Towards an Ontology-Driven Clinical Experience Sharing Ecosystem: Demonstration with Liver Cases

María del Mar Roldán-García [a,*,1] Suzan Uskudarli [b,1] Neda B. Marvasti [c] Burak Acar [c] José F. Aldana-Montes [a]

[a] *Dept. de Lenguajes y Ciencias de la Computación, University of Málaga, Campus de Teatinos, Málaga - 29071, Spain*

[b] *Department of Computer Engineering, Bogazici University, Istanbul, Turkey*

[c] *Electrical & Electronics Engineering Department, Bogazici University, Istanbul, Turkey*

Abstract

Past medical cases, hence clinical experience, are invaluable resources in supporting clinical practice, research, and education. Medical professionals need to be able to exchange information about patient cases and explore them from subjective perspectives. This requires a systematic and flexible methodology to case representation for supporting the exchange of processable patient information. We present an ontology based approach to modeling patient cases and use patients with liver disease conditions as an example. To this end a novel ontology, LiCO, that utilizes well known medical standards is proposed to represent liver patient cases. The utility of the proposed approach is demonstrated with semantic queries and reasoning using data collected from real patients. The preliminary results are promising in regards to the potentials of ontology based medical case representation for building case-based search and retrieval systems, paving the way towards a Clinical Experience Sharing platform for comparative diagnosis, research, and education.

Key words: Ontology driven application; Medical case representation; Knowledge representation; Biomedical resources and systems.

1. Introduction

Historically, the discipline of medicine has been developed on the collective experience gathered over centuries, if not millennia. This experience has been passed over generations in a master-apprentice relation and later through written accounts of past cases, until modern medicine endeavored to explain clinical outcomes by means of causal relations. Despite the leap forward that

* Corresponding author.

Email addresses: mmar@lcc.uma.es (María del Mar Roldán-García), suzan.uskudarli@boun.edu.tr (Suzan Uskudarli), neda.marvasti@boun.edu.tr (Neda B. Marvasti), acarbu@boun.edu.tr (Burak Acar), jfam@lcc.uma.es (José F. Aldana-Montes).

[1] Equally contributing authors.

[2] This work is partially funded by Grants TIN2014-58304 (MINECO), Regional projects P11-TIC-7529/P12-TIC-1519, TÜBİTAK ARDEB 1001 Programme Grant # 110E264, Boğaziçi University BAP Grant # 5324, and COST Action # 1302 (KEYSTONE), and the Turkish Ministry of Development under the TAM Project number DPT2007K120610.

this has brought, knowledge of past cases have not lost their importance as much more remains to be explained by modern medicine.

To capture and convey such knowledge, an efficient *Clinical Experience Sharing (*CES*)* ecosystem that would allow efficient representation and sharing tools for medical cases, is required (Barzegar Marvasti et al., 2013). A CES ecosystem is essentially composed of three main components, CES applications, semantic representation, and a repository. The most critical component in such an ecosystem is the *semantic representation*, which bridges the semantic gap between the data and the users. The semantic representation should capture the meaning of the data in a machine-readable format so that the computers may process/analyze the data in accordance with user needs. Considering a CES application for search/retrieval of similar cases, the user requests may be subjective. For example, to assist treatment decisions, similar symptoms with different treatments may be of interest. Alternatively, to assist medical education, cases with similar diagnoses but different symptoms may be needed. Considering a structured reporting CES application, on the other hand, the user will need to be guided in filling a medical report, which is dynamically adapted based on the observations/annotations reported. A CES ecosystem infrastructure, as outlined above, will facilitate the development of a multitude of CES applications.

The semantic representation that is at the center of the CES ecosystem, is a case-centric domain model based on medical ontologies and lexicon. It aims to represent an integrated and complete model of medical cases (e.g. patients). The trivial rationale behind this is the integrity of cases, as put by the father of modern medicine Dr. William Osler, "The good physician treats the disease; the great physician treats the patient who has the disease". Despite the recognition of this need, the majority of work on medical ontologies have so far concentrated on sub-disciplines of medicine, such as radiology (e.g. RadLex(Langlotz, 2006)), model anatomy (e.g. FMA (Rosse & Mejino Jr, 2008)), laboratory tests (e.g. LOINC(McDonald et al., 2003)), etc. To the best of our knowledge, there has been no work on developing a case-centric ontological domain model so far.

In this study, we present an ontology for liver cases (LiCO- *Li*ver *C*ase *O*ntology) that integrates several ontologies and/or lexicons developed for sub-disciplines of medicine, from a case-centric point of view. Figure 1 depicts the overview of the system that is grounded by a set of medical lexicons and ontologies. Here, LiCO serves to represent liver patient cases. In a full fledged system the case ontology would obviously have to cover a wide range of medical conditions. The aim of this work is to examine an ontological approach to *case* representation for a CES ecosystem in a non-trivial context, therefore significant effort was expended to model this approach with considerable detail. The focus has been on the integration of existing vocabularies to unify the representation for a given patient, for whom medical data is collected at different times and under different circumstances. The processing potential of cases represented in this manner is demonstrated with queries, reasoning, and rules that serve as building blocks of CES systems. Non-trivial SPARQL [3] queries and reasoning over liver cases are presented to demonstrate how they may be tailored to subjective clinical hypotheses of the users. Human-generated and readable queries are beyond the scope of this work and is left to future work related to designing and developing user experience aspects of the CES environments.

The main contributions of this paper are:

- LiCO, an ontology for liver cases that utilizes well known medical ontologies/lexicons (available for download [4]),
- validation of LiCO in terms of expressing and processing real patients data with a repository of real liver patient cases (sample RDF representations available on website.) ,
- presentation of a proof of concept for CES search/retrieval using semantic reasoner over the real patient repository.

The rest of the paper is organized as follows: Section 2 briefly presents the existing ontologies/lexicons of sub-disciplines of medicine and provides an overview of their applications, Section 3 describes the medical information related to liver patient cases, Section 4 provides a detailed description of LiCO as an approach to represent and process patient cases, Section 5 demonstrates the application of this approach using data collected for real liver patient cases, Section 6 evaluate LiCO in comparison to other ontologies, discusses the insights and limitations of the current version of LiCO and speculates about the future directions for using LiCO and extending case-centric domains. Finally Section 7 concludes the paper.

[3] The *W3C* (World Wide Web Consortium) standard for semantic querying.

[4] LiCO is available at `http://vavlab.ee.boun.edu.tr/pages.php?p=research/CARERA/preDownloadForm.php`

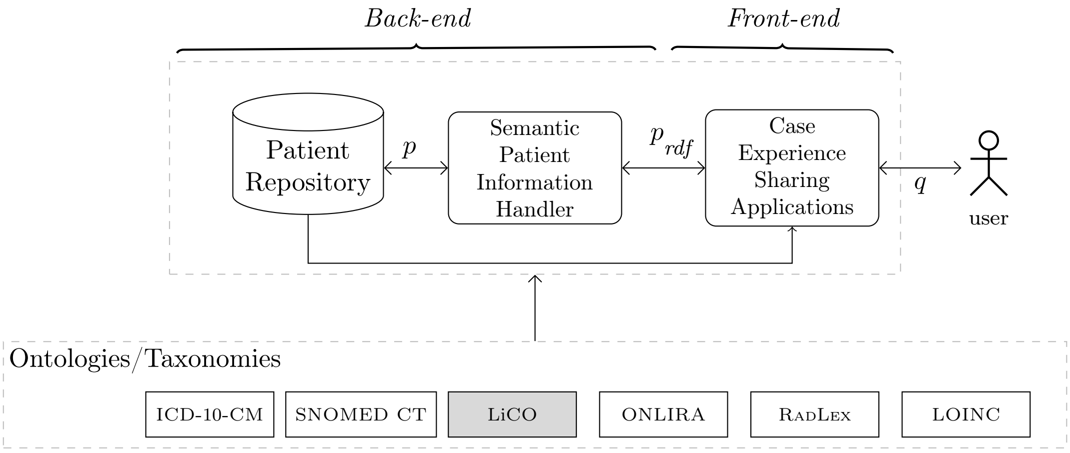

Figure 1. The main components of a Clinical Experience Sharing (CES) ecosystem. LiCO is an ontology for representing patient cases with reference to well known medical vocabularies.

2. Background and Literature Overview

This section describes the main concepts related to this work and other works that make use of ontologies and lexicons.

2.1. *Semantic Web Technologies*

The approach proposed in this work relies on the use of medical vocabulary, ontologies, and related semantic technologies for representing and processing patient cases. This section briefly describes these standards and technologies.

- *Ontology* is a formal representation of a domain of knowledge in terms of concepts and relationships to data (attributes) and other concepts (Gruber, 1993). These definitions provide information about the meaning of data in that domain and constraints that assure logical consistency.

- *RDF* is a graph-based language used to represent information about web resources by defining their properties using RDF statements, which are triples consisting of a subject, predicate, and object (Staab & Studer, 2009). The RDF Schema (RDFS) semantically extends RDF to enable the reference to the classes and properties of resources themselves (Staab & Studer, 2009).

- SPARQL is a query language recommended by W3C for extracting and manipulating information stored in RDF format. It uses graph-matching to match a query pattern against some data, where the values that satisfy a match constitute the result of the query (Pérez, Arenas, & Gutierrez, 2009). It supports networked queries over web resources identified via URIs.

- OWL is a semantic markup language proposed by the W3C ontology working group for publishing and sharing ontologies (Dean & Schreiber, 2004). It extends RDF and RDFS with vocabulary for describing classes and properties, such as relationships between classes (e.g. disjointedness), cardinality (e.g. exactly one), equality, richer property types, characteristics of properties (e.g. symmetry), and enumerated classes (McGuinness & Harmelen, 2004). From a formal point of view, OWL is equivalent to a very expressive description logic (Gruber, 1993).

- OWL-DL is a syntactic variant of the SHOIN (D) description logic (Haase & Stojanovic, 2005) that supports data values, data types and data type properties. It restricts OWL in two distinct ways (Horrocks & Patel-Schneider, 2003): first, it does not support some syntactic constructs like recursive descriptions; second, it requires that all classes, individuals and properties be disjoint. The formal description of semantic model proposed in this work is expressed with the OWL-DL syntax, which is described in Table 1.

2.2. *Clinical Lexicons and Ontologies*

The first step towards a standardized and structured information representation, which is a prerequisite for semantic analysis in medical big data, is to develop common terminology. Several large scale and highly successful projects have been initiated to respond to this need for specific sub-domains of medicine. Some of them are describes here.

- RadLex*(Radiology Lexicon)* is a controlled terminology developed by the Radiological Society of North America (RSNA) for the purpose of providing a unified source of radiology terms for the practice, education, and research(Langlotz, 2006; Kundu et al., 2009). It has been converted into the Web Ontology Language (OWL) format to improve its accessibility and usability(Rubin, 2008).

Table 1
Basic OWL-DL semantic syntax used to formally define the proposed ontology

Descriptions	Abstract Syntax	DL Syntax
Operators	$intersection(C_1, C_2, \cdots, C_n)$	$C_1 \sqcap C_2 \sqcap \cdots C_n$
	$union(C_1, C_2, \cdots, C_n)$	$C_1 \sqcup C_2 \sqcup \cdots C_n$
Restrictions	for at least 1 value V from C	$\exists V.C$
	for all values V from C	$\forall V.C$
	R is Symmetric	$R \equiv R^-$
Class Axioms	A $partial(C_1, C_2, \cdots, C_n)$	$A \sqsubseteq C_1 \sqcap C_2 \sqcap \cdots C_n$
	A $complete(C_1, C_2, \cdots, C_n)$	$A \equiv C_1 \sqcap C_2 \sqcap \cdots C_n$

- LOINC *(The Logical Observation Identifier Names and Codes)* is a coding system for documenting laboratory and clinical observations (Forrey et al., 1996; McDonald et al., 2003; Huff et al., 1998). It consists of more than 25,000 laboratory tests including the Health Level Seven (HL7) [5] observation messages that are used to automatically match the slots of Electronic Health Record (EHR) systems. LOINC is useful in standardizing health terms to support interoperability, however its utility is limited to the lexicon level.

- ICD-10-CM *(International Classification of Diseases, Tenth Revision, Clinical Modification)*, published by The World Health Organization (WHO), includes approximately 15,000 codes for diseases, signs and symptoms, abnormal findings, complaints, social circumstances, and external causes of injury or diseases (Organization et al., 2004; Organization, n.d.). It was endorsed in 1990 and has been globally adopted with several local/national extensions.

- SNOMED CT *(Systematized Nomenclature of Medicine, Clinical Terms)* covers a broad range of clinical terminology (*SNOMED CT website*, n.d.) that aims to be the standard used in EHR (Electronic Health Records) systems. EHRs are annotated with SNOMED CT concepts and relationships to codify patient information.

The abundance of medical data requires tools and techniques to unify concepts that are referred to in a myriad of manners so that information can be shared for the benefit of health care and advances in medical knowledge. The use of controlled vocabulary enables applications to process documents using synonyms, more general concepts (query expansion), or more specific concepts (query restriction) to retrieve more useful information. The explicit definition of concepts enables multiple languages to be represented to include documents across the globe.

