patients with stomach cancer.

3) *Semantics of Objective Graphs*

An objective graph $\mathbb{G}$ can be regarded as a concept denoted by $C(\mathbb{G})$ and modified by other concepts and properties that are denoted by $L(\mathbb{G})$. If each concept is identified with its extension, i.e., the set of instances of the concept, an objective graph can be identified with a subset of the set denoted by $C(\mathbb{G})$ that is obtained from $C(\mathbb{G})$ by restricting it by sets and functions denoted by $L(\mathbb{G})$. To clarify this identification, we define an interpretation of an objective graph, as follows.

Definition 2. For an objective graph $\mathbb{G}$, we define a set $[[\mathbb{G}]]$, which is called the *interpretation* of $\mathbb{G}$, as follows.

Case 1. Let $\mathbb{G}$ be an atomic objective graph. Then, $[[\mathbb{G}]] := \{c \in C(\mathbb{G}) | c.A_1 = a_1 \wedge \ldots \wedge c.A_n = a_n\}$, where $c.A_i$ is the value of the attribute A_i on c, and symbol $\wedge$ denotes the logical connective symbol "and."

Case 2. Let $\mathbb{G}$ be a complex objective graph. Then, $[[\mathbb{G}]] := \{x_0 \in [[\mathbb{G}_0]] | \exists x_1 \in [[\mathbb{G}_1]], \ldots, \exists x_n \in [[\mathbb{G}_n]]$ $(\bigwedge_{i=1,\ldots,n} R^i(x_i, x_0)) \wedge (\bigwedge_{i,j=0,\ldots,n} (\bigwedge_{k=1,\ldots,n(i,j)} R^{i,j}{}_k(x_i, x_j)))\}$.

Case 3. Let $\mathbb{G}$ be an OR-type objective graph. Then, $[[\mathbb{G}]] := [[\mathbb{G}_0]] \cap ([[\mathbb{G}_1]] \cup \ldots \cup [[\mathbb{G}_n]])$.

Note that for all concepts C and D we assume that if D is a subclass of C then the extension of D is a subset of that of C.

4) *Segments of an Objective Graph*

To define some quantifying concepts in the next section, we will use an objective graph $\mathbb{G}^*$, which is a kind of a subgraph of $\mathbb{G}$. We call such an objective graph a *segment* of $\mathbb{G}$ and describe it by $\mathbb{G}^*$. $\mathbb{G}^*$ is obtained from $\mathbb{G}$ by reducing some conditions of $\mathbb{G}$ that are represented by nodes or edges in $\mathbb{G}$. In fact, $\mathbb{G}^*$ satisfies that $[[\mathbb{G}]] \subseteq [[\mathbb{G}^*]]$.

Due to space limitations, we omit the definition of segments, and just give an example in Fig.6, which is a segment of the objective graph $\mathbb{G}$ in Fig.5.

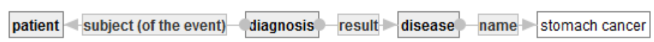


Figure 6. Segment $\mathbb{G}^*$ of $\mathbb{G}$.

C. *Graph Representation of Quality Indicators*

The main purpose of this study is to introduce transformation of MSO to relational data models and a method to transform quality indicators in QI-RS to queries on the relational data models. However, we will focus on only the transformations of MSO and objective graphs (see Sections IV and V for the reason). Thus, we explain quantification concepts and quality indicators in QI-RS very briefly.

1) *Quantifying Concepts*

A quantifying concept is a function that extracts the value from a given target that is expressed to be an objective graph. The function may have input data that are some attributes of the base concept of a given objective graph and another function on the values of the above attributes. Here, each concept is regarded as a set obtained from the concept under some situation. For example, in a hospital *A*, the concept "inpatient" is regarded as the set of inpatients in *A*.

