A FHIR-based System for the Generation and Retrieval of Clinical Documents

Crescenzo Diomaiuta, Mario Sicuranza, Mario Ciampi and Giuseppe De Pietro

*Institute of High Performance Computing and Networking, National Research Council of Italy,
Naples, Italy*

Keywords: FHIR, Terminology Service, Interoperability, Information Exchange, Health Information Systems.

Abstract: Clinical information exchange among heterogeneous systems is a complex process. It is necessary to use dictionaries, coding standards (e.g. LOINC and SNOMED), and terminology services, in order to: i) produce machine-interpretable information and coded data, ii) reduce information exchange complexity, and iii) allow the fulfillment of semantic interoperability among several systems. This work presents a system to support healthcare professionals in the creation of clinical documents using appropriate standards and in the information reading and retrieval of interest on clinical documents. The system allows healthcare professionals to use standard codes easily (via terminology server) during the creation of documents. Similarly, it also permits the decoding of health standard codes, during the read documents process, to facilitate the obtaining of information. In summary, it makes use of a simple and intuitive graphical user interface and a Fast Healthcare Interoperability Resources (FHIR) terminology server to support healthcare professionals in diagnosing and treating patients.

1 INTRODUCTION

The use of information and communication technology (ICT) in healthcare has enabled the creation of the Health Information System (HIS). A HIS is a complex system that captures, stores, manages, and transmits health information relating to patients (Haux, 2006). In the past few years, many HIS, each with its own characteristics and satisfying local requirements, have been defined (Esposito et al., 2012). This has meant that these systems tend to be heterogeneous (Minutolo et al., 2011). In fact, they use proprietary solutions and do not structure and standardize the managed information in the same manner. This failure to use the same standards makes the distribution and management of clinical information extremely diverse. An important HIS based on communication and interoperability among several systems is the Electronic Health Record system (EHR), which is a system that allows a complete management of the clinical events of the patients (Garde et al., 2007). To improve care processes in healthcare it is essential that different HIS are able to communicate with each other to exchange patient health information, and that they have a common semantic meaning in all systems. Interoperability has been defined by the Insti-

tute of Electrical and Electronics Engineers (IEEE) as the "ability of two or more components to exchange information and to use the information that has been exchanged" (IEEE, 1990). In this perspective, four different levels of interoperability can be distinguished (Walker et al., 2005):

- Level 1 - non-electronic data: no use of Information Technology (IT) to share information.
- Level 2 - machine transportable data: transmission of unstructured information through basic IT services.
- Level 3 - machine organizable data or syntactic interoperability: transmission of structured messages containing non-standardized data. This requires an incoming data translation, because sending an organization's vocabulary and receiving an organization's vocabulary are different functions.
- Level 4 - machine-interpretable data or semantic interoperability: transmission of structured messages containing standardized and coded data. In this case all systems exchange information using the same formats and vocabularies.

This paper presents a very pragmatic system capable of enabling healthcare professionals to write and read

135

clinical documents in accordance with HL7 FHIR standard.

1.1 Standards

This section presents certain widespread clinical standards. There are different health standards for the structured representation of clinical information. The most widely used standards in healthcare have been defined by an international non-profit association, termed Health Level Seven (HL7). Over the years, there has been an evolution of these standards (Lamprinakos et al., 2014). There follows a brief description of the evolution of the standards defined by HL7 for structuring information. In particular, the first standard described is **HL7 version 2** (**HL7 v2**). This messaging standard permits information exchange by the use of messages encoded in ASCII and delimited by a series of escape characters. The message is composed of one or more segments and must have a precise structure. For example, to represent clinical observations the OBX segment is used, as shown below.

```
OBX|3|TX|88304&MDT|1|ANOMALY BLOOD SUGAR
```