The Medical Subject Headings (MeSH) (Lipscomb, 2000) is a is a taxonomy of medical terminology curated by The U.S. National Library of Medicine (NLM) for the purpose of indexing and cataloging over $5,400$ of the world's leading biomedical journals for the MEDLINE / PubMed database. The MeSH terms are used to expand user specified queries in order to retrieve relevant documents. For example, the query "lesion blood vessel" searches for articles that include these terms as well as those that include corresponding MeSH terms (in this case "blood vessel" has numerous corresponding MeSH terms). While such extension are very useful in retrieving articles that use different terminology and sub-concepts, they are limited in returning results of semantic queries. For example, the query "liver lesions close to a vein " retrieves relevant terms, however not always the desired context, such as livers that have lesions and congested central veins. To retrieve such results, queries and content need to be processed according to domains of interest.

The unified medical language system (UMLS) (Bodenreider, 2004) was initiated by the U.S. National Library of Medicine in 1986 as a system for merging and mapping vocabulary from over 130 different sources to promote creation of more effective and inter-operable biomedical information systems and services. It aims to bridge medical terms across systems to improve retrieval performance of machine-readable information by unifying vocabulary from NCBI (national center for biotechnology information, US) taxonomy, Gene ontology, MeSH, Online Mendelian Inheritance in Man (OMIM) and the digital, anatomist symbolic knowledge base. It is used in EHR, classification tools, dictionaries and language translators. Its most important contribution is in linking medical terms, drug names, and billing codes across different computer systems to facilitate search engine retrieval, data mining, public health statistics reporting, and terminology research. It also provides a *semantic network* consisting of a set of broad subject categories, which they call *semantic types* for a consistent categorization of all concepts in its metathesaurus. Furthermore, they provide *semantic relations* between these semantic types to express five types of relations: physically, spatially, temporally, functionally, and conceptually. Similar to other controlled vocabularies, UMLS is limited in supporting domain-specific semantic processing.

The integration and alignment of controlled vocabular-

[5] The HL7 framework for sharing and integrating electronic health information: `http://www.hl7.org`

ies in the medical domain and tools that utilize them have gained increasing attention over the past couple of decades due to the potentials offered by digital exchange and processing of patient information towards providing better health care and advancing knowledge. As such, all the efforts mentioned here as well as others typically support some form of mapping to a variety of other vocabularies.

2.3. *Ontology Based Applications*

In order to address the need for capturing the meaning of medical concepts and how they relate to others, various ontologies in a wide range of domains have been developed. Ontologies play a valuable role in medical applications such as for decision support, diagnosis, annotation and retrieval systems, education, and sharing (Ivanović & Budimac, 2014). Ontologies have been used to determine the type of cancer and treatments using ontologies that were developed for specific types of cancer, such as liver (Alfonse, Aref, & Salem, 2014), breast (Salem & Alfonse, 2009), and lung (Salem & Alfonse, 2007).

Alternatively, some ontologies focus on recommendations to support diagnosis and treatment. For example, semantic queries based on a liver cancer ontology – that models disease types, symptoms, risk factors, treatments and diagnosis – is used to extract the type of liver cancer based on symptoms in order to generate diagnosis and treatment suggestions (Kaur & Khamparia, 2015). A patient-oriented system to support medication prescription based on an ontology that includes the side-effects and contraindications of drugs is proposed by (Chen, Huang, Bau, & Chen, 2012).

Ontology driven search and retrieval systems define domain specific semantic concepts to be used to increase retrieval performance. For example, a content based image retrieval (CBIR) system that automatically annotates medical images with semantic terms from an ontology is proposed by (Seifert et al., 2011), which retrieves medical images based on similarity of semantic as well as appearance of anatomical and pathological characteristics. A similar CBIR system combines image and semantic features (extracted from RadLex relations) to compute a similarity measure to automatically predict the values of image features. Similarly, the Foundation Model of Anatomy Ontology (Allampalli-Nagaraj & Bichindaritz, 2009), which utilizes UMLS, models the human body for search and retrieval tasks for low level image features. ONLIRA (Ontology of Liver for Radiology) describes radiological image observations of liver CT scans (Kokciyan et al., 2014) to support semantic processing tasks, such as identifying similar patients.

In order to make queries more accessible to end users, a search and retrieval system for radiological reports that augments an ontology with natural language processing toolkit is proposed in (Lacson, Andriole, Prevedello, & Khorasani, 2012), where queries are expanded using RadLex and National Cancer Institute Thesaurus (NCIT). The Virtual Imaging Platform (VIP) integrates existing ontologies (Foundational Model of Anatomy (FMA), the Mouse Pathology ontology (MPATH), the Phenotypic Attribute and Trait Ontology (PATO), RadLex and the Chemical Entities of Biological Interest ontology (ChEBI)) to annotate images to facilitate sharing and reuse (Gibaud et al., 2014). The concepts and relations in the ontology are used to create simulation objects, to semantically annotate, and to browse and query models.

With such a variety of ontologies it becomes important to facilitate integration and interoperability of tools. OntoNeuroBase (Temal, Dojat, Kassel, & Gibaud, 2008) proposes an approach to structuring ontologies where the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) (Masolo et al., 2003) serves as the foundation. This approach maintains overall consistency, while domain-specific ontologies provide different levels of abstractions.

Medical education is another important application of ontologies, since they model the domain of knowledge. Quinn et al. (Quinn, Bond, & Nugent, 2017) presents patient information that is generated based on the individual characteristics and health objectives of patients themselves. The personalization is achieved using rule-based reasoning. The main concepts of the ontology are patient, medical conditions, physical activities. The educational content is modeled in the management layer and presented to the patient on a web-based system. An ontology-driven educational system is proposed for nursing towards better health care (Khobreh, Ansari, Dornhöfer, & Fathi, 2013). Yet another system is for teaching anatomy with intelligent visualization (Warren, Agoncillo, Franklin, & Brinkley, 2006).

SUGGESTED REVISED LAST PARAGRAPH:
In line with the concept of *personalized* medicine, LiCO endeavors to provide a methodology for systematic representation of medical (liver) cases. It unifies the heterogeneous data sources (each of which could have been represented with individual ontologies) under a "case" ontology serving as an umbrella. As such, LiCO is not only expandable with the addition of new data sources (i.e. sub-domains) but also with updates in any single sub-domain. LiCO is aimed to serve as an proof-of-concept framework to represent heterogeneous data sources with their inter-relation that can further be adopted as a methodology for other specific patient groups, such as cardiac patients.

OLD TEXT:
With all the benefits that ontologies and standardized vocabularies offer, their successful use calls for their orchestration in delivering meaningful services. For this reason, LiCO focuses on exploring an approach towards representing a person's medical information (patient case) in terms of integrating the multitude of data collected from a variety of sources over a different times in her life. And, in this way to represent patient information as it pertains to an individual as a whole, so that this information can be processed and explored by those who are seeking to learn and to provide the best health care for individuals.

3. Liver Patients Cases

Information about a patient is typically gathered from different specialists, such as internist and radiologists, based on check-ups, tests, and treatments. Such information is often scattered across different systems, representations, and institutions. Therefore, it is difficult to get the whole picture of a person. In order to integrate information related to an individual patient, a representation that integrates all the medical information is required. The use of digital standards in data representation enables the information to be processed with a variety of automated applications.

This section provides the background information useful in understanding the semantic representation of liver cases proposed in this work (Section 4).

Cases consist of information about the patient, such as demographic, medical history, laboratory test results, drug use, and radiological image observations. Knowledge about liver patients was elicited through numerous face to face visits with doctors and radiologists who provided in depth explanations on how CT scans are inspected to document the state of the liver and its anomalies. The remainder of this section describes the data collected for liver patients.

3.1. *The Patient*

Patients are described by providing general information, medical history, and information about the current intervention. The general information consists of demographics and the following:

- Diseases that the patient has previously had or currently suffers from.
- Surgeries that the patient has undergone. They are important in interpreting medical examination results, as well as determining treatment alternatives. For instance, liver patients who have had their gallbladder removed are interpreted differently.
- Regular Drug Use are the drugs that are being taken on a regular basis. Drugs can affect laboratory tests and are important to know for correct interpretation.

3.2. *Study*

A study corresponds to an investigation triggered based on a patient visit due to some complaint. A study includes:

- Physical Examination, which based on the examination of the patient. Typical findings report *blood pressure, pulse, oedema, jaundice,* and *ascites.*
- Laboratory Results are obtained from analyzing the patient's blood and/or urine for, among others, *hepatitis* markers, *bilirubin, iron, kentonbodies, cholesterol, albumin, ammonia,* and *amylase.*
- Diagnoses describes liver related preliminary (pre) and final diagnosis. A pre-diagnosis is recorded before inspecting the symptoms through various tests.
- Current (Non-regular) Drug Use are drugs taken on a non-regular basis.

3.3. *Series/Images*

In medical imaging *series* represent sequence of images resulting from imaging procedures performed to observe features inside the body. For example, CT (computerized tomography) and MRI (Magnetic resonance imaging) imaging modalities are used in diagnosing diseases or injuries as well as planning treatment.

When radiologists inspect medical images they seek to identify and describe anomalies. Part of this task includes identifying the location of the anomaly, for which a radiologist navigates through a series of images to identify a particular slice and a region of interest (ROI) with abnormal characteristics. Hence, the relevant image along with the imaging observations must be documented for future reference.

The characteristics associated with the imaging technology used to observe the lesions, such as contrast pattern and contrast uptake, must be documented for accurate interpretation.

3.4. *Liver*

Liver is a vital organ whose main function is to filter the blood that comes from the digestive tract before it travels to other parts of the body. It is involved in many functions such as making proteins and blood clotting factors, producing triglycerides and cholesterol, glycogen synthesis, and bile production. Hence, liver function is significant in assessing a patient's health condition. The condition of a liver is examined through various laboratory tests, physical examinations, and medical imaging.

Since the liver is a complex organ, numerous conventions for referring to its regions have been developed. In this paper, the most common liver concepts and region reference approaches are considered. A liver consists of three *lobes*: *left*, *right*, and *caudate*. To provide more accurate location descriptions, the liver is divided into eight *segments* that are enumerated in a clockwise manner starting from the caudate lobe as *Segment 1*. The segments $2-4$ are located in the left lobe and $5-8$ are located in the right lobe of the liver. In practice, these segments are referred to using roman numerals, i.e. segment $I-VIII$. The liver is also divided into four *regions*: *anterior*, *posterior*, *lateral*, and *medial*. Finally, the segments are located in following regions: I and IV are in medial, II and III are in lateral, V and VIII are in anterior, and VI and VII are in posterior regions. The liver's vein system is influential in defining such localities:

- The *Right hepatic* vein divides the right lobe into *anterior* and *posterior* regions,
- The *Middle hepatic* vein marks the division between left and right lobes ,
- The *Portal* vein divides the liver into upper and lower segments.

Radiologists describe imaging observations using these references, for instance *"a hypo-dense lesion in segment VII of the liver"* or *"a well-circumscribed hypo-dense area in the right lobe of the liver"*.

The *Parenchyma* of the liver refers to the functional part of the liver as opposed to the connective and supporting tissue. The main characteristics of the parenchyma are its density type and a change in its density, which may reflect areas of abnormality.

The *Vasculature* of the liver refers to the hepatic vascularity. There are two main types of vessels *artery* and *vein*, each of which has a specific name, a diameter, and a type. Changes in vascularity are indicative of potential abnormalities. Veins are also significant in terms of proximity to region of interest. For instance, a vein may have been invaded by a lesion. For portal veins, it is important to document whether cavernous transformation was observed, which refers to the formation of venous channels within or around a previously thrombosed portal vein.

3.5. *Lesion*

A lesion is a pathological damage to an organ or tissue due to injury or disease. It may be manifested as an abscess, a tumor, an ulcer, or a wound. A lesion has a wall, a composition, possible components, debris, a leveling type, and a location. Aside of the location, the values of these attributes are restricted to a particular set. For example, the leveling type may be *fluid-fluid*, *fluid-gas*, *fluid-solid*, or *gas-solid*. Many attributes have boolean values to indicate the presence or absence of an observation (i.e. calcification). For some observations of an anomaly, further information is necessary. For instance, if leveling is observed, the type of the leveling must be described.

4. Semantic Representation

One of the main goals of this work is to develop a semantic model for liver patient case representation. To this end, an ontology, Liver Case Ontology (LiCO), to describe the main characteristics of liver patient case is developed.

The ontology is built with OWL (W3C Web Ontology Language) following the standard Ontology 101 development process of seven steps (Noy & McGuinness, 2001):

(i) **Determine the domain and scope of the ontology**. The scope of our ontology is the liver patients case as described in Section 3. The ontology will describe the liver patient information including demographic, medical history, laboratory test results, drug use, and radiological findings.

(ii) **Consider reusing existing ontologies**. A revised version of ONLIRA (Kokciyan et al., 2014) is used to describe the radiological information. Concepts in ONLIRA are mapped to RadLex. ICD-10-CM categories are used for patient diseases. Laboratory results are represented with LOINC codes. Finally, SNOMED CT codes are mapped to ontology classes whenever possible (i.e. physical examination concepts).

(iii) **Enumerate important terms in the ontology**. Important concepts were identified during the elicitation phase of liver patients case representation. This included field experts and literature search. The key terms include patient, disease, surgery, drug, study, physical examination, laboratory results, diagnoses, liver, lobe, segment, lesion, and vascularity.