There are generally three types of quantifying concepts:

a) *Total Numbers*

For a finite set *S*, the summation of numbers obtained from elements of *S* is called the total number of *S*. For example, if each element is assigned 1 denoting its existence, then the total number is the same as the cardinality of *S*. The quantifying concept ≪cardinality≫ is regarded as a function that has an objective graph $\mathbb{G}$ as input data and that outputs the cardinality of $[[\mathbb{G}]]$.

For a concept *S* and attributes $A_1, \ldots, A_n$ of *S*, a real-valued function on the product set of (extensions of) $A_1,\ldots, A_n$ is called an *attribute-quantifying function*. Moreover, for a concept *S*, attributes $A_1,\ldots, A_n$ of *S*, and an attribute-quantifying function f of $A_1, \ldots, A_n$, the summation $\Sigma_{s \in S} f(s.A_1, \ldots, s.A_n)$ is called the *total attribute number* of S with respect to $A_1, \ldots, A_n$ and f, where $s.A_i$ denotes the value of an instance s with respect to A_i.

The quantifying concept ≪total attribute number≫ is regarded as a function that has the following data as input data:

1. an objective graph $\mathbb{G}$,
2. attributes $A_1, \ldots, A_n$ of C($\mathbb{G}$), and
3. an attribute-quantifying function f of $A_1, \ldots, A_n$.

Moreover, the ≪total attribute number≫ outputs the total attribute number of [[$\mathbb{G}$]] with respect to $A_1, \ldots, A_n$ and f.

b) Rate

For a finite set S and its subset S^*, the rate of the total number of S^* among the total numbers of S is obtained in the same way as that of calculating the total number of S^* is called a rate of S^* among S. In particular, the rate of the cardinality of S^* among that of S is called the cardinality rate of S^* among S. Moreover, the rate of the total attribute number of S^* with respect to $A_1, \ldots, A_n$ and f among that of S with respect to the same attributes and the same attribute-quantifying function is called the total attribute number rate.

The quantifying concept ≪cardinality rate≫ is regarded as a function that has the following data as inputs:

1. an objective graph $\mathbb{G}$, and
2. a segment $\mathbb{G}^*$ of $\mathbb{G}$.

In contrast, the quantifying concept ≪total attribute number rate≫ is regarded as a function that has the following data as inputs:

1. an objective graph $\mathbb{G}$,
2. a segment $\mathbb{G}^*$ of $\mathbb{G}$,
3. attributes $A_1, \ldots, A_n$ of C($\mathbb{G}$), and
4. an attribute quantifying function f of $A_1, \ldots, A_n$.

Moreover, ≪total attribute number rate≫ outputs the rate of the total attribute number of [[$\mathbb{G}$]] with respect to $A_1, \ldots, A_n$ and f among that of [[$\mathbb{G}^*$]] with respect to the same attributes and the same attribute-quantifying function.

c) Average

For concept S, attributes $A_1, \ldots, A_n$ of S, and the attribute-quantifying function f, the ratio of the total attribute number of S with respect to $A_1, \ldots, A_n$ and f and the cardinality of S is called the average of the value of S with respect to $A_1, \ldots, A_n$ of f. The quantifying concept ≪average≫ is regarded as a function that has the same input data as that of ≪total attribute number≫ and that outputs the average of the value of S with respect to $A_1, \ldots, A_n$ of f.

2) Outline of Quality Indicator Graphs

We now explain the graph representation of a quality indicator in QI-RS by showing examples of quality indicator graphs. We call such a graph-represented quality indicator a *quality indicator graph*.

In Figs. 6 and 7, we showed an objective graph $\mathbb{G}$ representing five-year surviving patients with stomach cancer and its segment $\mathbb{G}^*$, respectively. Now, we connect them with a quantifying concept ≪cardinality rate≫ in Fig.7. Then, the labeled graph that consists of $\mathbb{G}$, $\mathbb{G}^*$, and the new node is a quality indicator graph that represents the quality indicator in Section II.A.