Newer releases of HL7 v2 use the XML language to define the message. Nevertheless, as a limitation of this standard it does not have an explicit information model and is characterized by numerous optionalities for any change of or addition to the format. This aspect undermines the interoperability, and, indeed, version 2 provides multiple ways of performing the same action. This is one of the reasons which led to the implementation of the standard **HL7 version 3** (**HL7 v3**). This standard, in contrast to version 2, is based on a message development framework. It involves the use of different models, for example, the *Reference Information Model (RIM)*, which is a set of base classes that can be used to model any entity in the health domain. One of the issues of this standard is its high complexity of use. There is no compatibility between HL7 v2 and HL7 v3. Another standard is the **HL7 Clinical Document Architecture (CDA)**, derived from the HL7 v3 RIM. This is based on the XML standard, used to specifiy the encoding, structure, and semantics of clinical documents, for the exchange between heterogeneous systems. This standard also presents some limitations. In particular, it is quite complex to read a CDA document, and, furthermore it does not provide granular access to data. Subsequently, the HL7 association has created a new standard called **Fast Healthcare Interoperability Resources** (**FHIR**). This standard combines the best features of HL7 v2, HL7 v3, and CDA with the addition of the use of web standards, such as HTTP, Atom, XML, and JSON. FHIR is published as

a Draft Standard for Trial Use (DSTU) and its adoption is growing rapidly (Bender and Sartipi, 2013). FHIR designers employ an incremental and iterative approach to develop standards that reflect the requirements of industry regarding the design of complex systems. FHIR is built on a set of modular components called "Resources". Each resource is a unique entity, with peculiar properties and an identity. Currently, it is allowed to represent about 90 resources. FHIR supports a RESTful (REpresentational State Transfer) architecture and seamless exchange of information, using messages or documents. The FHIR RESTful API provides a consistent set of HTTP services to find (through search operation) and manage resources (through read, create, update, and delete operations). It uses an open standard for authorization, Open Authorization (OAuth), Atom for query results, and XML or JSON for data representation (Luz et al., 2015). The FHIR specifications give support to provide a terminology service, offering a conformance statement, through three main concepts: *code systems*, *value sets*, and *concept maps*.

1.2 Context

In this paper, an architecture is proposed to support healthcare professionals in creating or reading structured clinical documents. The system is based on a terminology service, compliant with the FHIR standard, and uses clinical coding systems. Using the proposed architecture, the healthcare professional will be able to structure the information about a patient. With this system, he/she will be able to easily generate structured clinical documents according to the FHIR standard and will be able to extract rapidly the content of interest from structured documents. Moreover, the use of the standard promotes integration and sharing across heterogeneous systems. This represents a basis to construct a more complex systems to achieve clinical interoperability. In literature, there are several proposals, relating to the creation of clinical documents, making use of the terminology, data exchanging, and standardizing of the information contained in documents. jTerm is an open source terminology server, that supports the use of a wide variety of existing terminological systems. It uses an abstract model (Hogarth et al., 2003). The system is released in the form of a J2EE application, which supports a web front-end to browse terminologies. It is a multiplatform system, which uses a wide variety of terminology systems, accessed by a command/query mechanism. The work proposed by Navas et al. concerns a restrictive user interface, implemented in Java, with the objective of improving user data acquisition (Na-
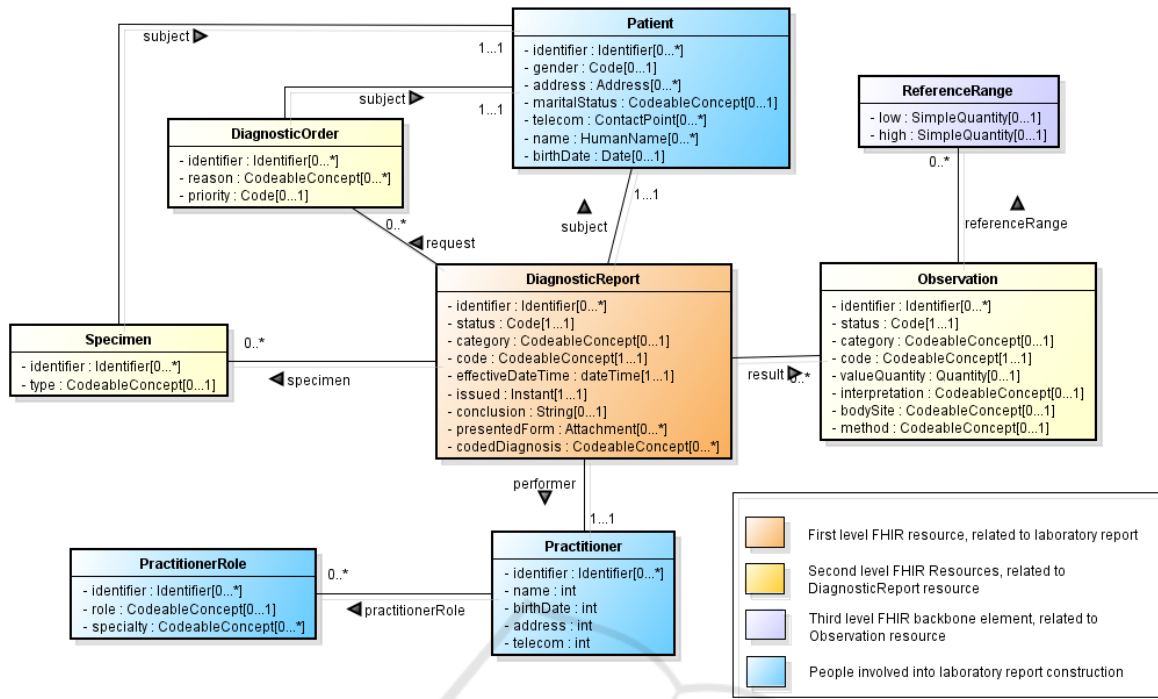
Figure 1: The DiagnosticReport FHIR resource, and its link to other interested resources.