(iv) **Define classes and the class hierarchy**. The classes correspond the the terms identified for liver patient cases. Figure 2 shows the main classes and their relationships to other classes. For instance, the `onlira`:HepaticVascularity class has several subclasses, including `onlira`:HepaticArtery, `onlira`:HepaticPortalVein, `onlira`:HepaticVein and `onlira`:VenacavaInferior. Likewise, the `onlira`:Lobe class has the `onlira`:CaudateLobe, `onlira`:LeftLobe and `onlira`:RightLobe subclasses.

(v) **Define the properties of classes and slots**. Classes have relationships to other classes and their attributes. Like classes, the relations were identified based on the liver patient case representation. Examples of object properties include: `lico:hasStudy` with domain `lico`:Patient and `lico`:Study, `lico:hasGeneticDisease` with domain `lico`:Patient and `lico`:Disease, `lico:hasLiver` with domain `lico`:Image and range `onlira`:Liver, and `onlira:hasLobe` with domain `onlira`:Liver and range `onlira`:Lobe. Examples of data type properties include: `lico:hasAge` with domain `lico`:Patient and range datatype `xsd`:nonNegativeInteger and `lico:hasAST` with domain `lico`:LaboratoryResults and range `xsd`:double. Furthermore, `lico:hasAST` is declared equivalent to `loinc:1920-8`.

(vi) **Define the facets of the slots**. This step includes the definition of cardinality constraints and value restrictions. Value restrictions are used in our ontology to restrict the liver segments in each liver lobe or liver region. For instance, caudate lobe is segmented by only segment I, and segment II is located only in Lateral region. Cardinality restrictions are used to specify the number of segments in each Lobe. For instance, caudate lobe, left lobe and right lobe are segmented by one, three and four segments, respectively.

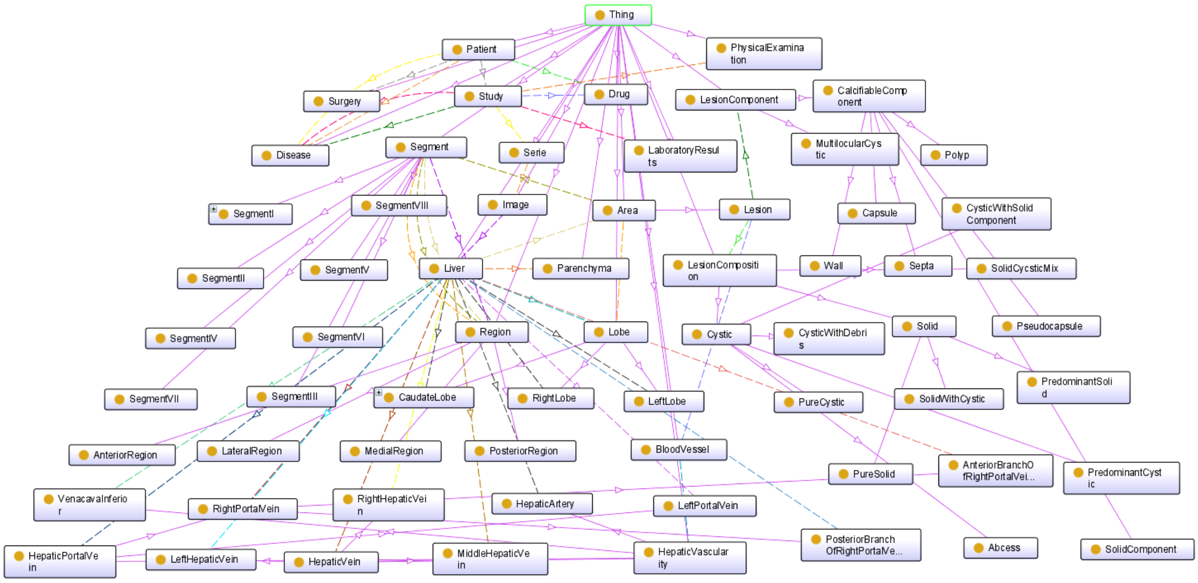

Figure 2. General overview of the LiCO ontology. Solid arrows mark subclass relation while dotted arrows indicate specific properties.

(vii) **Create instances**. Instances (individuals in OWL) correspond to the specific data obtained from a specific patient. Individuals are obtained from medical doctors through the `CaReRa` project. Doctors manually insert the data using an web based forms generated according to classes and properties of the ontology. Then, the collected data are exported to RDF.

In the remainder of this paper, several figures are used to describe aspects of the ontology and our approach. In these figures classes are represented with rectangles, data with parallelograms, individuals with ellipses, and properties with arrows.

The prefixes *onlira* and *lico* are used to refer ONLIRA and LiCO ontologies. As mentioned earlier, ONLIRA is an existing ontology from previous work, whereas LiCO is defined in this paper. Due to space limitations, in some figures the prefixes `o` and `l` are used to refer to `onlira` and `lico`, respectively.

4.1. *Ontology Model*

The liver patient case is modeled with the Liver Case Ontology (LiCO) multiple inter-related concepts, such as demographics, medical history, laboratory test results, drug usage, and radiological findings. The proposed ontology, resulting from the development process described earlier, has a total of 93 classes (groups of individuals sharing the same attributes), 36 object properties (binary relationships between individuals), 119 data properties (individual attributes), and 474 logical axioms [6].

In this paper, a subset of key classes and properties are defined. As described in Section 3 there are several aspects related to describing the state of a patient. In this case, the focus is a liver patient, thus the case must include the description of the patient as well as all the medical investigations related to the study and series related to images. Figure 3 shows the high level relationships between the classes beginning from the patient to the images where the liver observations are made.

The main classes are: *Patient*, *Study*, *Serie*, *Image*, *Liver*, *Area* and *Lesion*.

A set of properties related to classes are required in order to capture the conceptualization.

Patient is the main focus of LiCO. As listed in Table 2, a patient has an age (`lico:hasAge`), a gender, a name, and an ID. Important features of patients are their genetic (`lico:hasGeneticDisease`) and nonliver diseases provided with ICD-10-CM codes, regular drugs taken, and previous surgeries.

Study. Each Patient has at least one study. A study has a date, an ID, and possible complaints. Furthermore, it has a pre-diagnosis, non-regular drugs that are consumed, a final diagnosis (also ICD-10-CM), laboratory results and physical examinations. Table 3 shows the object and

[6] LiCO can be downloaded from http://www.vavlab.ee.boun.edu.tr/pages.php?p=research/CARERA/preDownloadForm.php

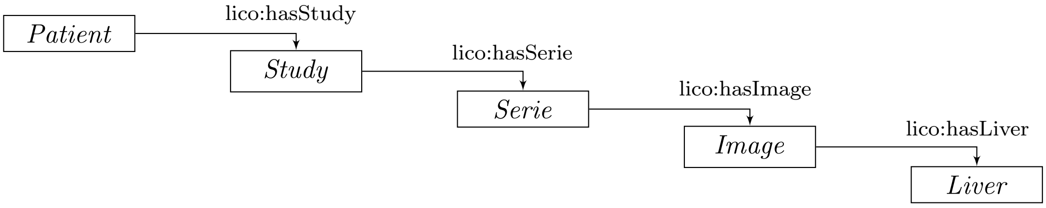

Figure 3. The relationships among the patient, study, serie, image, and liver that is usually triggered upon a patient visiting a doctor based on some complaint.

Table 2
lico:Patient object and datatype properties

Object Properties	Description Logic
lico:hasGeneticDiseases	$\top \sqsubseteq \forall$ lico:hasGeneticDisease.lico:Disease
lico:hasNonLiverDisease	$\top \sqsubseteq \forall$ lico:hasNonLiverDisease.lico:Disease
lico:hasRegularDrug	$\top \sqsubseteq \forall$ lico:hasRegularDrug.lico:Drug
lico:hasSurgery	$\top \sqsubseteq \forall$ lico:hasSurgery.lico:Surgery
lico:hasStudy	$\top \sqsubseteq \forall$ lico:hasStudy.lico:Study
Datatype Properties	**Description Logic**
lico:hasAge	$\top \sqsubseteq \forall$ lico:hasAge.Datatype nonNegativeInteger
lico:hasGender	$\top \sqsubseteq \forall$ hasGender:.{"F"^^string, "M"^^string}
lico:hasName	$\top \sqsubseteq \forall$ lico:hasName.Datatype string
lico:hasPatientID	$\top \sqsubseteq \forall$ lico:hasPatientID.Datatype string

Table 3
lico:Study object and datatype properties

Object Properties	Description Logic
lico:hasLiverPreDiagnosis	$\top \sqsubseteq \forall$ lico:hasLiverPrediagnosis.lico:Disease
lico:hasNonRegularDrugs	$\top \sqsubseteq \forall$ lico:hasNonRegularDrugs.lico:Drug
lico:hasLaboratoryResults	$\top \sqsubseteq \forall$ lico:hasLaboratoryResults.lico:LaboratoryResults
lico:hasPhysicalExamination	$\top \sqsubseteq \forall$ lico:hasPhysicalExamination.lico:PhysicalExamination
lico:hasFinalDiagnosis	$\top \sqsubseteq \forall$ lico:hasFinalDiagnosis.lico:Disease
lico:hasSerie	$\top \sqsubseteq \forall$ lico:hasSerie.lico:Serie
Datatype Properties	**Description Logic**
lico:hasComplaints	$\top \sqsubseteq \forall$ lico:hasComplaints.Datatype string
lico:hasDate	$\top \sqsubseteq \forall$ lico:hasDate.Datatype date
lico:hasStudyID	$\top \sqsubseteq \forall$ lico:hasStudyID.Datatype string

datatype properties for the lico:Study class. Datataype properties of the lico:PhysicalExamination class are presented in Table 4. These datatype properties are mapped to SNOMED CT codes. Table 5 shows the most representative datatype properties related to lico:LaboratoryResults class. These datatype properties are based on LOINC codes.

Finally, a study has series, each of which consists of a set of images. The image includes, among other things, the liver.

Liver. This is key concept in LiCO since the focus is on liver patients. This concepts allows the description of various liver properties. These properties are imported from the most recent version of ONLIRA [7]. Among other prop-

[7] https://bioportal.bioontology.org/ontologies/ONLIRA

Table 4
lico:PhysicalExamination datatype properties

Datatype Properties	Description Logic
lico:hasAscites	≡ snomed:389026000
	$\top \sqsubseteq \forall$ lico:hasAscites.{"no"^^string, "not specified"^^string, "yes"^^string}
lico:hasBloodPressure	≡ snomed:75367002
	$\top \sqsubseteq \forall$lico:hasBloodPressure.Datatype string
lico:hasJaundice	≡ snomed:18165001
	$\top \sqsubseteq \forall$ lico:hasJaundice. {"no"^^string, "not specified"^^string, "yes"^^string}
lico:hasOedema	≡ snomed:423666004
	$\top \sqsubseteq \forall$ lico:hasOedema. {"no"^^string, "not specified"^^string, "yes"^^string}
lico:hasPulse	≡ snomed:8499008
	$\top \sqsubseteq \forall$ lico:hasPulse.DatatypeRestriction
	(FacetmaxInclusive"180"^^nonNegativeInteger, FacetminInclusive"30"^^nonNegativeInteger)

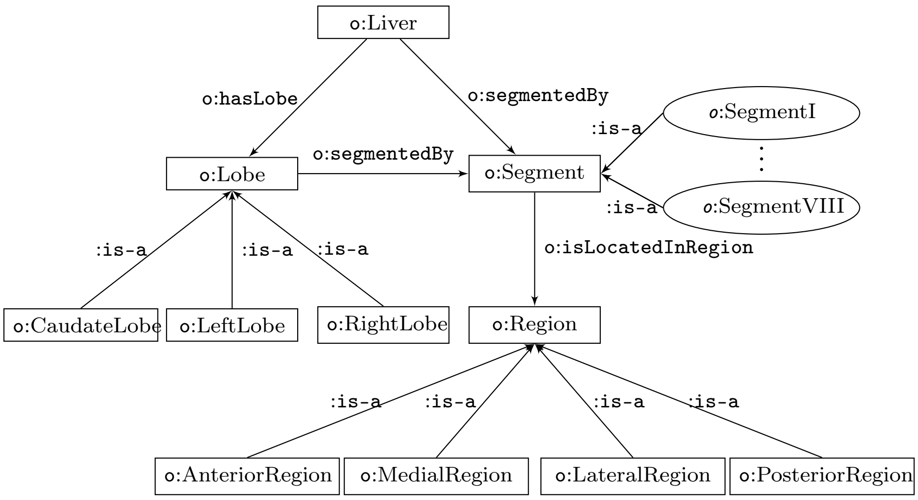

Figure 4. Various regional definitions used to refer to parts of a liver.

erties, the liver has a size, a density, and a contour. These datatype properties are given in Table 6. For instance, the *craniocaudal* dimension of the liver is defined with **onlira:hasCraniocaudalDimension** property whose range is an integer corresponding to the size in millimeters. Moreover, the change in liver size over a period of time is a medically significant characteristic and is defined with the **onlira:hasSizeChange** property, which can be associated with a value like *increased* (Table 6).

