



*Consiglio Nazionale delle Ricerche  
Istituto di Calcolo e Reti ad Alte Prestazioni*

Interoperabilità in sistemi informativi  
medicali eterogenei  
attraverso l'uso di ontologie

G. De Pietro – M. Esposito

**RT-ICAR-NA-06-07**

**03-2006**



Consiglio Nazionale delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR)  
– Sede di Napoli, Via P. Castellino 111, I-80131 Napoli, Tel: +39-0816139508, Fax: +39-  
0816139531, e-mail: [napoli@icar.cnr.it](mailto:napoli@icar.cnr.it), URL: [www.na.icar.cnr.it](http://www.na.icar.cnr.it)



*Consiglio Nazionale delle Ricerche  
Istituto di Calcolo e Reti ad Alte Prestazioni*

## Interoperabilità in sistemi informativi medicali eterogenei attraverso l'uso di ontologie

G.. De Pietro<sup>1</sup> – M. Esposito<sup>1</sup>

***Rapporto Tecnico N.:***  
**RT-ICAR-NA-06-07**

***Data:***  
**03-2006**

---

<sup>1</sup> Istituto di Calcolo e Reti ad Alte Prestazioni, ICAR-CNR, Sede di Napoli, Via P. Castellino 111, 80131 Napoli

*I rapporti tecnici dell'ICAR-CNR sono pubblicati dall'Istituto di Calcolo e Reti ad Alte Prestazioni del Consiglio Nazionale delle Ricerche. Tali rapporti, approntati sotto l'esclusiva responsabilità scientifica degli autori, descrivono attività di ricerca del personale e dei collaboratori dell'ICAR, in alcuni casi in un formato preliminare prima della pubblicazione definitiva in altra sede.*

# Interoperabilità in sistemi informativi medicali eterogenei attraverso l'uso di ontologie

G. De Pietro<sup>1</sup> and M. Esposito<sup>1</sup>

<sup>1</sup>ICAR-CNR, Via Castellino 111, 80131 Napoli, Italy  
{giuseppe.depietro, massimo.esposito@na.icar.cnr.it}

**Abstract:** Le moderne infrastrutture informative presenti nelle realtà medicali consistono di molti sistemi eterogenei con molteplici e differenti meccanismi per la gestione dei dati sottostanti. Le informazioni di un singolo paziente possono essere distribuite attraverso molti di questi sistemi. Appare palese, quindi, il bisogno di accedere ai dati dei pazienti da locazioni diverse, il che richiede un'architettura necessariamente e fortemente basata sull'interoperabilità. Lo scopo di questo technical report è descrivere un prototipo architetturale, realizzato nell'ambito del Progetto Telemedicina, basato su nodi di frontiera che accedono alle informazioni disponibili nei vari sistemi sanitari e che adoperano un modello basato su ontologie come strumento per la loro rappresentazione per garantire l'interoperabilità semantica.

## Introduzione

Le moderne organizzazioni sanitarie si compongono di numerosi sotto-sistemi autonomi, con molteplici specializzazioni, sedi fisiche distribuite e principi di base spesso divergenti fra loro. La gestione delle informazioni relative ai pazienti attraverso tali sistemi risulta essere un processo molto complesso, dovuto sia a differenze di tipo tecnologico, sia ad una diversa rappresentazione dei dati dei pazienti, sia alla parcellizzazione dei dati di un singolo paziente in più strutture sanitarie.

Il *Progetto Telemedicina* fornisce una possibile soluzione a questo problema attraverso la definizione e la realizzazione di un portale come strumento in grado di effettuare l'interscambio di dati di interesse medico e l'integrazione di archivi provenienti da differenti Enti locati su territorio regionale campano per creare un sistema unificato di consultazione.

Il servizio di consultazione remota di studi radiologici è una delle applicazioni pilota utilizzate per sperimentare la reale ed immediata applicabilità dei risultati previsti con il Progetto Telemedicina. I principali soggetti partecipanti alla sperimentazione, ossia il Secondo Policlinico della Università Federico II, il Primo Policlinico della Seconda Università di Napoli, l'Istituto Pascale e l'Azienda Ospedaliera Monaldi, hanno attivato i propri sistemi informativi *PACS* per l'archiviazione e catalogazione di immagini radiologiche e *RIS* per la gestione informatizzata di tutte le attività tipiche di un servizio di radiologia.