vas et al., 2007). This system also allows the selection of already coded terms, in the context of an in-patient electronic medical record. The proposed software interacts online with a terminology server for the Interactive Coding of Discharge Summaries. This solution has the limitation of being topic-specific, because it is related only to discharge summaries. An approach to ensure semantic interoperability among Electronic Health Records was proposed by Sànchez-Caro et al. (Sànchez-Caro et al., 2014). They present a proposal to facilitate the exchange of Cardiovascular Risk Stratification (CVRS) information, prototyped and focused on Heart Rate Turbulence (HRT), using archetypes and the SNOMED-CT ontology. An archetype is defined by the OpenEHR Foundation as follows: it "is a computable expression of a domain content model in the form of structured constraint statements, based on some reference models". More precisely, archetypes are models of clinical (or other domain specific) concepts. From a technical point of view, archetypes are formal specifications of clinical content, whereas, from a clinical point of view, archetypes are the basis to intuitively define, discuss and present clinical content. To create an HRT archetype they use openEHR Archetype Editor. Finally, Honko et al. designed a Wellness Warehouse Engine (W2E) that provides interfaces to different data sources and makes data available via REST API to other services (Honko et al., 2015). It includes Unifier, which

is an engine that transforms input data into generic units, and Analyzer, which is an engine for the advanced analysis of input data. The described architecture is used to ensure the semantic interoperability of consumer health data. In particular, their architecture offers solutions for the collection and storing of data, from several sensors, and for the visualization of the results. Furthermore, the W2E solution performs data unification where possible. These solutions are lacking for semantic interoperability, in particular they do not make use of standards to represent information (e.g. HL7 v2, CDA, and etc). In this way, integration and data exchange with existing systems is complex.

2 THE PROPOSAL

In this section our proposal will be presented, including a description of the rationale that led to the creation of the system, and of the technological decisions made for its implementation.