Each liver has a onlira:Parenchyma class identified with the onlira:hasParenchyma property. The lobes of the liver are represented with: onlira:CaudateLobe, onlira:RightLobe, and onlira:LeftLobe classes. The onlira:hasLobe property relates a liver to its Lobes. Lobes and livers are segmented by segments and segments are located in regions (see Figure 4). Finally, some object properties describe the vascularity of a liver (Table 7). All of the datatype properties of onlira:Liver and the onlira:hasParenchyma property are defined as *functional properties* to ensure that only one property value is asserted for a given liver.

ONLIRA models the segments that segment each lobe by means of the *owl:allValuesFrom* restriction. For example, to specify that onlira:LeftLobe may only be segmented by onlira:SegmentII, onlira:SegmentIII and onlira:SegmentIV ONLIRA asserts $LeftLobe \sqsubseteq \forall isSegmentedBy.(SegmentII$ $mentIII \sqcup SegmentIV)$. However, this is not sufficient for a reasoner to detect errors, since these segments

Table 5
lico:LaboratoryResults datatype properties

Datatype Properties	Description Logic
lico:hasHepatitisMarkers	$\top \sqsubseteq \forall$ lico:hasHepatitisMarkers.Datatype real
lico:hasIndirectBilirubin	$\equiv$ loinc:1971-1
	$\top \sqsubseteq \forall$ lico:hasIndirectBilirubin.DatatypeRestriction
	(FacetmaxInclusive"26.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasDirectBilirubin	$\equiv$ loinc:1968-7
	$\top \sqsubseteq \forall$ lico:hasDirectBilirubin.DatatypeRestriction
	(FacetmaxInclusive"24.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasTotalBilirubin	$\equiv$ loinc:1975-2
	$\top \sqsubseteq \forall$ lico:hasTotalBilirubin.DatatypeRestriction
	(FacetmaxInclusive"50.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasGGT	$\equiv$ loinc:2324-2
	$\top \sqsubseteq \forall$ lico:hasGGT.DatatypeRestriction
	(FacetmaxInclusive"800.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasLDLCholesterol	$\equiv$ loinc:18262-6
	$\top \sqsubseteq \forall$ lico:hasLDLCholesterol.DatatypeRestriction
	(FacetminInclusive"3.0"^^double, FacetmaxInclusive"1100.0"^^double)
lico:hasHDLCholesterol	$\equiv$ loinc:2085-9
	$\top \sqsubseteq \forall$ lico:hasHDLCholesterol.DatatypeRestriction
	(FacetmaxInclusive"150.0"^^double, FacetminInclusive"3.0"^^double)
lico:hasTotalCholesterol	$\equiv$ 2093-3
	$\top \sqsubseteq \forall$ lico:hasTotalCholesterol.DatatypeRestriction
	(FacetminInclusive"50.0"^^double, FacetmaxInclusive"1800.0"^^double)
lico:hasTriglyceride	$\equiv$ loinc:2571-8
	$\top \sqsubseteq \forall$ lico:hasTriglyceride.DatatypeRestriction
	(FacetmaxInclusive"2000.0"^^double, FacetminInclusive"15.0"^^double)
lico:hasALT	$\equiv$ loinc:1742-6
	$\top \sqsubseteq \forall$ lico:hasALT.DatatypeRestriction
	(FacetmaxInclusive"20000.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasAST	$\equiv$ loinc:1920-8
	$\top \sqsubseteq \forall$ lico:hasAST.DatatypeRestriction
	(FacetmaxInclusive"20000.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasALP	$\equiv$ loinc:6768-6
	$\top \sqsubseteq \forall$ lico:hasALP.DatatypeRestriction
	(FacetmaxInclusive"20000.0"^^double, FacetminInclusive"0.0"^^double)
lico:hasAmylase	$\equiv$ loinc:1798-8
	$\top \sqsubseteq \forall$ lico:hasAmylase.DatatypeRestriction
	(FacetmaxInclusive"195000.0"^^double, FacetminInclusive"0.0"^^double)

Table 6
`onlira`:Liver datatype properties

Datatype Properties	Description Logic
`onlira:hasCraniocaudalDimension`	$\top \sqsubseteq \forall$ `onlira`:hasCraniocaudalDimension.Datatype sizeMM
`onlira:hasDensity`	$\top \sqsubseteq \forall$ `onlira:hasDensity`.{"heterogeneous"^^string, "homogeneous"^^string, "other"^^string}
`onlira:hasLiverContour`	$\top \sqsubseteq \forall$ `onlira:hasLiverContour`.{"irregular"^^string, "lobulated"^^string, "nodular"^^string, "other"^^string, "regular"^^string}
`onlira:hasLiverDensityChange`	$\top \sqsubseteq \forall$ `onlira:hasLiverDensityChange`.{"decreased"^^string, "increased"^^string, "normal"^^string, "other"^^string}
`onlira:hasLiverPlacement`	$\top \sqsubseteq \forall$ `onlira:hasLiverPlacement`.{"downward displacement"^^string, "leftward displacement"^^string, "normal placement"^^string, "other"^^string, "upward displacement"^^string}
`onlira:hasSizeChange`	$\top \sqsubseteq \forall$ `onlira:hasSizeChange`.{"decreased"^^string, "increased"^^string, "normal"^^string, "other"^^string}

Table 7
`onlira`:Liver: object properties

Object Properties	Description Logic
`onlira:hasAnteriorBranchOfRightPortalVein`	$\top \sqsubseteq \forall$ `onlira:hasAnteriorBranchOfRightPortalVein`. `onlira`:AnteriorBranchOfRightPortalVein
`onlira:hasHepaticArtery`	$\top \sqsubseteq \forall$ `onlira:hasHepaticArtery.onlira`:HepaticArtery
`onlira:hasHepaticPortalVein`	$\top \sqsubseteq \forall$ `onlira:hasHepaticPortalVein.onlira`:HepaticPortalVein
`onlira:hasHepaticVein`	$\top \sqsubseteq \forall$ `onlira:hasHepaticVein.onlira`:HepaticVein
`onlira:hasLeftHepaticVein`	$\top \sqsubseteq \forall$ `onlira:hasLeftHepaticVein.onlira`:LeftHepaticVein
`onlira:hasLeftPortalVein`	$\top \sqsubseteq \forall$ `onlira:hasLeftPortalVein.onlira`:LeftPortalVein
`onlira:hasParenchyma`	$\top \sqsubseteq \forall$ `onlira:hasParenchyma.onlira`:Parenchyma
`onlira:hasPosteriorBranchOfRightPortalVein`	$\top \sqsubseteq \forall$ `onlira:hasPosteriorBranchOfRightPortalVein`. `onlira`:PosteriorBranchOfRightPortalVein
`onlira:hasRightHepaticVein`	$\top \sqsubseteq \forall$ `onlira:hasRightHepaticVein.onlira`:RightHepaticVein
`onlira:hasRightPortalVein`	$\top \sqsubseteq \forall$ `onlira:hasRightPortalVein.onlira`:RightPortalVein
`onlira:hasVenacavaInferior`	$\top \sqsubseteq \forall$ `onlira:hasVenacavaInferior.onlira`:VenacavaInferior
`onlira:isSegmentedBy`	$\top \sqsubseteq \forall$ `onlira:isSegmentedBy.onlira`:Segment
`onlira:hasArea`	$\top \sqsubseteq \forall$ `onlira:hasArea.onlira`:Area

are not disjoint. Therefore, LiCO asserts the disjointness of these segments with one another. Thus, if an instance of `onlira`:LeftLobe is segmented by a segment other than `onlira`:SegmentII, `onlira`:SegmentIII or `onlira`:SegmentIV the reasoner returns an error. Furthermore, to specify that all these segments do segment the `onlira`:LeftLobe, ONLIRA asserts *LeftLobe* $\sqsubseteq \forall$ *isSegmentedBy.SegmentII*, *LeftLobe* $\sqsubseteq \forall$ *isSegmentedBy.SegmentIII*, and *LeftLobe* $\sqsubseteq \forall$ *isSegmentedBy.SegmentIV*.

To avoid inconsistencies these axioms are replaced with the assertion in LiCO: *LeftLobe* $\sqsubseteq = 3$ *isSegmentedBy.Segment*, where the cardinality of the *isSegmentedBy* property when the domain is *LeftLobe* is specified as *exactly 3*. Similar specifications are defined for the other *lobes*.

Given the knowledge of where a lesion is located (i.e. *Segment II*), it is useful to know where else it is located in terms of other regional references (i.e. left lobe and lateral region). Each segment is located in a specific region (Section 3.4). However, ONLIRA does not model this relationship, which is modeled in LiCO with the `owl:allValuesFrom` and `owl:cardinality` restrictions as well as asserting the *disjointness* of the four regions. For example, to specify that *Segment II* is located in the lateral region, LiCO asserts that:

$SegmentII \sqsubseteq \forall\ isLocatedInRegion.LateralRegion$

$SegmentII \sqsubseteq =1\ isLocatedInRegion.Region$

Furthermore, LiCO introduces SWRL[8] (Semantic Web Rule Language) rules to infer new relationships for the*Segment*, *Lobe* and *Region* classes. SWRL uses the familiar logical expression "Antecedent ⇒ Consequent" to represent semantic rules. Based on the knowledge about lobes in the ontology, the following rule (*isLocatedInLobe*) infers the lobe in which an area is observed, given the segment in which it is observed:

```
Area(?area)
^ Lobe(?lobe)
^ isLocatedInSegment(?area, ?seg)
^ isSegmentedBy(?lobe, ?seg)
-> isLocatedInLobe(?area, ?lobe)
```

Similarly, the region where an area is located can be inferred with the rule *isLocatedInRegion*:

```
Area(?area)
^ Region(?reg)
^ Segment(?seg)
^ isLocatedInRegion(?seg, ?reg)
^ isLocatedInSegment(?area, ?seg)
-> isLocatedInRegion(?area, ?reg)
```

Area and Lesion. An abnormal area observed in an image of the liver can be described with various properties as shown in Table 8. The `onlira:hasAreaShape` describes the shape of an area, which may take the values: *band*, *fusiform*, *linear*, *nodular*, *ovoid*, *serpiginious*, and *other*. The `onlira:hasAreaDensity` describes the density of the area that may take the values *hyperdense*, *hypodense*, or *isodense*. The `onlira:isCalcified` property indicates whether an area is calcified and if it is, then `onlira:hasCalcification` property specifies the type of calcification which can be *coarse*, *focal*, *millimetric*, *punctate*, or *scattered*. The location of an area is specified with `onlira:isLocatedInSegment` and `onlira:isLocatedInLobe` (Table 8). A particularly important type of an area is a lesion. The `onlira`:Lesion class is as a subclass of `onlira`:Area.

Lesions have additional properties to those defined for Area. For example, lesions may have components and a composition. The object properties for the *Lesion* are summarized in Table 9.

The proximity of a lesion with respect to a blood vessel is significant and is modeled differently in LiCO than in ONLIRA. In order to handle the description of the proximity to a specific vessel the `lico`:BloodVessel class is introduced (Figure 5). It has two properties: an object property to specify the blood vessel that the lesion is close to (`lico:hasBloodVesselName`) whose range is the `onlira:HepaticVascularity` and a datatype property (`onlira:hasLesionBloodVesselProximity`) to specify value of the proximity (*adjacent, adjunct to contact, bended, circumscribed, invaded* and *other*). The datatype properties for `onlira`:Lesion (Table 9) include `onlira:hasDebrisLocation`, `onlira:isDebrisObserved` and `onlira:isLevelingObserved`. All datatype properties of *Area* and *Lesion* are functional properties .

The LiCO object and datatype properties formalize not only the radiologists description of imaging observations, but also the patient's medical information. The observation *"Hypo-dense lesion in Segment VII of the Liver"* is formalized with LiCO as:

```
onlira:isSegmentedBy(liver32,liver32SegmentVII).
onlira:isLocatedInSegment(lesion5,liver32SegmentVII).
onlira:hasAreaDensity(lesion5,"hyperdense").
```

where *liver32* and *lesion5* correspond to specific liver and lesion instances (corresponding to a patient).

Furthermore, the genetic diseases, age, iron, and pulse of the patient with the lesion (*lesion5*) are formally expressed as:

```
lico:hasStudy(patient56,patient56st1).
lico:hasSerie(patient56st1,patient56st1se1).
lico:hasImage(patient56st1se1,patient56st1se1img1).
lico:hasLiver(patient56st1se1img1,liver32).
lico:hasGeneticDiseases(patient56,C32).
lico:hasAge(patient56,60).
lico:hasLaboratoryResults(patient56,patient56lr1)
lico:hasGGT(patient56lr1,176.0).
lico:hasPhysicalExamination(patient56,patient56pex1).
lico:hasPulse(patient56pex1,100).
```

5. Validation

To validate the proposed approach, 46 real liver patient cases were used to express using LiCO and to perform queries to examine the utility of the proposed case representation in revealing useful information. The liver patient data patient data was acquired during the `CaReRa` project in a heterogeneous manner, while some critical data was collected based on the ONLIRA ontology, some other data was collected in more standard methods. All data was collected based on accepted medical standards.