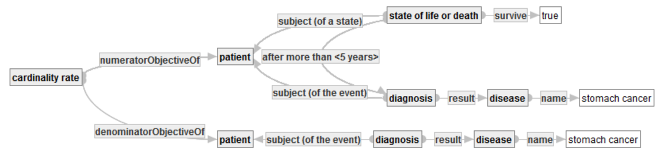


Figure 7. Quality indicator graph of five-year survival rate of stomach cancer patients.

IV. TRANSFORMATION OF MSO TO GDM

In this section, we introduce a transformation of MSO to GDM.

A. GDM and Transformation of MSO to GDM

GDM is a relational data model [10] that is obtained from MSO by transforming MSO with a standard technique, where MSO is regarded as an entity relationship model (ERM) [11] in the following manner.

1) Classification of Concepts in MSO

Concepts in MSO are grouped into *entity-type* and *dataset-type*, as follows. Let C be a concept in MSO.

i. If C has some attribute(s), then C is an entity-type concept.

ii. Otherwise, C is a dataset-type concept.

Basically, a concept that is an objective of quantification is an entity-type concept. For example, subclasses of [actor] (essentially [patient]), [event], and [state (of a patient)] are entity-type concepts. In fact, there exist entity-type concepts that are not objectives of quantification but are used for defining quality indicators, but we do not explain them in this paper. For example, the concept [disease] in Fig.8 below is an auxiliary concept that has attributes, and hence, it is an entity-type concept.

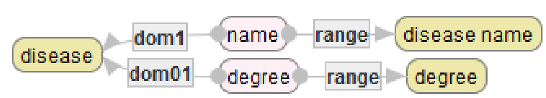


Figure 8. Concepts related to diseases.

On the other hand, a dataset-type concept is basically defined as a range of attributes of other concepts. Dataset-type concepts are further classified into *general dataset-type* and *medical knowledge-type*. For example, the concept [patient] has four attributes ⟨name⟩, ⟨birth⟩, ⟨sex⟩, and ⟨blood type⟩, while [disease] has two attributes ⟨name⟩ and ⟨degree⟩. The ranges of the attributes are concepts [human name], [birth], [sex], [blood type], [disease name], and [degree], respectively. These concepts are dataset-type. Specifically, [blood type] and [disease name] are medical knowledge-type concepts, while the others are general dataset-type concepts. The extension of a dataset-type concept is the same across hospitals. For example, the extensions of the above concepts are the same in every hospital. On the other hand, the extension of an entity-type concept such as [patient] is considered different among hospitals.

2) Classification of Properties in MSO

 ISBN: 978-1-61208-230-1

Properties in MSO are grouped into *relationship-type* and *attribute-type*, as follows. Let P be a property in MSO.

i. If P is a relation defined in Section III.A.4, then P is a relationship-type.
ii. If P is an attribute of a concept C in MSO and for every instance c contained in the extension of C c has at most one value with respect to P, then P is an attribute-type property.
iii. Otherwise, P is a relationship-type property.

From the MSO that can be regarded as the ERM by the above groupings, one can obtain a relational data model, which we describe by $\mathrm{GDM}_{\mathrm{MSO}}$ or simply GDM, using the standard transformation of ERMs to relational data models [11]. In fact, the relational data model GDM can be obtained from MSO by the following transformation of concepts and properties in MSO. We describe the transformation by | - |.

1. An entity-type concept C is transformed to an entity table $|C|$. Here, we set the primary key of $|C|$, which we describe by C-ID.
2. A dataset-type concept D, which has some concept(s) having attributes whose range is D, is transformed to a dataset (data-type) $|D|$.
3. A relationship-type property R between concepts C_1 and C_2 is transformed to a relationship table $|R|$. Here, we set a pair of C_1* and C_2* as the primary key of $|R|$, where, for $i = 1$ or 2, C_i* is C_i-ID (if C_i is an entity-type concept) or $|C_i|$ (if C_i is a dataset-type concept).
4. An attribute-type property A, which has a concept C having A as its attribute, is transformed to an attribute $|A|$ of the entity table $|C|$. Here, the data-type of $|A|$ is the dataset $|D|$ of the range D of A (if D is a dataset-type concept) or the set of values of the primary key of $|D|$ (if D is an entity-type concept).