2.1 Overview

The rationale for the realization of this proposal comprised several factors, among which one of the most important was the aim of supporting healthcare pro-

```
<contained>
  <Observation>
    <id value="r1"/>
    <status value="final"/>
    <code>
      <coding>
        <system value="http://loinc.org"/>
        <code value="718-7"/>
        <display value="Hemoglobin [Mass/volume] in Blood"/>
      </coding>
      <text value="Haemoglobin"/>
    </code>
    <subject>
      <reference value="Patient/A.SMITH"/>
    </subject>
    <performer>
      <reference value="Organization/Lab"/>
      <display value="Laboratory, Inc"/>
    </performer>
    <valueQuantity>
      <value value="170"/>
      <unit value="g/L"/>
      <system value="http://unitsofmeasure.org"/>
      <code value="g/L"/>
    </valueQuantity>
    <referenceRange>
      <low>
        <value value="135"/>
        <unit value="g/L"/>
        <system value="http://unitsofmeasure.org"/>
        <code value="g/L"/>
      </low>
      <high>
        <value value="180"/>
        <unit value="g/L"/>
        <system value="http://unitsofmeasure.org"/>
        <code value="g/L"/>
      </high>
    </referenceRange>
  </Observation>
</contained>
```

Figure 2: Clinical document focusing on the Observation resource.

fessionals, in creating and reading structured and co-ded clinical documents. In this way, the managing, sharing, and the use of information can be facilitated. The proposed architecture uses a healthcare standard to represent concepts. Therefore, it will be easy to integrate it with new systems, and add new features. The production of structured data via standards provi-des the capability of data sharing with other HIS, and allows the collection of clinical data from other health information systems. The defined system offers healt-hcare professionals two main features: *the creation of new clinical documents* enriched with diagnoses and personal observations and structured according to the standard; and the possibility of *decoding existing cli-nical documents*. In particular, the system allows he-althcare professionals to read the detailed information contained in clinical documents. The encoding and decoding operations are transparent to the user, and performed automatically by the system. The archi-tecture is modular, in this way providing the ability to add new features by adding modules. The modu-les interact with each other, and communicate through an existing terminology server, by means of a classic client-server communication paradigm. The standard used for the clinical document representation (for cre-ation and retrieval) is FHIR. For the codes-concepts association a terminology server is used with code sy-stems, such as LOINC, SNOMED, UMLS, ICD-10, etc.

Before presenting the system in detail, the following

paragraph shows the mapping between FHIR resour-ces and reality concepts. In particular, to structure the information contained within the clinical document according to the FHIR standards, it is necessary to represent concepts using appropriate FHIR resources.

2.2 Mapping Between Real Concepts and FHIR Resources

To represent the resources through FHIR, it was ne-cessary to identify the minimal set of concepts used in real cases to describe it, and then map them with FHIR resources. In particular, a clinical document contains information about the patient's health and other personal data. Healthcare professionals record in the documents the necessary and correct informa-tion for the planning, arrangement, provision, and monitoring of patient care. Only relevant informa-tion is entered into patient documents. For exam-ple, in the representation of a laboratory report the following information was selected: patient, healt-hcare professional, prescription order and accompa-nying reasons, and measurements and simple asser-tions about a patient. Considering a particular tre-atment situation, more precisely, the problem of the formulation of laboratory reports enriched with ob-servations and diagnoses, the following concepts have been identified: patient, healthcare professional, pres-cription order and accompanying reasons, and mea-surements and simple assertions about a patient. The choice of the "laboratory report" was not random, in that, in fact, it represents a very large domain of inte-rest. Furthermore, it contains a lot of information of interest such as *conclusion*, which refers to a clinical interpretation of test results. Following the identifi-cation of the concepts, the associated FHIR resources were determined. Figure 1 shows the UML diagram of the identified FHIR resources, listing for each one, the cardinality and attributes. Only the most signi-ficant have been selected (LIS, 2016). The resource "DiagnosticReport" is associated to the laboratory re-port concept. It references other resources for mo-deling concepts with a hierarchical dependence. The classification starts from the first level and progresses to the third. In particular, the following resources are included:

- **DiagnosticReport**: this is the main resource of the architecture and it represents the "laboratory report" concept. It contains the interpretation of diagnostic tests on patients or on groups of pa-tients. The report includes a set of information that is typically provided by a diagnostic service, e.g. a mixture of atomic results, images, textual and coded interpretations, and formatted repre-
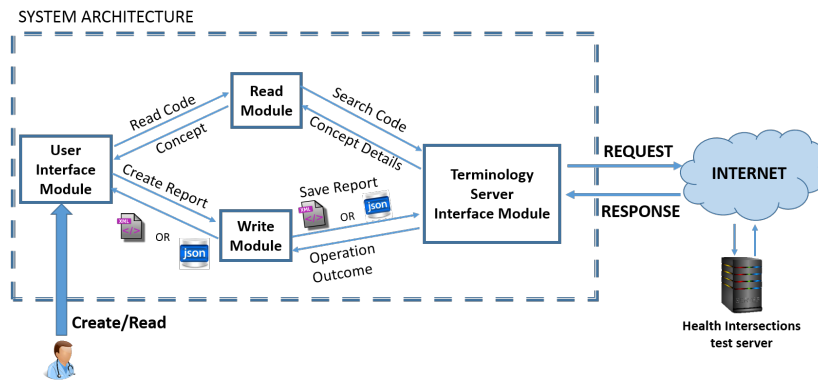
Figure 3: System architecture with a client request and server response.

sentations of diagnostic reports. It can be used in diagnostic contexts such as hematology and microbiology, etc.

- **Patient**: this resource models demographic information about a patient who needs medical treatment or other health-related services. As can be seen in figure 1, this resource is connected to the "Specimen" and "DiagnosticOrder" resources.

- **Practitioner**: this is used to represent a person who is directly or indirectly involved in the provision of healthcare, such as a healthcare professional. It is connected to extra information relative to *role* (e.g. doctor, nurse, etc.) and *specialization* (e.g. cardiologist, dentist, etc.).

- **DiagnosticOrder**: this can be used to record a request for a diagnostic investigation service to be performed, such as a *reservation*.

- **Observation**: this is designed to contain specific information on non-complex concepts in relation to a patient. It is the central element of any healthcare system based on FHIR, used to support diagnosis, monitor progress, and determine baselines and patterns.

- **Specimen**: this can be used to support the collection of useful samples for an analysis process, for example to indicate samples from biological entities, living or dead.

Figure 2 shows an example of the system's output. It is a structured report compliant with the FHIR standard, which is both human-readable and machine-readable. The report's format representation is XML (also available in JSON), and highlights the most important and required fields for the "Observation" resource description. In particular, it shows the resources and attributes needed to structure the result, relative to a patient's blood sample.

2.3 Architecture

The proposed system is shown in figure 3. It consists of several modules:

- **the user interface module**: this module is responsible for managing input/output. It offers an input command/query mechanism and shows output replies, interfacing between the system and healthcare professionals. In particular, it allows the creation or reading of observations and clinical documents, through the use of a simple graphical user interface. In this way the healthcare professional with a few clicks can request the services which he/she is interested in.

- **the read module**: this module helps healthcare professionals to read clinical documents. It is used in a transparent way, when the healthcare professional wants to read the clinical document. The module enables the reading of the codes associated with the diagnosis and/or observation contained in the clinical document. In particular, the module takes the codes entered by the healthcare professional, directly from the *User Interface Module*. Subsequently, it sends several requests to the *Terminology Server Interface Module* for code searching. The answers obtained contain the textual description of the concepts associated with the codes. The latter are sent to the *User Interface Module* for the presentation of information to the user. The read module uses a concept look-up function, offered by a terminology service compliant with the FHIR standard.