The initial task was to create the patient cases from this data. These cases are exported to RDF with functions that map the patient and imaging observation data to the desired triple format. More specifically, they are serialized in RDF format with code developed using the Apache Jena RDF API (McBride, 2001), which supports handling RDF graphs. These functions instantiate the classes that describe each patient and relate them to their object and datatype properties in accordance with LiCO. The RDF

[8] https://www.w3.org/Submission/SWRL/

Table 8
onlira:Area object and datatype properties

Object Properties	Description Logic
onlira:isLocatedInSegment	$\top \sqsubseteq \forall$ onlira:isLocatedInSegment.onlira:Segment
onlira:isLocatedInLobe	$\top \sqsubseteq \forall$ onlira:isLocatedInLobe.onlira:Lobe
onlira:isLocatedInRegion	$\top \sqsubseteq \forall$ onlira:isLocatedInRegion.onlira:Region
Datatype Properties	**Description Logic**
onlira:hasAreaContrastPattern	$\top \sqsubseteq \forall$ onlira:hasAreaContrastPattern. {"central"^^string, "early uptake then wash out"^^string, "fixing contrast in late phase"^^string, "heterogeneous"^^string, "homogeneous"^^string, "spokes wheel"^^string, "undecided"^^string, "other"^^string}
onlira:hasAreaDensity	$\top \sqsubseteq \forall$ onlira:hasAreaDensity.{"hyperdense"^^string,"hypodense"^^string, "isodense"^^string, "other"^^string}
onlira:hasAreaLengthFirst	$\top \sqsubseteq \forall$ onlira:hasAreaLengthFirst.Datatype sizeMM
onlira:hasAreaLengthSecond	$\top \sqsubseteq \forall$ onlira:hasAreaLengthSecond.Datatype sizeMM
onlira:hasAreaMarginType	$\top \sqsubseteq \forall$ onlira:hasAreaMarginType.{"geographical"^^string,"ill defined"^^string, "irregular"^^string, "lobular"^^string, "other"^^string, "serpiginious"^^string, "spiculative"^^string, "well defined"^^string}
onlira:hasAreaShape	$\top \sqsubseteq \forall$ onlira:hasAreaShape.{"band"^^string, "fusiform"^^string, "irregular"^^string, "linear"^^string, "nodular"^^string, "other"^^string, "ovoid"^^string, "round"^^string, "serpiginious"^^string}
onlira:hasCalcification	$\top \sqsubseteq \forall$ onlira:hasCalcification.{"coarse"^^string, "focal"^^string, "millimetric-fine"^^string, "other"^^string, "punctate"^^string, "scattered"^^string}
onlira:hasDensity	$\top \sqsubseteq \forall$ onlira:hasDensity.{"heterogeneous"^^string, "homogeneous", "other"^^string}
onlira:isCalcified	$\top \sqsubseteq \forall$ onlira:isCalcified Datatype boolean
onlira:isCentralLocalized	$\top \sqsubseteq \forall$ onlira:isCentralLocalized Datatype boolean
onlira:isContrasted	$\top \sqsubseteq \forall$ onlira:isContrasted Datatype boolean
onlira:isGallbladderAdjacent	$\top \sqsubseteq \forall$ onlira:isGallbladderAdjacent Datatype boolean
onlira:isPeripheralLocalized	$\top \sqsubseteq \forall$ onlira:isPeripheralLocalized Datatype boolean
onlira:isSubcapsularLocalized	$\top \sqsubseteq \forall$ onlira:isSubcapsularLocalized Datatype boolean

representations of liver cases are stored using the *Stardog* [9] RDF repository and reasoner, which provides a SPARQL endpoint for executing queries. Figure 6 illustrates the acquisition of data, its transformation to RDF, and the creation of an endpoint to enable the querying of cases. We found that LiCO is able to express the cases.

5.1. *The Liver Patient Data*

Patient data was collected from existing patient records at Istanbul University Medical School, Department of Radiodiagnostics, with the approval of Istanbul University Medical School, Ethics Assessment Committee [10] . The CT data was anonymized at the hospital site. The printed patient records were only accessed by the authorized medical personnel on site and the diagnostic information (exclud-

[9] http://www.stardog.com/

[10] Ethics committee approval #: 09/06/2010 - 01

Table 9
onlira:Lesion object and datatype properties

Object Properties	Description Logic
onlira:hasLesionComponent	⊤ ⊑ ∀ onlira:hasLesionComponent.onlira:LesionComponent
onlira:hasLesionComposition	⊤ ⊑ ∀ onlira:hasLesionComposition.onlira:LesionComposition
lico:isCloseToBloodVessel	⊤ ⊑ ∀ lico:isCloseToBloodVessel.lico:BloodVessel
Datatype Properties	**Description Logic**
onlira:hasDebrisLocation	⊤ ⊑ ∀ onlira:hasDebrisLocation.{"floating inside"^^string, "located on dependent position"^^string, "other"^^string }
onlira:hasLesionContrastUptake	⊤ ⊑ ∀ onlira:hasLesionContrastUptake.{"dense"^^string, "heterogeneous"^^string, "homogeneous"^^string, "minimal"^^string, "moderate"^^string, "other"^^string}
onlira:hasLesionQuantity	⊤ ⊑ ∀ onlira:hasLesionQuantity.{"1"^^integer, "2"^^integer, "3"^^integer, "4"^^integer, "5"^^integer, "multiple"^^string}
onlira:isDebrisObserved	⊤ ⊑ ∀ onlira:isDebrisObserved Datatype boolean
onlira:isLevelingObserved	⊤ ⊑ ∀ onlira:isLevelingObserved Datatype boolean

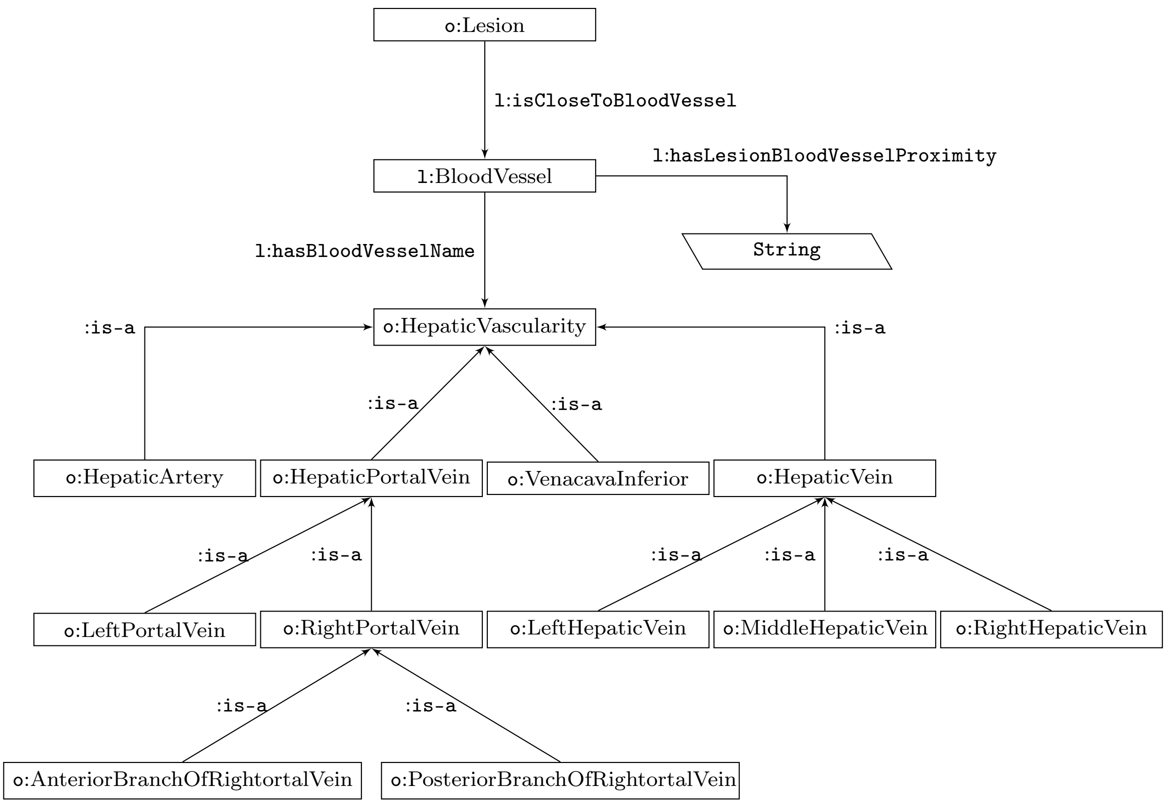

Figure 5. The modeling of proximity to a particular vessel in LiCO.

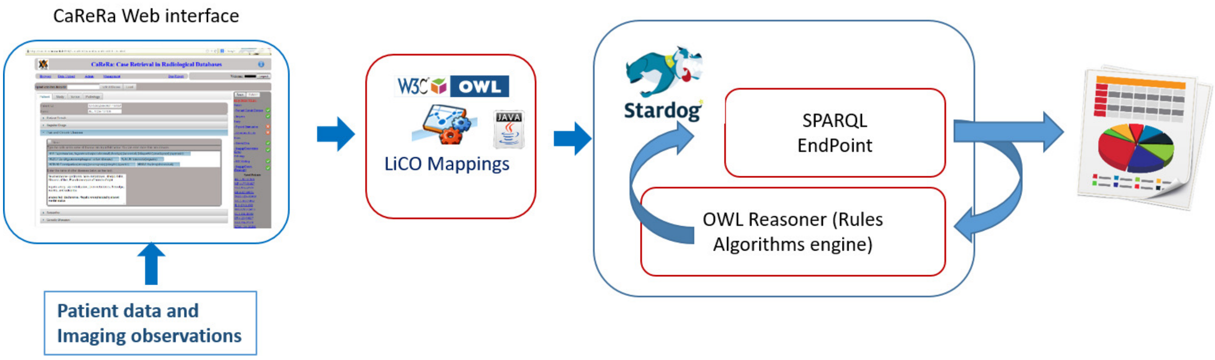

Figure 6. Overview of the semantic approach for querying liver patient cases.

ing all demographic data except gender and age) was input manually to a secure database via an in-house web tool. The data consists of healthy patients as well as those who have liver diseases.

Patient data consists of demographic, laboratory results, physical examination, and imaging observation data and recorded according to well known standards.

The imaging observations include general information about the CT scan performed such as *Slice Thickness* (i.e. $1.0mm$) and *KVP/Current* (i.e. $120/300v/ma$). The imaging observations describe the characteristics of the liver (i.e. *Density Type*: `heterogeneous` and *Right Portal Vein Lumen Diameter*: `decreased`). The pathology details are specified for an image selected from the series of a scan. Given an image, a region of interest is specified and observations related to it are documented. Figure 7 shows an screen shot of such an entry where *Cluster Size*: 1, *Shape*: `round`, *Lesion Compoition*: `PureSolid`, *Segment*: *SegmentIII, SegmentIV*, *height*: $64mm$ and *width*:60.

`CaReRa-Web` [11] is an ontology driven data collection and retrieval tool for liver patients. It exports patient information expressed with LiCO as RDF triples, which are used to perform semantic queries and reasoning (Section 5.2).

Figure 8 shows the part of a patient's data (Patient id: $Capa - 00029$) that is related to the imaging observations of a lesion in the liver. The stacked ellipses indicate that a patient may have multiple studies (corresponding to each investigation) and each study may have several series (for each scan).

RDF individuals for modeling the liver anatomy are created with Jena. Therefore, RDF individuals for each patient's liver, hepatic vascularity, segments, lobes, regions, etc. are generated. These individuals are related with the patient's liver by means of the LiCO object properties described in Table 7, such as `onlira:hasPortalVein`, `onlira:isSegmentedBy`, etc. For example, for patient $Capa - 00029$, the $liver579$ is created. RDF triples for relating the $liver579$ with its anatomy are among others:
`o:liver579 o:isSegmentedBy o:SegmentII.`
`o:liver579 o:hasPortalVein o:liver579PortalVein`

To allow reasoning with these individual and to avoid inconsistencies, we need to generate different individuals for different parts of the liver's anatomy in different livers.

Then RDF individuals for modeling the relationship between Patient and Liver shown in Figure 3 are created, that is, RDF triples starting from $Capa - 00029$ to $liver579$ are asserted for the example patient case illustrated in Figure 8. Also RDF triples are generated to connect the patient's studies with the corresponding Laboratory results, Physical examinations and Final diagnosis. Finally, each patient's liver is linked to its lesions and the corresponding imaging observations are assigned to them by means of the corresponding LiCO datatype properties presented in Table 6, Table 8 and Table 9.

Examples of these RDF triples are (see Figure 8):
`l:Study521 l:hasPhysicalExamination l:Study521PhysicalExam.`
`l:Study521PhysicalExam l:hasPulse 78.`
`o:liver579 o:hasArea o:lesion117.`
`o:lesion117 o:hasAreaDensity "hypodense".`

5.2. *Semantic Queries*

In this section we present several SPARQL queries to demonstrate the usefulness of our approach. These queries have been evaluated using the *Stardog* SPARQL endpoint. The *Stardog* RDF repository stores the liver cases described in Section 5.1.