In questo lavoro è proposto e descritto il modello architetturale ideato per la realizzazione di tale applicazione pilota e sono evidenziate le scelte progettuali che ne hanno determinato l'infrastruttura, fondamentalmente mirate a realizzare tre livelli di interoperabilità per garantire l'interscambio di informazioni a cui far aderire i vari Enti partecipanti alla sperimentazione, ponendo in particolare l'accento sull'approccio ontologico utilizzato per garantire l'interoperabilità semantica dei dati..

## **Modello architetturale**

Il modello architetturale proposto dal Progetto Telemedicina prevede un nodo centrale rappresentato da un portale che fruisce servizi di telemedicina, tra cui figura anche l'applicazione pilota di consultazione remota di immagini radiologiche, ed una serie di nodi periferici di frontiera che si interfacciano con i molteplici domini sanitari.

L'applicazione pilota di consultazione di immagini diagnostiche e referti, di cui ci si occupa in questo lavoro, ha come scopo quello di consentire ad un utente, in base a determinate politiche di autenticazione e autorizzazione, di consultare, attraverso il portale di accesso ai servizi, le informazioni relative ad un esame o studio effettuato da un paziente presso un determinato dipartimento radiologico di un dominio sanitario.

I nodi periferici di frontiera, essendo relativi alla sola applicazione in questione, hanno, quindi, il compito di relazionarsi non con l'intera realtà medica ma solo con i reparti radiologici, che rappresentano le realtà sanitarie che si occupano delle problematiche relative alla diagnostica per immagini.

L'informatizzazione dei dipartimenti radiologici ha seguito gli orientamenti proposti dall'IHE (Integrating the Healthcare Enterprise) [1] ed in particolare è stato adottato un modello architetturale di riferimento per i sistemi informativi sanitari che prevede che PACS e RIS rappresentino i due sottosistemi che consentono di gestire il dipartimento di radiologia in maniera completa ed adeguata.

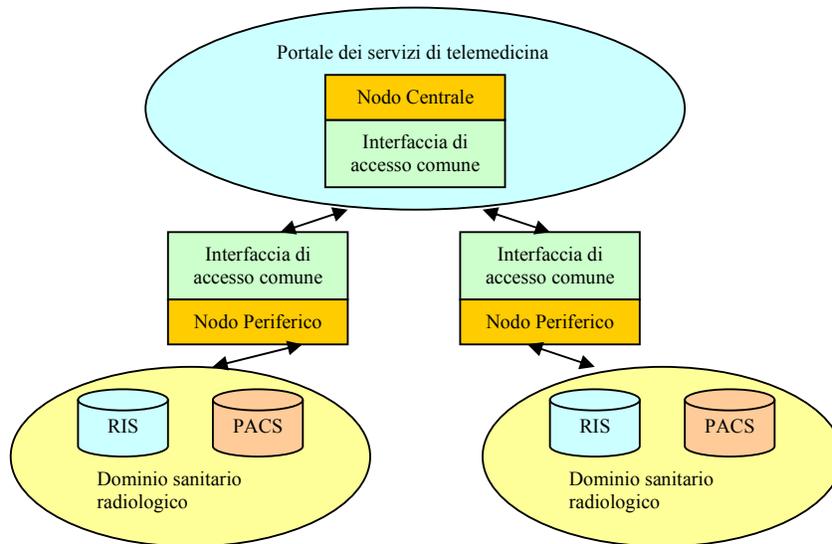
Il RIS (Radiology Information System) è il sistema informativo centrale del dipartimento ed ha lo scopo di automatizzare e gestire i compiti clinico-amministrativi al suo interno. Il PACS (Picture Archive and Communication System) è invece il sistema che si occupa della gestione nel suo complesso delle immagini radiologiche.

L'utilizzo di standard come DICOM 3 e HL7, da adottare rispettivamente da PACS e RIS, completano l'orientamento proposto dall'IHE.

DICOM 3 [2] regola tutte le problematiche connesse al trattamento di immagini a partire dalla loro rappresentazione fino ad arrivare ai protocolli necessari allo scambio delle stesse. HL7 [3] cerca, invece, di risolvere le problematiche relative allo scambio di informazioni tra sistemi informativi all'interno di domini sanitari formalizzando la formattazione dei messaggi utilizzati.

Tuttavia, sebbene siano degli standard, essi lasciano molti margini di libertà che hanno portato a differenti loro implementazioni da parte delle diverse case produttrici di RIS e PACS. Tali implementazioni presentano sfumature che spesso compromettono

l'interoperabilità della comunicazione tra diverse strutture radiologiche sia per quanto riguarda i protocolli di comunicazione che per quanto concerne il formato dei dati scambiati.