- **the write module**: this module is responsible for the construction of a new clinical document or the enrichment of an existing one with new observations. It includes two sub-features. The first invokes the *Terminology server Interface Module* to obtain the textual descriptions and codes associated with the concepts stored on the termino-
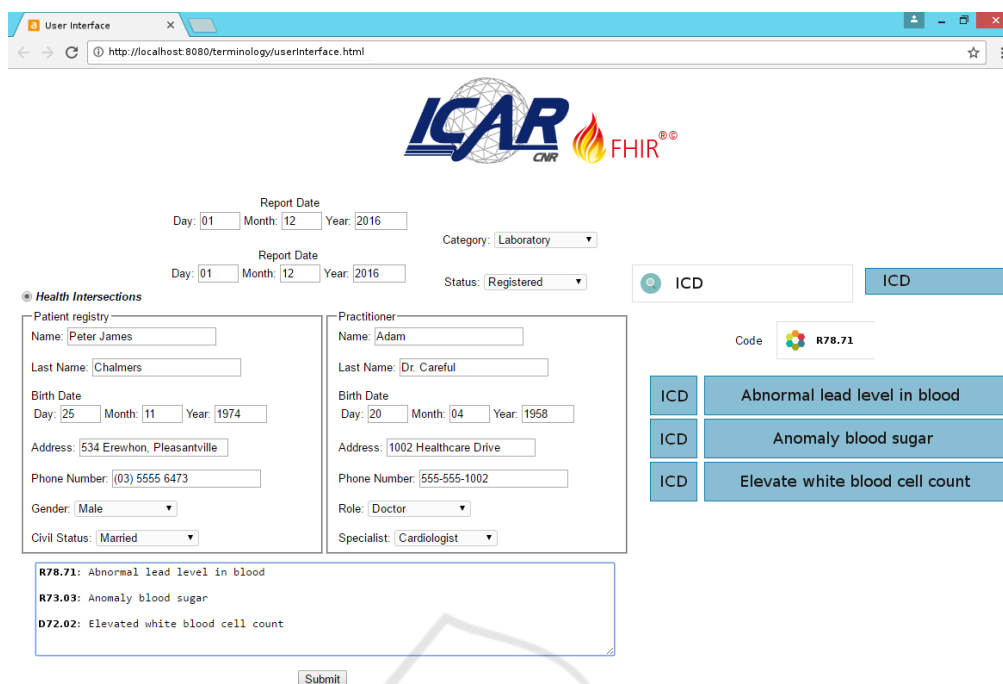
139

Figure 4: User interface form.

logy server. These data are then used to pre-fill the form presented to the user, via the *User Interface Module*. The textual descriptions are associated with the codes belonging to the coding systems (e.g., LOINC), and refer to observations, diagnoses, clinical reports, status, etc. The second feature is related to the structured and coded representation of the information entered and/or selected by the healthcare professional. Specifically, it refers to the document representation according to the FHIR standards and the mapping shown above. The created document will contain standardized and coded data, structured in XML or JSON format. The write module uses the concept look-up and expansion functions, offered by a terminology service compliant to the FHIR standard.

- **the interface terminology server module**: this is the main module. It is responsible for interfacing with the terminology server. In particular, it is used whenever the other modules need to interact with a terminology server. The communication is performed by sending a request and waiting for a reply. Requests are made using the RESTful API, while the received responses are in XML or JSON format. Requests and responses travel on the Internet network.

The final clinical document contains the following information: demographic information about a patient;

demographic and professional information about a healthcare professional; and "atomic" observations, narrative text and coded interpretations, written after the examination. The next subsection explains the technical details of the solutions chosen for the architecture implementation. Figure 4 shows the developed user interface that allows to the healthcare professionals, through interaction with defined architecture, to search descriptions of the codes by search form, and obtain the standard system code, description and terminology used.