Query 1 (Figure 10) finds patients who have lesions close to some vessel, using the lesion object properties `lico:isCloseToBloodVessel` and `lico:hasBloodVesselName`. The range of the property `lico:hasBloodVesselName` is the class `onlira`:HepaticVascularity which has several subclasses as explained in Section 3 . Therefore, a query regarding "*all lesions close to any blood vessel*" returns the specific vessels that the lesions are close to without

[11] A demo of the `CaReRa-Web` can be accessed at: `https://193.140.195.124:5904/CareraWeb2/`

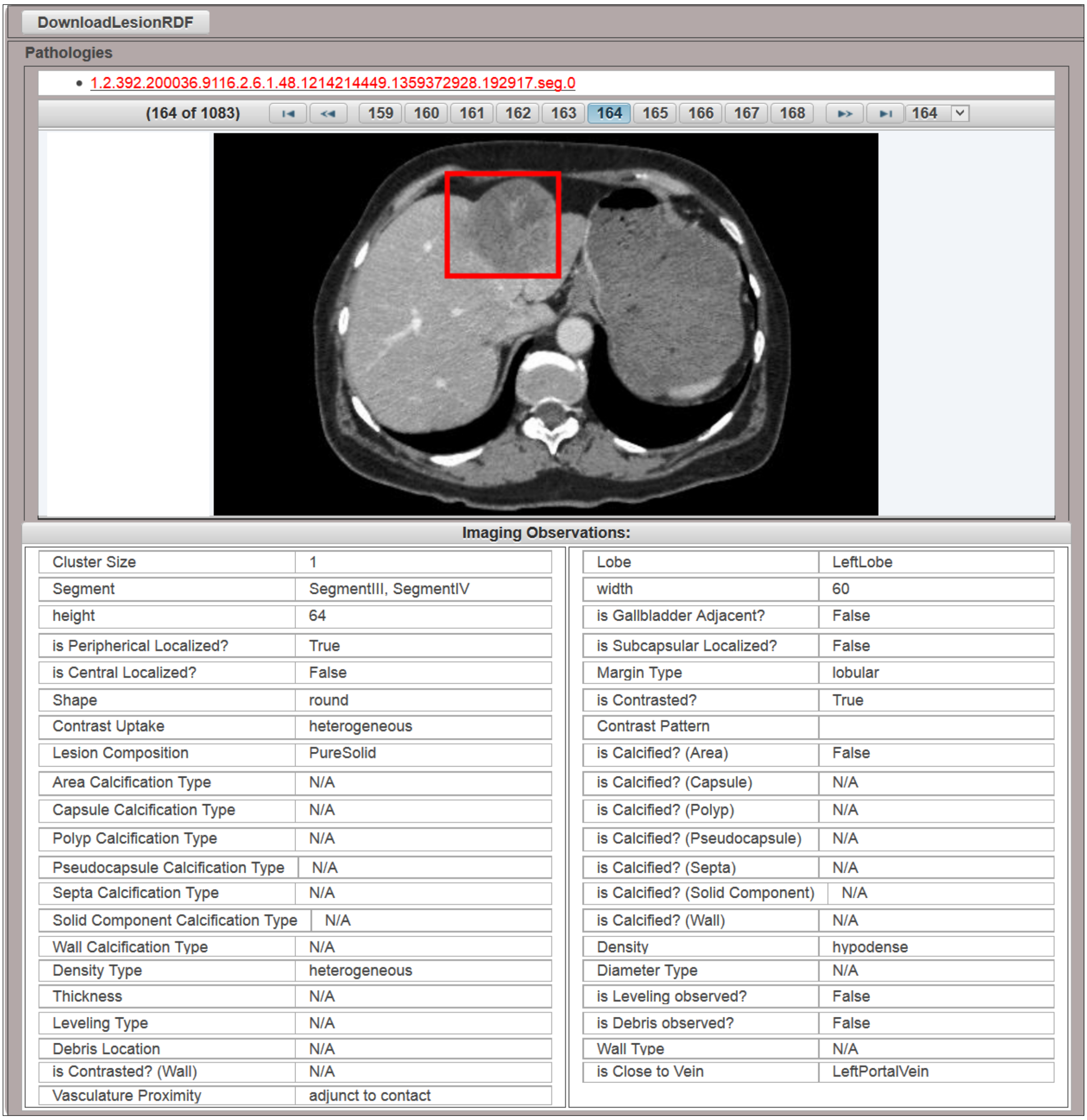

Figure 7. Ontology driven tool for collecting radiological image observations for liver patient cases.

having to specify the name of all possible blood vessels, thanks to the semantic reasoner (see Figure 9).

The final diagnosis of these patients can be queried by adding the {?study l:hasFinalDiagnosis ?disease} triple pattern to the previous query as illustrated in Query 2 where part of the result of this query also can be seen in Figure 11. As can be seen the query returns diseases in ICD-10-CM codes, such as $Q50.5$ (Hydatid cyst), $K75.9$ (Hepatitis), and $C22.0$ (Hepatocarcinoma). We can observe that some patients have a lesion close to more than one blood vessel, such as `l:Capa-00024`, or more than one lesion as `l:Capa-00342`. On the other hand, some patients, like `l:Capa-00050`, have more than one final diagnosis.

Query 3 (Figure 12) adds triple patters to obtain the

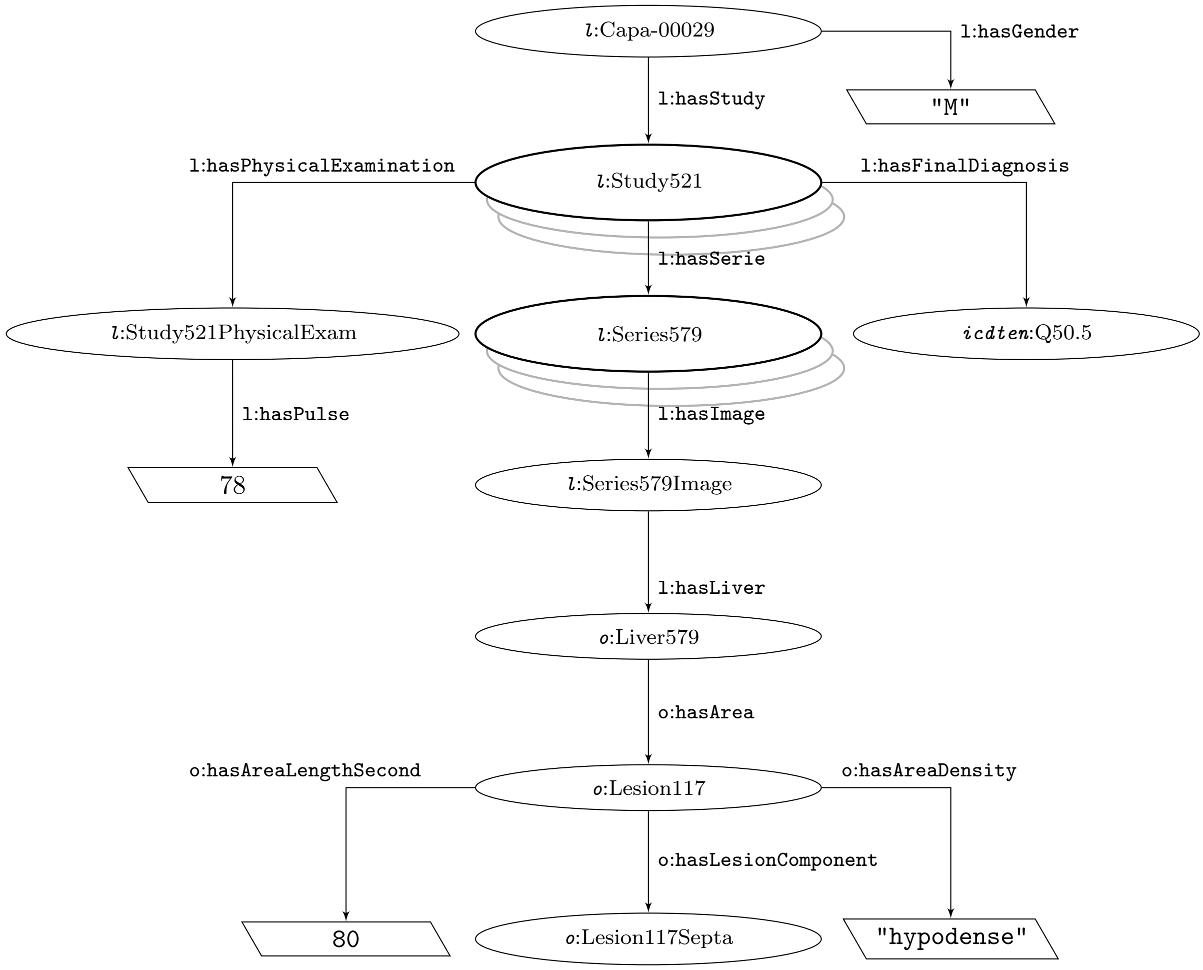

Figure 8. Selected case information for patient with id: *Capa-00029* based on RDF graphs of patient data gathered with `CaReRa-Web` tool. Ellipses denote individuals and parallelograms denote data values.

study ID, as well as the LHD [12] and pulse values, for patients retrieved by Query 2. With this kind of queries a Doctor can obtain information from integrated patient data, i.e. imaging observation, laboratory results, and physical examination which are usually allocated in different repositories.

Figure 13 shows some of the results for Query 4, which returns the patient, lesion or wall, final diseases, contrast patterns, and contrast uptake (only in case of lesion). The contrasting behavior is valuable in diagnosis and treatment. Both lesion and wall can be contrasted (Wall is a type of lesion component). Therefore, with a single query a Doctor can examine patients who have contrasted elements.

Query 5 (Figure 14) shows how to exploit the use of ICD-10-CM codes to represent diseases in LiCO to specify more complex queries. This query searches for patients that have a final diagnosis of “Malignant Neoplasm of Liver". The ICD-10-CM code of “Malignant Neoplasm of Liver" is *C22*. Therefore, the triple pattern {?study l:hasFinalDiagnosis icd10:C22} is used to return the corresponding patients. However, although there are several patients with final diagnosis C22.0 “Hepatocarcinomama” (see Figure 11) and a “Hepatocarcinomama” is a kind of “Malignant Neoplasm of Liver”, Query 5 does not return any result, because LiCO does not include the ICD-10-CM hierarchy. The ICD-10-CM *C22* has several sub-categories, such as *C22.0* “Hepatocarcinomama”, *C22.1* “Cholangiocarcinoma” and *C22.2* “Hepatoblastom”. To obtain the desired results, the subclasses of *C*22 are retrieved using Dione(Roldán-García, García-Godoy, &

[12] Lactate dehydrogenase (LDH)

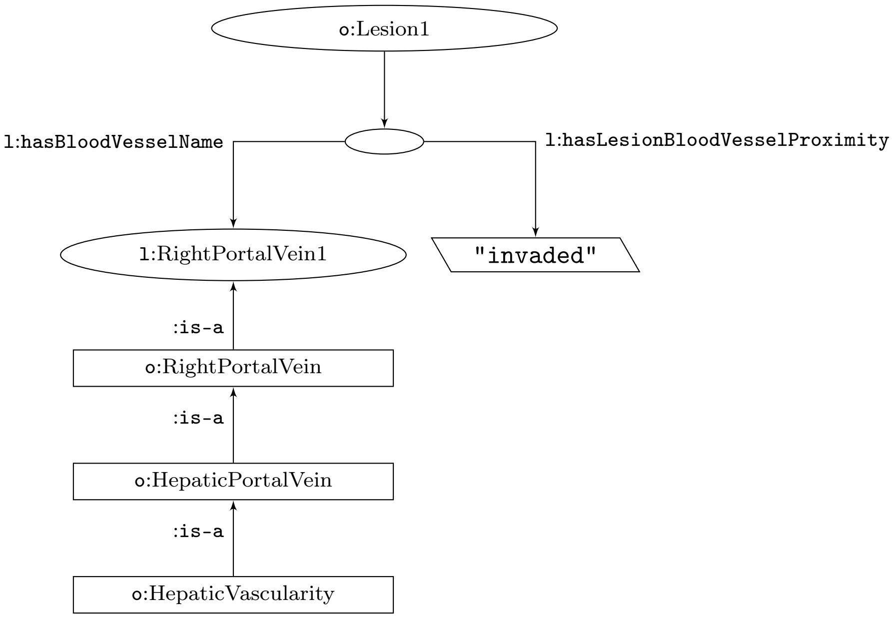

Figure 9. The modeling of proximity to a particular vessel in LiCO.

Query 1:

```
select ?patient ?lesion ?vessel
where {
  ?patient l:hasStudy ?study.
  ?study l:hasSerie ?serie.
  ?serie l:hasImage ?img.
  ?img l:hasLiver ?liver.
  ?liver o:hasArea ?lesion.
  ?lesion l:isCloseToBloodVessel ?bv.
  ?bv l:hasBloodVesselName ?vessel
}
```

Figure 10. Patients who have a liver lesion close to a hepatic blood vessel.

Aldana-Montes, 2016), which is an OWL representation of ICD-10-CM codes. Thereafter, the triple {?study l:hasFinalDiagnosis icd10:C22.} is transformed to:

```
{
  ?study l:hasFinalDiagnosis icd10:C22.0.
  ?study l:hasFinalDiagnosis icd10:C22.1.
  ...
  ?study l:hasFinalDiagnosis icd10:C22.9
}
```

Finally, with our approach, radiologists or doctors can use the RadLex codes in queries if they feel more comfortable. For example, the code RID66 (SegmentVI) can be employed to find study ID, ldh and pulse for patients with a lesion localed in the SegmentVI of liver. This is possible because the semantic reasoner is able to make owl:equivalentClass based inferences. The corresponding query (Query 6) is shown in Figure 15.