3) Example of Tables in GDM

We consider the concepts [patient] in Fig. 2 and [diagnosis] in Fig.9 below.

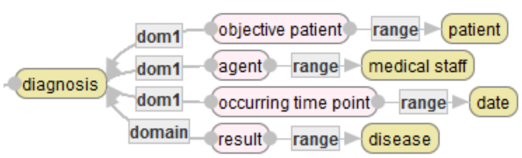


Figure 9. Diagnosis-related Concepts.

In Fig.9, the label “dom1” of the edge between ⟨objective patient⟩ and [diagnosis] indicates that each diagnosis, which is an instance d of [diagnosis], has a single patient that is the value of d with respect to ⟨objective patient⟩. Therefore, ⟨objective patient⟩ is an attribute-type property. Similarly, ⟨agent⟩ and ⟨occurring time point⟩ are attribute-type properties. On the other hand, ⟨result⟩ is a relationship-type property, since the label “domain” indicates that it is possible that some diagnosis has multiple diseases. On the other hand, [diagnosis], [patient] [medical staff] and [disease] are entity-type concepts, while [date] is a dataset-type concept.

From the above concepts and properties, one obtains the following entity and relationship types.

(Entity table of [diagnosis])
<u>diagnosis-ID</u>
objective patient: patient-ID
agent: medical staff-ID
occurring time point: |date|

(Entity table of [patient])
<u>patient-ID</u>
name: |name|
sex: |sex|
birth date: |date|
blood type: |blood type|

(Entity table of [medical staff])
<u>Medical staff-ID</u>
name: | human name|
sex: |sex|
birth date: |date|
affiliation: |department|

(Entity table of [disease])
<u>disease-ID</u>
name: |disease name|
degree: |degree|

(Relationship table of [result])
<u>diagnosis-ID</u>
<u>disease-ID</u>

Here, the keys with an underscore denote primary keys of tables. Moreover, “objective patient: patient-ID” indicates that the attribute “objective patient” has the data-type that is the set of values of patient-ID as a foreign key, while “occurring time point: |date|” indicates that the attribute “occurring time point” has the data-type |date| that is the dataset obtained from [date]. Other attributes have similar data-types.

In what follows, we often abbreviate brackets “[” and “]”. For example, we abbreviate “[patient]-ID” as “patient-ID” and “|[date]|” as “|date|.”

B. Transformation of QI to Queries on GDM

In this sub-section, we define a method of generating queries from quality indicators in QI-RS based on the transformation in the previous section and the semantics of objective graphs in Definition 2. A quality indicator graph consists of (an) objective graph(s) and a quantifying concept, and an objective graph represents a set of patients or events that is a target of the quantification represented by the quantifying concept. On the other hand, a quantifying concept represents a method of calculating numerical values from sets of patients or events represented by objective graphs. The calculation method is simple and independent of objective graphs. Thus, the transformation of a quantifying concept to the corresponding algorithm can be performed uniquely. Due to limitations of space, we define only a transformation of an objective graph to a query on GDM.

Definition 3. For an objective graph $\mathbb{G} = (\mathrm{N}, \mathrm{R}, \mathrm{E}, \mathrm{L}, \mathrm{C})$, we define an SQL query on GDM, which is described by $Q_{\mathbb{G}}$, as follows.

Case 1. Let $\mathbb{G}$ be an atomic objective graph with attributes $A_1,\ldots, A_n$ of C and values $a_1,\ldots, a_n$ of $A_1,\ldots, A_n$, respectively. Then, $Q_{\mathbb{G}}$:= SELECT ∗ FROM $|C(\mathbb{G})|$
WHERE $cond_1$ AND … AND $cond_n$.