2.4 Technical Details

The modules have been developed using the open source web framework "AngularJS", using Javascript scripting language. The graphical user interface, was made with HTML5 and CSS. Server requests are made by the AJAX (Asynchronous JavaScript and XML) method invocation, written according to the REST FHIR syntax, and using the HTTP protocol over the Internet. The server side used a public FHIR terminology server (defined for testing purposes). In particular, the server used was the Health Level Seven International FHIR server (GRA, 2016). This test server is based on FHIR version 1.5.0., and it is available in DSTU2 or DSTU3. The server has been installed locally, and uses the endpoint *http:localhost:960/open/*. A FHIR supported terminology systems list, used for the association between
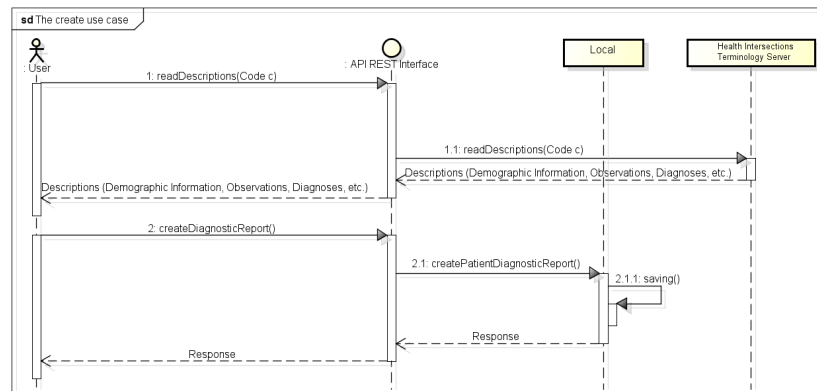
Figure 5: Sequence diagram to construct a structured clinical document.

modeled concepts and codes, can be found here (LST, 2016). The technologies used are very light from the computational point of view. In fact, they are simple web interfaces and AJAX requests, which contain the syntax of the REST APIs. Therefore, the proposed system is particularly suitable for mobile devices. In fact, this solution is much lighter compared to solutions, such as the SOAP technology. In this way, this system can be executed on terminals with few resources and a low performance.

3 USE CASE

The proposal has the purpose of structuring and encoding information in clinical documents. In this section two use cases of interest are examined. In particular, the "read use case", for the reading of a specific clinical document, and the "create use case", for the creation of a new clinical document, are considered.

3.1 The Read Use Case

This use case assumes that the healthcare professional has to read a clinical document. In particular, the patient provides a laboratory report and wants to have a consultation. The healthcare professional uses the system to read the document. The system permits a fast retrieval of information by decoding the health standard codes. The healthcare professional indicates the codes and the information contained in the document, and the system automatically sends the request to the terminology server. Finally, the responses from the server are compared and integrated to prepare the response for the user. In this way, the healthcare professional is able to formulate a more complete and fast diagnosis. For example, if in the document there is the ICD code *R78.71*, this means that the patient

presents an *abnormal lead level in blood*, as shown in figure 4. These operations are performed with a few clicks, through a web form provided by the system. The user enters the codes using the appropriate form, following which the system sends the requests to the terminology server. Finally, the information is presented to the user.

3.2 The Create Use Case

This use case assumes that the healtcare professional who is treating a patient, has to create a new clinical document. This document should contain demographic data of both the patient and the healthcare professional, and also information about possible diagnoses and observations relating to the patient. These data are contained within the pre-filled fields, and can be selected by the user through the corresponding web form. In particular, these descriptions are associated with the codes belonging to the coding systems (e.g. LOINC), and refer to observations, diagnoses, clinical reports and status, etc. Once the document is complete, it is saved by the user. The result is a clinical document, structured and standardized in a manner compliant with FHIR standards. The corresponding use case diagram is shown in figure 5, while an example of the system output is presented in figure 2. In particular, the first part is related to pre-compiled data populating, achieved by interrogating the terminology server. The second part refers to the creation of the document, representing the concepts through the FHIR resources listed above, and storing locally the final structured document.

4 CONCLUSIONS AND FUTURE WORK

In this paper we have proposed an architecture allowing healthcare professionals read and create clinical documents. In particular, the healthcare professional is able to produce structured, standardized and coded information relating to a patient. This architecture has the advantage of using medical standards, both encodings and concepts representations, and it permits to define user-friendly interface simply. The interface allows the healthcare professional to quickly translate a code into a description and back again. In particular, this system allows the healthcare professional to produce clinical documents that can be machine readable. This is its main advantage and is only a step, suggesting several possible future developments. One possibility may involve a complex or particular treatment situations. In fact, we only focus on a laboratory reports representation, while a complex case, might also refer to other information and other concepts, such as radiological or generic reports, etc. A second possibility concerns the integration of the developed system in other existing architectures, achieving clinical interoperability, which is the ability for two or more clinicians in different care teams to exchange patient data. For example, interfacing the proposed system with EHR systems that are based on different platforms, in order to retrieve and update clinical documents directly from an EHR.