6. Discussion and Future work

LiCO has been developed to serve as an example of a holistic approach to modeling medical cases. Towards this end liver patients have been modeled relying on established vocabularies (SNOMED CT, ICD-10-CM, LOINC, and RadLex) to represent knowledge whenever possible. Elicitation sessions with a radiologist that spanned several months were invaluable in gaining understanding of liver patients and CT-scan imaging observations. It also provided great insight into the practice that includes various

Query 2:

```
select ?patient ?lesion ?vessel ?finalDiagnosis
where {
   ?patient l:hasStudy ?study.
   ?study l:hasFinalDiagnosis ?finalDiagnosis;
      l:hasSerie ?serie.
   ?serie l:hasImage ?img.
   ?img l:hasLiver ?liver.
   ?liver o:hasArea ?lesion.
   ?lesion l:isCloseToBloodVessel ?bv.
   ?bv l:hasBloodVesselName ?vessel
}
```

Result:

patient	lesion	vessel	finalDiagnosis
l:Capa-00342	o:Lesion213	o:Liver849VenacavaInferior	icdten:Q50.5
l:Capa-00342	o:Lesion213	o:Liver849RightPortalVein	icdten:Q50.5
l:Capa-00342	o:Lesion212	o:Liver849VenacavaInferior	icdten:Q50.5
l:Capa-00342	o:Lesion212	o:Liver849RightPortalVein	icdten:Q50.5
l:Capa-00050	o:Lesion122	o:Liver589HepaticVein	icdten:K75.9
l:Capa-00050	o:Lesion122	o:Liver589HepaticVein	icdten:C22.0
l:Capa-00024	o:Lesion115	o:Liver576RightHepaticVein	icdten:Q50.5
l:Capa-00024	o:Lesion115	o:Liver576MiddleHepaticVein	icdten:Q50.5
l:Capa-00024	o:Lesion115	o:Liver576RightPortalVein	icdten:Q50.5
l:Capa-00241	o:Lesion178	o:Liver748RightPortalVein	icdten:Q50.5
l:Capa-00230	o:Lesion174	o:Liver738PosteriorBranchOfRightPortalVein	icdten:C22.0

Figure 11. The patient ID, lesion, vessel and final diagnosis for patients having a lesion close to a blood vessel

preferences, and subjectivity in expression and interpretation. These observations enforced our views regarding the significance applications that are based on explicit knowledge representations that can be utilized to provide alternative access to the vast amount of data being collected in medical institutions. Therefore, we are encouraged by the results and potentials that LiCO demonstrates for case based reasoning.

As LiCO was designed as a proof of concept, it is far from a comprehensive case representation. A true case representation will require intense amount of collaborative work. Like any standardization work that would likely take several years. However, it would be very interesting to extend this work to include other imaging techniques and organs.

The `CaReRa-Web` tool that is used to collect and search for data must be extended to support semantic searches as shown in Section 5.2. Physicians or researchers should be able formulate and save semantic queries. Likewise, the results of the queries would need to be clear to interpret and manipulate. In other words, the power of semantic querying and results must be delivered to the end users who may be physicians, researchers, or students. Both of these tasks would require serious user experience (UX) design, implementation, and user studies.

In medical institutions various established vocabularies are used. Physicians familiar with these terms, which are often codes, may use them when querying. Query 7 (shown in Figure 16) highlights how SNOMED CT concepts could be used in queries. This query searches for patients who have a lesion close to the portal vein. Doctors refer to this concept with various terms, such “portal vein", “main portal vein", and “hepatic portal vein". Regardless of how it is referred to, it has a unique SNOMED

Query 3:

```
select ?patient ?study ?lesion ?vessel ?ldh ?pulse
where {
   ?patient l:hasStudy ?study.
   ?study l:hasSerie ?serie;
      l:hasLaboratoryResults ?lb;
      l:hasPhysicalExamination ?fe;
   ?serie l:hasImage ?img.
   ?img l:hasLiver ?liver.
   ?liver o:hasArea ?lesion.
   ?lesion l:isCloseToBloodVessel ?bv.
   ?bv l:hasBloodVesselName ?vessel.
   ?lb l:hasLDH ?ldh.
   ?fe l:hasPulse ?pulse
}
```

Result:

patient	study	lesion	vessel	ldh	pulse
l:Capa-00342	l:Study780	o:Lesion212	o:Liver849VenacavaInferior	3.07E2	84
l:Capa-00342	l:Study780	o:Lesion212	o:Liver849RightPortalVein	3.07E2	84
l:Capa-00050	l:Study522	o:Lesion122	o:Liver589HepaticVein	4.66E2	78
l:Capa-00024	l:Study509	o:Lesion115	o:Liver576RightHepaticVein	3.85E2	88
l:Capa-00024	l:Study509	o:Lesion115	o:Liver576MiddleHepaticVein	3.85E2	88
l:Capa-00024	l:Study509	o:Lesion115	o:Liver576RightPortalVein	3.85E2	88
l:Capa-00352	l:Study790	o:Lesion216	o:Liver859LeftPortalVein	2.95E2	70
l:Capa-00352	l:Study790	o:Lesion216	o:Liver859RightPortalVein	2.95E2	70
l:Capa-00021	l:Study507	o:Lesion135	o:Liver574PosteriorBranchOfRightPo...	3.2E2	92
l:Capa-00241	l:Study679	o:Lesion178	o:Liver748RightPortalVein	2.73E2	80

Figure 12. The study ID, lesion, vessel, ldh and pulse for patients having a lesion close to a blood vessel.

CT code [13]. These codes can be used in the SPARQL queries when `onlira`:PortalVein is defined to be equivalent to `snomed`:32764006. Results in Figure 16 highlights how the use of a semantic reasoner enhances the query results. As `onlira`:RightPortalVein and `onlira`:LeftPortalVein are subclasses of `onlira`:PortalVein, lesions close to right and left portal veins are also included in the query results.

While the current version of LiCO maps the ONLIRA concepts to the corresponding RadLex terms, it does not yet map them to SNOMED CT. Only the class `onlira`:PortalVein has been mapped to the class `snomed`:32764006 to evaluate Query 7. These mapping along with other refactoring work are among the immediate future work. Such mappings are important and being undertaken by other vocabulary for similar reasons. For example, SNOMED CT and LOINC are in the process of being mapped involving thousands of terms [14].

We have carried out an evaluation of the proposed semantic queries to demonstrate the usefulness of LiCO as a mechanism for representing and querying integrated medical information related to individual liver patients, providing shared medical knowledge to users. For this purpose we identified the kinds of functionality that is required to

[13] https://phinvads.cdc.gov/vads/ViewCodeSystemConcept.action?oid=2.16.840.1.113883.6.96&code=32764006

[14] See `http://www.snomed.org/snomed-ct/mapping-to-other-terminologies/loinc` for current status

Query 4:

select ?patient ?finalDiagnosis ?lesion ?wall ?contrastPattern ?contrastUptake

where

```
{
  ?patient l:hasStudy ?study.
  ?study l:hasFinalDiagnosis
             ?finalDiagnosis;
         l:hasSerie ?serie.
  ?serie l:hasImage ?img.
  ?img   l:hasLiver ?liver.
  ?liver o:hasArea ?lesion.
  ?lesion o:hasLesionComponent ?wall.
  ?wall o:isContrasted true.
  optional { ?lesion
           o:hasLesionContrastPattern
           ?contrastPattern}.
       optional {?lesion
           o:hasLesionContrastUptake
           ?contrastUptake }
  }
union{
  ?patient l:hasStudy ?study.
  ?study l:hasFinalDiagnosis
             ?finalDiagnosis;
         l:hasSerie ?serie.
  ?serie l:hasImage ?img.
  ?img l:hasLiver ?liver.
  ?liver  o:hasArea ?lesion.
  ?lesion o:isContrasted true;
          o:hasLesionContrastPattern
  ?contrastPattern;
          o:hasLesionContrastUptake
  ?contrastUptake}
}
```

order by asc(?finalDiagnosis) ?wall

Result:

patient	finalDiagnosis	lesion	wall	contrastPattern	contrastUptake
l:Capa-00021	icdten:C22.0	o:Lesion135		homogeneous	homogeneous
l:Capa-00021	icdten:K75.9	o:Lesion135		homogeneous	homogeneous
l:Capa-00230	icdten:C22.0	o:Lesion174		heterogeneous	heterogeneous
l:Capa-00050	icdten:C22.0	o:Lesion122		homogeneous	minimal
l:Capa-00050	icdten:K75.9	o:Lesion122		homogeneous	minimal
l:Capa-00130	icdten:C22.0	o:Lesion140		peripheric	heterogeneous
l:Capa-00130	icdten:K75.9	o:Lesion140		peripheric	heterogeneous
l:Capa-00323	icdten:C22.0	o:Lesion208		peripheric	minimal
l:Capa-00323	icdten:K75.9	o:Lesion208		peripheric	minimal
l:Capa-00029	icdten:Q50.5	o:Lesion117	o:Lesion117Wall		
l:Capa-00352	icdten:Q50.5	o:Lesion216	o:Lesion216Wall	peripheric	

Figure 13. The patient ID, lesion or wall ID, final diagnosis, contrast pattern and contrast uptake values for patients who have a lesion or a wall that is contrasted.

Query 5:

```
PREFIX icd10: <http://purl.bioontology.org/ontology/ICD-10/>
select ?patient
where
{
    ?patient l:hasStudy ?study.
    ?study l:hasFinalDiagnosis icd10:C22
}
```

Figure 14. Patients who have “Malignant Neoplasm of Liver" as final diagnosis, using ICD-10-CM codes

Query 6:

```
PREFIX radlex: <http://www.radlex.org/RID/#>
select ?patient ?study ?lesion ?ldh ?pulse
where{
      ?patient l:hasStudy ?study.
      ?study l:hasSerie ?serie.
         l:hasLaboratoryResults ?lb;
         l:hasPhysicalExamination ?fe;
      ?serie l:hasImage ?img.
      ?img l:hasLiver ?liver.
      ?liver o:hasArea ?lesion.
      ?lesion o:isLocatedInSegment ?sg.
      ?sg rdf:type radlex:RID66.
      ?lb l:hasLDH ?ldh.
      ?fe l:hasPulse ?pulse
}
```

Result:

patient	study	lesion	ldh	pulse
l:Capa-00152	l:Study599	o:Lesion149	4.93E2	88
l:Capa-00024	l:Study509	o:Lesion115	3.85E2	88
l:Capa-00021	l:Study507	o:Lesion135	3.2E2	92
l:Capa-00321	l:Study759	o:Lesion207	4.06E2	88
l:Capa-00230	l:Study669	o:Lesion174	3.17E2	70
l:Capa-00255	l:Study693	o:Lesion185	2.73E2	70
l:Capa-00273	l:Study711	o:Lesion194	2.99E2	70
l:Capa-00267	l:Study705	o:Lesion192	3.52E2	80
l:Capa-00055	l:Study525	o:Lesion125	9.4E2	80
l:Capa-00072	l:Study532	o:Lesion127	3.89E2	80
l:Capa-00354	l:Study792	o:Lesion217	3.58E2	88
l:Capa-00306	l:Study744	o:Lesion202	3.31E2	88

Figure 15. The patient, study, lesion, ldh and pulse values for patients who have a lesion in segment VI, using the RadLex code R66 (segment VI).

Query 7:

```
PREFIX snomed: <http://snomed.info/sct/>
select ?patient ?lesion ?vessel
where {
  ?patient l:hasStudy ?study.
  ?study l:hasSerie ?serie.
  ?serie l:hasImage ?img.
  ?img l:hasLiver ?liver.
  ?liver o:hasArea ?lesion.
  ?lesion l:isCloseToBloodVessel ?bv.
  ?bv l:hasBloodVesselName ?vessel.
  ?vessel rdf:type snomed:32764006
}
```

Result:

patient	lesion	vessel
l:Capa-00342	o:Lesion212	o:Liver849RightPortalVein
l:Capa-00342	o:Lesion213	o:Liver849RightPortalVein
l:Capa-00352	o:Lesion216	o:Liver859LeftPortalVein
l:Capa-00352	o:Lesion216	o:Liver859RightPortalVein
l:Capa-00021	o:Lesion135	o:Liver574PosteriorBranchOfRightPortalVein
l:Capa-00230	o:Lesion174	o:Liver738PosteriorBranchOfRightPortalVein
l:Capa-00255	o:Lesion185	o:Liver762RightPortalVein
l:Capa-00014	o:Lesion133	o:Liver571LeftPortalVein
l:Capa-00035	o:Lesion119	o:Liver583AnteriorBranchOfRightPortalVein

Figure 16. Patients who have a liver lesion close to the portal vein using the SNOMED CT code 32764006 "Portal Vein"

retrieve the kinds of queries we demonstrated. And, examined LiCO in comparison to other approaches.