Here, $cond_i$ is defined as follows.

i. If $|A_i|$ is an attribute of $|C(\mathbb{G})|$, then $cond_i$ is defined to be $(|A_i| = a_i)$.

ii. Otherwise, $cond_i$ is defined to be
(a_i IN (SELECT $|D_i|$ FROM $|A_i|$)),
where $|A_i|$ is a relationship table obtained from A_i and $|D_i|$ is a data-type obtained from the range of A_i.

Case 2. Let $\mathbb{G}$ be a complex objective graph with the same components as those in Case 2 of Definition 1. Then,

$Q_{\mathbb{G}}$:= SELECT ∗ FROM $|C(\mathbb{G})|$
WHERE
EXISTS(SELECT ∗ FROM $|Q_{\mathbb{G}1}|,\ldots,|Q_{\mathbb{G}n}|$
WHERE (
$cond_1$ AND ⋯ AND $cond_n$
AND
($cond_{0,0,1}$ AND ⋯ AND $cond_{0,0,n(0,0)}$)
AND
($cond_{0,1,1}$ AND ⋯ AND $cond_{0,1,n(0,1)}$)
AND ⋯ AND
($cond_{n,n,1}$ AND ⋯ AND $cond_{n,n,n(n,n)}$))).

Here,
$|Q_{\mathbb{G}i}|$ is the table obtained by the query $Q_{\mathbb{G}i}$,
$cond_i$:= ($|C(\mathbb{G})|.\,C(\mathbb{G})\text{-ID} = |R^i|.\,C(\mathbb{G})\text{-ID}$ AND
$|Q_{\mathbb{G}i}|.\,C(\mathbb{G}_i)\text{-ID} = |R^i|.\,C(\mathbb{G}_i)\text{-ID}$)
and
$cond_{i,j,k}$:= ($|Q_{\mathbb{G}i}|.\,C(\mathbb{G}_i)\text{-ID} = |R^{i,j}_k|.\,C(\mathbb{G}_i)\text{-ID}$ AND
$|Q_{\mathbb{G}j}|.\,C(\mathbb{G}_j)\text{-ID} = |R^{i,j}_k|.\,C(\mathbb{G}_j)\text{-ID}$).

Case 3. Let $\mathbb{G}$ be an OR-type objective graph with the same components as those in Case 3 of Definition 1. Then,

$Q_{\mathbb{G}}$:= SELECT ∗ FROM $|C(\mathbb{G})|$
WHERE
EXISTS(SELECT ∗ FROM $|Q_{\mathbb{G}1}|,\ldots,|Q_{\mathbb{G}n}|$
WHERE ($cond_1$ OR …OR $cond_n$)).

Here, $|Q_{\mathbb{G}i}|$ is the table obtained from the query $Q_{\mathbb{G}i}$ and $cond_i$:= ($|C(\mathbb{G})|.\,C(\mathbb{G})\text{-ID} = |Q_{\mathbb{G}i}|.\,C(\mathbb{G}_i)\text{-ID}$).

Note that $C(\mathbb{G}) = C(\mathbb{G}_0)$ in the above Cases 2 and 3.

Remark. The relation of time ordering defined in Section III.A.4 has a parameter of time length. Thus, in Case 2 of Definition 3, if R^i or $R^{i,j}_k$ is a relation of time ordering, $cond_i$ or $cond_{i,j,k}$ may have the third condition for such a parameter. For example, if $|R^i|$ is a relation of time ordering with a year parameter, then $cond_i$ may have an additional condition
$|R^i|$.|year parameter| = p,
where p has the value of year length.

1) Example of queries on GDM

We now construct a query on GDM from the objective graph $\mathbb{G}$ in Example 0 by Definition 3. By the transformation | - |, we obtain the following tables from the concept [state of life or death] and relations ⟨subject (of an event)⟩, ⟨state object⟩, and ⟨after more than <p>⟩ (Section 3.A.4), which are used to compose $\mathbb{G}$.