**Figura 1: Modello architetturale**

In virtù di queste problematiche, per favorire l'interscambio e la comunicazione tra nodo centrale e domini sanitari radiologici, e quindi, in altri termini, tra portale e i sotto-sistemi informativi radiologici, esautorando il portale dall'incombenza di doversi adattare alle diverse realtà, sono stati inseriti i nodi periferici ai quali è stato assegnato il compito di dover garantire almeno tre livelli di interoperabilità:

- interoperabilità del formato dei file contenenti immagini e referti: essa viene fornita, nel caso in cui gli standard previsti risultassero incompatibili, attraverso la conversione dei file in formati di rappresentazione di uso diffuso, come JPEG per le immagini e RFT per i referti.
- interoperabilità del protocollo di comunicazione e delle interfacce di accesso: i nodi periferici sono stati implementati adoperando la tecnologia dei Web Services [4], il che permette al nodo centrale di poter comunicare con essi attraverso protocolli standard quali SOAP [5] su HTTP, evitando così possibili problemi di sicurezza connessi alla presenza di firewall e disinteressandosi completamente della complessità e della eventuale incompatibilità tra protocolli medicali. Alle funzionalità offerte al nodo centrale dalle interfacce di accesso dei nodi periferici è stato previsto di potervi accedere adoperando procedure comuni

che utilizzino protocolli e tecnologie standard adottate per la programmazione di applicazioni web (HTTP, HTML, XML, Java).

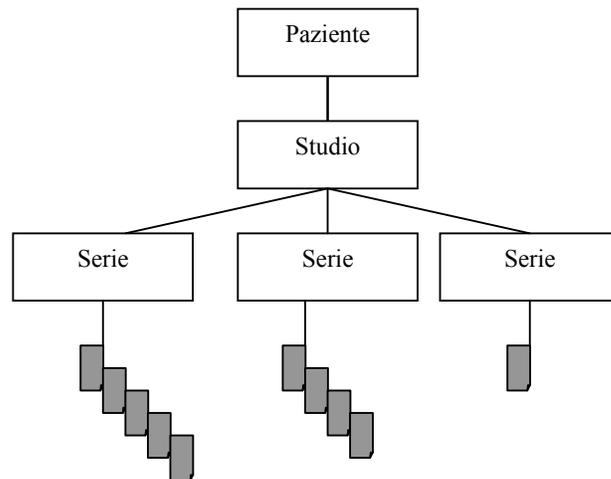
- interoperabilità dei dati sanitari scambiati relativi ai pazienti: di questo se ne parla più in dettaglio nei paragrafi seguenti.

### **Modello dei dati sanitari relativi a studi radiologici di un paziente**

Il modello dei dati sanitari relativi ad uno studio radiologico di un paziente è organizzato gerarchicamente:

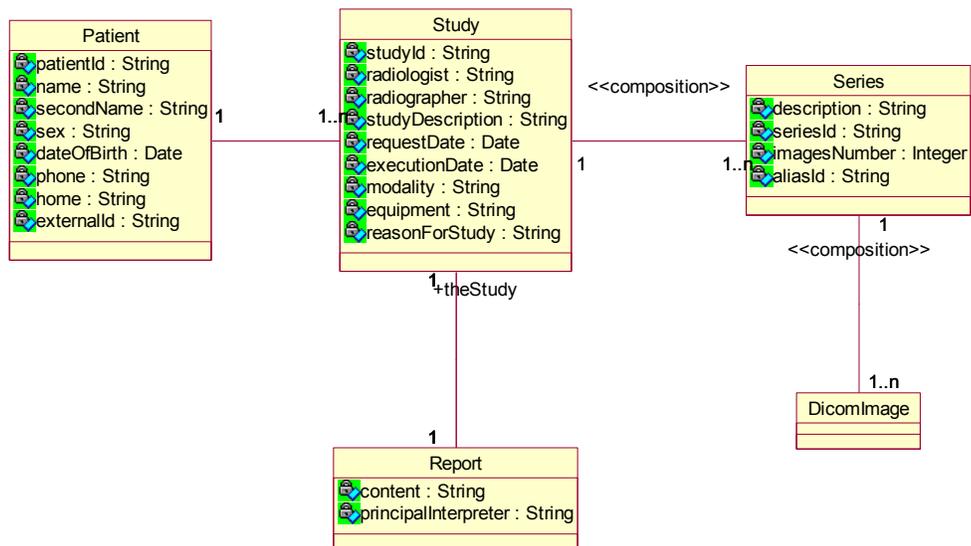
- al livello più alto ci sono le informazioni relative al *paziente*: riguardano i dati demografici per l'identificazione del paziente a cui si riferisce uno studio radiologico;
- successivamente ci sono le informazioni relative agli *studi* effettuati dal paziente e i relativi *referti*: tali informazioni provengono da un insieme di sorgenti, quali ad esempio il RIS che identifica lo studio, le modalità effettuanti lo studio che aggiungono informazioni relative al tempo in cui la prestazione viene effettuata, i referti che vengono inseriti a posteriori attraverso stazioni di refertazione.
- ci sono, poi, le informazioni sulle serie di immagini: sono le prime completamente generate dalla modalità effettuante l'esame e riguardano il modo in cui una serie viene generata, la persona coinvolta, la parte del corpo esaminata, l'identificazione della serie stessa, il nome ed il tipo di modalità effettuante l'esame, il posizionamento delle immagini e così via.
- infine ci sono le informazioni relative alle *immagini*: esse descrivono i dettagli relativi all'immagine in relazione ai pixel che la compongono, la dimensione, i colori, l'ordine con cui i dati delle immagini sono ordinati in diversi piani e così via.