Table 10 describes the functionality required to solve the SPARQL queries presented in Section 5.2. Each functionality is given an acronym for reference purposes to be used in the comparison table. For example, to solve Query 1, a functionality to search for lesions that are close to a blood vessel is required (LVB); for Query 2, an additional functionality to retrieve the final diagnosis of a patient is needed (LBV and FD); and for Query 6 a functionality to translate RadLex codes to other terminology (RD2STH) is needed – provided the system evaluating the query does not.

For each query, the number of functionalities required to solve it with the terminologies/ontologies described in Section 2.2 is determined. Table 11 summarizes this information in terms of the number of and the specific functionalities required. For example, to solve Query 1, RadLex needs an external functionality to retrieve the lesions close to a blood vessel. Whereas, ONLIRA is able to evaluate Query 1 without any external functionality. This evaluation is conducted optimistically, in that when when in doubt whether a terminology provides a specific functionality, it is assumed that it does.

As a result, we can see that LiCO is able to evaluate almost all queries without any external functionality, except for Query 5, where LiCO needs to fetch the subclasses of ICD-10-CM classes. ONLIRA, which generally performs better than the others in such queries, still needs external functionality regarding patient information. These results demonstrate that LiCO is a representation that integrates patient medical information, addressing one of the most important requirements of a CES platform.

Furthermore, reasoning can be employed to classify patients of interest in order to examine patients according to certain characteristics. Subjective queries by physicians, researchers, and students is an important feature for CES. This can be achieved with SWRL rules that describe the

Table 10
Functionalities needed to solve the SPARQL queries

Acronym	System functionality
LBV	Lesion close to Blood Vessel
FD	Final Diagnosis
LR	laboratory results
PEX	Physical EXamination
LCP	Lesion ComPonents
LCT	Lesion ConTrast
CCT	Lesion Component Contrast
ICD2STH	ICD-10-CM to SomeThing
LLS	Lesion Located in Segment
RD2STH	RadLex to SomeThing
SN2STH	SNOMED CT to SomeThing

desired patient characteristics. With such rules doctors are easily able to refer to patients they interested in analyzing. For example, the SWRL code:

```
lico:Patient(?x)
^ lico:hasStudy(?x, ?s)
^ lico:hasLaboratoryResults(?s, ?lr)
^ lico:hasAlbumin(?lr, ?alb)
^ lico:hasSerie(?s, ?se)
^ lico:hasImage(?se, ?im)
^ lico:hasLiver(?im, ?li)
^ onlira:hasArea(?li, ?lesion)
^ onlira:isContrasted(?lesion, true)
^ greaterThan(?alb, 4.5)
-> lico:PatientCLA(?x)
```

creates a class of patients called *PatientCLA* that have a contrasted lesion image and an albumin value greater than 4.5.

A system that includes user experience design to deliver the search and rule definition functionality to end users is required for a meaningful CES ecosystem. Such patient class description rules, could themselves be saved for sharing experiences at a higher level.

7. Conclusions

An ontology, LiCO was developed for the domain of liver patients as a proof of concept for composite medical case based representation and reasoning. LiCO maps several established medical vocabularies (SNOMED CT, LOINC, RadLex, and ICD-10-CM) to relevant classes. It is demonstrated on real patient data that LiCO successfully represents the liver cases and that the *Stardog* reasoner can perform non-trivial semantic case queries over this representation, via exploiting the relations with other vocabularies.

Furthermore, SWRL rules were proposed to demonstrate the flexibility and extensibility of the proposed approach in supporting complex relationships that physicians, researchers, and students with differing interests and preferences require. As such LiCO can effectively accommodate the subjectivity of medical case search and retrieval.

In conclusion, the presented results are highly encouraging for the adoption of semantic web technologies for modeling medical cases, which are far more complex than other domains. Future research would include the extension of LiCO to other organs (medical subdomains) and development of intuitive user interfaces that would allow end-users to build subjective medical case queries in an effective and efficient way.

Acknowledgement

We thank Dr. Rüstü Türkay for sharing his domain knowledge and Abdülkadir Yazıcı for his support in extending `CaReRa-Web` to support RDF.

References

Alfonse, M., Aref, M. M., & Salem, A.-B. M. (2014). An ontology-based system for cancer diseases knowledge management. *International Journal of Information Engineering and Electronic Business*, *6*(6), 55.

Allampalli-Nagaraj, G., & Bichindaritz, I. (2009). Automatic semantic indexing of medical images using a web ontology language for case-based image retrieval. *Engineering Applications of Artificial Intelligence*, *22*(1), 18–25.

Barzegar Marvasti, N., Akgül, C. B., Acar, B., Kökciyan, N., Üsküdarlı, S., Yolum, P., ... Bakır, B. (2013). Clinical experience sharing by similar case retrieval. In *Proceedings of the 1st acm international workshop on multimedia indexing and information retrieval for healthcare* (pp. 67–74). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2505323.2505335 doi: 10.1145/2505323.2505335

Bodenreider, O. (2004). The unified medical language system (umls): integrating biomedical terminology. *Nucleic acids research*, *32*(suppl_1), D267–D270.

Table 11
Number and type of external system functionalities needed to solve each proposed query.

Query	RadLex	ONLIRA	SNOMED CT	ICD-10-CM	LCO (Kaur & Khamparia, 2015)	LiCO
Query 1	**1** (LBV)	**0**	**1** (LBV)	**1** (LBV)	**1** (LBV)	**0**
Query 2	**2** (LBV,FD)	**1** (FD)	**1** (LBV)	**1** (LBV)	**1** (LBV)	**0**
Query 3	**3** (LBV,LR,PEX)	**2** (LR,PEX)	**2** (LBV,LR)	**3** (LBV,LR,PEX)	**3** (LBV,LR,PEX)	**0**
Query 4	**4** (FD,LCP,LCT,CCT)	**1** (FD)	**0**	**4** (FD,LCP,LCT,CCT)	**3** (LCP,LCT,CCT)	**0**
Query 5	**2** (FD,ICD2STH)	**2** (FD,ICD2STH)	**1** (ICD2STH)	**0**	**2** (FD,ICD2STH)	**1** (ICD2STH)
Query 6	**3** (LR,PEX, LLS)	**2** (LR,PEX)	**3** (LR,LLS, RD2STH)	**4** (LR,PEX, LLS, RD2STH)	**0**	**0**
Query 7	**2** (LBV, SN2STH)	**1** (SN2STH)	**1** LBV	**2** (LBV, SN2STH)	**2** (LBV, SN2STH)	**0**

Chen, R.-C., Huang, Y.-H., Bau, C.-T., & Chen, S.-M. (2012). A recommendation system based on domain ontology and swrl for anti-diabetic drugs selection. *Expert Systems with Applications*, *39*(4), 3995–4006.

Dean, M., & Schreiber, G. (2004). *Owl web ontology language reference* (Tech. Rep.). W3C Recommendation, 10 February 2004. Retrieved from `http://www.w3.org/TR/2004/REC-owl-ref-20040210/`

Forrey, A. W., Mcdonald, C. J., DeMoor, G., Huff, S. M., Leavelle, D., Leland, D., ... others (1996). Logical observation identifier names and codes (loinc) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clinical Chemistry*, *42*(1), 81–90.

Gibaud, B., Forestier, G., Benoit-Cattin, H., Cervenansky, F., Clarysse, P., Friboulet, D., ... others (2014). Ontovip: An ontology for the annotation of object models used for medical image simulation. *Journal of biomedical informatics*, *52*, 279–292.

Gruber, T. R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition,*, *5*(2), 199-220.

Haase, P., & Stojanovic, L. (2005). Consistent evolution of owl ontologies. In A. Gómez-Pérez & J. Euzenat (Eds.), *The semantic web: Research and applications* (Vol. 3532, p. 182-197). Springer Berlin Heidelberg.

Horrocks, I., & Patel-Schneider, P. (2003). Reducing owl entailment to description logic satisfiability. In D. Fensel, K. Sycara, & J. Mylopoulos (Eds.), *The semantic web - iswc 2003* (Vol. 2870, p. 17-29). Springer Berlin Heidelberg.

Huff, S. M., Rocha, R. A., McDonald, C. J., De Moor, G. J., Fiers, T., Bidgood Jr, W. D., ... others (1998). Development of the logical observation identifier names and codes (loinc) vocabulary. *Journal of the American Medical Informatics Association*, *5*(3), 276–292.

Ivanović, M., & Budimac, Z. (2014). An overview of ontologies and data resources in medical domains. *Expert Systems with Applications*, *41*(11), 5158–5166.

Kaur, P., & Khamparia, A. (2015). diagnosis of liver cancer ontology using sparql. *international journal of applied engineering research*, *10*(69), 2015.

Khobreh, M., Ansari, F., Dornhöfer, M., & Fathi, M. (2013). An ontology-based recommender system to support nursing education and training. In *Lwa* (pp. 237–244).

Kokciyan, N., Turkay, R., Uskudarli, S., Yolum, P., Bakir, B., & Acar, B. (2014). Semantic description of liver ct images: An ontological approach. *IEEE journal of biomedical and health informatics*, *18*(4), 1363–1369.

Kundu, S., Itkin, M., Gervais, D. A., Krishnamurthy, V. N., Wallace, M. J., Cardella, J. F., ... Langlotz, C. P. (2009). The {IR} radlex project: An interventional radiology lexicon - a collaborative project of the radiological society of north america and the society of interventional radiology. *Journal of Vascular and Interventional Radiology*, *20*(7, Supplement), S275 - S277. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1051044309003297`

Lacson, R., Andriole, K. P., Prevedello, L. M., & Khorasani, R. (2012). Information from searching content with an ontology-utilizing toolkit (iscout). *Journal of digital imaging*, *25*(4), 512–519.

Langlotz, C. P. (2006). *Radlex: a new method for indexing online educational materials.* Radiological Society of North America.

Lipscomb, C. E. (2000). Medical subject headings (mesh). *Bulletin of the Medical Library Association*, *88*(3), 265.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A., & Schneider, L. (2003). Dolce: a descriptive ontology for linguistic and cognitive engineering. *WonderWeb Project, Deliverable D17 v2*, *1*, 75–105.

McBride, B. (2001). Jena: Implementing the rdf model and syntax specification. In *Proceedings of the second international conference on semantic web-volume 40* (pp. 23–28).

McDonald, C. J., Huff, S. M., Suico, J. G., Hill, G., Leavelle, D., Aller, R., . . . others (2003). Loinc, a universal standard for identifying laboratory observations: a 5-year update. *Clinical chemistry*, *49*(4), 624–633.

McGuinness, D., & Harmelen, F. (2004). *Owl web ontology language overview* (Tech. Rep.). W3C Recommendation.

Noy, N. F., & McGuinness, D. L. (2001). *Dontology development 101: A guide to creating your first ontology* (Tech. Rep.). tanford University Knowledge Systems Laboratory Technical Report KSL-01-05. Retrieved from `http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html`

Organization, W. H. (n.d.). *Who | international classification of diseases.* `http://www.who.int/classifications/icd/en/`. (Accessed: 2017-08-05)

Organization, W. H., et al. (2004). Icd-10: international statistical classification of diseases and related health problems: tenth revision.

Pérez, J., Arenas, M., & Gutierrez, C. (2009, September). Semantics and complexity of sparql. *ACM Trans. Database Syst.*, *34*(3), 16:1–16:45.

Quinn, S., Bond, R., & Nugent, C. (2017). Ontological modelling and rule-based reasoning for the provision of personalized patient education. *Expert Systems*, *34*(2).

Roldán-García, M. M., García-Godoy, M. J., & Aldana-Montes, J. F. (2016). Dione: An OWL representation of ICD-10-CM for classifying patients' diseases. *J. Biomedical Semantics*, *7*, 62.

Rosse, C., & Mejino Jr, J. L. (2008). The foundational model of anatomy ontology. In *Anatomy ontologies for bioinformatics* (pp. 59–117). Springer.

Rubin, D. L. (2008). Creating and curating a terminology for radiology: ontology modeling and analysis. *Journal of digital imaging*, *21*(4), 355–362.

Salem, A.-B. M., & Alfonse, M. (2007). Building web-based lung cancer ontology. In *Proceedings of 1st national symposium on e-health and bioengineering-ehb* (pp. 177–182).

Salem, A.-B. M., & Alfonse, M. (2009). Ontological engineering approach for breast cancer knowledge management. *Proc. of Med-e-Tel the International eHealth, Telemedicine and Health ICT for Education, Networking and Business, Luxembourg*, 320–324.

Seifert, S., Thoma, M., Stegmaier, F., Hammon, M., Kramer, M., Huber, M., . . . Comaniciu, D. (2011). Combined semantic and similarity search in medical image databases. In *Spie medical imaging* (Vol. 7967, pp. 79670301–79670310).

SNOMED CT website. (n.d.). http://www.snomed.org/snomed-ct/. (Accessed: 17-August-2017)

Staab, S., & Studer, R. (2009). *Handbook on ontologies.* International Handbooks on Information Systems. Springer.

Temal, L., Dojat, M., Kassel, G., & Gibaud, B. (2008). Towards an ontology for sharing medical images and regions of interest in neuroimaging. *Journal of Biomedical Informatics*, *41*(5), 766–778.

Warren, W., Agoncillo, A., Franklin, J., & Brinkley, J. (2006). Intelligent web-based whole body visualization for anatomy education. In *Amia annual symposium proceedings* (Vol. 2006, p. 1136).