(Entity table of [state of life or death])
<u>state of life or death-ID</u>
subject (of a state): patient-ID
starting event: short term event-ID
terminating event: short term event-ID
starting time point: |time point|
terminating time point: |time point|
survive: |truth|

(Relationship table of ⟨subject (of an event)⟩)
<u>diagnosis-ID</u>
<u>patient-ID</u>

(Relationship table of ⟨state object⟩)
<u>state of life or death-ID</u> (We abbreviate it as "LorD-ID.")
<u>patient-ID</u>

(Relationship table of ⟨after more than <p>⟩)
<u>diagnosis-ID</u>
<u>state of life or death-ID</u>
year parameter: |number|

By using the above tables and Example 1, we can obtain the query $Q_{\mathbb{G}}$ as follows.

$Q_{\mathbb{G}}$:= SELECT ∗ FROM |patient|
WHERE
EXISTS(
SELECT ∗ FROM $|Q_{\mathbb{G}1}|, |Q_{\mathbb{G}2}|$
WHERE
|patient|.patient-ID=|subject (of event)|.patient-ID
AND
$|Q_{\mathbb{G}1}|$.diagnosis-ID=|subject (of event)|.diagnosis-ID
AND
|patient|.patient-ID=|state object|.patient-ID
AND
$|Q_{\mathbb{G}2}|$.LorD-ID=|state object|.LorD-ID
AND
$|Q_{\mathbb{G}1}|$.diagnosis-ID
=|after more than <p>|.diagnosis-ID
AND
$|Q_{\mathbb{G}2}|$.LorD-ID=|after more than <p>|.LorD-ID
AND
|after more than <p>|.|year parameter|=5)

$Q_{\mathbb{G}1}$:= SELECT * FROM |diagnosis|
WHERE
EXISTS(
SELECT * FROM | $Q_{\mathbb{G}11}$ |
WHERE
|diagnosis|.diagnosis-ID=|result|.diagnosis-ID
AND
| $Q_{\mathbb{G}11}$ |.disease-ID=|result|.disease-ID)

$Q_{\mathbb{G}11}$:= SELECT * FROM |disease|
WHERE name = “stomach cancer”

$Q_{\mathbb{G}2}$:= SELECT * FROM |state of life or death|
WHERE survive = “true”

C. Elimination of Relationship Tables Obtained from MSO-Relations

Let us call a relationship table in GDM that is obtained from a relation in MSO an *MSO-relation-based table*. In this sub-section, we explain how MSO-relation-based tables can be eliminated. In fact, one can give a relation in MSO a canonical definition by using the attributes of domain concepts, as follows.

1. *Relations of patients and events* are determined by using the attribute ⟨objective patient⟩ of events.
2. *Relations of patients and states* are determined by using the attribute ⟨subject (of a state)⟩ of patient states.
3. *Relations of time ordering* are determined by using the attributes of starting, terminating, and occurring time points of events and patient states.
4. *Belonging relations of events* are determined by the attributes of time points, objective patients, and types of events.

Thus, one can regard MSO-relation-based tables as view tables defined by using other tables in GDM, and hence, for a query $Q_{\mathbb{G}}$ in Definition 3, one can replace all conditions in Q defined using relation-based tables from other conditions without an MSO-relation-based table.

However, it is still meaningful to consider MSO-relation-based tables. The reason is that some hospitals may need their own definitions of such relations. Especially, it is possible that the definitions of belonging relations of events have slight differences across hospitals that need to modify the conditions for the table to contain an instance. To calculate proper indicators of medical service quality, it would be useful to consider a framework of both canonical ways and special ways to deal with it.