ACKNOWLEDGMENTS

REFERENCES

(2016). Fhir resource list. http://www.hl7.org/implement/standards/fhir/resourcelist.html. Accessed: 2016-07-14.

(2016). Fhir terminologies systems. http://www.hl7.org/fhir/terminologies-systems.html. Accessed: 2016-07-14.

(2016). Health level seven international, welcome to the fhir server (2015). http://fhir2.healthintersections.com.au/. Accessed: 2016-06-15.

Bender, D. and Sartipi, K. (2013). Hl7 fhir: An agile and restful approach to healthcare information exchange. *IEEE 26th International Symposium*, pages 326–331.

Esposito, C., Ciampi, M., De Pietro, G., and Donzelli, P. (2012). Notifying medical data in health information systems. In *Proceedings of the 6th ACM International Conference on Distributed Event-Based Systems*, pages 373–374. ACM.

Garde, S., Knaup, P., Hovenga, E. J. S., and Heard, S. (2007). Towards semantic interoperability for electronic health records. *Methods of Information in Medicine*, 46:332–343.

Haux, R. (2006). Health information systems - past, present, future. *International Journal of Medical Informatics*, 75:268–281.

Hogarth, M. A., Stone, J., and Gorin, F. A. (2003). jterm: A server for terminological systems. *AMIA Annual Symposium Proceedings*, 2003:861.

Honko, H., Andalibi, V., Aaltonen, T., Parak, J., Saaranen, M., Viik, J., and Korhonen, I. (2015). W2e - wellness warehouse engine for semantic interoperability of consumer health data. *IEEE Journal of Biomedical and Health Informatics*.

IEEE (1990). Institute of electrical and electronics engineers standard computer dictionary: A compilation of ieee standard computer glossaries. *IEEE*, New York.

Lamprinakos, G. C., Mousas, A. S., Kapsalis, A. P., Kaklamani, D. I., Venieris, I. S., Boufis, A. D., Karmiris, P. D., and Mantzouratos, S. G. (2014). Using fhir to develop a healthcare mobile application. In *Wireless Mobile Communication and Healthcare (Mobihealth), 2014 EAI 4th International Conference on*, pages 132–135. IEEE.

Luz, M. P., Cavalini, L. T., de Matos Nogueira, J. R., and Cook, T. W. (2015). Providing full semantic interoperability for the fast healtchare interoperability resources schemas with resource description framework. *International Conference on Healthcare Informatics*, pages 463–466.

Minutolo, A., Esposito, M., and De Pietro, G. (2011). A mobile reasoning system for supporting the monitoring of chronic diseases. In *International Conference on Wireless Mobile Communication and Healthcare*, pages 225–232. Springer.

Navas, H., Osornio, A. L., Baum, A., Gomez, A., Luna, D., and de Quiros, F. G. B. (2007). Creation and evaluation of a terminology server for the interactive coding of discharge summaries. *Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems*, 129:650–654.

Sànchez-Caro, A., Soguero-Ruiz, C., Mora-Jimènez, I., Lechuga, L., Ramos-Lòpez, J., Garcìa-Alberola, A., Serrano-Balazote, P., and Rojo-Àlvarez, J. L. (2014). Towards semantic interoperability for cardiovascular risk stratification into the electronic health records using archetypes and snomed-ct. *Computing in Cardiology Conference (CinC)*, 41:497–500.

Walker, J., Pan, E., Johnston, D., Adler-Milstein, J., Bates, D. W., and Middleton, B. (2005). The value of health care information exchange and interoperability. *Health Affairs published online January 19*.