Nella figura 2 è mostrato il mapping fra uno studio ed il modello di dati mostrato sopra: un paziente esegue uno studio che può essere eseguito su più modalità. Ogni modalità può produrre una o più serie, ciascuna costituita da una o più immagini.



**Figura 2 Modello dei dati sanitari di un paziente**

Ognuno di questi concetti, ossia paziente, studio e referto, serie, immagine, nella modellazione effettuata in questo lavoro, che in parte attinge dal modello di dati fornito da HL7 ed in parte da quello fornito da DICOM 3, possiede un insieme di proprietà che sono state accennate sopra ma che vengono definite più in dettaglio in figura 3, formalizzandole attraverso una rappresentazione UML in cui i concetti sopraccitati rappresentano le classi e le proprietà ad essi relativi gli attributi.



**Figura 3: Rappresentazione UML del modello dei dati sanitari**

## Interoperabilità semantica mediante l'uso di ontologie

L'insieme di concetti e proprietà ad essi relativi sopra mostrati rappresenta il contenuto informativo che nodo centrale e nodi periferici si scambiano fra loro. Tali concettualizzazioni tuttavia dovrebbero essere definite a livello di conoscenza senza alcuna dipendenza da un particolare linguaggio di codifica. Questo perché gli applicativi che sono in esecuzione sul nodo centrale e su quelli periferici potrebbero essere implementati facendo riferimento a sistemi di rappresentazione diversi.

Per ovviare a questo problema basterebbe rappresentare tali informazioni adoperando un linguaggio come XML [6] che è in grado di garantire l'interoperabilità sintattica dei dati, intesa come la facilità da parte degli applicativi di leggere dati e ottenerne una rappresentazione utilizzabile indipendentemente dal linguaggio di codifica in cui sono realizzati e riconosciuta da tutti in maniera univoca.

In questo lavoro si è cercato di ottenere un livello di interoperabilità dei dati ancora più spinto mirando ad ottenere l'interoperabilità a livello semantico ossia ad ottenere chiarezza, formalità e organizzazione dei dati, collegando le informazioni relative ai dati sanitari a concetti modellati gerarchicamente in ontologie, a loro volta descritte in meta-documenti, consentendo ai vari applicativi di poterne cogliere il contenuto semantico.

Tutto ciò è stato realizzato modellando un'ontologia, di cui è fornito un frammento nella figura 4, in linguaggio OWL-Lite [7], condivisa da nodo centrale e nodi periferici e da essi accettata e riconosciuta che rappresenta attraverso classi e proprietà le classi, gli attributi e le relazioni tra classi viste nella rappresentazione UML presentata nel paragrafo precedente.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf:RDF xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
  <owl:Ontology rdf:about="Ontologia dei dati sanitari di un paziente relativi a studi radiologici" />
  <owl:Class rdf:about="http://telemedicina#Patient">
    <owl:equivalentClass>
      <owl:Class>
        <owl:intersectionOf rdf:parseType="Collection">
          <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#name"/>
            <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
          </owl:Restriction>
          <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#secondName"/>
            <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
          </owl:Restriction>
          <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#dateOfBirth"/>
            <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
          </owl:Restriction>
          <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#externalId"/>
            <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
          </owl:Restriction>
          <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#sex"/>

```