V. TRANSFORMATION OF MSO TO A LOCAL DATA MODEL

Based on the transformation of MSO to GDM, we briefly explain a transformation of MSO to a data model on a medical database in a hospital. We call such a data model a *Local Data Model (LDM)*. To realize transformation of MSO to LDM, it is necessary to develop a mapping between GDM and LDM. From the remark in the previous section, one can effectively omit MSO-relation-based tables from GDM. Thus, to develop a mapping between GDM and LDM, we need to develop entity tables and relationship tables (besides MSO-relation-based tables) as view tables on LDM, which we call proper relationship tables. Through the queries on LDM that are definitions of the view tables, one can obtain queries on LDM to calculate the value of a given quality indicator in QI-RS. In fact, for a given $\mathbb{Q}$ consisting of one or two objective graph(s) $\mathbb{G}$ (and $\mathbb{G}^*$), one can obtain a query $Q^{+}_{\mathbb{G}}$ (and $Q^{+}_{\mathbb{G}^*}$), as follows.

1. First, obtain a query $Q_{\mathbb{G}}$ on GDM, which is defined based on entity tables and proper relationship tables.
2. Then, replace the tables above by sub-queries on LDM that define the tables as view tables on LDM.
3. Then, one can obtain a nested query $Q^{+}_{\mathbb{G}}$ on LDM to calculate the values of the objective graph $\mathbb{G}$.

On the other hand, one can easily define an algorithm that is obtained from the quantifying concept in $\mathbb{Q}$. For example, if the quantifying concept is ≪cardinality rate≫, the algorithm only counts the numbers of rows in the tables that are calculated by $Q^{+}_{\mathbb{G}}$ and $Q^{+}_{\mathbb{G}^*}$, respectively, and shows the ratio between the numbers. Such an algorithm can be defined independently from the input data, the objective graphs. Thus, one can obtain the value of $\mathbb{Q}$ from the algorithm above based on $Q^{+}_{\mathbb{G}}$ and $Q^{+}_{\mathbb{G}^*}$.

VI. RELATED WORKS

A. Transformation of Queries on GDM into Those on LDM

The concept of a GDM and the transformation of GDM into LDMs in this paper have already been developed in previous research on distributed databases. Especially in the 1990s, many productive algorithms were developed for the transformation of queries on a GDM into those on LDMs (see, for example, chapter 9 of [12] or [13]). However, we need a representation system of quality indicators that satisfies compatibility of formality and understandability of quality indicator representation. Moreover, to guarantee the formality of the representation system, one needs to establish a concrete way to calculate values of quality indicators represented by the system. This paper shows a solution to the problem by transforming quality indicators (more precisely, objective graphs) in QI-RS into queries on LDMs.

B. Ontology-Based Information Retrieval

Ontology-based information retrieval has been actively investigated. In particular, research results on transformations between ontologies in RDF and relational data models on RDBs (see, for example, [14] or [15]) are closely related to the results in this paper. Moreover, ontology-matching (see, for example, [16] or [17]) has been investigated as the basis of ontology-based information retrieval.

The results of this paper are based on the transformation of MSO into GDM, and previous research can be consulted to reproduce these results. However, in order to actually realize a transformation from ontology such as MSO, which is developed for a special purpose such as the assessment of

 ISBN: 978-1-61208-230-1

medical services, to a suitable data model, one still requires special techniques. For example, although D2RQ [18] is one of most useful tools to connect ontologies in RDF with relational data models, we found that it requires considerable customization and extension of the functions to translate MSO into GDM. The results in this paper can be regarded as the special requirements based on common techniques of transformation on RDF-ontologies into relational data models to transform MSO into GDM (or LDMs).

Moreover, to calculate the value of a quality indicator in QI-RS based on medical databases, one also needs to deal with the transformation of objective graphs and quantifying concepts, particularly the transformation of objective graphs into queries on GDM (and LDMs). This paper solves this problem by employing the semantics (interpretations) of objective graphs (Definitions 3 and 4).

VII. CONCLUSION AND FUTURE WORK

This paper introduces a method to transform a quality indicator represented by Medical Service Ontology (MSO) to queries on a virtual relational data model called a Global Data Model (GDM). To this end, concepts and properties in MSO are grouped and accordingly transformed to entity and relationship tables in GDM. Moreover, based on a mapping from GDM to each relational data model on a medical database in a hospital, this study extends the method to transform a quality indicator to queries on the data models on medical databases. Thus, the value of the quality indicator based on MSO is automatically calculated based on data in the medical databases.

We have still not explained any method to perform mapping between GDM and the relational data model on a given medical database in detail. Moreover, we need to develop a method of generating queries on relational data models that are efficient from the viewpoint of computation. These issues will be solved in future works.

ACKNOWLEDGMENT

This work was supported by a JSPS KAKENHI Grant Number 24500167.

REFERENCES

[1] OECD: Health Care Indicators Project Initial Indicators Report, OECD Health Working Papers 22 (2006).

[2] Nihon Hospital Organization: Clinical Indicators 2009, http://www.hosp.go.jp/7,7018,61.html, (in Japanese), [retrieved: 9, 2012].

[3] B. T. Collopy: Clinical indicators in accreditation: An effective stimulus to improve patient care, International Journal for Quality in Health Care 12(3) (2000), pp.211-216.

[4] J. Mainz, B. R. Krog, B, Bjørnshave, and P. Bartels: Nationwide continuous quality improvement using clinical indicators: The Danish national indicator project, International Journal for Quality in Health Care 16(1) (2004), pp.i45-i50.

[5] J. Mainz, A. M. Hansen, T. Palshof, and P. D. Bartels: National quality measurement using clinical indicators: The Danish national indicator project, Journal of Surgical Oncology 99(8) (2009), pp.500-504.

[6] O. Takaki, I. Takeuti, K. Takahashi, N. Izumi, K. Murata, and K. Hasida: Representation system of quality indicators towards accurate evaluation of medical services based on medical databases, Proceedings of the 4th International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED2012) (2012), pp.249-255.

[7] O. Takaki, I. Takeuti, K. Takahashi, N. Izumi, K. Murata, and K. Hasida: Representation system for quality indicators by ontology, Semantics - Advances in Theories and Mathematical Models, InTech, April (2012), pp.193-212.

[8] O. Takaki, I. Takeuti, K. Takahashi, N. Izumi, K. Murata, M. Ikeda, and K. Hasida: Evaluation of a representation system of quality indicators, Proceedings of the 10th Conference on Knowledge-Based Software Engineering (JCKBSE2012), Frontiers in Artificial Intelligence and Applications 240, IOS Press, (2012), pp.144-153.

[9] K. Hasida: Introduction to Semantic Editor, http://i-content.org/semauth/intro/index.html, (in Japanese), [retrieved: 9, 2012].

[10] E. F. Codd: A relational model of data for large shared data banks, Commun. ACM 26, 1 (1983), pp.64-69.

[11] P. P. S. Chen: The entity-relationship model: A basis for the enterprise view of data, In Proceedings of National Computer Conference (AFIPS '77), ACM, (1977), pp.77-84.

[12] M. T. Özsu and P. Valduriez: Principles of Distributed Database Systems (3rd Ed.), Springer, (2011).

[13] A. Y. Halevy: Answering queries using views: A survey, The VLDB Journal 10 (2001), pp.270-294.

[14] N. Konstantinou, D.-E. Spanos, and N. Mitrou: Ontology and database mapping: A survey of current implementations and future directions. J. Web Eng., 7(1) (2008), pp.1-24.

[15] S. S. Sahoo, W. Halb, S. Hellmann, K. Idehen, T. Thibodeau Jr, S. Auer, J. Sequeda, and A. Ezzat: A Survey of Current Approaches for Mapping of Relational Databases to RDF, http://www.w3.org/2005/Incubator/rdb2rdf/RDB2RDF_SurveyReport.pdf, (2009), [retrieved: 9, 2012].

[16] N. F. Noy: Semantic integration: A survey of ontology-based approaches. SIGMOD Record, 33(4) (2004), pp.65-70.

[17] J. Euzenat and P. Shvaiko: Ontology Matching. Springer (2007).

[18] D2RQ: Accessing Relational Databases as Virtual RDF Graph, http://d2rq.org, [retrieved: 9, 2012].