```

        <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
    </owl:Restriction>
    <owl:Restriction>
        <owl:onProperty rdf:resource="http://telemedicina#patientId"/>
        <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger">1</owl:cardinality>
    </owl:Restriction>
    </owl:intersectionOf>
</owl:Class>
</owl:equivalentClass>
</owl:Class>
.....
.....
<owl:DatatypeProperty rdf:about="http://telemedicina#name">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:about="http://telemedicina#studyDescription">
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
<owl:ObjectProperty rdf:about="http://telemedicina#hasStudy"/>
.....
.....
<owl:Class rdf:about="http://telemedicina#Patient">
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:onProperty rdf:resource="http://telemedicina#hasStudy"/>
            <owl:someValuesFrom>
                <owl:Class rdf:about="http://telemedicina#Study"/>
            </owl:someValuesFrom>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
</rdf:RDF>

```

**Figura 4:Frammento dell'ontologia condivisa descrivente il modello dei dati di un paziente**

Il nodo centrale gestisce l'ontologia definita attraverso un componente locale, chiamato ontology server, che si occupa di offrire funzionalità per caricare e verificare la coerenza di ontologie e memorizzarne i concetti in essa espressi, il tutto utilizzando una Knowledge Base. Tra i vari motori, che implementano meccanismi logici e forniscono una KB è stato scelto FaCT (Fast Classification of Terminologies) [8], un classificatore di concetti open source molto popolare ed efficiente.

Ogni volta che i nodi periferici vogliono comunicare col nodo centrale per scambiarsi informazioni sanitarie, modellano tali informazioni come istanze dei concetti generali definiti nell'ontologia condivisa e li formalizzano in un'altra ontologia. Tale ontologia viene spedita come documento owl al nodo centrale che a sua volta la invia all'ontology server. Quest'ultimo verifica la correttezza sintattica del documento, lo carica poi nella KB da lui gestita e verifica la coerenza del contenuto in esso espresso con l'ontologia generale.

Attraverso questa fase di validazione è possibile quindi evitare, ad esempio, che vi siano pazienti senza nome, più referti diversi associati ad uno stesso studio (si parte dall'ipotesi che ve ne sia uno solo per studio), che siano associate serie di immagini a pazienti senza che esse si relazionino ad uno studio e così via.

Dopo questo processo il nodo centrale entra in possesso delle informazioni richieste, avendone quindi non solo riconosciuto il formato sintattico ma anche verificato e compreso la semantica. Quindi anche se i dati sono rappresentati in maniera diversa presso i vari domini sanitari radiologici e se ad esempio dati di uno stesso

paziente sono distribuiti presso più domini, attraverso la loro formalizzazione ontologica operata dai nodi periferici è possibile ricostruirne l'insieme senza ambiguità e senza incoerenze e mancanze.

### **Conclusioni**

La scelta di integrare tecnologie del web semantico all'interno di realtà medicali rappresenta una possibile soluzione per risolvere i problemi di interoperabilità tra le entità che ne fanno parte.

L'utilizzo, quindi, di ontologie come vocabolari condivisi che descrivono l'insieme di informazioni medicali relative ai dipartimenti radiologici e di Owl, e nello specifico della sua versione Lite, come linguaggio capace di fornire un formalismo universalmente interpretabile e semanticamente non ambiguo rappresentano gli aspetti chiave dell'approccio presentato.

Il potere espressivo di cui è fornito OWL consente, inoltre, di utilizzare le ontologie non solo come strumento di conoscenza statico ma di poter effettuare inferenze sul contenuto informativo che esse offrono risolvendo così possibili incoerenze logiche.

Altri aspetti come la sicurezza e la privacy, che assumono una maggiore rilevanza in sistemi sanitari dove le informazioni scambiate rappresentano l'elemento chiave, potranno essere gestiti con ontologie, anche se nei limiti delle potenzialità dei motori inferenziali esistenti e delle funzionalità offerte dai linguaggi ontologici in relazione a politiche di sicurezza e alla possibilità di limitare o meno la visibilità di parte del contenuto informativo scambiato.

## Bibliografia

- [1] “*IHE Integrating the Healthcare Enterprise*”, <http://www.ihe.net/>
- [2] “*DICOM 3 Digital Imaging and Communication in Medicine*”,  
<http://medical.nema.org/>
- [3] “*HL7 Health Level Seven*”, <http://medical.nema.org/>
- [4] W3C, “*Web Services*”, <http://www.w3.org/2002/ws/>
- [5] W3C, “*SOAP specifications*”, <http://www.w3.org/TR/soap/>
- [6] W3C, “*Extensible Markup Language XML*”, <http://www.w3.org/XML/>
- [7] W3C, “*OWL Web Ontology Language*”, <http://www.w3.org/TR/owl-features/>
- [8] Horrocks, “*The FaCT System*”, <http://www.cs.man.ac.uk/~horrocks/FaCT/>